9-1-2011

Upscaling fluxes from towers to regions, continents and global scales using datadriven approaches

Jingfeng Xiao
University of New Hampshire, Durham, j.xiao@unh.edu

Kenneth J. Davis
Pennsylvania State University

Nathan M. Urban
Pennsylvania State University

Klaus Keller
Pennsylvania State University

Nicanor Z. Saliendra
University of Maryland at Baltimore County

Recommended Citation
Xiao, J. (2010). Upscaling fluxes from towers to regions, continents and global scales using datadriven approaches. FluxLetter, The Newsletter of FLUXNET, vol. 3, No. 3, December, 2010. https://dx.doi.org/10.1029/2010JG001568

Upscaling carbon fluxes from towers to the regional scale: Influence of parameter variability and land cover representation on regional flux estimates

Jingfeng Xiao,[1,2] Kenneth J. Davis,[1] Nathan M. Urban,[3,4] Klaus Keller,[3] and Nicanor Z. Saliendra[5]

Received 30 September 2010; revised 7 May 2011; accepted 25 May 2011; published 1 September 2011.

[1] Quantifying the current carbon cycle of terrestrial ecosystems requires that we translate spatially sparse measurements into consistent, gridded flux estimates at the regional scale. This is particularly challenging in heterogeneous regions such as the northern forests of the United States. We use a network of 17 eddy covariance flux towers deployed across the Upper Midwest region of northern Wisconsin and Michigan and upscale flux observations from towers to the regional scale. This region is densely instrumented and provides a unique test bed for regional upscaling. We develop a simple Diagnostic Carbon Flux Model (DCFM) and use flux observations and a data assimilation approach to estimate the model parameters. We then use the optimized model to produce gridded flux estimates across the region. We find that model parameters vary not only across plant functional types (PFT) but also within a given PFT. Our results show that the parameter estimates from a single site are not representative of the parameter values of a given PFT; cross-site (or joint) optimization using observations from multiple sites encompassing a range of site and climate conditions considerably improves the representativeness and robustness of parameter estimates. Parameter variability within a PFT can result in substantial variability in regional flux estimates. We also find that land cover representation including land cover heterogeneity and the spatial resolution and accuracy of land cover maps can lead to considerable uncertainty in regional flux estimates. In heterogeneous, complex regions, detailed and accurate land cover maps are essential for accurate estimation of regional fluxes.

Citation: Xiao, J., K. J. Davis, N. M. Urban, K. Keller, and N. Z. Saliendra (2011), Upscaling carbon fluxes from towers to the regional scale: Influence of parameter variability and land cover representation on regional flux estimates, *J. Geophys. Res.*, *116*, G00J06, doi:10.1029/2010JG001568.

1. Introduction

[2] Terrestrial ecosystems play an important role in regulating atmospheric carbon dioxide (CO_2) concentrations and the climate. Net ecosystem carbon exchange (NEE), the difference between photosynthetic uptake and release of CO_2 by respiration from autotrophs (plants) and heterotrophs (e.g., microbial decomposition), represents the net exchange of CO_2 between terrestrial ecosystems and the atmosphere. Quantifying NEE over regions can improve our understanding of the feedbacks between the terrestrial biosphere and the atmosphere.

[3] Several techniques have been used to estimate NEE. Atmospheric inversions [e.g., *Butler et al.*, 2010], ecosystem models [e.g., *Xiao et al.*, 2009], and inventory approaches [e.g., *State of the Carbon Cycle Report*, 2007] have been used to infer net exchange of CO_2 and provide aggregated information on NEE over large areas during the past two decades. The eddy covariance technique provides an alternative approach to estimate NEE. Eddy flux measurements of carbon fluxes are based on the covariance of high-frequency fluctuations in vertical velocity and CO_2 concentration [*Baldocchi et al.*, 1988]. Eddy covariance flux towers provide continuous measurements of ecosystem-level CO_2 exchange. However, these measurements only represent the fluxes from the scale of the tower footprint with longitudinal dimensions ranging between a hundred meters and several kilometers [*Schmid*, 1994]. To quantify

[1]Department of Meteorology, Pennsylvania State University, University Park, Pennsylvania, USA.
[2]Now at Complex Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, New Hampshire, USA.
[3]Department of Geosciences, Pennsylvania State University, University Park, Pennsylvania, USA.
[4]Now at Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, New Jersey, USA.
[5]Center for Urban Environmental Research and Education, University of Maryland at Baltimore County, Baltimore, Maryland, USA.

0148-0227/11/2010JG001568

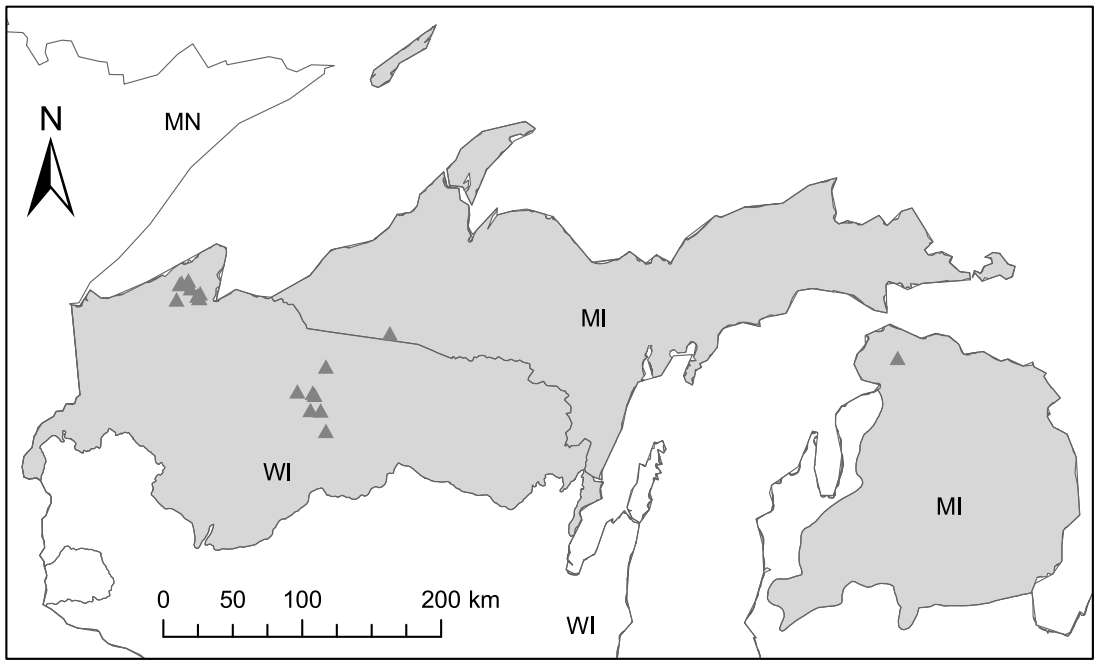

Figure 1. The study region and the location of eddy covariance flux towers. Symbols are the eddy flux sites. The shaded area is the northern forests ecoregion across northern Wisconsin (WI) and Michigan (MI). The dotted line stands for state boundaries.

the net exchange of CO_2 between the terrestrial biosphere and the atmosphere, we need to upscale these observations from towers to regions [*Davis*, 2008; *Xiao et al.*, 2008].

[4] Significant progress has been made in upscaling flux observations from towers to regional or continental scales during the last several years. Several different approaches have been used for the upscaling of eddy flux observations, including machine learning approaches [e.g., *Xiao et al.*, 2008; *Jung et al.*, 2009; *Zhang et al.*, 2011], light use efficiency models [e.g., *Mahadevan et al.*, 2008; *Cook et al.*, 2009], and empirical or process-based ecosystem models [e.g., *Sun et al.*, 2011]. Some of these studies have produced continuous flux fields and examined the terrestrial carbon dynamics over broad regions [e.g., *Xiao et al.*, 2010, 2011]. Significant uncertainty remains, however, regarding the impacts of parameter variability and land cover representation in regional upscaling.

[5] Plant functional type (PFT) is a key factor controlling terrestrial carbon fluxes. Many empirical and process-based diagnostic models simulate carbon fluxes of mature ecosystems. These models often use data from a single site to estimate the parameters of a given PFT. In reality, however, stands within a PFT have a range of stand age, disturbance history, and aboveground biomass. A single site may not represent the full range of stands within a PFT, and this type of model parameterization or parameter estimation may lead to significant uncertainties in the resulting flux estimates. Land cover is another key determinant of terrestrial carbon fluxes and is highly dependent on climate and human activities [*Hurtt et al.*, 2002; *Foley et al.*, 2005]. Land cover maps used in carbon cycle modeling are largely based on satellite data, and the properties and limitations of these maps will lead to flux uncertainties [*Reich et al.*, 1999; *Quaife et al.*, 2008]. Few studies have examined the impacts of parameter variability and land cover representation on regional flux estimates.

[6] The Upper Midwest region of northern Wisconsin and Michigan, United States, is a highly heterogeneous mixture of upland forests and lowland wetlands (Figure 1). This region is densely instrumented with eddy covariance flux towers as a result of the Chequamegon Ecosystem-Atmosphere Study (ChEAS). The ChEAS began with flux tower measurements collected at the WLEF tall tower [*Davis et al.*, 2003]. Since 1998, eddy flux tower systems have been deployed at 16 different sites spanning a range of ecosystem types and stand ages, in addition to the WLEF tall tower [*Cook et al.*, 2004; *Desai et al.*, 2005, 2008; *Gough et al.*, 2008; *Noormets et al.*, 2008; *Sulman et al.*, 2010]. The high density of sites covering most of the land cover types in the region makes the ChEAS region a unique test bed for the development and testing of the inferences and assumptions needed to diagnose and predict ecosystem carbon exchange at regional scales [*Chen et al.*, 2008].

[7] We used the ChEAS array of eddy covariance measurements, a simple diagnostic NEE model, and a data assimilation approach to produce gridded flux estimates at the regional scale. The objectives of this study are to (1) develop a simple Diagnostic Carbon Flux Model (DCFM) that integrates eddy covariance flux measurements, satellite observations, and climate data; (2) estimate model parameters using flux observations and a data assimilation approach; (3) produce continuous flux estimates across the region; and (4) assess the impacts of uncertainties of parameter variability and land cover representation on regional flux estimates. We hypothesize that (1) model parameters vary not only across PFTs but also within a

Table 1. Location and Site Characteristics of Eddy Covariance Flux Sites in the Chequamegon Ecosystem-Atmosphere Study (ChEAS) Region Across Northern Wisconsin (WI) and Upper Peninsula of Michigan (MI)[a]

PFT	Site	ID	State	Lat	Lon	Data Period	Stand Age (years)	Dominant Cover	Reference
Evergreen forests (EF)	Intermediate Red Pine	IRP	WI	46.687	−91.153	2003	30	Red pine	*Noormets et al.* [2008]
	Mature Red Pine	MRP	WI	46.739	−91.166	2002–2005	70	Red pine, aspen	*Noormets et al.* [2008]
	Red Pine Clearcut	RPC	WI	46.649	−91.069	2005	7	Red pine	*Noormets et al.* [2008]
	Young Jack Pine	YJP	WI	46.619	−91.081	2004–2005	22	Jack pine	*Noormets et al.* [2008]
	Young Red Pine	YRP	WI	46.619	−91.081	2002	17	Red pine, jack pine	*Noormets et al.* [2008]
Deciduous forests (DF)	Intermediate Hardwood	IH	WI	46.730	−91.233	2003	26	Aspen	*Noormets et al.* [2008]
	Riley Creek	RC	WI	45.910	90.116	2005–2006	10	Aspen	This study
	Thunder Creek	TC	WI	45.671	90.053	2005–2006	7	Aspen	This study
	Willow Creek	WC	WI	45.806	−90.080	2000–2006	70	Sugar maple, basswood, green ash	*Cook et al.* [2004]
	Young Hardwood Clearcut	YHC	WI	46.722	−91.252	2002	13	Aspen, red maple	*Noormets et al.* [2008]
Mixed forests (MF)	Park Falls/WLEF	WLEF	WI	45.946	−90.272	2000–2005	~45	Northern hardwoods, aspen	*Davis et al.* [2003]
	Sylvania Wilderness Area	SWA	MI	46.242	−89.348	2001–2006	200	Eastern hemlock, sugar maple, birch	*Desai et al.* [2005]
	University of Michigan Biological Station	UMBS	MI	45.560	−84.714	2000–2003	79	Aspen, white pine, red oak, sugar maple	*Gough et al.* [2008]
Shrublands (Sh)	Pine Barren 1	PB1	WI	46.625	−91.298	2002–2003		Sweet fern, black cherry, willow, red pine	*Noormets et al.* [2008]
Woody wetlands (WW)	Lost Creek	LC	WI	46.083	−89.979	2001–2006	45	Alder-willow shrubs	*Sulman et al.* [2010]
Herbaceous wetlands (HW)	Wilson Flowage	WF	WI	45.817	90.172	2005–2006		Sedges and marsh grass	*Sulman et al.* [2010]
	South Fork	SF	WI	45.925	90.131	2005–2006		Sphagnum bog with Labrador Tea and LeatherLeaf	*Sulman et al.* [2010]

[a]The complexity of the WLEF flux footprint [e.g., *Desai et al.*, 2008] makes it difficult to assign a single stand age. This value should be viewed with caution.

given PFT, suggesting that multiple flux towers improve characterization of a PFT; and (2) land cover representation such as land cover heterogeneity and the spatial resolution and accuracy of land cover maps can lead to significant uncertainties in regional flux estimates.

2. Data

2.1. Study Region and Site Descriptions

[8] The Upper Midwest region of northern Wisconsin and Michigan, United States (Figure 1) is an area of temperate/subboreal forests and glaciated landforms with many small glacial lakes and wetlands. The majority of upland forests consist of mature northern hardwood forests (e.g., maple, basswood, birch, and ash) and younger fast growing aspen (*Populus termulouides*) forests; coniferous species include red pine, jack pine, eastern hemlock and white pine forests cover smaller areas [*Desai et al.*, 2008]. Around 1/3 of the region is lowland wetlands, including forested wetlands (e.g., black spruce, white cedar or tamarack), shrub wetlands (alder or willow species), and open meadows [*Desai et al.*, 2008].

[9] We used 17 eddy flux sites across the region (Table 1 and Figure 1): the 447m Park Falls/WLEF-TV tall tower [*Davis et al.*, 2003; *Ricciuto et al.*, 2008], three other AmeriFlux towers (Sylvania Wilderness Area [*Desai et al.*, 2005]; Lost Creek [*Sulman et al.*, 2010]; and Willow Creek [*Cook et al.*, 2004]), four "roving" towers where flux instruments were shared between pairs of sites for a 2 year field campaign (two herbaceous wetland and two young aspen), all located in the Medford–Park Falls district of the Chequamegon-Nicolet National Forest in north central Wisconsin, seven upland chronosequence sites in the Washburn District of the Chequamegon-Nicollet National Forest in northwestern Wisconsin and an additional pine barrens site in the Upper Peninsula of Michigan [e.g., *Noormets et al.*, 2008], and the University of Michigan Biological Station's AmeriFlux tower (UMBS) in the northern portion of Lower Michigan [e.g., *Gough et al.*, 2008; *Su et al.*, 2008]. These 17 sites were grouped into the following PFTs: evergreen forests (EF, five sites), deciduous forests (DF, five sites), mixed forests (MF, three sites), shrublands (Sh, one site), woody wetlands (WW, one site), and herbaceous wetlands (HW, two sites) (Table 1).

2.2. Eddy Flux Observations

[10] We used the half-hourly data from the AmeriFlux Level 4 product for all sites (Table 1) except the roving towers: Riley Creek (RC), Thunder Creek (TC), Wilson Flowage (WF), and South Fork (SF). A negative sign denotes carbon uptake, and a positive sign denotes carbon release. We aggregated the gap-filled half-hourly NEE and climate data to daily values. Only days with no less than 75% of original half-hourly measurements were used in our analysis.

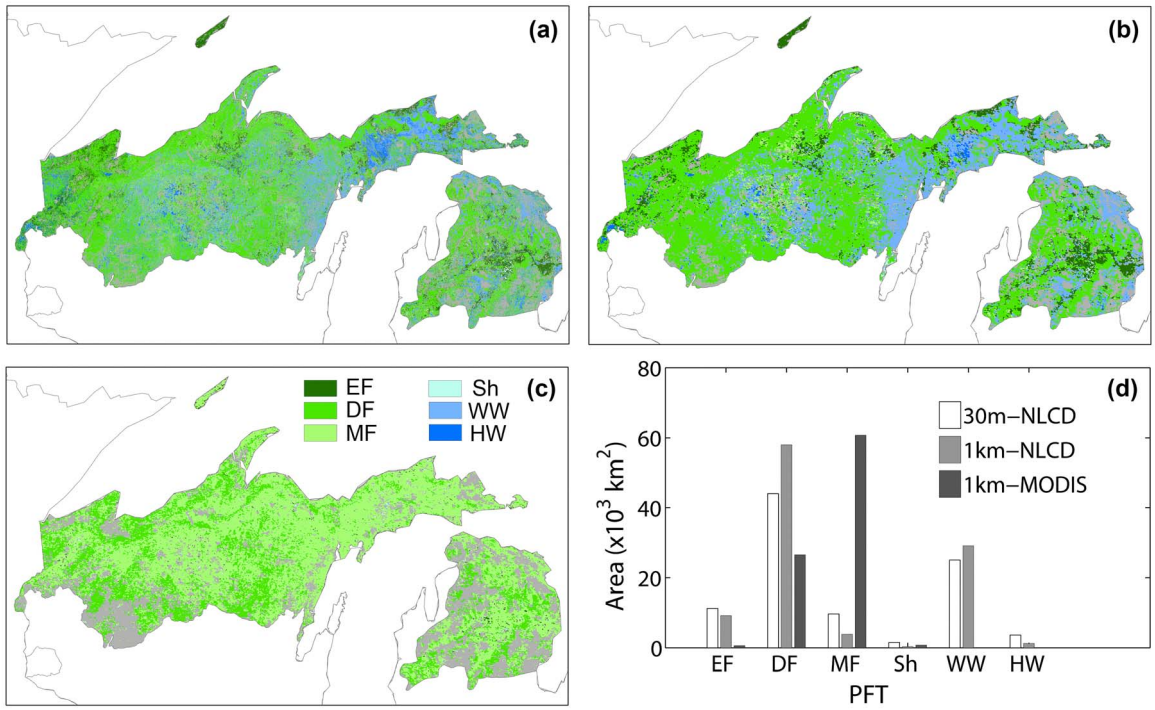

Figure 2. Land cover maps used for the study region: (a) 30 m NLCD (30m-NLCD); (b) 1 km NLCD (1km-NLCD); (c) 1 km MODIS (1km-MODIS); and (d) total area of each plant functional type (PFT): evergreen forests (EF), deciduous forests (DF), shrublands (Sh), woody wetlands (WW), and herbaceous wetlands (HW) for each land cover map. Other land cover types (e.g., urban, croplands, and barren) are shown in gray.

[11] For the roving towers, we used the mean diurnal variation (MDV) method [*Falge et al.*, 2001] to fill data gaps and then aggregated the half-hourly data to daily values. In this method, a missing observation is replaced by the mean of observations for that time interval (half hour) from adjacent days. A window size of 8 days was chosen for averaging in this study. We used the MDV method because of its simplicity and consistent performance [*Moffat et al.*, 2007]. In addition, we only used days with no less than 75% of good half-hourly measurements, which minimized the effects of gap-filling.

2.3. MODIS Data

[12] We used the vegetation indices (MOD13A2) [*Huete et al.*, 2002] and surface reflectance (MOD09A1) [*Vermote and Vermeulen*, 1999] products derived from MODIS observations. For each eddy flux site, we obtained MODIS ASCII subsets (Collection 5) for both products from the Oak Ridge National Laboratory's Distributed Active Archive Center (ORNL DAAC). The MODIS ASCII subsets consist of 7 km × 7 km regions centered on the flux tower. For each variable, we extracted average values for the central 3 km × 3 km area within the 7 km × 7 km cutouts to better represent the flux tower footprint [*Xiao et al.*, 2008].

[13] For the entire ChEAS region, we obtained MODIS surface reflectance (MOD09A1) and vegetation indices (MOD13A2) from the Earth Observing System (EOS) Data Gateway for each 8 day interval over the period 2000–2007. For each variable, we determined the quality of the value of each pixel within the area using the quality assurance (QA) flags included in the products [*Xiao et al.*, 2008], and replaced the bad value using a linear interpolation approach [*Zhao et al.*, 2005].

2.4. Land Cover Maps

[14] To examine the influence of land cover representation on regional flux estimates, we used land cover maps with different spatial resolution and from two different sources. The ChEAS region is characterized by heterogeneous landscapes, and the dominant scale for landscape variability is about a few hundred meters. We used the NLCD (National Land Cover Data Set) 2001 land cover map [*Homer et al.*, 2004] with 30 m resolution (Figure 2a) as the main land cover map for the study, and this land cover map is referred to as 30m-NLCD hereafter. We also aggregated the 30 m land cover map to 1 km spatial resolution (Figure 2b) by selecting the dominant PFT for each 1 km cell, and the resulting 1 km NLCD land cover map is referred to as 1km-NLCD. These two maps exhibited significant differences in the areas of PFTs (Figure 2d). For instance, the areas of deciduous forests on 1km-NLCD is 31.8% higher than that of 30m-NLCD, while the area of mixed forests is 60.5% lower than that of 30m-NLCD.

[15] We also used the 1 km MODIS land cover map (Figure 2c) with the IGBP (International Geosphere-Biosphere Programme) classification scheme [*Friedl et al.*, 2002]. This land cover map is referred to as 1km-MODIS hereafter. The MODIS land cover map was derived from moderate-resolution MODIS data and a global land cover classification algorithm. The NLCD and MODIS land cover

maps were used to examine how regional flux estimates are influenced by the classification accuracy and spatial resolution of land cover maps and land cover heterogeneity. We used the following PFTs for all land cover maps: evergreen forests, deciduous forests, mixed forests, shrublands, woody wetlands, and herbaceous wetlands.

[16] The NLCD and MODIS land cover maps exhibited considerable differences in the distribution and areas of PFTs (Figure 2). The dominant PFT of NLCD maps is deciduous forests, while the dominant PFT of 1km-MODIS is mixed forests. 1km-MODIS shows that wetlands only account for a negligible portion (~0.01%) of the region, while the NLCD land cover maps show that nearly 1/3 of the vegetated area is wetlands, mainly woody wetlands. Most mixed forests on 1km-MODIS are classified as woody wetlands on the NLCD maps. In addition, 1km-MODIS shows only 620 km^2 of evergreen forests, while both NLCD land cover maps show nearly 10,000 km^2 of evergreen forests.

2.5. Climate Data

[17] We obtained daily air temperature data from the NCEP/NCAR Reanalysis data set (~2.5 degree resolution) [*Kalnay et al.*, 1996]. The NCEP/NCAR reanalysis project data set was created by assimilating climate observations from a wide variety of sources, such as weather stations, ships, aircrafts, and satellites. We also used daily 0.5° resolution PAR data from the Surface Radiation Budget (SRB) project, as generated by the National Oceanic and Atmospheric Administration (NOAA), National Environmental Satellite, Data and Information Service (NESDIS) [*Pinker et al.*, 2002]. The surface radiation fluxes were derived from Geostationary Operational Environmental Satellites (GOES).

2.6. Independent Flux Estimates

[18] We used independent regional flux estimates for comparison purposes. We obtained the MODIS GPP product (MOD17A3) [*Running et al.*, 2004; *Zhao et al.*, 2005]. This product is based on a light use efficiency model driven by MODIS and meteorological data. It provides GPP estimates at a spatial resolution of 1 km at the global scale. We obtained MODIS GPP data for the ChEAS region over the period 2001–2006.

[19] We also used the EC-MOD flux fields derived from eddy covariance (EC) flux measurements and MODIS data [*Xiao et al.*, 2008, 2010, 2011]. EC-MOD consists of continuous GPP and NEE estimates with 1 km spatial resolution and 8 day time step for the conterminous U.S. over the period 2000–2006. EC-MOD was developed from flux observations from 42 eddy covariance flux towers encompassing a wide range of ecosystem and climate types and wall-to-wall MODIS data streams using a data-driven approach [*Xiao et al.*, 2008, 2010]. We calculated annual GPP and NEE from 8 day EC-MOD flux estimates for the ChEAS region over the period 2001–2006.

3. Methods

3.1. Model Framework

[20] We developed a simple Diagnostic Carbon Flux Model (DCFM) for the estimation of NEE. In our model, NEE is the difference of two carbon fluxes – gross primary productivity (GPP) and ecosystem respiration (R_e):

$$NEE = -\varepsilon_{\max} \times PAR \times fPAR \times W_s \times T_s + \left(R'_{ref} + \gamma \times AGB + \lambda \times GPP\right) \times e^{E_0\left(1/\left(T_{ref}-T_0\right)-1/\left(T-T_0\right)\right)} \quad (1)$$

where $\varepsilon_{\max}$ is the maximum light use efficiency (LUE) (g C m^{-2} MJ^{-1} APAR), PAR is the incident photosynthetically active radiation (MJ m^{-2}) per time period (e.g., day or month), fPAR is the fraction of PAR absorbed by vegetation canopies, W_s is the water scalar, T_s is the temperature scalar, R'_{ref} is a parameter associated with the rate of respiration at the reference temperature, T_{ref} is the reference temperature, E_0 is an activation energy parameter that determines the temperature sensitivity, and T_0 is a constant regression parameter. W_s and T_s represent the limiting effects of water availability and temperature on GPP, respectively, and both scalars vary from 0 to 1. T_{ref} is set to 10°C, and T_0 is kept constant at −46.02°C as by *Lloyd and Taylor* [1994]. Negative NEE values denote carbon uptake, while positive values denote carbon release to the atmosphere.

[21] We used a LUE approach to estimate GPP. The LUE or "radiation use efficiency" logic is one of the most frequently applied concepts for modeling GPP [e.g., *Prince and Goward*, 1995; *Ruimy et al.*, 1996; *Landsberg and Waring*, 1997; *Veroustraete et al.*, 2002; *Running et al.*, 2004; *Xiao et al.*, 2004; *Cook et al.*, 2009; *Mahadevan et al.*, 2008]. This logic was first proposed by *Monteith* [1972, 1977], suggesting that the NPP of well-watered and fertilized annual crop plants was linearly related to the amount of absorbed photosynthetically active solar radiation (PAR).

[22] We used enhanced vegetation index (EVI) to approximate fPAR. EVI is calculated as [*Huete et al.*, 1997]:

$$EVI = 2.5\frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + (6\rho_{red} - 7.5\rho_{blue}) + 1} \quad (2)$$

where ρ_{red}, ρ_{nir}, and ρ_{blue} are the visible-red, near-infrared, and blue reflectance, respectively. A linear transformation of EVI was used to approximate fPAR:

$$fPAR = \alpha + \beta \times EVI \quad (3)$$

where α and β are empirical constants.

[23] We used the normalized difference water index (NDWI) [*Gao*, 1996] to calculate W_s. *Gao* [1996] developed the NDWI from satellite data to measure vegetation liquid water. The NDWI was shown to be strongly correlated with leaf water content [*Jackson et al.*, 2004] and soil moisture [*Fensholt and Sandholt*, 2003] over time. W_s is calculated as [*Xiao et al.*, 2004]:

$$W_s = \frac{1 + NDWI}{1 + NDWI_{\max}} \quad (4)$$

where $NDWI_{\max}$ is the maximum NDWI for each individual site or pixel over the period that MODIS data are available. NDWI and W_s were calculated using band 2 and band 6 of the MODIS surface reflectance product.

Table 2. Values of Parameters Associated With the Temperature Scalar for Each Plant Functional Type (PFT)[a]

Parameter	T_{min} (°C)	$T_{opt,min}$ (°C)	$T_{opt,max}$ (°C)	T_{max} (°C)
Evergreen forests (EF)	−1.0	17.0	30.0	34.0
Deciduous forests (EF)	0	17.0	30.9	34.0
Mixed forests (MF)	−0.5	17.0	30.5	34.0
Shrublands (Sh)	1.0	15.1	35.1	44.0
Woody wetlands (WW)	0	19.1	33.1	38.0
Herbaceous wetlands (HW)	0	13.0	32.7	38.0

[a]These parameters are fixed in parameter estimation.

[24] We used the temperature scalar implemented in the Terrestrial Ecosystem Model (TEM) [*Tian et al.*, 1999; *Xiao et al.*, 2009]:

$$T_s = \begin{cases} 0, & T < T_{\min} \\ \dfrac{(T-T_{\min})(T-T_{\max})}{(T-T_{\min})(T-T_{\max}) - \left(T-T_{opt,\min}\right)^2}, & T_{\min} \le T < T_{opt,\min} \\ 1 & T_{opt,\min} \le T < T_{opt,\max} \\ \dfrac{(T-T_{\min})(T-T_{\max})}{(T-T_{\min})(T-T_{\max}) - \left(T-T_{opt,\max}\right)^2}, & T_{opt,\max} \le T \le T_{\max} \\ 0, & T > T_{\max} \end{cases} \quad (5)$$

where T is the mean air temperature (C), $T_{\min}$ and $T_{\max}$ are the maximum and minimum constraints for GPP, and $T_{opt,\min}$ and $T_{opt,\max}$ represent the range of temperature for optimal carbon uptake. $T_{\min}$, $T_{\max}$, $T_{opt,\min}$ and $T_{opt,\max}$ values were obtained from TEM (Table 2).

[25] Ecosystem respiration (R_e) has been widely modeled as an exponential function of either air or soil temperature (T) [*Lloyd and Taylor*, 1994]:

$$R_e = R_{ref} e^{E_0\left(1/\left(T_{ref}-T_0\right)-1/(T-T_0)\right)} \quad (6)$$

where R_{ref} is the rate of respiration at the reference temperature. This exponential function, however, could not explain the variability of carbon pools within a PFT. Autotrophic respiration can be empirically modeled as a function of air temperature and tissue carbon [*Tian et al.*, 1999; *Xiao et al.*, 2009]. Plant respiration is also a relatively constant proportion of GPP [*Chapin et al.*, 2002]. In our NEE model, we modified this function for the estimation of R_e by including aboveground biomass and GPP as explanatory variables (equation (1)). The rate of respiration at the reference temperature in our model is $R'_{ref} + \gamma \times AGB + \lambda \times GPP$.

[26] Several lines of evidence showed that vegetation indices integrated over the growing season exhibited moderate to strong relationships with aboveground biomass for a variety of PFTs, such as forests [e.g., *Myneni et al.*, 2001], grasslands [e.g., *Tucker et al.*, 1985], crops [e.g., *Persson et al.*, 1993], and tundra [e.g., *Boelman et al.*, 2003]. We used integrated EVI over the growing season as a proxy for aboveground biomass so that we could produce flux estimates for the entire study region.

[27] DCFM is run using a daily time step. The choice of this simple model structure was motivated by the ability to evaluate the variability of parameters both within and across PFTs using a high density and range of eddy flux observations and global optimization methods. This simple diagnostic model does not require a detailed mechanistic understanding of complex ecosystem processes, and thus can be easily optimized using eddy flux measurements. We used the following broad PFTs: evergreen forests, deciduous forests, mixed forests, shrublands, woody wetlands, and herbaceous wetlands. Mixed forests are treated as a PFT here as this vegetation type is the dominant land cover type for the MODIS land cover map.

[28] Disturbance is an important factor controlling the sizes of forest carbon pools and the quantity of litterfall. The legacy of disturbance thus affects heterotrophic respiration of forest ecosystems. In our simple diagnostic model, we use aboveground biomass to account for the spatial variability of respiration over space within a given PFT. However, aboveground biomass cannot account for the pools of litterfall and soil carbon. We thus conducted an experiment by adding stand age as an additional variable for the estimation of R_e to examine the effects of disturbance legacy on the estimation of R_e and NEE. For forests, the modified model for this experiment can be written as:

$$\begin{aligned} NEE = & -\varepsilon_{\max} \times PAR \times fPAR \times W_s \times T_s \\ & + \left(R'_{ref} + \gamma \times AGB + \delta \times Age + \lambda \times GPP\right) \\ & \times e^{E_0\left(1/\left(T_{ref}-T_0\right)-1/(T-T_0)\right)} \end{aligned} \quad (7)$$

where *Age* is stand age in years. This modified model was only used to examine whether the addition of stand age can improve the performance of the model.

3.2. Parameter Estimation

[29] We used the differential evolution (DE) algorithm [*Price et al.*, 2006] to estimate the parameters of DCFM. DE is a stochastic, population-based optimization algorithm. One key advantage of DE is that it seeks the global minimum of a multidimensional and multimodal (i.e., exhibiting more than one minimum) function relatively fast and with high reliability. We used the DEoptim package [*Ardia*, 2009] implemented in the R statistical package. DEoptim is a R-vectorized variant of the DE algorithm.

[30] In this analysis, seven parameters ($\varepsilon_{\max}$, α, β, R'_{ref}, γ, λ, and E_0) were allowed to vary in parameter estimation. The lower and upper bounds of $\varepsilon_{\max}$ were determined from the range of $\varepsilon_{\max}$ (0.39 − 2.75 g C m MJ^{-1} APAR) used in LUE models [*Prince and Goward*, 1995; *Ruimy et al.*, 1996; *Landsberg and Waring*, 1997; *Veroustraete et al.*, 2002; *Xiao et al.*, 2004; *Potter et al.*, 2007]. The initial ranges of α and β for the approximation using EVI here were specified as [−0.5, −0.3] and [1.5, 1.85], respectively. The bounds of other parameters were largely based on conventional knowledge of their possible values. The initial range of γ was specified as −1.0 − 1.0 g C m^{-2} d^{-1}. λ is dimensionless, and its initial range was specified as [−1.0, 1.0]. We specified the initial range of E_0 as 0–500 °C. R'_{ref} is nonnegative, and was assigned the range 0–20 g C m^{-2} d^{-1}.

[31] For each PFT, we conducted cross-site (or joint) optimization using eddy flux observations from all sites within the PFT. The resulting model parameters were used for the estimation of regional carbon fluxes. We also con-

Table 3. Estimated Parameter Values for Each Pant Functional Type (PFT)

PFT	ε_{max} (g C m^{-2} MJ^{-1} APAR)	α	β	R'_{ref} (g C m^{-2} d^{-1})	γ (g C m^{-2} d^{-1})	λ	E_0 (°C)
Evergreen forests (EF)	2.75	−0.30	1.85	13.14	−0.099	0.71	68.78
Deciduous forests (DF)	2.39	−0.47	1.85	0	0.011	0.47	142.77
Mixed forests (MF)	0.71	−0.31	1.80	2.07	−0.005	−0.01	277.80
Shrubland (Sh)	1.24	−0.40	1.85	14.43	−0.143	0.07	355.95
Woody wetland (WW)	0.95	−0.31	1.83	4.40	−0.024	0.18	214.27
Herbaceous wetland (HW)	2.75	−0.30	1.85	0	0.003	0.78	77.26

ducted leave-one-out model optimization for each forest PFT to examine the performance of the model and to examine how model parameters vary across sites within a PFT. For each PFT, one site was excluded at a time and data from all other sites within the PFT were used to estimate the parameters.

[32] Similarly, we used DE to estimate the parameters of the modified model accounting for disturbance effects (equation (7)). We examined whether the addition of stand age can improve the performance of the model. This modified model was not applied to the region to create regional flux estimates because there is no spatially explicit information on stand age available for the region to date.

3.3. Regional Flux Prediction

[33] We applied the optimized model to the study region to produce regional flux estimates. To reduce the computational complexity of the spatial prediction, the 30m-NLCD (Figure 2a) was downgraded to 1 km spatial resolution. For each 1 km cell, the percent cover of each PFT was calculated. Carbon fluxes were estimated for each PFT within each 1 km cell, and total fluxes for each cell were calculated by summing up the fluxes for different PFTs weighted by their percent cover. We produced continuous estimates of carbon fluxes with 1 km spatial resolution and daily time step for the period 2001–2007.

[34] To examine the influence of land cover representation on regional flux estimates, we also used 1km-NLCD (Figure 2b) and 1km-MODIS (Figure 2c) to specify the PFT of each pixel and produce continuous estimates of carbon fluxes for the region, respectively. We then examined how land cover representation including land cover heterogeneity and the spatial resolution and accuracy of land cover maps affect regional flux estimates.

[35] To examine the influence of parameter variability on regional flux estimates, we also used parameter estimates resulting from leave-one-out model optimizations to produce regional estimates of carbon fluxes. That is, we estimated model parameters for each PFT multiple times by removing one site at a time from the collection of flux towers within that PFT, and then conducted multiple model simulations for evergreen forests, deciduous forests, and mixed forests using the multiple sets of parameter values.

4. Results

4.1. Parameter Estimation and Variability

[36] The estimated parameter values are given in Table 3. The parameter behaviors generally fell into two categories: well-constrained or edge-hitting. Well-constrained parameters exhibited optimal values within the range of their initial values. All parameters for mixed forests, shrublands, and woody wetlands and most parameters for other PFTs fell into this category. Three parameters, γ, λ, and E_0, were well constrained for all PFTs. The remaining parameters, ε_{max}, α, β, and R'_{ref}, however, hit the edge of their prior ranges for some PFTs. This means that the estimated values were equal to one of the edges of their initial ranges, either the lower or upper bounds. For example, the optimized values of ε_{max} for evergreen forests and herbaceous wetlands were equal to the upper bounds of the initial values; R'_{ref} of deciduous forests and herbaceous wetlands exhibited values equal to the lower bounds of their initial values. When edge-hitting occurred, widening the range on the edge-hit side within reasonable bounds did not lead to well-constrained parameters but simply shift the estimated value in that direction.

[37] The estimated values of α and β were relatively consistent among PFTs. The slope of the fPAR-EVI relationship, β, was around 1.8. Most parameters, however, varied substantially with PFT. For example, the estimated value of ε_{max} of mixed forests was generally much lower than those of evergreen forests and deciduous forests. Our results showed that ε_{max} converged within our prior bounds for deciduous forests, mixed forests, shrublands, and woody wetlands, and hit the upper bound for evergreen forests and herbaceous wetlands. The estimated value of R'_{ref} varied with PFT, and this parameter hit the lower bounds for deciduous forests and herbaceous wetlands. The parameter γ converged within bounds for all PFTs and exhibited negative values for all PFTs except deciduous forests and herbaceous wetlands. The parameter λ also converged within bounds for all PFTs. λ exhibited positive values for all PFTs except mixed forests. The activation energy parameter, E_0, exhibited a large range across PFTs and was within bounds for all PFTs.

[38] The leave-one-out model optimization showed that parameter values varied not only across PFTs but also across sites within a given PFT (Table 4). For each PFT, one site was excluded at a time and data from all other sites within the PFT were used to estimate the parameters; the optimized model was then used to predict NEE for the site excluded. The R^2 of the leave-one-out verifications of evergreen forests varied between 0.48 and 0.66 except the parameter estimation with Intermediate Red Pine (IRP) excluded. The verification for mixed forests also exhibited intermediate correlations with R^2 between 0.37 and 0.55. The R^2 for deciduous forests, however, showed larger variability (0 to 0.66). Within a given PFT, most parameters were generally consistent among leave-one-out optimizations; some para-

Table 4. Leave-One-Out Parameter Estimation and Model Verification for Three Plant Functional Types (PFTs): Evergreen Forests, Deciduous Forests, and Mixed Forests[a]

Site	ε_{max} (g C m^{-2} MJ^{-1} APAR)	α	β	R'_{ref} (g C m^{-2} d^{-1})	γ	λ (g C m^{-2} d^{-1})	E_0 (°C)	Slope	Intercept (g C m^{-2} d^{-1})	R^2	N
Evergreen forests (EF)											
IRP	2.75	−0.30	1.85	13.12	−0.099	0.71	69.02	0.55	−0.40	0.04	4
MRP	1.15	−0.16	1.48	0.011	0.002	0.20	283.72	0.41	0.31	0.49	276
RPC	2.75	−0.30	1.85	16.85	−0.123	0.71	68.75	0.46	0.34	0.57	51
YJP	2.75	−0.30	1.85	15.77	−0.116	0.62	93.21	0.99	−2.89	0.48	131
YRP	2.05	0	1.85	13.33	−0.010	0.73	60.65	0.44	−1.36	0.66	97
Deciduous forests (DF)											
IH	2.70	−0.52	1.97	0	0.010	0.57	102.07	0.37	−1.73	0.61	20
RC	1.93	−0.42	1.83	0	0.013	0.37	181.51	−0.13	2.21	0.16	20
TC	2.01	−0.40	1.77	0	0.013	0.37	180.63	0.02	2.14	0.00	21
WC	0.61	−0.65	1.66	0	0.003	0.04	500.00	0.21	0	0.66	518
YHC	2.36	−0.55	1.92	0	0.009	0.48	105.07	1.02	−2.40	0.33	68
Mixed forests (MF)											
SWA	0.65	−0.34	1.86	2.61	−0.009	−0.06	309.85	0.44	0.12	0.44	390
UMBS	0.71	−0.31	1.79	2.07	−0.005	−0.01	277.81	0.56	0.02	0.55	110
WLEF	0.90	−0.35	1.74	7.01	−0.045	0.06	230.37	0.61	−0.47	0.37	771

[a]The estimated parameter values and statistics (slope, intercept, R^2, and the number of days of observations – N) for the correlation between predicted and observed net ecosystem exchange (NEE) are given here.

meters, however, varied substantially with optimization. For example, the estimated value of ε_{max} with Mature Red Pine (MRP) site excluded was much lower than that of any other leave-one-out optimization within evergreen forests; the estimated value of ε_{max} with Willow Creek (WC) excluded was much lower than that of any other optimization within deciduous forests. Mixed forests exhibited the least within-PFT variability in ε_{max}. The two parameters α and β generally exhibited little within-PFT variability, and had consistent values among leave-one-out optimizations.

[39] To examine the effects of disturbance on parameter estimation and flux estimation, we included stand age as an additional variable for the prediction of ecosystem respiration (equation (7)). Similarly, we estimated the parameters of this modified model using DE. The addition of stand age improved the performance of the model for estimating NEE for deciduous forests and evergreen forests. For evergreen forests, the inclusion of stand age increased R^2 from 0.45 to 0.69 and reduced the root mean squared error (RMSE) by 23.3%. For deciduous forests, the addition of stand age increased R^2 from 0.61 to 0.65 and reduced RMSE by 6.0%. For mixed forests, however, the modified model did not significantly improve the model performance, and both the increase in R^2 and the reduction in RMSE were negligible.

4.2. Influence of Parameter Variability and Land Cover Representation on Regional Flux Estimates

[40] We produced daily flux estimates across the region for the period 2001–2007 using 30m-NLCD to specify the PFTs and percent cover within each cell, and then calculated annual GPP and NEE for each year. Figure 3 shows mean annual GPP and NEE over the 7 year period. Annual GPP

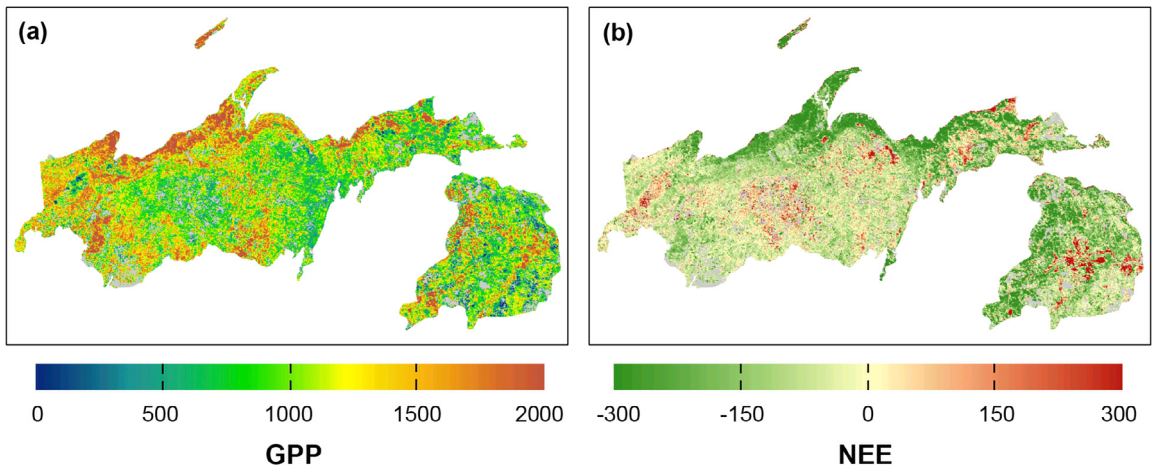

Figure 3. Mean annual carbon fluxes for the study region over the period 2001–2007: annual (a) GPP; and (b) NEE. The units are g C m^{-2} yr^{-1}. The estimation of these carbon fluxes is based on the 30m NLCD land cover map (30m-NLCD).

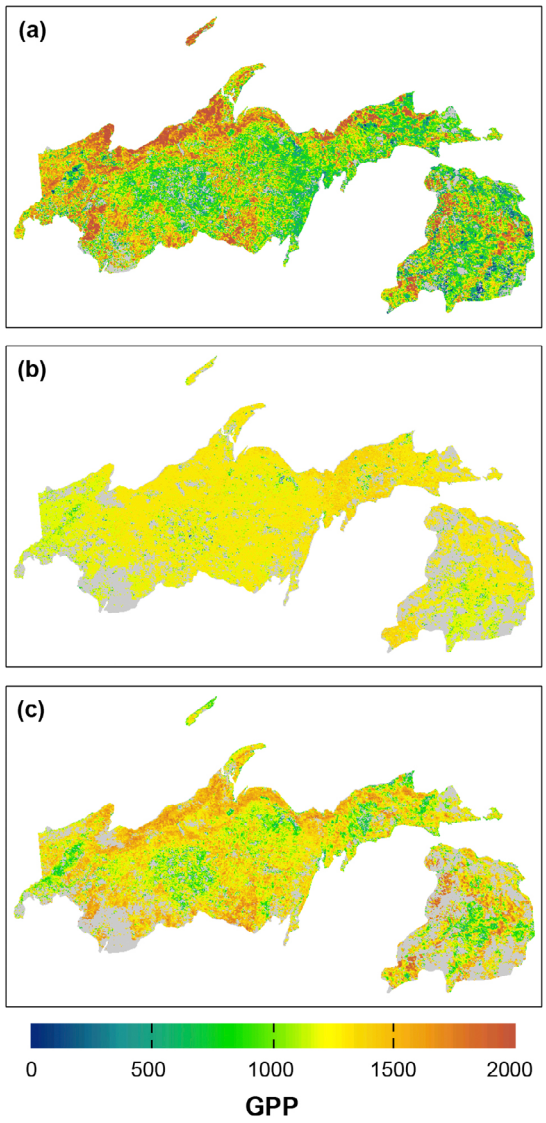

Figure 4. Comparison of mean annual GPP over the period 2001–2006 for the ChEAS region: (a) our GPP estimate from this study; (b) MODIS GPP product; and (c) EC-MOD GPP. The units are g C m^{-2} yr^{-1}.

varied substantially over space across the northern forests ecoregion. The northern portion of northern Wisconsin and Upper Peninsula of Michigan and some areas in the central and southern part of the region and Lower Peninsula of Michigan exhibited high annual GPP, while the remaining region showed relatively low to moderate annual GPP. Mean annual NEE also varied substantially over space. The northern portion of Wisconsin and Upper Peninsula of Michigan absorbed carbon, while areas along the border between Wisconsin and Upper Michigan and central Lower Michigan released carbon into the atmosphere; the remaining part of the study region was nearly carbon neutral.

[41] We compared our mean annual GPP with estimates derived from MODIS GPP product and EC-MOD over the period 2001–2006 (Figure 4). MODIS GPP exhibited much less spatial variability than our estimate and EC-MOD. The spatial patterns of our estimate and EC-MOD GPP were generally similar to each other. For deciduous and evergreen forests, our estimate had slightly higher GPP than EC-MOD; for mixed forests and other PFTs, our estimate had lower GPP than EC-MOD. We also compared our annual NEE with EC-MOD NEE (Figure 5). Similar to GPP, the spatial patterns of annual NEE generally agreed with each other. In the north of the study region, both estimates exhibited large carbon uptake with annual NEE of approximately −300 g C m^{-2} yr^{-1}. In the southern portion of the region, however, our estimate showed that ecosystems were nearly carbon neutral, while EC-MOD showed large carbon uptake with annual NEE of approximately −300 g C m^{-2} yr^{-1}. Both estimates showed that other areas were nearly carbon neutral or released carbon into the atmosphere.

[42] We calculated the total annual GPP and NEE of the entire region and total annual fluxes for each PFT (Figure 6). Total annual GPP over the ChEAS region was estimated to be 120.3 Tg C yr^{-1}, which was largely contributed by deciduous (57.5%) and evergreen forests (17.7%). Woody wetlands had intermediate GPP (16.7 Tg C yr^{-1}), and the remaining PFTs (mixed forests, shrublands, and herbaceous wetlands) accounted for 10.9% of the regional GPP. The regional annual NEE was estimated to be −9.8 Tg C yr^{-1}. Similar to annual GPP, annual NEE also varied substantially with PFT. Deciduous forests had the highest net carbon uptake, followed by shrublands and evergreen forests.

[43] To examine the impacts of land cover representation on regional flux estimates, we also produced regional estimates of carbon fluxes over the period 2001–2007 using

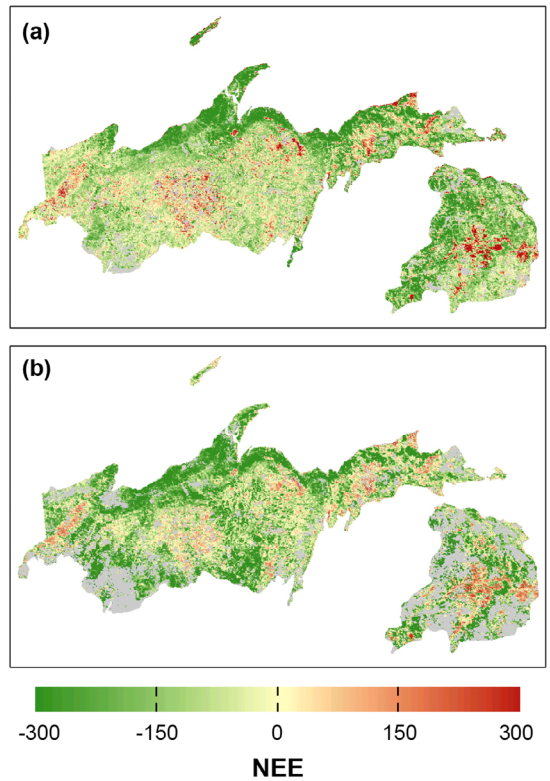

Figure 5. Comparison of mean annual NEE over the period 2001–2006 for the ChEAS region: (a) our estimate from this study; and (b) EC-MOD NEE. The units are g C m^{-2} yr^{-1}.

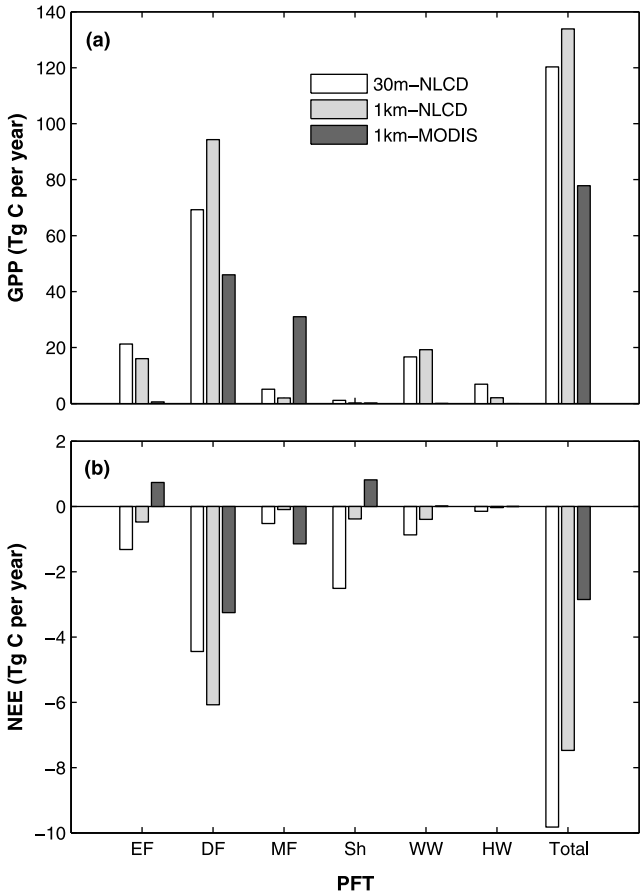

Figure 6. Regional annual carbon fluxes averaged over the period 2001–2007 for each plant functional type (PFT) based on different land cover representations (30m-NLCD, 1km-NLCD, and 1km-MODIS): annual (a) GPP; and (b) NEE. The units are Tg C yr^{-1}. The PFTs are evergreen forests (EF), deciduous forests (DF), mixed forests (MF), shrublands (Sh), woody wetlands (WW), and herbaceous wetlands (HW).

1km-NLCD and 1km-MODIS to specify the PFT of each cell. The differences in land cover representation among these land cover maps resulted in significant differences in regional flux estimates (Figure 6). The carbon fluxes integrated over the region based on 1km-NLCD were different from those from 30m-NLCD (Figure 6). For deciduous forests, for example, annual GPP of 1km-NLCD was 36.2% higher than that of 30m-NLCD; similarly, in absolute magnitude, annual NEE of 1km-NLCD was 36.7% higher than that of 30m-NLCD. Integrated across all PFTs over the region, 1km-NLCD led to significantly higher annual GPP (11.3%) and NEE (23.9% in absolute magnitude) than 30m-NLCD.

[44] The carbon fluxes integrated over the region based on 1km-MODIS were substantially different from those from the NLCD land cover maps (Figure 6). For deciduous forests, the annual GPP based on 1km-MODIS was 51.2% and 33.6% lower than those of 30m-NLCD and 1km-NLCD, respectively; the annual NEE based on 1km-MODIS was 46.5% and 26.8% lower than those of 30m-NLCD and 1km-NLCD, respectively. For mixed forests, annual GPP resulting from 1km-MODIS was 31.0 Tg C yr^{-1}, while the annual GPP based on 30m-NLCD and 1km-NLCD were only 5.1 and 2.0 Tg C yr^{-1}, respectively. Integrated across PFTs over the region, the annual GPP based on 1km-MODIS was 41.9% and 35.5% lower than those of 30m-NLCD and 1km-NLCD, respectively; the annual NEE of 1km-MODIS was 61.9% and 71.0% lower than those of 30m-NLCD and 1km-NLCD, respectively.

[45] We also examined the impacts of parameter variability on regional flux estimates for evergreen forests, deciduous forests, and mixed forests (Figure 7) by conducting model runs using parameter sets resulting from leave-one-out model optimizations (Table 4). The three land cover maps were also used for each model run to further examine the impacts of land cover representation on regional flux estimates. For evergreen forests, the five model runs resulting from leave-one-out model optimizations exhibited large variability in annual GPP with a standard deviation of 19.4 and 14.8 Tg C yr^{-1} for 30m-NLCD and 1km-NLCD, respectively; the model runs also showed large variability in annual NEE with a standard deviation of 1.4 and 0.5 Tg C yr^{-1} for 30m-NLCD and 1km-NLCD, respectively, and little variability in GPP and larger variability in NEE for 1km-MODIS (Figures 7a and 7b). For each PFT, the mean annual fluxes of these model runs, however, were very close to the annual fluxes based on the cross-site (or joint) optimization using flux observations from all the sites within the PFT. For deciduous forests, the model runs also exhibited large variability in annual fluxes for each land cover map, while the mean annual fluxes of these model runs had larger differences from fluxes based on the cross-site optimization using flux observations from all sites (Figures 7c and 7d). For mixed forests, the model runs showed little variability in annual fluxes for both 30m-NLCD and 1km-NLCD and large variability in annual NEE for 1km-MODIS.

5. Discussion

5.1. Parameter Estimation and Variability

[46] Eddy covariance flux observations are increasingly used to estimate the parameters of carbon cycle models [e.g., *Braswell et al.*, 2005; *Knorr and Kattge*, 2005; *Mahadevan et al.*, 2008; *Mo et al.*, 2008; *Ricciuto et al.*, 2008]. The assimilation of flux observations can help determine parameter values and reduce their uncertainties, leading to reduced uncertainties in estimated carbon fluxes. However, many of these studies used data from a single site to constrain model parameters for a given PFT [e.g., *Braswell et al.*, 2005; *Knorr and Kattge*, 2005; *Mahadevan et al.*, 2008; *Mo et al.*, 2008; *Ricciuto et al.*, 2008]. Few studies used observations from multiple sites encompassing a range of climate and disturbance history for the parameter optimization of a given PFT. The high density and range of eddy flux measurements in the ChEAS region make it possible to assess the variability of parameters both within and across PFTs, and to produce model parameters that are representative of the PFTs.

[47] The large variability of most model parameters among PFTs confirmed the well-known ecologically different behaviors among PFTs. The large variability of ε_{max} among PFTs is generally consistent with the range shown by

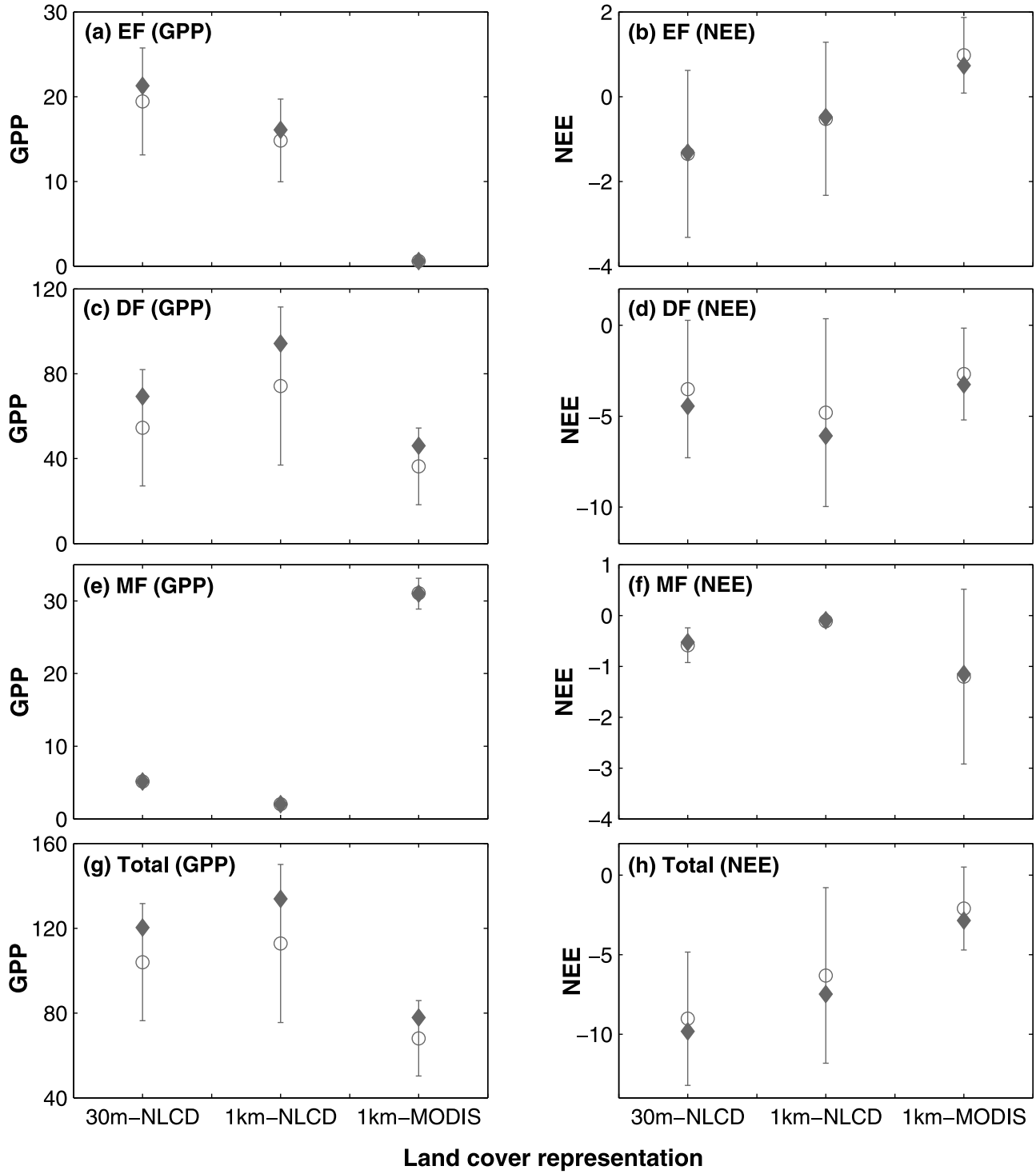

Figure 7. Variability of regional annual carbon fluxes averaged over the period 2001–2007 resulting from parameter variability based on different land cover representations: annual (a) GPP and (b) NEE for evergreen forests (EF); annual (c) GPP and (d) NEE for deciduous forests (DF); annual (e) GPP and (f) NEE for mixed forests (MF); and total annual (g) GPP and (h) NEE of all PFTs (evergreen forests, deciduous forests, mixed forests, shrublands, woody wetlands, and herbaceous wetlands) over the entire region. The units are Tg C yr^{-1}. The diamond symbol stands for fluxes based on parameters derived from cross-site (or joint) optimization using flux observations from all sites within each PFT. Open circle stands for fluxes averaged from model runs with parameters derived from leave-one-out model optimization, and the error bars on the open circles represent the standard deviation (or variability) of the regional annual fluxes from the model runs.

previous studies [e.g., *Ruimy et al.*, 1994; *Goetz and Prince*, 1996; *Gower et al.*, 1999; *Heinsch et al.*, 2003], and is inconsistent with the assumption in some previous studies [e.g, *Potter et al.*, 2007] that ε_{max} is a constant regardless of PFT. As the magnitude of EVI can be smaller than fPAR, the direct use of EVI as an approximation for fPAR [e.g., *Xiao et al.*, 2004; *Potter et al.*, 2007; *Mahadevan et al.*, 2008] can lead to substantial overestimation of ε_{max} in optimization. Most PFTs exhibited negative λ values, indicating that R_e is inversely proportional to aboveground biomass. Low-biomass, recently disturbed ecosystems may have higher litter, and thus have higher heterotrophic respiration and R_e.

[48] The general consistency of most parameters among leave-one-out optimizations shows that most sites within a given PFT exhibited similar behavior. For a given PFT, different sites can have a range of stand age, disturbance history, microclimate, and edaphic properties, and thus have different parameter values. The estimated parameter values resulting from a single site does not encompass the full range of variability of parameter values within a PFT. The within-PFT variability of parameters can be partly explained by the differences of stand age and disturbance history among the sites. For instance, stand age of evergreen forest sites varies from 3 to 63 years. Mature Red Pine is a mature stand, while all other sites are recently disturbed young stands. The large within-PFT variability in stand age and associated aboveground biomass and disturbance history likely contribute to the large within-PFT variability in parameters of evergreen forests. The same can be said of the role of Willow Creek within the deciduous forests. By contrast, mixed forests are either mature (UMBS, WLEF) or old-growth (SWA) ecosystems with similar stand age, which likely leads to the similarity of optimized parameters within this PFT.

[49] Many empirical or process-based ecosystem models use data from a single site to estimate the parameters of a given PFT [e.g., *Raich et al.*, 1991; *Mahadevan et al.*, 2008]. The large within-PFT variability of some model parameters indicates that it is inadequate to use data from a single site to estimate the parameters of a given PFT for regional applications. This traditional parameter estimation approach may not capture the variability of ecological and biophysical properties within a PFT, and could introduce biases to the resulting regional flux estimates. Ideally, observations from multiple sites encompassing a range of site conditions (e.g., stand age, disturbance history, and aboveground biomass) should be used for parameter estimation of a given PFT.

[50] The five factors controlling R_e such as biomass and soil carbon can vary substantially within a PFT. To account for the spatial variability of R_e within PFTs, we introduced aboveground biomass into the model. R_e consists of autotrophic respiration (R_a) and heterotrophic respiration (R_h). R_a can be empirically modeled as a function of air temperature and tissue carbon (foliage, stem, roots), while R_h is often modeled as a function of substrate availability. For instance, maintenance respiration is modeled as a direct function of plant biomass while R_h is a function of soil carbon storage in TEM [*Tian et al.*, 1999; *Xiao et al.*, 2009]. For forest ecosystems, aboveground biomass is also significantly related to stand age. Thus, our model can partly account for the variability of biomass and disturbance history within a PFT.

[51] Aboveground biomass, however, cannot account for the sizes of soil carbon pools and litterfall. Our results show that the inclusion of stand age significantly improves the performance of the model for estimating NEE for evergreen forests and deciduous forests. Thus the addition of stand age can improve the estimation of R_e and thus NEE. However, the modified model with stand age included for forested sites was not used to produce regional flux estimates because there is no spatially explicit information on stand age available for the ChEAS region. The development of regional stand age maps and their incorporation into modeling are expected to improve regional flux estimates.

5.2. Influence of Parameter Variability and Land Cover Representation on Regional Flux Estimates

[52] Our results show that the variability of parameters within a PFT can result in considerable uncertainty in regional flux estimates. The estimated parameter values resulting from a single site do not encompass the full range of variability of parameter values within a given PFT. The annual fluxes based on the cross-site (or joint) optimization using flux observations from all sites were very close to the mean annual fluxes resulting from the model runs based on leave-one-out model optimizations, showing that for a given PFT, cross-site optimization using flux observations from all sites can lead to more robust flux estimates.

[53] The parameterization of ecosystem models for a given PFT is typically based on observations from a single site [e.g., *Potter et al.*, 2007; *Xiao et al.*, 2009]. The optimization of model parameters based on flux observations and data assimilation (or model-data fusion) techniques also typically use observations from a single site for a given PFT [e.g., *Braswell et al.*, 2005; *Knorr and Kattge*, 2005; *Mahadevan et al.*, 2008; *Mo et al.*, 2008]. Our results indicate that this type of parameterization or optimization can result in significant uncertainty in flux estimates; for a given PFT, observations from multiple sites should be used for the parameterization and optimization of ecosystem models to minimize the uncertainty of parameter variability on flux estimates. There are currently over 500 eddy flux towers over the globe, and these sites encompass a large range of ecosystem and climate types. The availability of flux observations from these sites makes it possible to examine the variability of parameters within and across PFTs over broader spatial domains or other geographical regions.

[54] Our results show that land cover representation including land cover heterogeneity and the spatial resolution and accuracy of land cover maps can result in large uncertainties in regional flux estimates in heterogeneous regions, although such uncertainties could be much smaller in more homogeneous regions [*Quaife et al.*, 2008]. The aggregation of 30 m NLCD data to 1 km spatial resolution results in the subgrid representation of PFT fractions within each 1 km grid cell. The differences in regional fluxes obtained using these two land cover maps were substantial, demonstrating that land cover heterogeneity and the spatial resolution of land over maps can result in significant uncertainty to flux estimates. The MODIS land cover is based on moderate-resolution MODIS data, and cannot capture the spatial

details and resolve the proportions of PFTs within each grid cell. Compared to NLCD land cover maps, the MODIS land cover map also fail to detect the substantial presence of wetlands that are ecologically distinct from upland forest ecosystems. In addition, 1km-MODIS is based on a global classification algorithm, while 30m-NLCD is based on 30 m Landsat data and regional training sites. NLCD land cover maps, therefore, are likely to have higher classification accuracy than 1km-MODIS. The considerable differences in regional flux estimates based on the NLCD and MODIS land cover maps demonstrate that regional flux estimates can be significantly affected by the accuracy of land cover maps and land cover heterogeneity. Our results also imply that intercomparison studies of different ecosystem models should consider the differences in the underlying land cover maps. Ideally, the same land cover map with the same spatial resolution should be used for model intercomparison studies.

[55] One key remaining challenge is to produce gridded, probabilistic flux estimates based on the probability distribution functions (PDFs) of land cover and forest structure and the merging of uncertainties from input data and parameter PDFs, and to assess the importance of sources of uncertainty and the reduction of uncertainty. Bayesian approaches, Markov chain Monte Carlo (MCMC) methods in particular [e.g., *Braswell et al.*, 2005], have been used to obtain posterior distributions of model parameters that combine information from the data and from assumed prior parameter distributions. The uncertainty of model parameters and other input data can be propagated through modeling to quantify the uncertainties of flux estimates. The large range of eddy flux measurements in the Upper Midwest region and MCMC make it promising to optimize LUE-based or process-based ecosystem models and to produce such gridded, probabilistic flux estimates.

6. Conclusions

[56] We used eddy flux observations and a data assimilation approach to estimate the parameters of the simple Diagnostic Carbon Flux Model (DCFM) and examined the influence of parameter variability and land cover representation on regional flux estimates. Our results show that some model parameters vary not only across PFTs but also within a given PFT. The within-PFT variability in parameters indicates that it is inadequate to use data from a single site to estimate the parameters of a given PFT for regional applications, and multiple sites encompassing a full range of site conditions (e.g., stand age, disturbance history, and climate) should be used. Our results show that parameter variability can result in substantial variability in regional flux estimates. Our results also demonstrate that land cover representation including land cover heterogeneity and the spatial resolution and accuracy of land cover maps can introduce considerable uncertainty to regional flux estimates. In heterogeneous, complex regions, detailed land cover maps are essential for accurate estimation of regional carbon fluxes.

[57] **Acknowledgments.** This study was supported by the National Aeronautics and Space Administration (NASA). Terrestrial Ecology Program. Data collection was supported by NASA's Terrestrial Ecology Program and the Department of Energy's Office of Biological and Environmental Research, Terrestrial Carbon Program, and the National Institute for Climatic Change Research. We thank T. Hilton for assistance with data assimilation and A. Desai for helpful discussion. We also thank J. Chen, K. Cherrey, P. Curtis, A. Desai, C. Gough, and A. Noormets for contributions to the flux observations used in this study. We also thank the two anonymous reviewers and D.D. Baldocchi for constructive comments on earlier versions of the manuscript.

References

Ardia, D. (2009), DEoptim: Differential Evolution Optimization in R. R package, version 1.3-3, http://CRAN.R-project.org/package=DEoptim.

Baldocchi, D. D., B. B. Hicks, and T. P. Meyers (1988), Measuring biosphere-atmosphere exchanges of biologically related gases with micrometeorological methods, *Ecology*, *69*, 1331–1340, doi:10.2307/1941631.

Boelman, N. T., M. Stieglitz, H. M. Rueth, M. Sommerkorn, K. L. Griffin, G. R. Shaver, and J. A. Gamon (2003), Response of NDVI, biomass, and ecosystem gas exchange to long-term warming and fertilization in wet sedge tundra, *Oecologia*, *135*, 414–421.

Braswell, B. H., B. Sacks, E. Linder, and D. S. Schimel (2005), Estimating ecosystem process parameters by assimilation of eddy flux observations of NEE, *Global Change Biol.*, *11*, 335–355, doi:10.1111/j.1365-2486.2005.00897.x.

Butler, M. P., K. J. Davis, A. S. Denning, and S. R. Kawa (2010), Using continental observations in global atmospheric inversions of CO2: North American carbon sources and sinks, *Tellus, Ser. B*, *62*, 550–572, doi:10.1111/j.1600-0889.2010.00501.x.

Chapin, F. S., P. A. Matson, and H. A. Mooney (2002), *Principles of Terrestrial Ecosystem Ecology*, 436 pp., Springer, New York.

Chen, J., K. J. Davis, and T. P. Meyers (2008), Ecosystem-atmosphere carbon and water cycling in the upper Great Lakes region, *Agric. For. Meteorol.*, *148*, 155–157, doi:10.1016/j.agrformet.2007.08.016.

Cook, B. D., et al. (2004), Carbon exchange and venting anomalies in an upland deciduous forest in northern Wisconsin, USA, *Agric. For. Meteorol.*, *126*, 271–295, doi:10.1016/j.agrformet.2004.06.008.

Cook, B. D., P. V. Bolstad, E. Næsset, R. S. Anderson, S. Garrigues, J. T. Morisette, J. Nickeson, and K. J. Davis (2009), Using LiDAR and QuickBird data to model plant production and quantify uncertainties associated with wetland detection and land cover generalizations, *Remote Sens. Environ.*, *113*, 2366–2379, doi:10.1016/j.rse.2009.06.017.

Davis, K. J. (2008), Integrating field measurements with flux tower and remote sensing data, in *Field Measurements for Landscape-Scale Forest Carbon Monitoring*, vol. XVIII, edited by C. M. Hoover, pp. 207–226, Springer, New York, ISBN:978-1-4020-8505-5.

Davis, K. J., P. S. Bakwin, C. Yi, B. W. Berger, C. Zhao, R. M. Teclaw, and J. G. Isebrands (2003), The annual cycles of CO_2 and H_2O exchange over a northern mixed forest as observed from a very tall tower, *Global Change Biol.*, *9*, 1278–1293, doi:10.1046/j.1365-2486.2003.00672.x.

Desai, A. R., P. V. Bolstad, B. D. Cook, K. J. Davis, and E. V. Carey (2005), Comparing net ecosystem exchange of carbon dioxide between an old-growth and mature forest in the Upper Midwest, USA, *Agric. For. Meteorol.*, *128*, 33–55, doi:10.1016/j.agrformet.2004.09.005.

Desai, A. R., et al. (2008), Influence of vegetation and seasonal forcing on carbon dioxide fluxes across the Upper Midwest, USA: implications for regional scaling, *Agric. For. Meteorol.*, *148*, 288–308, doi:10.1016/j.agrformet.2007.08.001.

Falge, E., et al. (2001), Gap filling strategies for defensible annual sums of net ecosystem exchange, *Agric. For. Meteorol.*, *107*, 43–69, doi:10.1016/S0168-1923(00)00225-2.

Fensholt, R., and I. Sandholt (2003), Derivation of a shortwave infrared water stress index from MODIS near- and shortwave infrared data in a semiarid environment, *Remote Sens. Environ.*, *87*, 111–121, doi:10.1016/j.rse.2003.07.002.

Foley, S., et al. (2005), Global consequences of land use, *Science*, *309*, 570–574, doi:10.1126/science.1111772.

Friedl, M. A., et al. (2002), Global land cover mapping from MODIS: algorithms and early results, *Remote Sens. Environ.*, *83*, 287–302, doi:10.1016/S0034-4257(02)00078-0.

Gao, B. C. (1996), NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space, *Remote Sens. Environ.*, *58*, 257–266, doi:10.1016/S0034-4257(96)00067-3.

Goetz, S. J., and S. D. Prince (1996), Remote sensing of net primary production in boreal forest stands, *Agric. For. Meteorol.*, *78*, 149–179, doi:10.1016/0168-1923(95)02268-6.

Gough, C. M., C. S. Vogel, H. P. Schmid, H.-B. Su, and P. S. Curtis (2008), Multi-year convergence of biometric and meteorological esti-

mates of forest carbon storage, *Agric. For. Meteorol.*, *148*, 158–170, doi:10.1016/j.agrformet.2007.08.004.

Gower, S. T., C. J. Kucharik, and J. M. Norman (1999), Direct and indirect estimation of leaf area index, f (APAR), and net primary production of terrestrial ecosystems, *Remote Sens. Environ.*, *70*, 29–51, doi:10.1016/S0034-4257(99)00056-5.

Heinsch, F. A., et al. (2003), User's Guide: GPP and NPP (MOD17A2/A3) Products. NASA MODIS Land Algorithm, version 2.0, 2 December.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan (2004), Development of a 2001 national landcover database for the United States, *Photogram. Eng. Remote Sens.*, *70*, 829–840.

Huete, A. R., H. Q. Liu, K. Batchily, and W. van Leeuwen (1997), A comparison of vegetation indices global set of TM images for EOS-MODIS, *Remote Sens. Environ.*, *59*, 440–451, doi:10.1016/S0034-4257(96)00112-5.

Huete, A., K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira (2002), Overview of the radiometric and biophysical performance of the MODIS vegetation indices, *Remote Sens. Environ.*, *83*, 195–213, doi:10.1016/S0034-4257(02)00096-2.

Hurtt, G. C., S. W. Pacala, P. R. Moorcroft, J. Caspersen, E. Shevliakova, R. A. Houghton, and B. Moore (2002), Projecting the future of the US carbon sink, *Proc. Natl. Acad. Sci. U. S. A.*, *99*, 1389–1394, doi:10.1073/pnas.012249999.

Jackson, T. J., D. Chen, M. Cosh, F. Li, M. Anderson, C. Walthall, P. Doriaswamy, and E. R. Hunt (2004), Vegetation water content mapping using Landsat data derived normalized difference water index fro corn and soybeans, *Remote Sens. Environ.*, *92*, 475–482, doi:10.1016/j.rse.2003.10.021.

Jung, M., M. Reichstein, and A. Bondeau (2009), Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, *Biogeosciences*, *6*, 2001–2013, doi:10.5194/bg-6-2001-2009.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Knorr, W., and J. Kattge (2005), Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling, *Global Change Biol.*, *11*, 1333–1351, doi:10.1111/j.1365-2486.2005.00977.x.

Landsberg, J. J., and R. H. Waring (1997), A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning, *For. Ecol. Manage.*, *95*, 209–228, doi:10.1016/S0378-1127(97)00026-1.

Lloyd, J., and J. A. Taylor (1994), On the temperature dependence of soil respiration, *Funct. Ecol.*, *8*, 315–323, doi:10.2307/2389824.

Mahadevan, P., S. C. Wofsy, D. M. Matross, X. Xiao, A. L. Dunn, J. C. Lin, C. Gerbig, J. W. Munger, V. Y. Chow, and E. W. Gottlieb (2008), A satellite-based biosphere parameterization for net ecosystem CO2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM), *Global Biogeochem. Cycles*, *22*, GB2005, doi:10.1029/2006GB002735.

Mo, X., J. M. Chen, W. Ju, and T. A. Black (2008), Optimization of ecosystem model parameters through assimilating eddy covariance flux data with an ensemble Kalman filter, *Ecol. Modell.*, *217*, 157–173, doi:10.1016/j.ecolmodel.2008.06.021.

Moffat, A. M., et al. (2007), Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes, *Agric. For. Meteorol.*, *147*, 209–232, doi:10.1016/j.agrformet.2007.08.011.

Monteith, J. (1972), Solar radiation and productivity in tropical ecosystems, *J. Appl. Ecol.*, *9*, 747–766, doi:10.2307/2401901.

Monteith, J. (1977), Climate and efficiency of crop production in Britain, *Philos. Trans. R. Soc. London, Ser. B*, 277–294.

Myneni, R. B., J. Dong, C. J. Tucker, R. K. Kaufmann, P. E. Kauppi, J. Liski, L. Zhou, V. Alexeyev, and M. K. Hughes (2001), A large carbon sink in the woody biomass of northern forests, *Proc. Natl. Acad. Sci. U. S. A.*, *98*, 14,784–14,789, doi:10.1073/pnas.261555198.

Noormets, A., et al. (2008), Moisture sensitivity of ecosystem respiration: comparison of 14 forest ecosystems in the upper Great Lakes region, USA, *Agric. For. Meteorol.*, *148*, 216–230, doi:10.1016/j.agrformet.2007.08.002.

Persson, P., K. Hall-Könyves, G. Sjöström, and S. Pinzke (1993), NOAA/AVHRR data for crop productivity estimation in Sweden, *Adv. Space Res.*, *13*, 111–116, doi:10.1016/0273-1177(93)90211-S.

Pinker, R. T., O. Laszlo, J. D. Tarpley, and K. Mitchell (2002), Geostationary satellite parameters for surface energy balance, *Adv. Space Res.*, *30*, 2427–2432, doi:10.1016/S0273-1177(02)80296-4.

Potter, C., S. Klooster, A. Huete, and V. Genovese (2007), Terrestrial carbon sinks for the United States predicted from MODIS satellite data and ecosystem modeling, *Earth Interact.*, *11*, 1–21, doi:10.1175/EI228.1.

Price, K. V., R. M. Storn, and J. A. Lampinen (2006), *Differential Evolution: A Practical Approach to Global Optimization*, Springer, New York.

Prince, S. D., and S. N. Goward (1995), Global primary production: a remote sensing approach, *J. Biogeogr.*, *22*, 815–835, doi:10.2307/2845983.

Quaife, T., S. Quegan, M. Disney, P. Lewis, M. Lomas, and F. I. Woodward (2008), Impact of land cover uncertainties on estimates of biospheric carbon fluxes, *Global Biogeochem. Cycles*, *22*, GB4016, doi:10.1029/2007GB003097.

Raich, J. W., E. Rastetter, J. Melillo, D. Kicklighter, P. Steudler, B. Peterson, A. Grace, B. Moore III, and C. Vorosmarty (1991), Potential net primary productivity in South America: application of a global model, *Ecol. Appl.*, *7*, 444–460.

Reich, P., D. Turner, and P. Bolstad (1999), An approach to spatially distributed modeling of net primary production (NPP) at the landscape scale and its application in validation of EOS NPP products, *Remote Sens. Environ.*, *70*, 69–81, doi:10.1016/S0034-4257(99)00058-9.

Ricciuto, D. M., M. P. Butler, K. J. Davis, B. D. Cook, P. S. Bakwin, A. Andrews, and R. M. Teclaw (2008), Causes of interannual variability in ecosystem-atmosphere CO_2 exchange in a northern Wisconsin forest using a Bayesian model calibration, *Agric. For. Meteorol.*, *148*, 309–327, doi:10.1016/j.agrformet.2007.08.007.

Ruimy, A., B. Saugier, and G. Dedieu (1994), Methodology for the estimation of terrestrial net primary production from remotely sensed data, *J. Geophys. Res.*, *99*, 5263–5283, doi:10.1029/93JD03221.

Ruimy, A., G. Dedieu, and B. Saugier (1996), TURC: a diagnostic model of continental gross primary productivity and net primary productivity, *Global Biogeochem. Cycles*, *10*, 269–285, doi:10.1029/96GB00349.

Running, S. W., R. R. Nemani, F. A. Heinsch, M. Zhao, M. Reeves, and H. Hashimoto (2004), A continuous satellite-derived measure of global terrestrial primary production, *BioScience*, *54*, 547–560, doi:10.1641/0006-3568(2004)054[0547:ACSMOG]2.0.CO;2.

Schmid, H. P. (1994), Source areas for scalars and scalar fluxes, *Boundary Layer Meteorol.*, *67*, 293–318, doi:10.1007/BF00713146.

State of the Carbon Cycle Report (2007), *The First State of the Carbon Cycle Report (SOCCR): the North American Carbon Budget and Implications for the Global Carbon Cycle*, edited by A. W. King et al., U.S. Clim. Change Sci. Program, Washington, D. C.

Su, H.-B., H. P. Schmid, C. S. B. Grimmond, C. S. Vogel, and P. S. Curtis (2008), An assessment of observed vertical flux divergence in long-term eddy-covariance measurements over two midwestern forest ecosystems, *Agric. For. Meteorol.*, *148*, 186–205, doi:10.1016/j.agrformet.2007.08.009.

Sulman, B. N., A. R. Desai, N. Z. Saliendra, P. M. Lafleur, L. B. Flanagan, O. Sonnentag, D. S. Mackay, and A. G. Barr (2010), CO_2 fluxes at northern fens and bogs have opposite responses to inter-annual fluctuations in water table, *Geophys. Res. Lett.*, *37*, L19702, doi:10.1029/2010GL044018.

Sun, G., et al. (2011), Upscaling key ecosystem functions across the conterminous United States by a water-centric ecosystem model, *J. Geophys. Res.*, *116*, G00J05, doi:10.1029/2010JG001573.

Tian, H., J. M. Melillo, D. W. Kicklighter, A. D. McGuire, and J. Helfrich (1999), The sensitivity of terrestrial carbon storage to historical climate variability and atmospheric CO_2 in the United States, *Tellus, Ser. B*, *51*, 414–452.

Tucker, C. J., C. L. Vanpraet, M. J. Sharman, and G. Van Ittersum (1985), Satellite remote sensing of total herbaceous biomass production in the Senegalese Sahel: 1980–1984, *Remote Sens. Environ.*, *17*, 233–249, doi:10.1016/0034-4257(85)90097-5.

Vermote, E. F., and A. Vermeulen (1999), MODIS Algorithm Technical Background Document: Atmospheric Correction Algorithm: Spectral Reflectances (MOD09), version 4.0, http://modis.gsfc.nasa.gov/data/atbd/atbd_mod08.pdf.

Veroustraete, F., H. Sabbe, and H. Eerens (2002), Estimation of carbon mass fluxes over Europe using the C-Fix model and Euroflux data, *Remote Sens. Environ.*, *83*, 376–399, doi:10.1016/S0034-4257(02)00043-3.

Xiao, J., et al. (2008), Estimation of net ecosystem carbon exchange of the conterminous United States by combining MODIS and AmeriFlux data, *Agric. For. Meteorol.*, *148*, 1827–1847, doi:10.1016/j.agrformet.2008.06.015.

Xiao, J., Q. Zhuang, E. Liang, A. D. McGuire, A. Moody, D. W. Kicklighter, X. Shao, and J. M. Melillo (2009), Twentieth century droughts and their impacts on terrestrial carbon cycling in China, *Earth Interact.*, *13*(10), 1–31, doi:10.1175/2009EI275.1.

Xiao, J., et al. (2010), A continuous measure of gross primary productivity for the conterminous U.S. derived from MODIS and AmeriFlux data, *Remote Sens. Environ.*, *114*, 576–591, doi:10.1016/j.rse.2009.10.013.

Xiao, J., et al. (2011), Assessing net ecosystem carbon exchange of U.S. terrestrial ecosystems by integrating eddy covariance flux measurements and satellite observations, *Agric. For. Meteorol.*, *151*, 60–69, doi:10.1016/j.agrformet.2010.09.002.

Xiao, X., D. Hollinger, J. Aber, M. Goltz, E. A. Davidson, Q. Zhang, and B. Moore III (2004), Satellite-based modeling of gross primary production in an evergreen needleleaf forest, *Remote Sens. Environ.*, *89*, 519–534, doi:10.1016/j.rse.2003.11.008.

Zhang, L., B. K. Wylie, L. Ji, T. G. Gilmanov, L. L. Tieszen, and D. M. Howard (2011), Upscaling carbon fluxes over the Great Plains grasslands: Sinks and sources, *J. Geophys. Res.*, *116*, G00J03, doi:10.1029/2010JG001504.

Zhao, M., F. A. Heinsch, R. R. Nemani, and S. W. Running (2005), Improvements of the MODIS terrestrial gross and net primary production global data set, *Remote Sens. Environ.*, *95*, 164–176, doi:10.1016/j.rse.2004.12.011.

K. J. Davis, Department of Meteorology, Pennsylvania State University, 512 Walker Bldg., University Park, PA 16802-5013, USA.

K. Keller, Department of Geosciences, Pennsylvania State University, 436 Deike, University Park, PA 16870, USA.

N. Z. Saliendra, Center for Urban Environmental Research and Education, University of Maryland at Baltimore County, 1000 Hilltop Circle, TRC Bldg., Rm. 102, Baltimore, MD 21250, USA.

N. M. Urban, Woodrow Wilson School of Public and International Affairs, Princeton University, 411A Robertson Hall, Princeton, NJ 08544, USA.

J. Xiao, Complex Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, 8 College Rd., Durham, NH 03861, USA. (j.xiao@unh.edu)