Task Name: Mesoscale Carbon Data Assimilation for NACP

PI Name: A. Scott Denning

Award Number: NNG05GD15G

RTOP Number: 622-94-51-01

Student Record:

Two M.S. level Graduate Students: Erica McGrath-Spangler and Nick Parazoo

Two Ph.D. level Graduate Students: Kathy Corbin and Andrew Schuh

One B.S. level student: John Heizer

Task Abstract: The North American Carbon Program (NACP) is a multi-year program of integrated research supported by many US agencies which seeks to quantify the current budget of CO_2, CO, and CH_4 over North America, to understand and predict the processes governing these fluxes, and to provide timely and practical information products to support management decisions. A major component of NACP is a greatly enhanced system for observing temporal and spatial variations for carbon gases in the atmosphere over North America and adjacent coastal oceans. After 2007, the dense in-situ network of atmospheric measurements for NACP will be augmented by hundreds of thousands of column CO_2 observations each day made from NASA's Orbiting Carbon Observatory (OCO). These new observations are intended both to provide an integral atmospheric constraint to upscaled ("bottom-up") models of carbon exchange processes, and to enable quantitative but process-agnostic estimates of regional monthly sources and sinks by ("top-down") transport inversion. Currently available analytical methods for flux estimation by inverse modeling involve assumptions about the spatial and temporal patterns of carbon fluxes that will be inappropriate to the much greater density of sampling by NACP and OCO.

We have developed a generalized framework for flux estimation from multiple streams of carbon observations. These include spectral vegetation and land cover imagery, eddy covariance flux observations, meteorological data, and both in-situ and remotely sensed observations of atmospheric carbon gases. This has been accomplished using Ensemble Data Assimilation (EnsDA) techniques applied to a fully coupled model of regional meteorology, ecosystem carbon fluxes, and biomass burning (SiB-CASA-RAMS). The coupled model simulates terrestrial carbon fluxes over North America due to photosynthesis, autotrophic respiration, decomposition, fires, and a "residual" time-mean source or sink. Unknown parameters related to light response, allocation, drought stress, phonological triggers, combustion efficiency, PBL entrainment, convective efficiency, and the time-mean sink will be estimated to obtain optimum consistency with a wide variety of ecological, meteorological, and trace gas observations.

The EnsDA method does not require the development of an adjoint of the coupled model, but rather applies an optimization method that involves a large ensemble of forward simulations. Unlike previous high-resolution inversions using transport model adjoint methods, we will not assume surface fluxes remain constant on monthly time scales, and we will not treat the transport model as "perfect." Parameters in the forward coupled model will be quantitatively estimated, as will transport model error. The model will be integrated on a 20-km grid over a domain including most of North America and adjacent oceans, with lateral boundary conditions specified from the output of a global model.

When this project ended in 2007, it was not renewed. Though instrumental in setting and promoting the overall agenda for NACP, and helping NASA's Earth Science program "fit in" to the larger interagency context, the PI is no longer a NASA NACP investigator. This PI is frankly disillusioned, and wonders whether this work is relevant to a NASA science program increasingly devoted to "SPAM in a can." We continue to contribute to NACP through NOAA, DOE, and NSF.

On Feb 24, 2009, OCO failed to reach orbit. How very sad for NASA, and especially for the US taxpayer. Please accept our sincere condolences.

Detailed Narrative:

1. Simple Biosphere Model Development

In our development of the Simple Biosphere Model (SiB), we aim to continually improve our ability to simulate fluxes of mass (including trace gases), energy and momentum between the atmosphere and terrestrial biosphere on short (diurnal) and long (annual) timescales. The goal is to do so while explicitly representing actual biophysical and biogeochemical mechanisms as opposed to simplified empirical relationships.

Explicit representation of sunlit and shaded canopy fraction for radiative transfer and photosynthesis calculation: For years, SiB has used what has been incorrectly called a 'big leaf' configuration for determining fluxes. In this configuration, flux is calculated for a single square meter of top leaf and then scaled to the entire canopy using logarithmic relationships between nitrogen content and canopy depth as described in Sellers et al (1992, 1996a). We prefer the moniker 'Continuous Canopy'. This provides an elegant solution to the leaf-to-canopy scaling problem inherent to landsurface modeling, in part due to its ability to incorporate remotely-sensed data (fractional absorption of Photosynthetically Active Radiation-fPAR) into the scaling formulation. However, the Continuous Canopy method is not without problems. It uses the daily-mean solar zenith angle in its absorption coefficients. Combined with upscaling from single leaf flux behavior, this results in a canopy light response that is unrealistically 'square'. In other words, photosynthesis in this model has a more binary response than is commonly observed.

Furthermore, separation of the photosynthesis upscaling from the canopy energetics scheme reveals a lack of internal consistency that can affect Bowen ratio in situations where the canopy is dense and non water-limited. We've adopted a scheme that explicitly resolves sunlit and shaded canopy fraction, following those of de Pury and Farquhar (1997), Wang and Leuning (1998) and Dai, Dickinson and Wang (2004). The biophysical characteristics of sunlit and shaded fractions are calculated using the illumination partitioning and canopy resource allocation based on observations (Aranda et al, 2004; Pandey and Kushawaha, 2005; Pons and Anten, 2004; Bond et al, 1999; Kull, 2002; Le Roux et al, 1999; Krujit et al, 2002). The result is a self-consistent equation set for calculating canopy energetics and photosynthesis that provides not only a more biophysically 'real' representation of canopy processes, but also produces a better match to observed fluxes. Furthermore, an explicit treatment of sunlit and canopy fraction provides a means to model canopy response to changing partitioning of light between beam and diffuse components such as has been described in Gu et al (1999, 2002, 2003).

Frost stress: Coniferous forests are known to require a recovery interval for photosynthesis following cold temperatures (Strand and Lundemark, 1995; Vogg et al, 1998). We've incorporated a photosynthesis limitation term in SiB that reproduces this feature. The effect is minimal at all but the highest latitudes, be we note a delay in the onset of photosynthesis in springtime in many Boreal forests, more in line with observed flux.

Soil/snow formulation: Following the Community Land Model (Dai et al, 2003), IAP94 (Dai and Qingcun, 1997) and SNTHERM (Jordan, 1991), we've modified the SiB soil and snow submodel to explicitly resolve liquid and ice water fraction in the soil and allow up to 5 snow models with individual density and liquid/ice fractions. The snow modifications provide remarkable simulation of snow accumulation and ablation rates, as well as capturing the inversion commonly seen with snowpack.

NDVI interpolation: The compositing of NDVI for SiB phenology has known problems (described previously). We have developed the ability to use shorter time period observations (15- or 8-day NDVI maps, as opposed to monthly) to help resolve leaf-out and die-off (Philpott et al, in preparation). Additionally, the shape of the curve can be used to help determine temporal siting of NDVI observations.

Autotrophic respiration: SiB ecosystem metabolism as described in Denning et al (1996a) links respiration to soil temperature and moisture, yet neglects autotrophic respiration associated with leaf-out. By partitioning total photosynthate into autotrophic and heterotrophic pools we are able to assess a carbon 'penalty' onto the leaf-out process, as well as more realistically representing the costs of canopy maintenance. This procedure does not effect energy and moisture fluxes, and the change to diurnal Net Ecosystem Exchange (NEE) is negligible. However, the annual cycle of NEE is favorably modified when compared to observations.

2. SiBCASA Model Development

We finished development and testing of the Simple Biosphere/Carnegie-Ames-Stanford Approach (SiBCASA) terrestrial carbon cycle model. SiBCASA combines the photosynthesis and biophysical calculations in SiB3 with the biogeochemistry from CASA to create a hybrid capable of estimating terrestrial carbon fluxes and biomass from subdiurnal to decadal time scales (Fig 1).

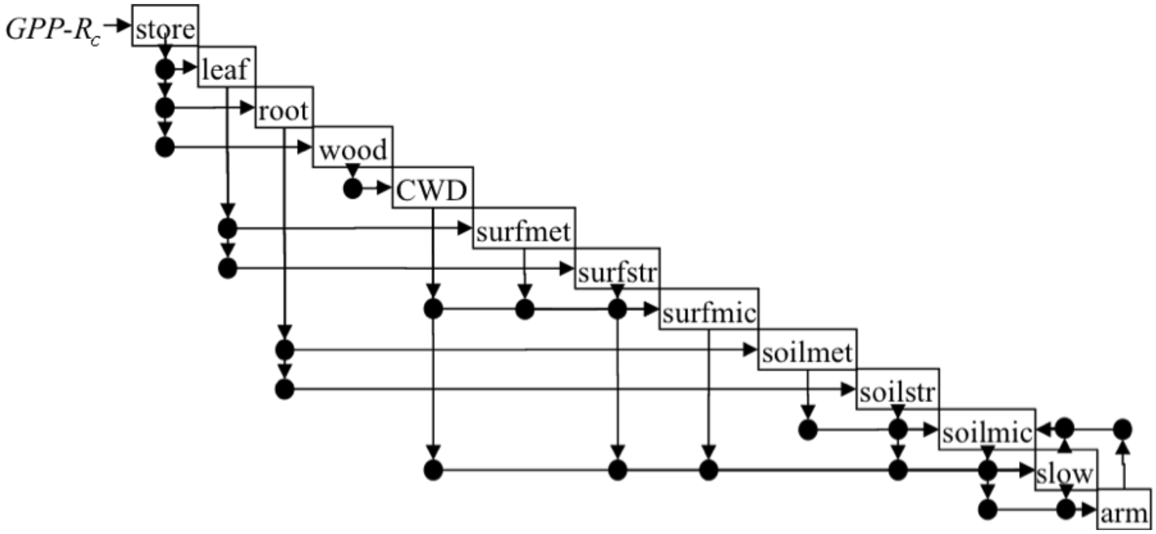

Figure 1: The SibCasa pool configuration. Dots are transfers between pools; the vertical arrows are losses from each pool; and horizontal arrows are inputs to each pool. The carbon generally flows from upper left to lower right. The primary input is canopy net assimilation (*GPP* – R_c, where *GPP* is Gross primary Productivity and R_c is canopy autotrophic respiration).

Adding CASA biogeochemistry allows SiBCASA to estimate biomass as a function of stand age (Fig 2), and thus long-term carbon sources and sinks. New features not in either SiB or CASA include dynamic allocation of Gross Primary Productivity for the growth and maintenance of leaves, roots, and wood, explicit calculation of autotrophic respiration, and leaf biomass prescribed using leaf area index (LAI). Simulated carbon fluxes and biomass are consistent with observations at selected eddy covariance flux towers in the AmeriFlux network representing a broad spectrum of biome types. We documented the SiBCASA model and testing results in a paper in *Global Biogeochemical Cycles* (Schaefer et al, 2008).

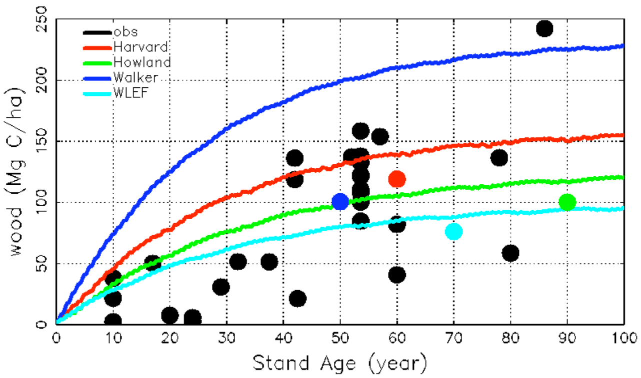

Figure 2: SiBCASA simulated wood biomass (color lines) and observed wood (color coded points) at four deciduous flux tower sites, and observed wood biomass from Forest Inventory Analysis plots (black points) indicate SiBCASA generally reproduces observed wood biomass as a function of stand age.

3. Boundary-Layer Entrainment Parameterization for Improved Tracer Transport in RAMS

At present the Regional Atmospheric Modeling System (RAMS) does not include any entrainment parameterization at the top of the planetary boundary layer. The entrainment of mass and energy into the PBL allows the boundary layer to grow through not only the simple addition of more air, but also through the addition of high potential temperature air into the boundary layer. Adding high potential temperature air into the boundary layer leads to an increase in the temperature and energy within the PBL and allows the boundary layer to grow deeper than without the inclusion of entrainment. We have modified RAMS to include the entrainment flux of energy and mass, producing a more accurate prediction of boundary layer height (McGrath-Spangler and Denning, 2009).

At the inversion interface, overshooting thermals inject cool, moist, CO_2-depleted turbulent boundary layer air into the overlying inversion and entrain warm, CO_2-rich free tropospheric air downward creating negative heat and carbon fluxes in the region of overshoot. Since the mixed layer is heated both from the surface due to solar heating and from the top of the boundary layer by entrainment, the profile of heat flux throughout the boundary layer is linear (Stull, 1976). Above the inversion, the perturbation vertical velocity is assumed zero, implying the heat flux is also zero.

This implies that the negative heat flux at the base of the capping inversion ($\overline{w'\theta_v'}|_{zi}$) is linearly proportional to the heat flux at the surface, leading to the closure assumption for the heat flux at the top of the boundary layer:

$$\overline{w'\theta_v'}|_{zi} = -\alpha\overline{w'\theta_v'}|_{s} \tag{1}$$

where α is the proportionality constant (Stull, 1988). Estimates of α from experimentation and theory range anywhere from zero to one with most published values being between 0.1 and 0.3 (Betts, 1973; Carson, 1973; Deardorff, 1974; Rayment and Readings, 1974; Willis and Deardorff, 1974; Stull, 1976; Sullivan et al., 1998).

We use this assumption to include a heat flux and a flux of other variables such as wind velocity and carbon dioxide concentration from overshooting thermals to alter the temperature, water vapor mixing ratio, winds, and CO_2 mixing ratios of the boundary layer and the lowest layer of the inversion. The heat flux can be used to define a time rate of change of potential temperature for the layers just above and just below the base of the inversion:

$$\frac{\partial \theta}{\partial t} = \frac{\alpha \overline{w'\theta_v'}\,|_s}{\Delta z}. \tag{2}$$

where Δz is the thickness of the layer. Through this heating at the top of the boundary layer, the whole mixed layer warms through turbulent mixing whereas the inversion layer cools. This process weakens the capping inversion and makes it easier for the boundary layer to grow.

The heat flux across the capping inversion can be used to define a mass flux representative of the amount of mixing between the layers surrounding the interface between the boundary layer and the capping inversion required to produce these temperature changes. This mass flux is given by:

$$M = \frac{\rho \alpha \overline{w'\theta_v'}\,|_s}{\Delta \theta_v} \tag{3}$$

where ρ is the density of the air and is computed from the total Exner function and ice-liquid potential temperature prognosed by RAMS (Medvigy et al., 2005). This gives the mass flux the units of kg m^{-2} s^{-1}. When multiplied by the specific humidity of the layer, q_v, this becomes a mass flux of water vapor. This can be used to find a time rate of change of the mixing ratio for the layers above and below the interface given by:

$$\frac{\partial r_v}{\partial t} = -\frac{\alpha q_v \rho \overline{w'\theta_v'}\,|_s}{\Delta \theta_v \rho_{dry} \Delta z} \tag{4}$$

where ρ_{dry} is the density of the dry air. This induces a drying of the whole boundary layer through turbulent mixing and a moistening of the capping inversion.

This same mass flux can be used to define time rates of change of the different components of the wind vector, TKE, and CO_2 concentration given in the equations below.

$$\frac{\partial u_i}{\partial t} = \frac{u_i \alpha \overline{w'\theta_v'}|_s}{\Delta\theta_v \Delta z} \tag{5}$$

$$\frac{\partial w}{\partial t} = -\frac{w \alpha \overline{w'\theta_v'}|_s}{\Delta\theta_v \Delta z} \tag{6}$$

$$\frac{\partial TKE}{\partial t} = -\frac{TKE \alpha \overline{w'\theta_v'}|_s}{\Delta\theta_v \Delta z} \tag{7}$$

$$\frac{\partial C}{\partial t} = \frac{C \alpha \overline{w'\theta_v'}|_s}{\Delta\theta_v \Delta z} \tag{8}$$

In these equations, u_i represents the two components of the horizontal wind vector. These equations represent the mixing across the overlying capping inversion and introduce negative fluxes of horizontal momentum and CO_2 concentration and positive vertical fluxes of vertical velocity and TKE.

Equations (2), (4), (5), (6), (7), and (8) are included in a parameterization within the shallow cumulus convection parameterization unique to BRAMS (Walko et al., 2002; Freitas et al., 2006). The boundary layer height is diagnosed from the virtual potential temperature profile as the height where the temperature increases by 0.5 K, defining the capping inversion. After the PBL height is determined by this method, the temperature, mixing ratio, wind velocity, TKE, and carbon dioxide concentration tendencies for the layers above and below Z_i are altered by the above equations. Before the end of the timestep, the seven variables are updated by their respective tendency arrays.

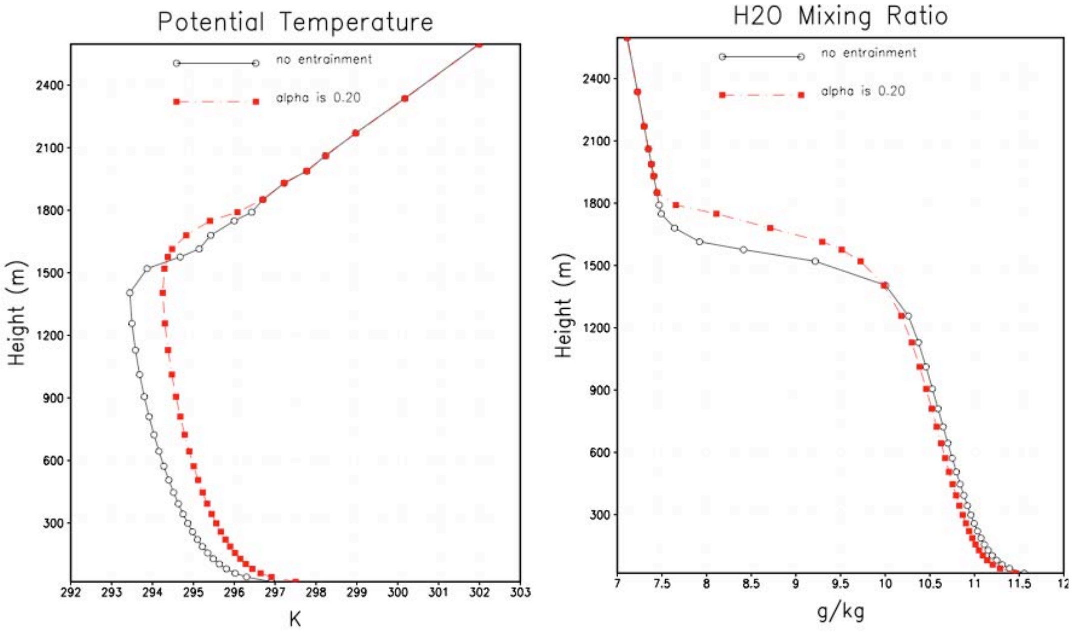

Figure 2: Effect of enhanced entrainment on temperature and moisture in the PBL as simulated by RAMS for an idealized case.

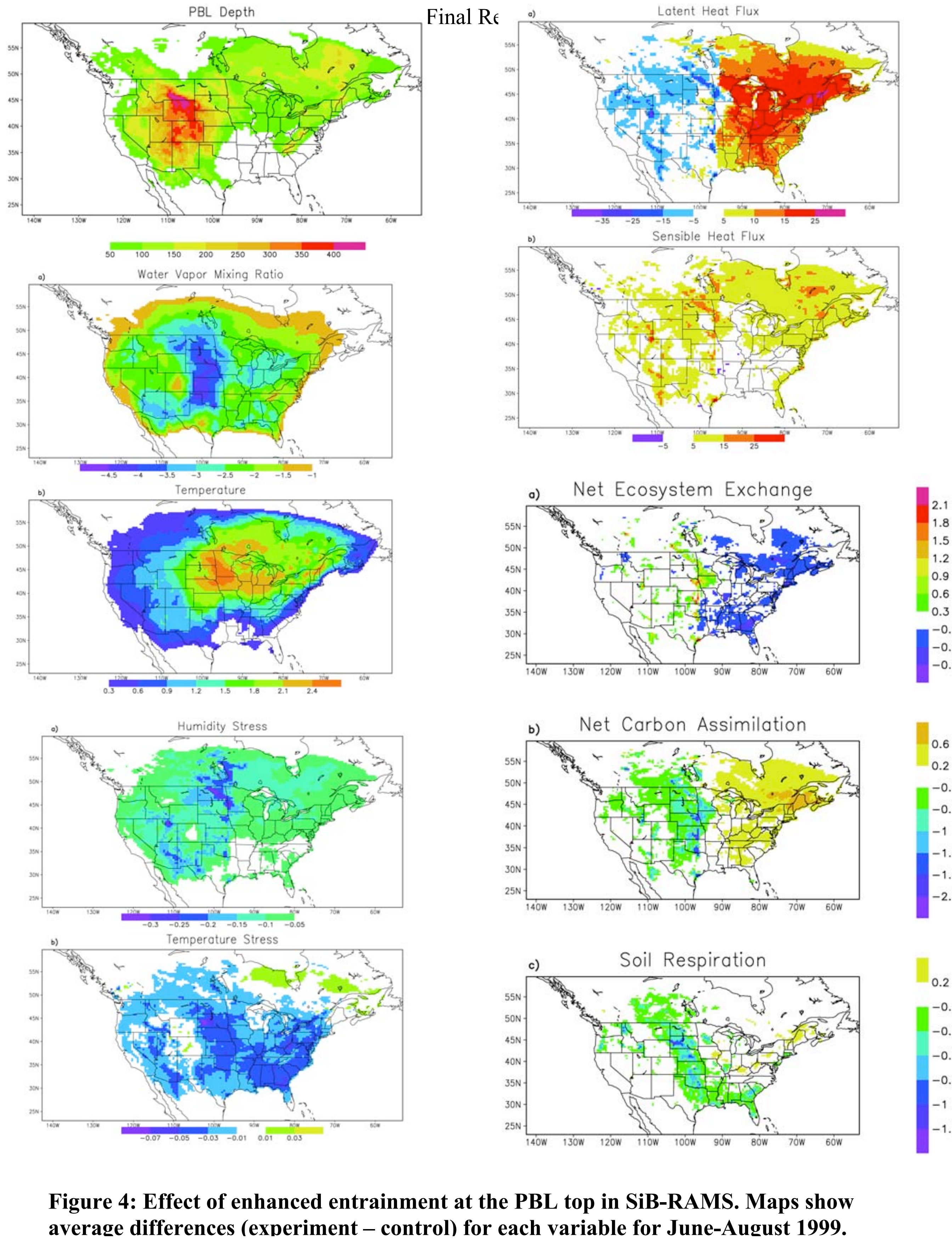

Figure 4: Effect of enhanced entrainment at the PBL top in SiB-RAMS. Maps show average differences (experiment – control) for each variable for June-August 1999.

In order to evaluate the new parameterization of PBL entrainment, we used SiB-RAMS to investigate a case for which we had excellent measurements of PBL depth in addition to other meteorological and CO2 data: the summer of 1999 at the WLEF tower in Wisconsin. As predicted by the idealized model, enhanced turbulent entrainment produced a warmer and drier PBL over most of the domain. Unlike the idealized model however, these changes propagated downwind, leading to altered weather, climate, and ecosystem function over much of the model domain. This resulted in substantial differences in physiological stress, latent and sensible heat fluxes, photosynthesis, respiration, and net ecosystem exchange (NEE) of CO2 even in the seasonal average. (Fig 4).

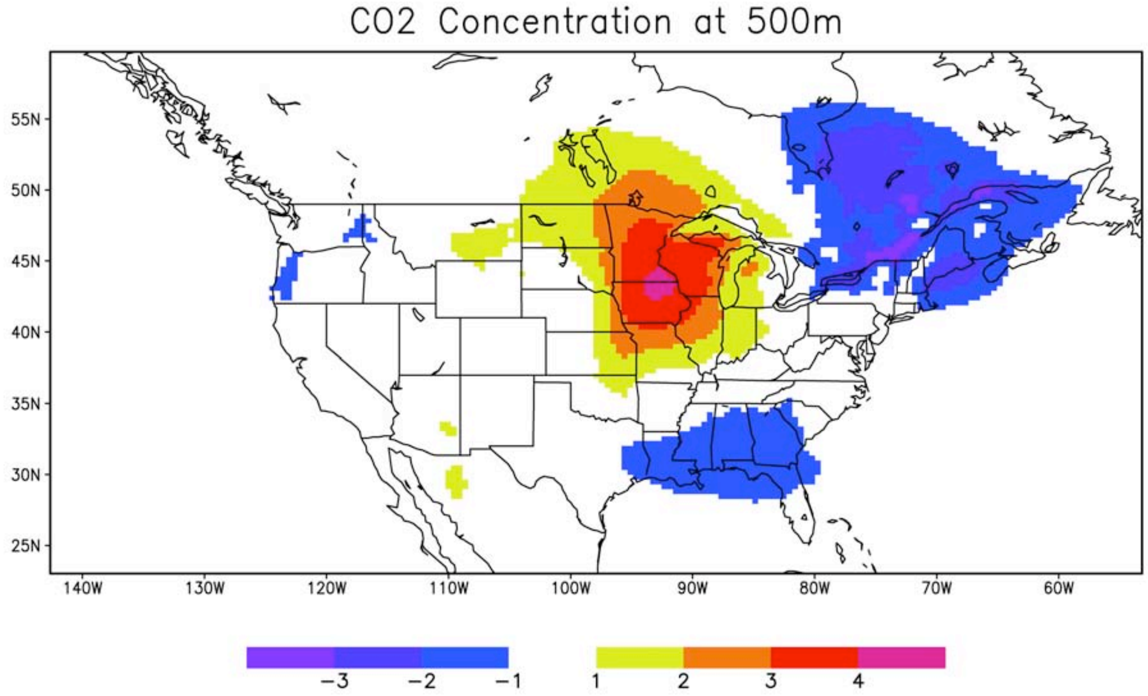

Figure 5: Simulated ifference in mean CO2 concentration at 500 m above local ground level averaged from June-August 1999 due to enhanced PBL-top entrainment.

The differences in simulated climate, ecosystem function, and NEE interacted with the deeper PBL to produce differences in atmos[pheric CO2 mixing ratio with a distinct spatiual pattern (Fig 5). Gradients in the CO2 concentration in the PBL were changed by almost 8 ppm from Iowa to Maritime Canada: a difference in seasonal mean CO2 twice as large as the annual difference from the Arctic to the Antarctic! These differences are certain to be important to source/sink inversions as data from the emerging NACP network become available.

How realistic is the new, deeper PBL simulated by SiB-RAMS? We evaluated the new parameterization by comparing simulated to observed PBL depth at the WLEF tower in Wisconsin during the summer of 1999, when a 915 MHz radar wind profiler was deployed (Denning et al, 2008). This site is unique because the depth of the nocturnal stable layer can be determined from CO2 and temperature gradients measured through the 400 m depth of instrumentation at the tower, while daytime convective boundary layer depth was estimated from changes in radar reflectivity. The control version of the model consistently underestimated daytime PBL depths, but the new version with enhanced entrainment more accurately captured diurnal and seasonal variations (Fig 6). Both models are reasonably successful in predicting the evening transition and the collapse of the PBL to form a shallow stable layer at the base of the daytime mixed layer, which grows deeper through the night under a residual layer aloft. If anything, the model may need even more turbulent entrainment to properly capture the depth of the daytime mixed layer.

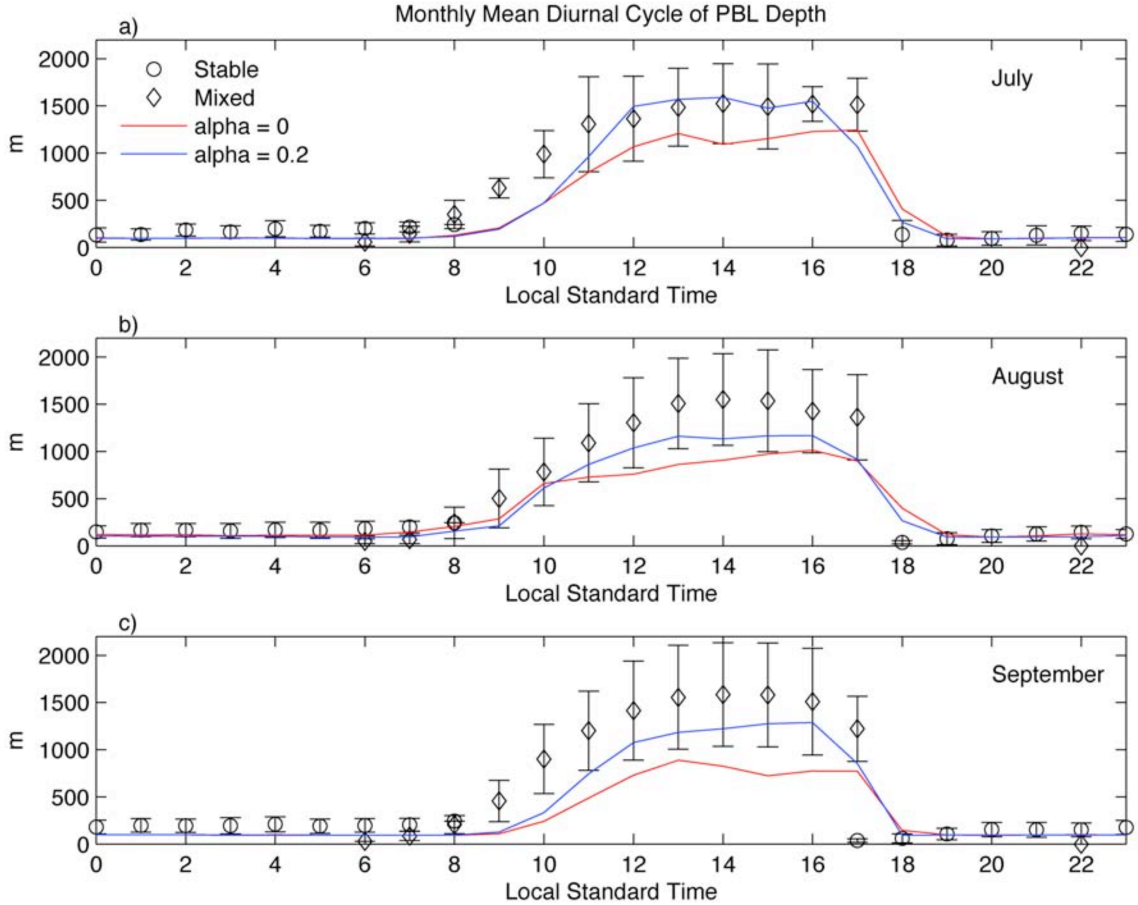

Figure 6: Monthly mean diurnal cycle of PBL depth (m) in (a) July, (b) August, and (c) September. Model estimates are averaged over non-precipitating hours only. Error bars are the standard deviation of daily values at each hour.

4. Further SiB-RAMS Model Development

SiB-RAMS has also undergone substantial development in the coupling of SiB3 to RAMS. The photosynthesis in SiB3-RAMS is driven by 1-km resolution LAI and fraction of fPAR data from NASA's MODIS satellite, and the vegetation cover is derived from 1-km MODIS land classification data (Zhao et al., 2005). We calculate the respiration factor, soil moisture fields, and initial values for prognostic variables in SiB3-RAMS for each grid cell from an offline SiB3 run. We have implemented the capability to initialize CO_2 concentrations and to nudge lateral boundaries to mixing ratios from global simulations of the Parameterized Chemical Transport Model (PCTM), which has 1^o x 1.25^o resolution, 20 vertical levels, and a 3-hourly timestep. SiB3-RAMS includes surface fluxes of carbon to the atmosphere due to fossil fuel combustion, cement production, and gas flaring, which were derived from the 1995 CO_2 emission estimates of Andres et al. (1998). Air-sea CO_2 fluxes included in SiB3-RAMS are monthly 1995 estimates from Takahashi et al. (2002). The soil classification is from a 5' map containing values of % sand and % clay generated by the SoilData System (Global Soil Data Task, 2000), which were converted to SiB3 soil classes using a triangulation program. The meteorological fields for RAMS are initialized and driven by the National Center for Environmental Prediction (NCEP) Mesoscale Eta-212 grid reanalysis with 40 km horizontal resolution; and lateral boundary and interior nudging is performed for horizontal wind speed, relative humidity, air temperature, and geopotential height.

To test these developments, we analyzed a SiB3-RAMS simulation from June 1 to July 10, 2004. This simulation had a coarse grid covering most of North America with 40 km grid spacing and a nested grid centered over the WLEF tower in Wisconsin. The nested grid has 62 x 62 grid cells with a 10 km grid increment. Both grids have 46 vertical levels and use the two-stream radiation scheme developed by Harrington (1997) and the Grell convective cumulus scheme, which is an entraining plume model based on the quasi-equilibrium assumption (Grell, 1993; Grell and Devenyi, 2002).

SiB3-RAMS simulates the observed net ecosystem exchange (NEE), sensible heat (SH), and latent heat (LE) fluxes at various towers well (Figure 7). The top left panel shows the coarse domain and the location of the towers used to evaluate the model, the model results are depicted by the red line, and the observations are the black lines. The model captures the day-to-day variability of the fluxes. However, SiB3-RAMS slightly overestimates the heat fluxes, as seen by the overestimation of SH at Western Peatland and WLEF and the overestimation of LE at Howland. The right panel shows additional NEE comparisons at a variety of locations with different biome classifications, and the nighttime respiration and daytime drawdown of CO_2 in the model closely matches the observations.

The resulting CO_2 concentrations from SiB3-RAMS compared to the observations is shown in Figure 8. Since the majority of the concentrations are measured near the surface, in order to exclude vertical mixing effects we have plotted the daytime minimum concentrations, except for the bottom right panel which shows observations at 396 m above the boundary layer. Overall the model does a reasonable job of capturing the synoptic events. The model captures the drop in CO_2

at the beginning of July at the Lethbridge, Western Peatland, and Berms sites. At

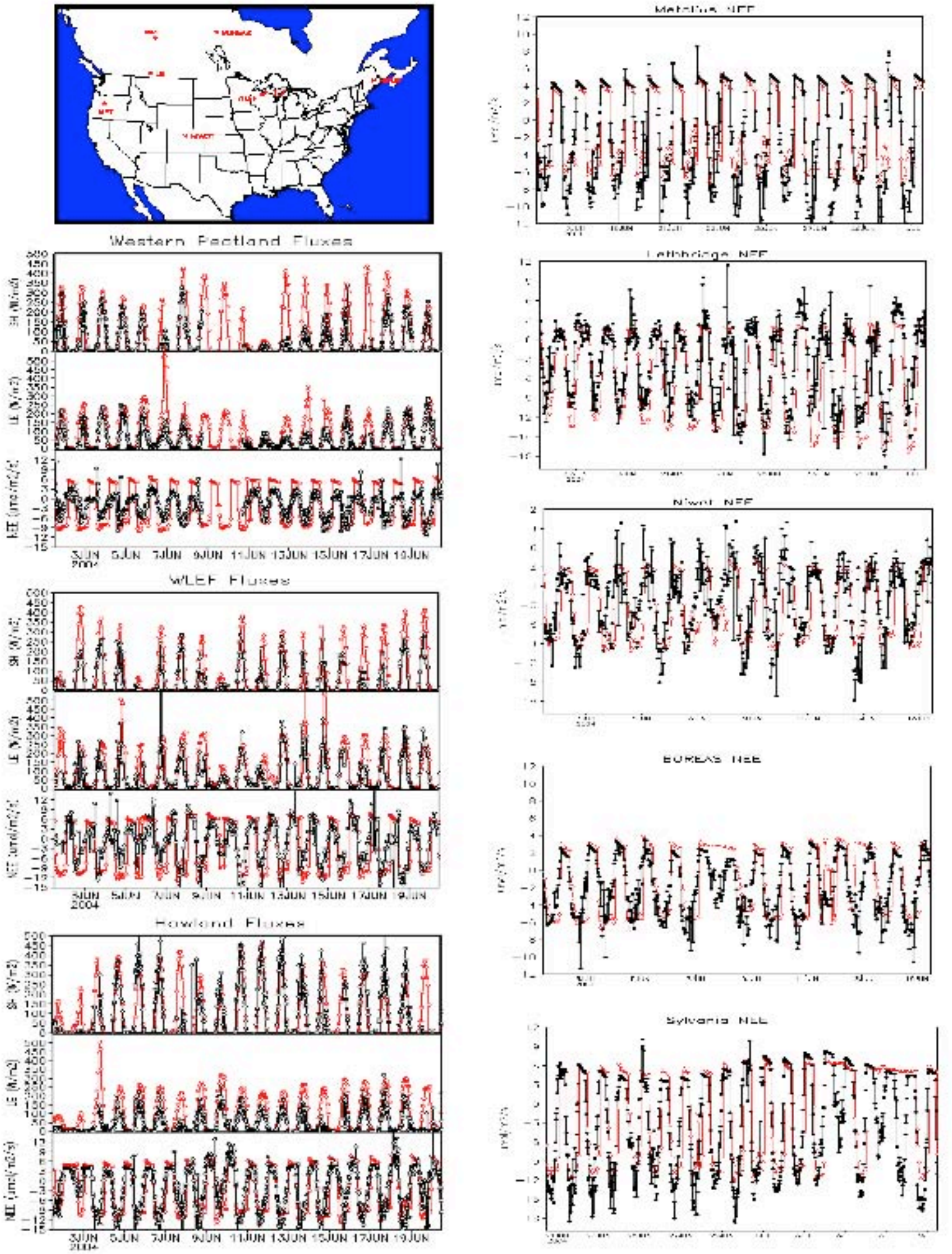

Figure 7: Evaluation of simulated energy and CO2 fluxes at Ameriflux sites.

Fraserdale, the model captures the CO_2 signal but tends to overestimate the magnitude, and at Harvard Forest SiB3-RAMS underestimates synoptic variability.

Looking above the boundary layer at the WLEF tower, SiB3-RAMS underestimates the concentration at the beginning of the simulation, but matches the observations quite well after June 9. With these recent model developments, SiB3-RAMS

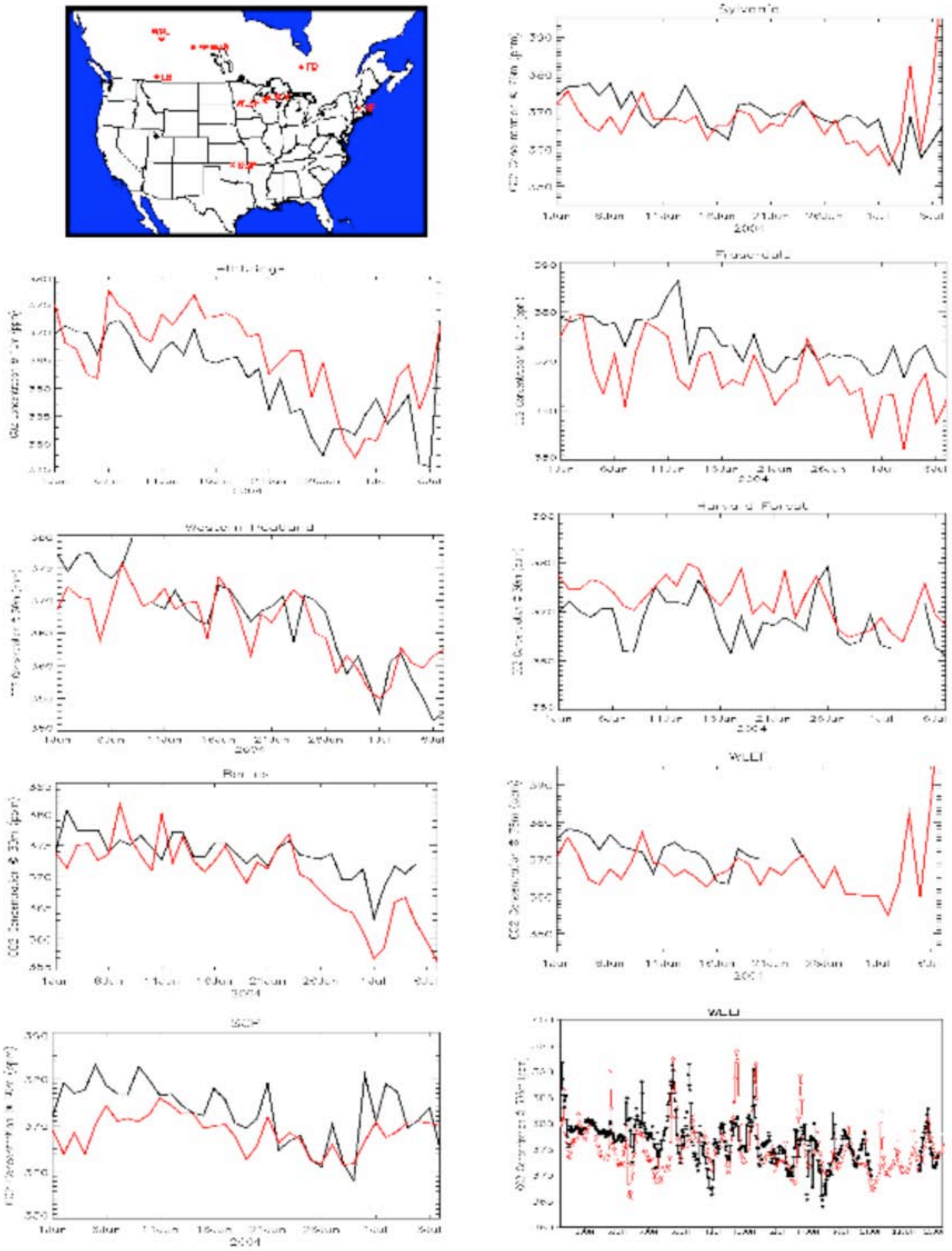

Figure 8: Evaluation of simulated CO2 in SiB-RAMS.

produces realistic CO_2 concentrations that can be used in inverse models to further interpret mixing ratio measurements.

5. Mesoscale Carbon Data Assimilation for 2004

Modeled carbon dioxide at the tower is calculated as the sum of 3 component fluxes convoluted by *time* and *tower* dependent transport.

$$CO_2(time,tower) = Transport_{time,\,tower}\left(\begin{array}{l} Boundary\ Inflow(x,y,time) \\ \quad + Fossil\ Fuel(x,y,time) \\ \quad + Domain\ Biogenic\ Fluxes(x,y,time \end{array}\right.$$

The boundary inflow component was calculated by convolving the influence functions from the LPDM model over boundary CO_2 fields derived using a global biosphere-transport model. At any point in time, the boundary inflow is the average of all upstream particles located in a 3 dimensional 40 km thick rectangular "ring" around the domain. CO_2 resulting from the transport of fossil fuels to the towers is calculated by convolving the influence functions from the LPDM model with surface fossil fuel flux estimates. In particular, the boundary CO_2 fields were calculated by combining transport from the parameterized chemistry transport model (PCTM) (Kawa et al., 2004; Parazoo, N.C. 2007) and pre-calculated archived hourly SiB3 fluxes (Baker et al., 2007) on a 1.25-degree by 1-degree global grid. The model was spun up for 2000-2004 and the CO2 was centered around the Northern Hemispheric mean CO2 for 2004. In addition to this, results from the CarbonTracker project, which provide globally optimized CO_2 concentration fields, are used for comparison purposes.

Fossil fuel fields were constructed using recently available high resolution Vulcan fossil fuel inventory fields (Gurney et al., 2008), at a 10 km horizontal spatial scale and hourly temporal scale. Previously available fossil fuel flux fields were derived by distributing country-level fossil fuel sources spatially as a function of population at a 1-degree resolution (Andres et al., 1995). The Vulcan fields provide many improvements including the incorporation of mobile emission sources and power plants, often located in areas distant from high density population centers, increased temporal resolution allowing the modeling of diurnal variability, and increased spatial resolution allowing better delineation of high density population centers. The sensitivity to the new fossil fuel fields is tested by running inversions using both the Vulcan fields as well as the Andres et al. 1995 fields.

SiB3 balances carbon annually by assuming that ER is in approximate balance with GPP on an annual time frame for each surface location and therefore annual NEE is zero for each surface location or grid cell (Raich et al., 1991; Denning et al., 1996). While this is accurate to a large degree and provides reasonable approximations of respiration on diurnal time scales, it ignores annual imbalances in carbon due to a number of external factors such as land use, fertilization effects, disturbance history, etc. For example, aerial photos and satellite images of the coastal mountains of Oregon show a patchwork of forest ages largely dictated by forest management practices. Under standard models of forest regrowth, a regenerating forest will eventually enter a long period in which carbon is being drawn from the atmosphere and stored in wood and roots, thereby providing a sink of carbon from the atmosphere to the biosphere. Conversely, recently burned forests usually enter a short-term period in which they represent a significant carbon source to the atmosphere. Processes such as these that are largely responsible for annual imbalances in NEE are not characterized in SiB3.

The effect of this on boundary inflow estimates is that the PCTM-SiB3 calculated boundary CO_2 fields lacks the effect of sources or sinks in 2004. Given the consensus opinion of an annual mean sink for carbon resulting from the biosphere, this means that the CO_2 fields used will be biased somewhat by the effect of not including this expected global sink. We investigate the effect of this by including a comparison of the inversion using CarbonTracker optimized CO_2 concentration fields for boundary inflow, which provides an estimate of sources/sinks. As of this time, carbon dioxide resulting from forest fires is not included in the global PCTM-SiB3 inflow or domain SiB3 runs, but is included in the CarbonTracker inflow providing one more contrast between the two fields.

Observational Data

Calibrated CO_2 observations were provided half-hourly at eight measuring sites (WLEF, Harvard Forest, ARM, BERMS, Fraserdale, Western Peatland, WKWT, and Argyle (ME)) for 2004 (Parazoo, 2007). Gerbig et al. [2003] found mean standard deviations on the order of 0.6 to 1 ppm when viewing morning and afternoon vertical profiles of CO_2 in the mixed layer. As a consequence, robust afternoon snapshot observations, at 12PM, 2PM, 4PM, and 6PM local time, are used to lessen the impact of low quality modeled measurements made during times of extremely stable and stratified nocturnal atmospheric conditions near the ground. One exception is the WKWT tower in Moody, TX. Data at this tower consistently showed high diurnally-influenced CO_2 concentrations in the 12PM records for most days. It is uncertain exactly what the cause of this is but it appears that it may be due to some kind of systematic late venting of nocturnal respiration-based CO_2 buildup. For this tower, mixed boundary layer conditions appeared to be better represented by snapshot observations shifted by 2 hours: 2PM, 4PM, 6PM, and 8PM. The first 10 days of the year are not comparable due to a lack of transport preceding 2004. In all there were 2433 missing observations, resulting in 4 (observations/day) * 8 (towers) * 355 (days) – 2433 (missing) = 8927 observations.

In a previous pseudo-data inversion using a very similar model (Zupanski et al., 2007), the errors on the observations were assumed to be 1 ppm for afternoon observations. Nevertheless, relative to the inversion techniques presented in the next section, the errors on these observations should include errors due to calibration error, mapping error, transport error, and representation error. For this inversion, transport error and representation error are likely the largest components which are notoriously tricky to quantify. Investigations into the sensitivity of inversion test results combined with initial maximum likelihood estimation results suggest errors in the range of 5-6 ppm are appropriate for this particular inversion. For the remaining inversions, the errors are assumed to be identical and independently distributed (i.i.d.) mean zero errors with standard deviation set to 5.5 ppm. It should be noted that while it is possible to run inversions with artificially low prescribed “observation” errors, this will generally manifest itself in a need to “over tighten” the a priori covariance structure.

Climatic Conditions for 2004

The 2004 year was the 6th wettest in the contiguous United States over the preceeding 110 years (1894-2004). It was also warmer than on average. Nevertheless, there was a great amount of variability in precipitation and temperature as a function of location and season. Drought continued in the west through the summer of 2004, essentially prolonging a multi-year period of drought conditions. The spring was also very dry for the southeast, extending a period of dry conditions from late in 2003. However, summer brought increased precipitation to the east and southeast, culminating in enormous amounts of rain in late summer and early fall due to an extremely active hurricane season. The south (Texas, Louisiana, Mississippi, Arkansas, Oklahoma, and Kansas) had the wettest summer on record and was much cooler than average. These conditions were important as they provided initial conditions for the inversion that involved soil moisture induced plant stress over large areas of the United States.

Inversion Technique

Standard multivariate Gaussian assumptions are made and data are assimilated using a modified Kalman Filter algorithm (Kalman 1960). In particular, for an initial length n CO_2 measurement vector $\boldsymbol{y}$ representing the first set of measurements, length m unknown CO_2 flux bias vector $\boldsymbol{\beta}$, $n \times n$ observation error covariance matrix

$\boldsymbol{\Sigma}$, *n x m Jacobian* transport matrix $\boldsymbol{G}$, length *m* prior flux estimate $\boldsymbol{\beta_0}$, and *m x m* model-prior mismatch covariance matrix $\boldsymbol{\Sigma_0}$, the Bayesian statistical assumptions are:

$$\begin{aligned} y \mid \beta, \Sigma &\sim N(G\beta, \Sigma) \\ \beta &\sim N(\beta_0, \Sigma_0) \end{aligned} \tag{4}$$

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

$$\begin{aligned} p(\beta \mid y, \Sigma) &\propto -\frac{1}{2}\left[(G\beta - y)^T \Sigma^{-1}(G\beta - y) + (\beta - \beta_0)^T \Sigma_0^{-1}(\beta - \beta_0)\right] \\ &\sim N\left(\left(\Sigma_0^{-1} + G^T\Sigma^{-1}G\right)^{-1}\left(\Sigma_0^{-1}\beta_0 + G^T\Sigma^{-1}y\right), \left(\Sigma_0^{-1} + G^T\Sigma^{-1}G\right)^{-1}\right) \end{aligned} \tag{5}$$

With a little bit of algebra, one can rewrite the mean of the posterior distribution of the mean, giving the Kalman-filter updating equation for the mean.

$$E[\beta] = \beta_0 + \left(G^T\Sigma^{-1}G + \Sigma_0^{-1}\right)G^T\Sigma^{-1}(y - G\beta_0) \tag{6}$$

The posterior mean and variance of $\boldsymbol{x}$ are then fed into the next filter step with a new set of measurements. This particular inversion estimates biases over 7-day periods using available data from that 7-day period of time. Therefore, bias estimates for both ecosystem respiration and GPP as well as corresponding variance estimates are available for all of 2004 with the bias estimates changing with a weekly resolution.

Two difficulties often arise when using filter-style correction schemes. The filter estimates can drift away from realistic values if the data are not plentiful or precise enough to constrain it. Secondly, the nature of the Kalman filter at each step is to create posterior variance estimates that are in general smaller than the prior estimates. This can essentially cause the filter to get "stuck" and also produce unrealistically small posterior variance estimates around the biases. There is generally no easy solution to this problem. Artificially inflating the posterior variance at each filter step is one method in which one can try to circumvent (Zupanski et al. 2007). This accommodates the fact the biases are likely to change in reality and it allows the filter to consider a wider range of possibilities for the bias factors. However, it does not necessarily constrain the biases to any particular "reasonable" region of values allowing the bias estimates to drift into unrealistic parameter space. Therefore, we have chosen to weight the filter at each step with a "grand" prior. This effectively handles both of the preceding problems. With respect to our inversion, there will be three pieces of information at each step, the grand prior which is derived from the forward SiB3-RAMS model with an error assumption, the local prior which is derived from the previous filter step's posterior flux bias distribution, and the data which forms the statistical likelihood function. In some sense, this new piece of the covariance structure provides a bound upon how much the inversion can "learn" about the bias structure.

In order to quantify, we denote the grand prior as a multivariate Gaussian distribution around $\boldsymbol{\beta_{grand}}$ with covariance matrix $\sigma_{\text{grand}}^2\Sigma_{\text{grand}}$, and additional weight factor ***w***, and we rewrite the expression given in (4) as:

$$p(\beta \mid y,\Sigma) \propto -\frac{1}{2}\left[(G\beta-y)^T \sigma_{obs}{}^{-2} I (G\beta-y) + (\beta-\beta_0)^T \sigma_0{}^{-2}\Sigma_0{}^{-1}(\beta-\beta_0) + (\beta-\beta_{grand})^T w\sigma_{grand}{}^{-2}\Sigma_{grand}{}^{-1}(\beta-\beta_{grand})\right] \tag{7}$$

Thus β is distributed as a multivariate Gaussian with parameters:

$$Mean(\beta) = E[\beta] = \left(w^{-1}\sigma_{grand}{}^{-2}\Sigma_{grand}{}^{-1} + \sigma_0{}^{-2}\Sigma_0{}^{-1} + G^T\sigma_{obs}{}^{-2}IG\right)^{-1}\left(w\sigma_{grand}{}^{-2}\Sigma_{grand}{}^{-1}\beta_{grand} + \sigma_0{}^{-2}\Sigma_0{}^{-1}\beta_0 + G^T\sigma_{obs}{}^{-2}Iy\right) \tag{8}$$

$$Variance(\beta) = E\left[\beta^2\right] - \left(E[\beta]\right)^2 = \left(\left(w\sigma_{grand}{}^{-2}\Sigma_{grand}{}^{-1} + \sigma_0{}^{-2}\Sigma_0{}^{-1} + G^T\sigma_{obs}{}^{-2}IG\right)\right)^{-1} \tag{9}$$

Eq. 7 specifically separates out the variance scalars, σ_{grand}^2 , σ_0^2, and σ_{obs}^2 from the covariance matrices, leaving the covariance matrices essentially scaled to 1. The ***w*** weight is a redundant factor and is simply included to facilitate easier interpretation of tightening/loosening of the grand prior covariance (around the SiB3 derived a priori carbon fluxes). Unless otherwise specified, this weight, ***w***, on the grand covariance matrix is set to 2. This means that the initial variance around the grand prior is increased, thus providing a weaker constraint. For the initial filter step, only the grand prior is used. After that point, there exist both a grand prior and a prior (from the posterior of the previous filter step). The inversion is further constrained by the assumption of spatially correlated errors in the grand prior, i.e. the covariance matrix Σ_{grand} will take on the following form.

$$\Sigma_{grand} = \begin{bmatrix} \Sigma_{Respg,prior} & 0 \\ 0 & \Sigma_{Assimn,prior} \end{bmatrix} \quad (10)$$

The respiration and GPP covariance matrices are each formed from the exponential covariance function, where $t_{i,j}$ is the distance between points β_i and β_j.

$$Cov(\beta_i, \beta_j) = \begin{cases} \sigma_0{}^2(1-\alpha_0)\exp\left(\frac{-t_{i,j}}{h_0}\right), i \neq j \\ \alpha_0\sigma_0{}^2, i = j \end{cases} \quad (11)$$

The h_0 parameter is the range, or decorrelation length scale parameter, giving the distance at which the covariance between two points is equal to $\sigma_0{}^2(1-\alpha_0)e^{-1}$. The σ^2 parameter is the scalar variance parameter and determines the variance of the marginal distribution of the particular flux component. The parameter α_0 controls what percentage of the covariance can be attributed to spatial covariance, as opposed to spatially independent errors.

Inversion techniques can be extremely sensitive to assumptions. It was shown in Chapter 3 that this inversion model is robust to small spatial scale random deviations in flux bias and that post-aggregated (in space) estimates can be very good even when using a fairly sparse network of towers observing CO_2. Nevertheless, given the unconstrained nature of the inversion problem, it is always important to assess the impact of varying certain unknown parameters in the inversion, such as spatial decorrelation length scales, the weight given to the "grand" prior, and the fixed CO_2 contributions from both the boundary inflow and fossil fuel sources.

Sensitivity

The inversion essentially guarantees some improvement in prediction of observed CO_2 (Eq. 5). However, when using a regression style approach in a heavily unconstrained environment, this improvement can often be overstated because of the great freedom the inversion has to fit the data. Therefore, it is often desirable to go beyond simply comparing observed carbon dioxide at the towers to model-based predicted carbon dioxide. Comparing model observations to independent observations not used in the inversion, comparing models which predict similar quantities, as well as testing the sensitivity of the model to variations in unknown parameters are all methods of generating more confidence in estimates.

We used a variety of different procedures to test the sensitivity of the inversion. Regional inversions have been shown to be very sensitive to boundary inflow variations. Therefore, we first test the sensitivity of the inversion to varying the inflow of CO_2 at the boundaries. To do this, we derive boundary inflow to the 8 towers using the LPDM model and optimized carbon dioxide concentration fields from the CarbonTracker project (Peters et al., 2007). Inversion results are then compared with the results derived from the LPDM model and the PCTM inflow. Secondly, we vary several different variance parameters and derive annual domain-summed NEE and tower observation based RMSE based upon the varied parameters. Thirdly, we use a re-sampling procedure in which we create 100 different observation data subsets by holding out a randomly selected 50% of the observation data for each. Each set of data is run through the weekly inversion scheme and the sensitivity of the predicted CO_2 at the towers and the estimated flux biases is explored. This provides estimates of the variability of the flux correction factors and can be used to assess the sensitivity of the source/sink to the constraint provided by the data. Using the held out data as independent evaluation data and the complementing data as training data for the inversion, one may also derive a more accurate estimate of root mean-squared error (RMSE) of the inversion-optimized fluxes. We test the impact of the high resolution Vulcan fossil fuel inventory on the inversion results by comparing inversion results relying upon Vulcan to those results utilizing the Andres et al. [1995] fossil fuel inventory.

SiB3 has been evaluated at many sites and over many time periods, nevertheless, the particular model run used for the a priori flux estimates was not optimized to fit the flux data at any site in particular. Even though there is a mismatch in representation, with the flux towers representing footprints of less than a square kilometer and the inversion results representing flux estimates on the scale of thousands of square kilometers, we believe that these comparisons are of value, especially in locations that are more spatially homogeneous than others, such as grasslands and large forest reaches. This is then the fourth comparison we make.

Results

As was indicated in the previous section, there are a number of variables that the inversion will likely be sensitive to and therefore the results are expected to be quite variable. For results, we choose to present one particular case with a fixed set of inversion inputs as an initial case study and then use it to compare the effect of varying the boundary inflow and the source of the domain fossil fuel fluxes. With reference to the preceding section and Eq. 7 in particular, the following values are used for these inversions: σ_{grand} = 0.15, σ_0 = 0.15, σ_{obs}^2 = 5.5 ppm, w = 2, h_0 = 250 km. In particular, a value of σ_{grand} = 0.15 would mean that we expect that approximately 68% of the GPP and ER biases are within +/- 15% of the original SiB3 estimated fluxes, with 95% within +/- 30%. This variation when combined with positive spatial correlations was shown to provide a reasonable a priori range of annual domain-summed NEE. These deviations must generally be kept to less than 30% - 40% to ensure that posterior ER and GPP fluxes are not reduced by more than 100 percent, which makes no conceptual sense. We then test the sensitivity of the results over a number of varying inversion inputs using the PCTM boundary conditions and the Vulcan fossil fuel flux field.

General Structure of Results

CO_2 can be predicted by invoking the relationship shown in Eq. 3. The predicted mean observed CO_2 is derived as $G\hat{x}$ where $\hat{x}$ represents one (for the prior fluxes) plus the inversion-optimized flux biases. Using the PCTM boundary conditions and the Vulcan fossil fuel inventory, a comparison of the inversion-corrected posterior predictions at the towers to the observations is shown in Fig. 9. For domain-summed temporal plots, NEE is calculated via Eq. 2 while ER and GPP are calculated via the two respective summands on right hand side of that equation. These domain-summed temporal results are shown in Fig. 10.

The observed carbon dioxide concentrations contain information that infers a dampening of the a priori annual GPP cycle, and hence the a priori annual ER cycle (due to the strong correlation of the annual sums of each). Since both GPP and ER are significantly dampened, it is not surprising that the NEE signal is dampened as well. Furthermore, the data suggest a weak temporal shift in the prior NEE signal. This manifests itself as a stronger, but more gradual onset of spring, followed by a weaker overall carbon sink over the middle and late summer periods.

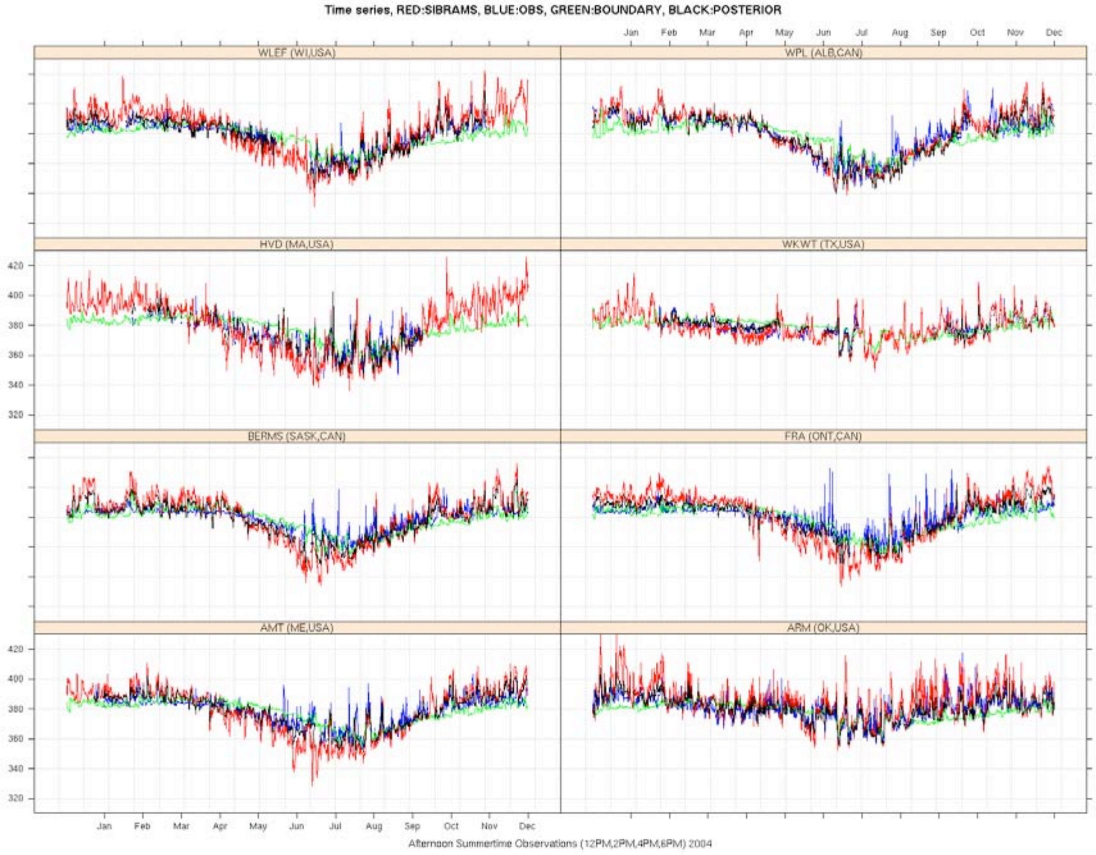

Figure 9: Time series plots of carbon dioxide observations (blue), SiBRAMS prior (red), boundary inflow estimate (green), and posterior prediction (black).

We use a resampling procedure to account for variability that might be associated with over fitting the model and which provides additional variability to the standard covariance estimates of the biases given in Eq. 6. One hundred different inversions are run, each based upon a different subsample of the observations. Assuming temporal independence of the errors in the filter, one may simulate properties of the annual NEE probability density functions (pdf) for each of these 100 inversions by using the posterior covariance provided at each step of the Kalman Filter for each inversion. A 95% confidence interval (CI) for the entire domain can be calculated at each step of the filter for each of the 100 inversions. The CI shown in Figure 10 then characterizes variability in the NEE by selecting the 95% CI of each set of 95% CIs for each weekly time step.

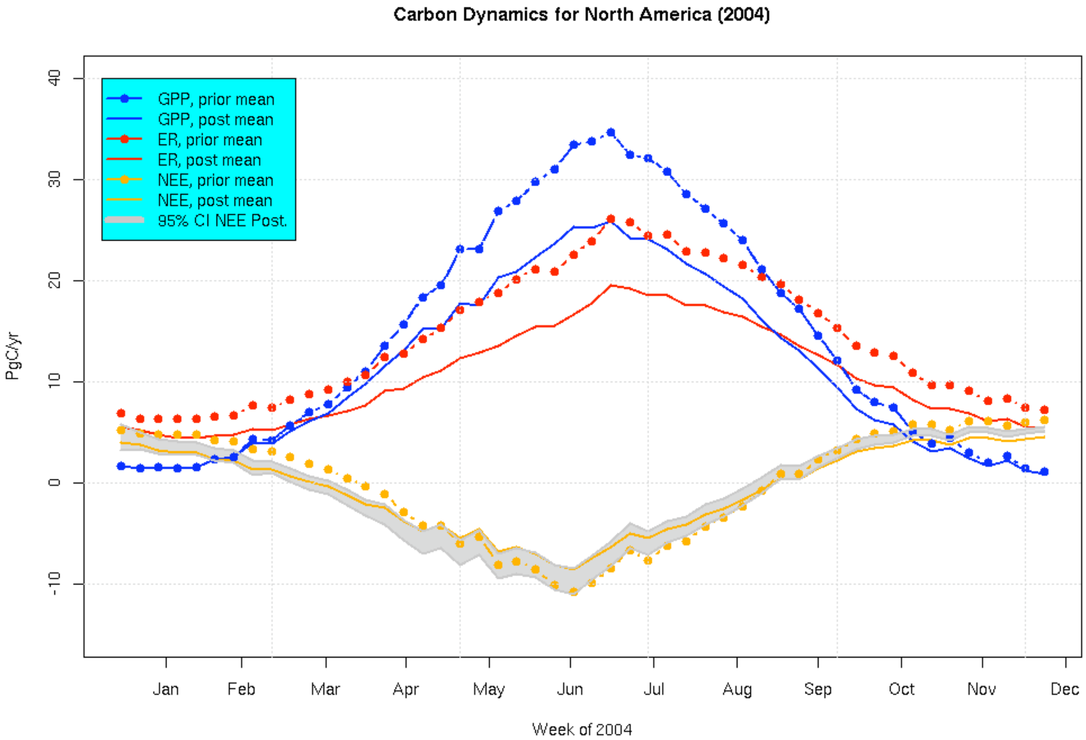

Figure 10: Plots of prior and posterior estimates for GPP, ER, and NEE. Results are shown for a single inversion while the confidence intervals are derived from an ensemble of 100 inversions.

The ensemble mean of the domain summed annual NEE flux is approximately -0.5 Pg/yr while the standard deviation of this estimate is about 0.15 Pg/yr. It is important to note that this standard deviation estimate does appear to be too small, giving tighter bounds on the flux than found in other inversion papers (Gurney et al. 2002; Peters et al. 2007). The spatial representation of these sources and sinks can be seen in the first panel of Fig. 13. Depictions of this variability in a spatial framework are shown in Fig. 11. This variability is partitioned into two pieces, variability associated with the spread of mean estimates over the 100 inversions (measure of over fitting) and variability associated with summing up the posterior variances at each filter step (regular KF variance) evaluated over all 100 inversions. Besides the spatial display of posterior variance information for NEE, which roughly tracks the convolution of the sampling footprint of the network and the prior ER/GPP signals, the results show that over fitting the model may provide a significant source of variability comparable to that which is normally constructed from each filter step's posterior covariance matrix.

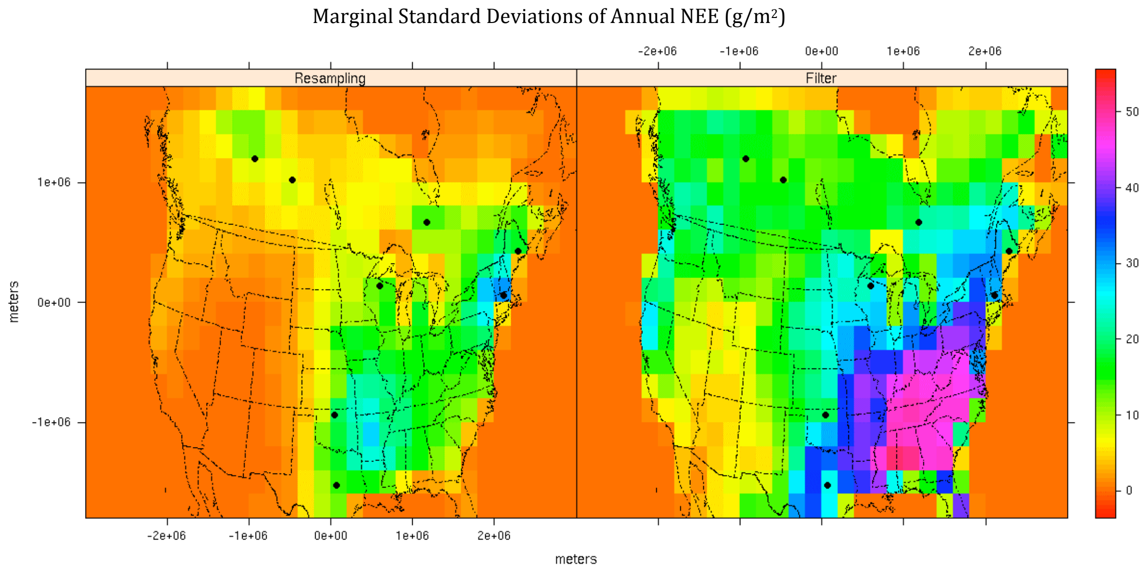

Figure 11: Uncertainty measures in annual NEE. The left panel is the result of running 100 inversions each using a randomly selected 50% of the data and then calculating the variance of each cell's mean estimate, over the 100 inversions, and summing over each of the weekly filter cycles. Finally, the square root of this summed variance (standard deviation) is displayed and is a measure of the uncertainty of the mean estimate due to model overfitting. For the right panel, the summed annual variance in NEE is calculated for each inversion, from the weekly filter estimates, and the the square root of this (standard deviation) is shown for each cell. These plots aim to provide a measure of the uncertainty of each cell's NEE estimate, incorporating the correlation between ER and GPP in each cell, but not incorporating the spatial correlation in the covariance matrices.

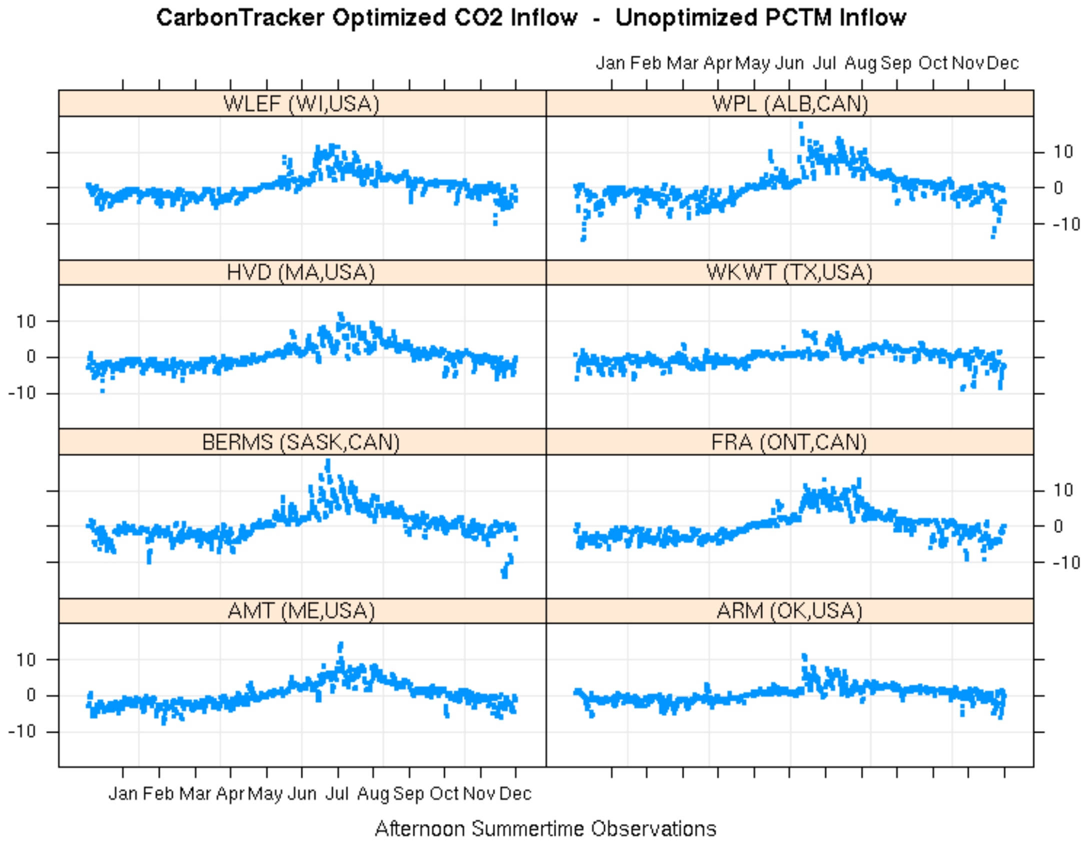

Figure 12: Difference of boundary inflows (1420 sequential '12/2/4/6PM' observation sequences for each of 8 towers.)

Sensitivity and Robustness of Results to Inflow

Inflow of CO_2 from the boundaries has typically been a large concern of regional models (Gerbig et al., 2003; Peylin et al., 2005). In extremely limited domain problems, the variance of the CO_2 coming in from the boundary can easily dwarf the changes inside the domain due to local biotic uptake and release. Therefore it is of interest to gauge the sensitivity of the inversion to varying boundary inflows. The boundary conditions included in this model were constructed from a global simulation using SiB3 and PCTM (Parazoo et al., 2007). The CarbonTracker project has provided CO_2 mixing ratio data based upon globally optimized fluxes (Peters et al., 2007). SiB3 has no annual source/sinks whereas CarbonTracker includes an annual source/sink estimated from observations of CO_2. A plot of the difference between the two inflows is shown in Fig. 12. The inflow annual mean and temporal pattern is very similar for PCTM and CarbonTracker with the main difference being a seasonally stronger cycle in the PCTM-SiB3 results, likely a result of the underlying biosphere model, SiB3, providing a stronger seasonal GPP/NEE signal than the corresponding CASA model used in CarbonTracker. In addition to running comparison inversions between these two CO_2 inflow estimates, we also run the inversion with a fixed inflow estimate of 378 ppm representing the annually averaged PCTM inflow over the period of the simulation in order to show the necessity of reasonable boundary inflow values in calculating reasonable source/sink estimates.

Fig 13 shows a comparison plot of maps of the annual mean NEE estimate based upon CarbonTracker (w/ CASA), PCTM (w/ SiB3), and the fixed inflow condition. The results are similar for the CarbonTracker and PCTM inflows but also surprising in their differences. The estimates have similar spatial and temporal characteristics and differ mainly in magnitude. However, the PCTM-based inversion results in a sink of 0.1 – 0.2 Pg/yr less than that of the CarbonTracker-based result, which is the opposite of what one might guess. Fig 5.6 shows inflow from CarbonTracker that is generally less than that of PCTM for the first 4 months of the year. During this period the inversion must add sinks to the model to account for a lower concentration of CO2 coming in from the boundaries. During the summer, when the situation is reversed the opposite occurs, the inversion must add sources to the model to compensate for the fact that CT inflow is higher than PCTM inflow. However, the magnitude of this adjustment does not seem to be a simple linear function of the concentration differences seen in Fig. 11, which would imply that larger sources are needed in the CT-inflow model during the summer time than in the winter. On the contrary, the relative carbon sources that are a result of CT inflow being higher during the summer months are actually less than the carbon sinks resulting from the CT inflow being less during the winter/spring months. This results in the CT-inflow based inversion having a larger annual sink estimate than the PCTM-inflow based inversion. The sink estimated with the PCTM inflow was 0.47 Pg/yr while the sink estimated with the CarbonTracker inflow was estimated at 0.58 Pg/yr. It does seem somewhat surprising that the results from the two inflows are so close, within approximately 20% of one another. This indicates that local observations may be affected significantly more by local fluxes than by larger scale fluxes in distant locations outside of the model boundary.

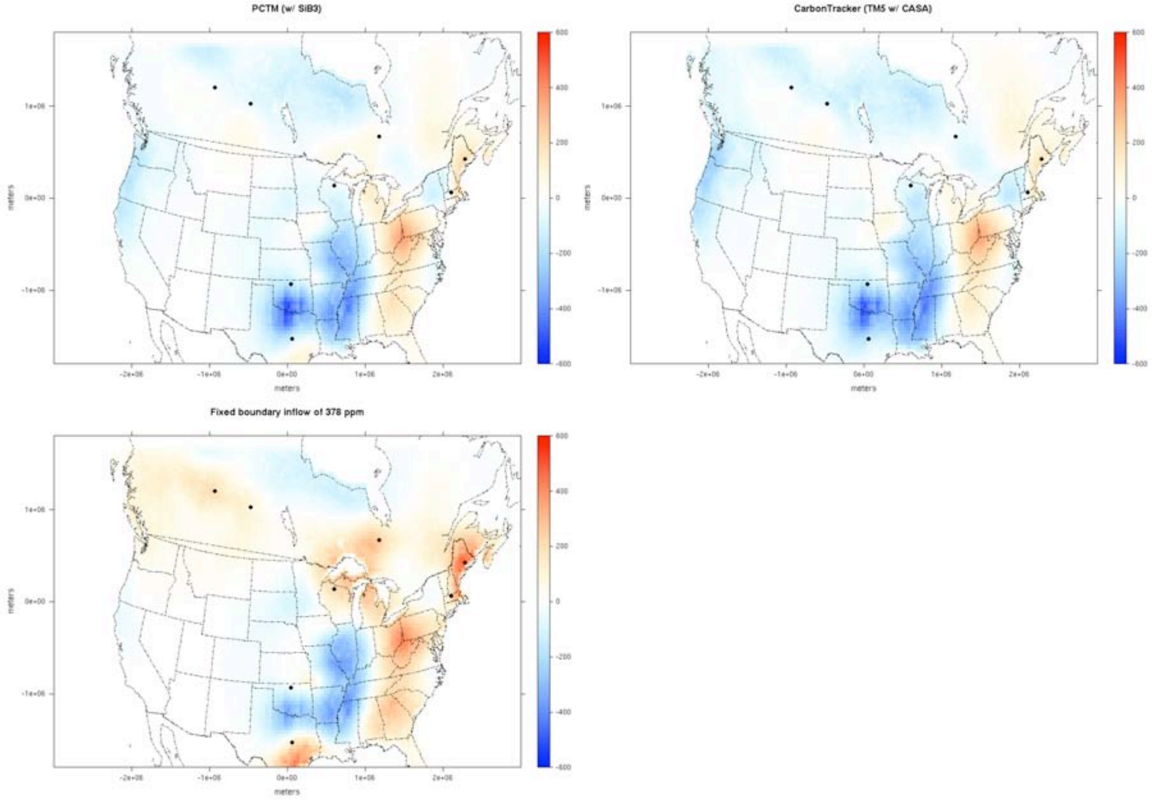

Figure 13: Inversion estimates for three different inflow scenarios, one without modeled annual source/sink (PCTM w/ SiB), one with modeled source/sink (CarbonTracker w/ CASA), and a uniform fixed 378 ppm inflow. Summed annual sinks are 0.47 PgC/yr, 0.59 PgC/yr, and 0.27 PgC/yr respectively.

Sensitivity of Results to Fossil Fuel Inventory

Until the release of the Vulcan fossil fuel inventory in 2008, most researchers were reliant upon the Andres et al. [1995] fossil fuel inventory, which was released at annual time scales and at a 1-degree resolution over the globe. For many large-scale inversion applications, this inventory is adequate. However, for higher resolution studies within the United States, the Vulcan fossil fuel inventory provides a dramatic improvement in both space and time accounting of fossil fuel fluxes. The main difference between these inventories is the redistribution of some fossil fuel sources from population centers to more distant locations representing mobile sources and power plants. The Vulcan fossil fuel flux estimates are at a much higher resolution in both time and space. Previous inversions had to grapple with the fact that some observing stations are located within enormous fossil fuel flux regions. For example, a semi-rural location like Harvard Forest would very likely be located in the same grid cell as the large metropolitan city of Boston. Given no sub-annual temporal resolution to the fossil fuel fluxes, an observing tower located at Harvard

Forest was often seeing a 24 hour continuous stream of fossil fuel fluxes arising from a city over 100 km away. However, the 10 km horizontal resolution of the Vulcan inventory allows these to be separated and additionally provides a diurnal and seasonal estimate of these fluxes, which is important for inversions based upon hourly observations.

In order to gauge the impact of incorporating the Vulcan data, we first contrasted the contributions to each of the 8 towers from each of the inventories. For many of the stations, the afternoon differences between the two were very small. Differences at the ARM site in Oklahoma, the WLEF site in Wisconsin, the Canadian sites, and the Argyle, Maine site were on the order of a few ppm. Differences at the Moody, Texas tower were in the range of -5 ppm to 5 ppm. While the differences across most towers were relatively small, the differences at Harvard Forest were between -25 ppm and 30 ppm!

The difference in the annual NEE estimate is shown in Fig. 14. The effect on the inversion is far from trivial with differences of up to 300 g/m2 per year recorded along the northeast coast of the United States. These differences are a result of coarse fossil fuel flux fields providing artificially high sources of CO_2 to the Harvard Forest tower which must be neutralized via a large local sink.

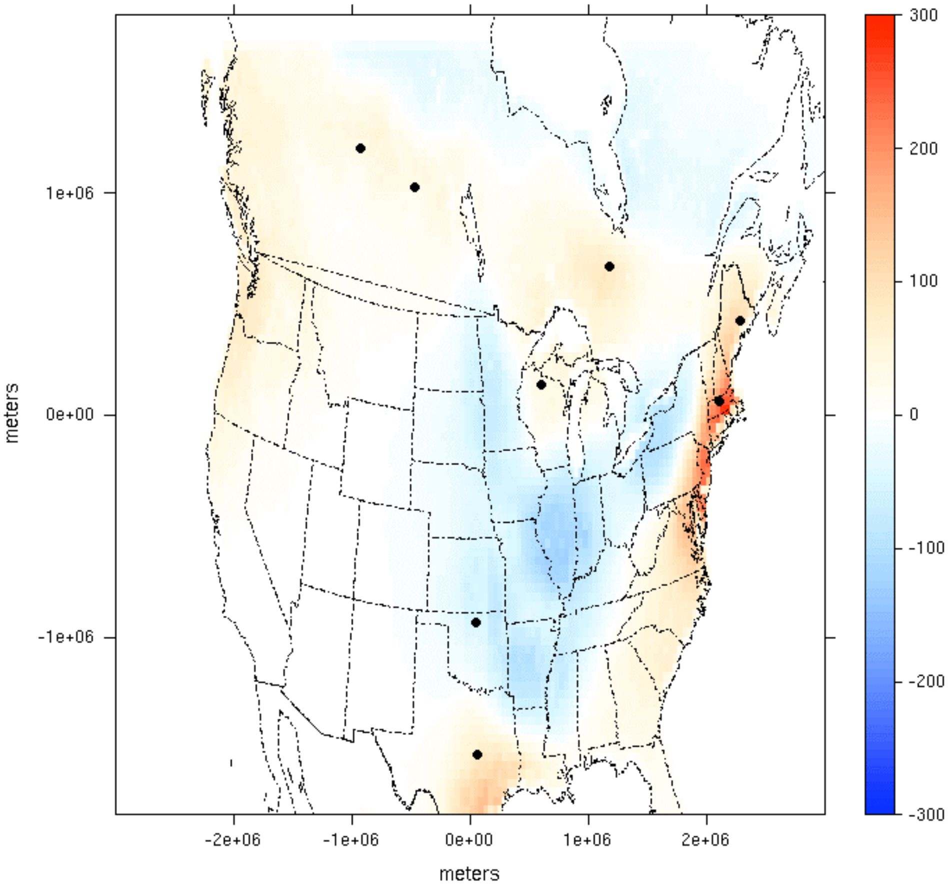

Figure 14: Difference in annual sink inferred by inversions based upon the Vulcan fossil fuel inventory and the Andres et al. [1995] fossil fuel inventory. Spatially-summed annual difference between Vulcan-based NEE estimate for 2004 and Andres[1995] based NEE estimate for 2004 is -0.05 PgC.

7. References Cited

Aranda, I., F. Pardo, L. Gil, J.A. Pardos, 2004: Anatomical Basis of the Change in Leaf Mass Per Area and Nitrogen Investment with Relative Irradiance Within the Canopy of Eight Temperate Tree Species. *Acta Oecologica*, **25**, 187-195.

Bond, B.J., B.T. Farnsworth, R.A. Coulombe, W.E. Winner, 1999: Foliage Physiology and Biochemistry in Response to Light Gradients in Conifers with Varying Shade Tolerance. *Oecologia*, **120**, 183-192.

Dai, Y., R.E. Dickinson, Y-P. Wang, 2004: A Two-Big-Leaf Model for Canopy Temperature, Photosynthesis and Stomatal Conductance. *Journal of Climate*, **17**(12), 2281-2299.

Dai, Y. and Qingcun, Z., 1997: A Land Surface Model (IAP94) for Climate Studies Part I: Formulation and Validation in Off-line Experiments. *Advances in Atmospheric Sciences*, **14**(4), 433-460.

de Pury, D.G.G. and G.D. Farquhar, 1997: Simple Scaling of Photosynthesis from Leaves to Canopies Without the Errors of Big-Leaf Models. *Plant, Cell and Environment*, **20**, 537-557.

Gu, L., D.D. Baldocchi, S.C. Wofsy, J.W. Munger, J.J. Michalsky, S.P. Urbanski, T.A. Boden, 2003. Response of a Deciduous Forest to the Mount Pinatubo Eruption: Enhanced

Photosynthesis. *Science*, **299**, 2035-2038.

Gu, L., D. Baldocchi, S.B.Verma, T.A. Black, T. Vesala, E.M.Falge, P.R. Dowty, 2002: Advantages of Diffuse Radiation for Terrestrial Ecosystem Productivity. *Journal of Geophysical Research*, **107** (D6), 10.1029/2001JD001242.

Gu, L. J.D. Fuentes, H.H. Shugart, R.M. Staebler, T.A. Black, 1999: Responses of Net Ecosystem Exchanges of Carbon Dioxide to Changes in Cloudiness: Results from Two North American Deciduous Forests. *Journal of Geophysical Research*, **104** (D24), 31421-31434.

Jordan, R., 1991: A One-Dimensional Temperature Model for a Snow Cover. United States Army Corps of Engineers, Cold Regions Research and Engineering Laboratory, Special Report 91-16.

Kull, O., 2002: Acclimation of Photosynthesis in Canopies: Models and Limitations. *Oecologia*, **133**, 267-279.

Le Roux, X., S. Grand, E. Dreyer, F-A. Daudet, 1999: Parameterization and Testing of a Biochemically Based Photosynthesis Model of Walnut (*Juglans regia*) Trees and Seedlings. *Tree Physiology*, **19**, 481-492.

Meir, P., B. Krujit, M. Broadmeadow, E. Barbosa, O. Kull, F. Carswell, A. Nobre, P.G. Jarvis, 2002: Acclimation of Photosynthetic Capacity to Irradiance in Tree Canopies in Relation to Leaf Nitrogen Concentration and Leaf Mass Per Unit Area. *Plant, Cell and Environment*, **25**, 343-357.

Pandey, S. and Kushawaha R., 2005: Leaf Anatomy and Photosynthetic acclimation in *Valeriana jatamansi* L. Grown Under High and Low Irradiance. *Photosynthetica*, **43**(1), 85-90.

Pons, T.L. and Anten, N.P.R., 2004: Is Plasticity in Partitioning of Photosynthetic Resources Between and Within Leaves Important for Whole-Plant Carbon Gain in Canopies? *Functional Ecology*, **18**, 802-811.

Strand, M. and Lundmark, T., 1995: Recovery of Photosynthesis in 1-Year-Old Needles of Unfertilized and Fertilized Norway Spruce (*Picea abies* (L.) Karst.) During Spring. *Tree Physiology*, **15**, 151-158.

Uliasz, M., 1994: Lagrangian particle dispersion modeling in mesoscale applications. In P. Zannetti, Editor, *Environmental Modeling II*, pages 71-102. Computational Mechanics Publications.

Vogg, G., R. Heim, J. Hansen, C. Schafer, E. Beck, 1998: Frost Hardening and Photosynthetic Performance of Scots Pine (*Pinus sylvestris* L.) Needles. I: Seasonal Changes in the Photosynthetic Apparatus and its Function. *Planta*, **204**, 193-200.

Wang, Y-P. and R. Leuning, 1998: A Two-Leaf Model for Canopy Conductance, Photosynthesis and Partitioning of Available Energy I: Model Description and Comparison with a Multi-Layered Model. *Agricultural and Forest Meteorology*, **91**, 89-111.

8 Publications

1. Corbin, K.D. and A.S. Denning, 2006: Using Continuous Data to Estimate Clear-Sky Errors in Inversions of Satellite CO_2 Measurements. *Geophysical Research Letters*, **33** (L12810), doi:10.1029/2006GL025910.
2. Wang, J.-W., A. S. Denning, L. Lu, I. T. Baker, K. D. Corbin, and K. J. Davis, 2007. Observations and simulations of synoptic, regional, and local variations in atmospheric CO_2. *J. Geophys. Res.,* 112, D04108, doi:10.1029/2006JD007410.
3. Peters, W., J.B. Miller, J. Whitaker, A.S. Denning, A. Hirsch, M.C. Krol, D. Zupanski, L. Bruhwiler, P.P. Tans, 2005. An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations. *Jour. Geophys. Res*. doi: 2005JD006157R.

4. Zupanski, D., A. S. Denning, M. Uliasz, M. Zupanski, A. E. Schuh, P. J. Rayner, W. Peters, and K. D. Corbin, 2007. Carbon flux bias estimation employing the Maximum Likelihood Esemble Filter (MLEF). *Jour. Geophys. Res.* **112**, D17107, doi:10.1029/2006JD008371.
5. Prihodko, L., A.S. Denning, N.P. Hanan, I. Baker, K. Davis, 2008. Sensitivity, uncertainty and time dependence of parameters in a complex land surface model. *Agric. Forest Meteorol.*, **148**, 268-287, doi:10.1016/j.agrformet.2007.08.006.
6. Corbin, K.D., A.S. Denning, J.-W. Wang, L. Lu, L. Prihodko, and I.T. Baker, 2008. Representation error in potential atmospheric CO_2 retrievals from satellites. *Jour. Geophys. Res.,* **113**, D02301, doi:10.1029/2007JD008716
7. Lauvaux, T., M. Uliasz, C. Sarrat, F. Chevallier, P. Bousquet, C. Lac, K. J. Davis, P. Ciais, A. S. Denning, and P. Rayner, 2008. Mesoscale inversion: first results from the CERES campaign with synthetic data. *Atmos. Chem. Phys.,* **8**, 3459-3471.
8. Schuh, A. E., A. S. Denning, M. Uliasz, K. D. Corbin, 2008. Seeing the Forest through the Trees: Recovering large scale carbon flux biases in the midst of small scale variability. Submitted to *Jour. Geophys. Res.*

9. Presentations

1. Wang, J.-W., A. S. Denning, L. Lu, I. T. Baker, and K.D. Corbin, 2005. Signals from synoptic CO2 variability and local ecosystem - a case study. Presented at the Seventh WMO International Carbon dioxide Conference, Boulder, USA.
2. Denning, A. S., Dusanka Zupanski, Marek Uliasz, Andrew Schuh, and Milija Zupanski , 2005. Regional carbon flux estimation using the Maximum Likelihood Ensemble Filter. Presented at the Seventh WMO International Carbon dioxide Conference, Boulder, USA.
3. Corbin, K.D., A.S. Denning, L. Lu, I. Baker, A. Wang, 2005. Using a high-resolution coupled ecosystem-atmosphere model to evaluate spatial, temporal, and clear-sky errors in satellite CO2 measurements. Presented at the Seventh WMO International Carbon dioxide Conference, Boulder, USA.
4. Schuh, A.E., M. Uliasz, S. Denning, and D. Zupanski, 2005. A case study in regional inverse carbon modeling. Presented at the Seventh WMO International Carbon dioxide Conference, Boulder, USA.
5. Baker, I.T., A.S. Denning, N. Hanan, J.A. Berry, G.J. Collatz, K.M. Schaeffer, A.W. Philpott, L. Prihodko, N.S. Suits, 2006: The Next Generation of the Simple Biosphere Model (SiB3): Model Formulation and Preliminary Results. Poster presented at iLEAPS 1st Science Conference, 21-26 January 2006, Boulder CO, USA.
6. Ciais, P, S. Piao, N. Viovy, C. Roedenbeck, P. Peylin, D. Baker, A. S. Denning, K. Davis, G. Nabuurs, M. Reichstein, R. Houghton, S. Zaehle, and M. Heimann, Top-down and bottom-up carbon budgets of North America, Europe and Asia. American Geophysical Union Fall 2006 Meeting, San Francisco CA, USA.
7. Denning, A.S., D. Zupanski, A. Schuh, K.D. Corbin, M. Uliasz, P. Rayner, M. Butler, S. Wofsy, S. Vay, 2006: Continental Carbon Cycle Data Assimilation in SiB-RAMS. American Geophysical Union Fall 2006 Meeting, San Francisco CA, USA.

8. Zupanski, D., 2006: Ensemble Kalman filter. *Advanced Numerics Seminar*, 8 March, 2006, Langen, Germany. (Invited)
9. Zupanski D., A.S. Denning, M. Uliasz, and A. E. Schuh, 2006: Carbon data assimilation using Maximum Likelihood Ensemble Filter (MLEF). *The Joint Workshop on NASA Biodiversity, Terrestrial Ecology, and Related Applied Sciences*, 21-25 August 2006, Adelphi, MD.
10. Zupanski D., S. Q. Zhang, A. Y. Hou, A.S. Denning, M. Uliasz,, L. Grasso and M. DeMaria, 2006: An ensemble-based approach for information content analysis. *The 2006 AGU Fall Meeting, A13 Data Assimilation session*. 11-15 December 2006, San Francisco, CA.
11. Zupanski D., A.S. Denning, and M. Uliasz,, 2006: Ensemble data assimilation applications to atmospheric and carbon cycle science. *Extended abstracts of The 86th AMS Annual Meeting, 10th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface,* 29 January – 2 February 2006, Atlanta, GA, 4pp.
12. Uliasz, M, Denning, S, Lu, L, Corbin, K, Zupanski, D, Miles, N, Richardson, S, and Davis, K J. Applications of Lagrangian Particle Transport Modeling in the Top-Down Regional CO2 Studies. Presented at the American Geophysical Union Fall 2007 Meeting, San Francisco CA, USA.
13. Schuh, A E, Denning, A S, Corbin, K D, Uliasz, M, Zupanski, D, and Parazoo, N C. Seeing the Forest Through the Trees: Investigating Signal to Noise Problems in Regional Atmospheric Inversions. Presented at the American Geophysical Union Fall 2007 Meeting, San Francisco CA, USA.