Ecological Modelling 336 (2016) 57–69

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Variation of parameters in a Flux-Based Ecosystem Model across 12 sites of terrestrial ecosystems in the conterminous USA

Qianyu Li[a], Jianyang Xia[b,c], Zheng Shi[d], Kun Huang[b,c], Zhenggang Du[b,c], Guanghui Lin[a,*], Yiqi Luo[a,d,**]

[a] Ministry of Education Key Laboratory for Earth System Modeling, Center for Earth System Science, Tsinghua University, Beijing 100084, China
[b] Tiantong National Forest Ecosystem Observation and Research Station, School of Ecological and Environmental Sciences, East China Normal University, Shanghai 200062, China
[c] Research Center for Global Change and Ecological Forecasting, East China Normal University, Shanghai 200062, China
[d] Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK 73019, USA

ARTICLE INFO

Article history:
Received 12 February 2016
Received in revised form 24 May 2016
Accepted 27 May 2016

Keywords:
Ecological model
Carbon cycle
Parameters
Data-model fusion
Bayesian optimization

ABSTRACT

Terrestrial ecosystem models have been extensively used in global change research. When a model calibrated with site-specific parameters is applied to another site, how and why the parameters have to be adjusted again in order to fit data well are pervasive yet underexplored issues. In this exploratory study, we examined how and why model parameters of a Flux-Based Ecosystem Model (FBEM) varied across different sites. Parameters were estimated from data at 12 eddy-covariance towers in the conterminous USA using the conditional inversion method. Results showed that optimized values of these parameters varied across sites. For example, the estimated coefficients in the Leuning model, g_l and D_0, exhibited high cross-site variation, but the ratio of internal to air CO_2 concentration (f_{Ci}) and canopy light extinction coefficient (k_n) varied little among these sites. Parameters greatly varied with ecosystem types at adjacent sites where climate conditions were similar. Five parameters (activation energy of carboxylation, E_{Kc}; activation energy of oxygenation, E_{Vm}; ecosystem respiration, R^0_{eco}; temperature sensitivity of respiration, Q_{10}; and stomatal conductance coefficient, D_0) were highly correlated with mean annual temperature and precipitation across sites, which were distributed in different climate regions of conterminous US. Our results indicate that individual parameters vary to different degrees across sites and parameter variation can be related to different biological factors (e.g., ecosystem types) and environmental conditions (e.g., temperature and precipitation). It is essential to further examine magnitudes of and mechanisms underlying the parameter variation in ecosystem models so as to improve model prediction.

1. Introduction

Ecological models that simulate responses of ecosystem photosynthesis and respiratory processes to elevated atmospheric CO_2 and increased temperature are fundamental to projecting carbon balance and impacts of global change on the biosphere (Long, 1991; Lloyd and Taylor, 1994; Sellers et al., 1997; Bernacchi et al., 2001). It has been well noticed that variation in parameters is one of the main sources of uncertainty in model predictions (Luo et al., 2011, 2015; Xiao et al., 2014). Most global biosphere models classify the terrestrial ecosystems with a small number of categories, referred to as plant functional types (PFTs). Parameters are often assigned to some fixed values for a given PFT and may vary among PFTs. For example, Sellers et al. (1996) found that some parameters such as maximum photosynthetic carboxylation rate and minimum stomatal conductance need better parameterizations because they vary strongly with PFTs at the global scale. Even within the same PFT, however, model parameters appeared more variable than assumed. For instance, Groenendijk et al. (2011) found better simulations of photosynthesis and transpiration using site-specific calibrated parameters compared with using fixed vegetation parameters across sites with the same PFT. Xiao et al. (2014) also identified variation of estimated values of parameters both within and across PFTs in a diagnostic carbon flux model, suggesting

* Corresponding author.
** Corresponding author at: Ministry of Education Key Laboratory for Earth System Modeling, Center for Earth System Science, Tsinghua University, Beijing 100084, China.
E-mail addresses: lingh@mail.tsinghua.edu.cn (G. Lin), yluo@ou.edu (Y. Luo).

http://dx.doi.org/10.1016/j.ecolmodel.2016.05.016
0304-3800/

the use of observations from multiple sites for a given PFT provided more representative estimates of parameter values. Although it is recognized that parameters may have to vary with sites, it remains unknown how the parameters vary across sites and what contributed to the parameter variation.

Some studies have shown that parameters vary with both biotic and abiotic factors. For example, maximum rate of carboxylation (i.e., $V_{c,max}$) is a key parameter in the leaf model of C_3 photosynthesis of Farquhar et al. (1980). Previous studies have identified four factors – species differences, light intensity, seasonal patterns, and water availability – that could cause variability in $V_{c,max}$ (normalized to 25 °C) (Wullschleger, 1993; Medvigy et al., 2013; Wilson et al., 2000). Despite various factors to influence $V_{c,max}$, its variation usually can be well represented by $V_{c,max}$—nitrogen (N) relationships as in many Earth system models (Rogers, 2014). Temperature sensitivity of soil heterotrophic respiration (i.e., Q_{10}) is another critical parameter regulating carbon-climate feedback. Although many ecosystem models commonly use a constant Q_{10} (Tian et al., 1999; Schimel et al., 2000; Chen and Tian, 2005), a small deviation of Q_{10} will significantly change the estimate of the total CO_2 efflux from soil to the atmosphere (Xu and Qi, 2001). In fact, a few studies have shown that Q_{10} varies at site scale with climate variables such as mean annual precipitation and temperature (Zhou et al., 2009; Peng et al., 2009). These examples all illustrate that parameters might have to vary temporally and spatially.

Recently developed techniques of data-model fusion and inversion analysis have emerged as useful tools to gain ecological knowledge about parameter variation in biogeochemical models. For instance, Shi et al. (2015) estimated C–N coupling parameters using Bayesian Markov Chain Monte Carlo (MCMC) technique under ambient and elevated CO_2 in Duke Forests, revealing that C-N coupling parameters exhibited significant changes in response to rising atmospheric CO_2. By using inversion analysis, Zhou and Luo (2008) discovered that estimated carbon residence times were highly heterogeneous over the conterminous United States. Wang et al. (2007) found better predictions of net ecosystem CO_2 exchange (NEE) with seasonally varying $V_{c,max}$ and maximum rate of electron transport at 25 °C (i.e., J_{max}) in CSIRO biosphere model using a nonlinear inversion approach from eight FLUXNET sites.

Most of the previous parameter estimation studies with data-model fusion technique have pointed out that only a few parameters in process-based ecosystem models could be well constrained by the measurements of NEE (Wang et al., 2001; Braswell et al., 2005; Xiao et al., 2014). Toward that end, Wu et al. (2009) developed a conditional Bayesian inversion method to maximize the number of constrained parameters by assimilating NEE into FBEM. The conditional inversion method increased the number of constrained parameters from 6 to13 out of a total of 16 eventually. In this study, we used conditional inversion method and separated NEE into gross primary production (GPP) and ecosystem respiration (*Reco*) to fully extract data information to constrain model parameters.

The overall objective of this study is to understand variation of model parameters across different sites. Eddy covariance measurement of carbon dioxide across the canopy-atmosphere interface (Goulden et al., 1996; Baldocchi, 2003), including GPP and *Reco*, from 12 sites in North America were used to constrain the parameters in FBEM. Simulations were carried out from 2003 to 2007 according to the availability of data for each site. With eddy-covariance data and conditional inversion, we attempted to address the following three questions: (1) whether and how the model parameters vary across the 12 sites? (2) Whether model parameters vary within each PFT? (3) Which environmental factors are important in regulating the cross-site variation of the key parameters?

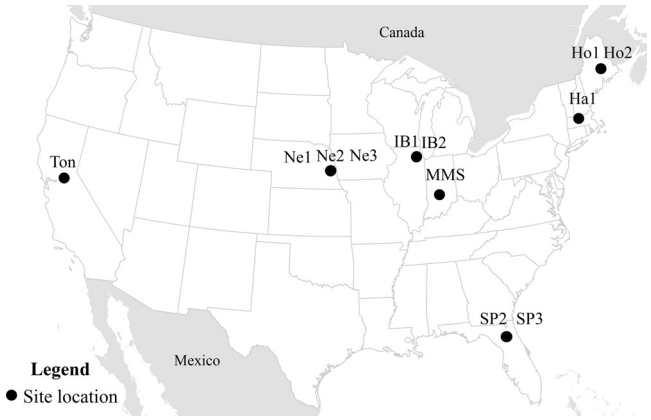

Fig. 1. Locations of the 12 sites in the conterminous United States. Among these sites, there were four clusters of geographically adjacent sites: Ne1, Ne2 and Ne3; Ho1 and Ho2; SP2 and SP3; IB1 and IB2. The sites within each cluster had distinct climate conditions, vegetation types or management strategies except Ho1 and Ho2.

2. Materials and methods

2.1. Site descriptions

The 12 sites in this study were: Harvard Forest (Ha1), Mead Irrigate (Ne1), Mead Irrigate Rotation (Ne2), Mead Rainfed (Ne3), Morgan Monre State Forest (MMS), Tonzi Ranch (Ton), Howland Forest (Ho1), Howland Forest West (Ho2), Mize (SP2), Donaldson Florida (SP3), Fermi Agricultural (IB1) and Fermi Prairie (IB2). Locations of those sites were shown in Fig. 1. The sites belonged to six ecosystem types: two for deciduous broad leaf forest (DBF), four for evergreen needle leaf forest (ENF), four for cropland (CRO), one for grassland (GRA) and one for savanna (SAV). The 12 sites covered four climate types, with mean annual temperature (MAT) varying from 5.13 to 20.25 °C, and mean annual precipitation (MAP) varying from 559 to 1314 mm. Information on the climate type, vegetation type, MAT, MAP and related reference of each site was listed in Table 1.

Among the 12 sites, there were four clusters of geographically adjacent sites. The first cluster included Ne1, Ne2 and Ne3, three fields located within 1.6 km of each other at the University of Nebraska Agricultural Research and Development Center near Mead, Nebraska. The vegetation type in Ne2 and Ne3 were maize soybean rotation while that in Ne1 was maize. Ne1 and Ne2 sites were irrigated while Ne3 site was rainfed. In Howland Forest, there were two eddy covariance towers Ho1 and Ho2 with nearly identical meteorological conditions, vegetation types and site histories. The third cluster included SP2 and SP3. SP2 site was once clearcut and planted with mixed genotype seedlings, and was covered mainly by evenly aged slash pine plantation. SP3 site was also slash pine plantations but established earlier than the new seedlings in SP2 site. The fourth cluster includes IB1 and IB2. IB2 had an eddy correlation system installed on a restored prairie, while there was a corn/soybean rotation agricultural field established in IB1.

2.2. Data sources

The datasets used in this study included climate data (i.e., air temperature at top canopy [*Ta*], photosynthetically active radiation [PAR] and relative humidity [RH]), biometric data (i.e., tower-collected leaf area index [LAI]), and eddy flux data (i.e., NEE, GPP and *Reco*). These datasets were downloaded from the AmeriFlux database at http://public.ornl.gov/ameriflux (AmeriFlux, 2007). Direct LAI measurements were not made or available in many sites.

Table 1
Site descriptions and relevant references.

Site	Type[a]	Climate[b]	Longitude[c]	Latitude[d]	Elevation (m)	MAT (°C)[e]	MAP (mm a^{-1})[f]	Reference
Ha1	DBF	Dfb	−72.18	42.54	340	6.62	1071	Wu et al. (2009)
Ne1	CRO	Dfa	−96.48	41.17	361	10.07	790	Owen et al. (2007)
Ne2	CRO	Dfa	−96.47	41.16	362	10.08	789	Kalfas et al. (2011)
Ne3	CRO	Dfa	−96.44	41.18	363	10.11	784	Owen et al. (2007)
MMS	DBF	Cfa	−86.41	39.32	275	10.85	1032	Curtis et al. (2002)
Ton	SAV	Csa	−120.97	38.43	177	15.8	559	Tang and Baldocchi (2005)
Ho1	ENF	Dfb	−68.74	45.2	60	5.27	1070	Hollinger and Richardson (2005)
Ho2	ENF	Dfb	−68.75	45.2	91	5.13	1064	Hollinger and Richardson (2005)
SP2	ENF	Cfa	−82.22	29.76	50	20.07	1314	Gholz and Clark (2002)
SP3	ENF	Cfa	−82.16	29.75	50	20.25	1312	Gholz and Clark (2002)
IB1	CRO	Dfa	−88.22	41.86	226.5	9.18	929	Matamala et al. (2008)
IB2	GRA	Dfa	−88.24	41.84	226.5	9.04	930	Jastrow (1987)

[a] DBF, deciduous broadleaf forests; ENF, evergreen needleleaf forests; CRO, croplands; GRA, grasslands; SAV, savanna ecosystem.
[b] Dfb, warm summer continental (signigicant precipitation in all seasons); Dfa, humid continental (humid with severe winter, no dry season, hot summer); Cfa, humid subtropical (mild with no dry season, hot summer); Csa, mediterranean (mild with dry, hot summer).
[c] Values indicate west longitude.
[d] Values indicate north latitude.
[e] MAT, mean annual temperature.
[f] MAP, mean annual precipitation.

Therefore we chose the above 12 sites which had direct and complete LAI measurements for analysis.

One year of data were used because this study was not designed to study inter-annual variability with multiple-year NEE data. According to the availability of data, data from 2006 to 2007 were used for Ha1, Ne1, Ne2, Ne3, Ho1, Ho2, IB1 and IB2. At MMS and Ton, we used data from 2005 to 2006 because no measured LAI data in 2007 were available for the two sites. For consistency, we used the data in 2006 to optimize the model, and data in another year to validate the model. For SP2 and SP3, data in 2004 were used for model optimization, and data in 2003 for model validation.

Data from AmeriFlux have been checked by each investigator and formatted by Carbon Dioxide Information Analysis Center (CDIAC). In this study, we used two CO_2 fluxes, i.e., GPP and *Reco*, to optimize and validate the model. GPP and *Reco* were estimated via the flux partitioning method described in Reichstein et al. (2005). This method extrapolated nighttime measurement of NEE (representing nighttime *Reco*) into daytime *Reco* using a short-term calibrated temperature response function. GPP was then calculated as the difference between *Reco* and NEE. We used hourly gap-filled Level 2 climate and flux data from sites of Ha1, Ne1, Ne2, Ne3 and MMS, and half-hourly gap-filled level 2 climate and flux data from sites of Ho1, Ho2 and Ton. Gap-filled Level 2 data were not available for sites of SP2, SP3, IB1 and IB2, so we used half-hourly Level 4 flux data products, which were gap-filled by Marginal Distribution Sampling (MDS) method. Hourly or half-hourly LAI data were linearly interpolated from LAI measurements. Information about instrument and method used to collect each variable was available in the references listed in Table 1.

2.3. Model description and parameters

The model used in this paper is a Flux-Based Ecosystem Model (FBEM), which was designed to simulate net CO_2 ecosystem exchange with a data-model fusion module (Wu et al., 2009; Yuan et al., 2012; Du et al., 2015). The model was driven by four variables, including *Ta*, PAR, RH and LAI. Flux data, GPP and *Reco*, were used to constrain parameter estimation. In brief, FBEM consists of two major carbon cycling processes: canopy-level photosynthesis (A_c) and ecosystem respiration (*Reco*), which are regulated by environmental variables (Fig. 2). Details about each part of the model and equations were listed as followed.

2.3.1. Leaf-level photosynthesis

Leaf-level photosynthesis is described by a model developed by Farquhar et al. (1980). For C_3 plants, gross leaf CO_2 uptake (A, μmol CO_2 m^{-2} s^{-1}) is calculated as

$$A = \min\left\{J_c, J_e\right\} \tag{1}$$

where J_c and J_e represent the rate limited by carboxylation enzymes and by light electron transport, respectively. The carboxylation processes (J_c, μmol CO_2 m^{-2} s^{-1}) are

$$J_c = V_m \times \frac{C_i - \Gamma_*}{C_i + K_c \times (1 + O_x/K_0)} \tag{2}$$

where C_i is the leaf internal CO_2 concentration (μmol CO_2 mol^{-1} air), expressed as

$$C_i = f_{C_i} \times C_a \tag{3}$$

C_a is ambient CO_2 concentration and f_{Ci} is ratio of leaf internal CO_2 to ambient air CO_2 concentration. O_x is oxygen concentration in the air (0.21 mol O_2 mol^{-1} air). V_m is maximum carboxylation rate (μmol CO_2 m^{-2} s^{-1}), which is related to canopy temperature T_k (K) by Arrhenius' equation:

$$V_m = V_m^{25} \times \exp\left(\frac{E_{V_m} \times (T_k - 298)}{R \times T_k \times 298}\right) \tag{4}$$

with an activation energy E_{V_m}, where V_m^{25} is maximum carboxylation rate at 25 °C and R is universal gas constant (8.314 J K^{-1} mol^{-1}). The CO_2 compensation point without dark respiration is represented as Γ_* (μmol CO_2 m^{-1}). It is also adjusted by Arrhenius' equation in

$$\Gamma_* = \Gamma_*^{25} \times \exp\left(\frac{E_{\Gamma_*^{25}} \times (T_k - 298)}{R \times T_k \times 298}\right) \tag{5}$$

where Γ_*^{25} is the CO_2 compensation point without dark respiration at 25 °C and $E_{\Gamma_*^{25}}$ describes the temperature dependence of Γ_*. Two Michaelis–Menten constants have a temperature dependence based on the Arrhenius' equation similar to V_m.

K_c, Michaelis–Menten constant for carboxylation (μmol mol^{-1}), was represented by

$$K_c = K_c^{25} \times \exp\left(\frac{E_{K_C} \times (T_k - 298)}{R \times T_k \times 298}\right) \tag{6}$$

with an activation energy E_{K_C}, where K_c^{25} is the Michaelis–Menten constant for carboxylation at 25 °C.

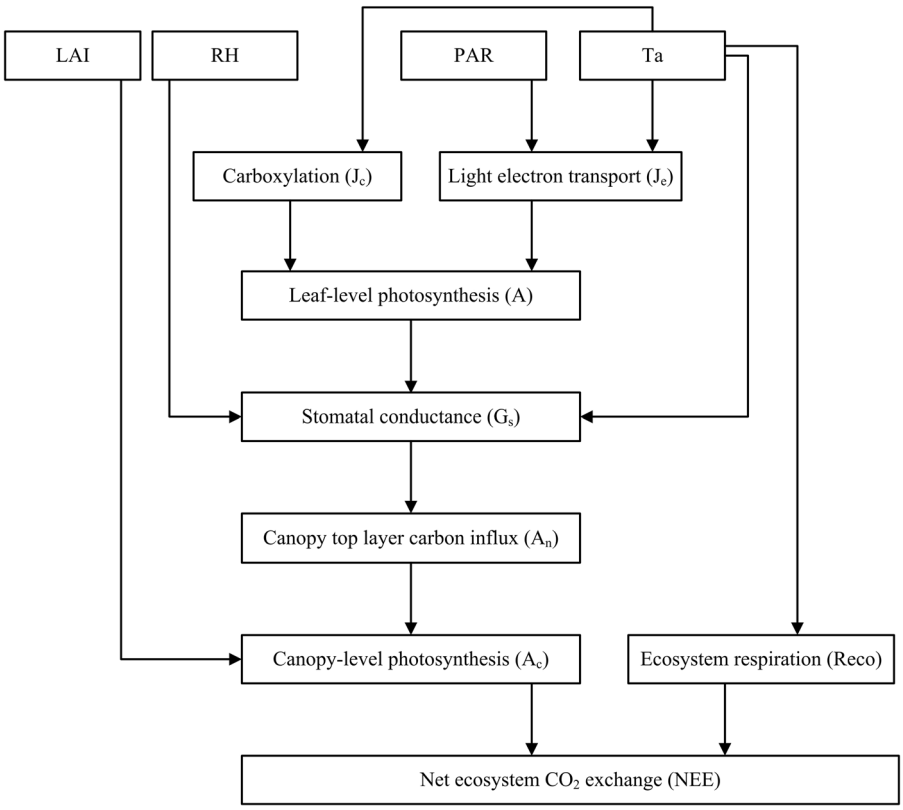

Fig. 2. Schematic diagram of FBEM for data-model fusion. This model is forced by four environmental variables: LAI, *Ta*, PAR and RH. LAI, leaf area index; *Ta*, air temperature; PAR, photosynthetically active radiation; RH, relative humidity. Leaf-level photosynthesis (A) is calculated by the rates of carboxylation enzymes (J_c) and light electron transport rates (J_e) according to a model developed by Farquhar et al. (1980). Then A is coupled with stomatal conductance (G_s) to calculate the carbon influx of the top leaf layer (A_n). Canopy level photosynthesis (A_c), which equals to gross primary production (GPP), is calculated by leaf area index (LAI) and the carbon influx of the top leaf layer (A_n). *Ta* influences J_c, J_e and G_s. PAR is involved in the calculation of J_e. RH is related to the calculation of G_s. Ecosystem respiration (*Reco*) is only regulated by air temperature.

K_o, Michaelis–Menten constant for oxygenation (μmol mol^{-1}), was represented as

$$K_o = K_o^{25} \times \exp\left(\frac{E_{K_o} \times (T_k - 298)}{R \times T_k \times 298}\right) \quad (7)$$

with an activation energy E_{K_o}, where K_o^{25} is the Michaelis–Menten constant for oxygenation at 25 °C.

The light electron transport processes (J_e, μmol CO$_2$ m^{-2} s^{-1}) are

$$J_e = \frac{\alpha_q \times I \times J_m}{\sqrt{{J_m}^2 + {\alpha_q}^2 \times I^2}} \times \frac{C_i - \Gamma_*}{4 \times (C_i + 2\Gamma_*)} \quad (8)$$

when I is absorbed PAR (μmol m^{-2} s^{-1}). α_q is quantum efficiency of photon capture (mol mol^{-1} photon) and J_m is maximum electron transport rate (μmol CO$_2$ m^{-2} s^{-1}). J_m depends on temperature and is computed by

$$J_m = r_{J_{m}V_m} \times V_m^{25} \times \exp\left(\frac{E_{V_m} \times (T_k - 298)}{R \times T_k \times 298}\right) \quad (9)$$

where $r_{J_{m}V_m}$ is the ratio of J_m to V_m^{25} at 2 °C.

2.3.2. Stomatal conductance

The stomatal conductance (G_s) is coupled with leaf-level photosynthesis by Leuning model (Leuning, 1995), so that the carbon influx of the top leaf layer (A_n) is estimated by

$$A_n = G_s \times (C_a - C_i) \quad (10)$$

$$G_s = g_l \times \frac{A}{(C_a - \Gamma_*) \times \left(1 + \frac{D}{D_0}\right)} \quad (11)$$

where g_l and D_0 (kPa) are empirical coefficients and D is vapor pressure deficit (kPa). Vapor pressure deficit is calculated by air temperature (T_k) and RH (in %) (Chang, 2006).

$$\ln e_s = 21.382 - \frac{5347.5}{T_k} \quad (12)$$

$$D = 0.1 \times e_s \times (1 - \mathrm{RH}) \quad (13)$$

where e_s is saturation vapor pressure (mbar).

2.3.3. Canopy-level photosynthesis

In order to scale up leaf-level photosynthesis to canopy-level photosynthesis, an approach of Sellers et al. (1992) was used to describe the relationship between the canopy photosynthesis (A_c) and the carbon influx of the top leaf layer, derived as

$$A_c = A_n \times \frac{1 - \exp(-k_n \times \mathrm{LAI})}{k_n} \quad (14)$$

where k_n is light extinction coefficient.

2.3.4. Ecosystem respiration

Ecosystem respiration (*Reco*) is modeled as a function of temperature (*Ta*, in °C) with the widely used van't Hoff equation (van't Hoff, 1899):

$$R_{eco} = R_{eco}^0 \times Q_{10}^{Ta/10} \quad (15)$$

where R_{eco}^0 is ecosystem respiration at 0 °C and Q_{10} is the relative increase ($Reco/R_{eco}^0$) in respiration for every 10 °C rise in temperature. NEE was calculated by:

$$\mathrm{NEE} = Reco - A_c \quad (16)$$

Table 2
Symbols, definitions, units, prior values, ranges of parameters, source and related equations that were used in data-model fusion.

Parameter	Defination	Unit	Value	Minimum	Maximum	Source	Equation
α_q	Canopy quantum efficiency of photon conversion	mol mol^{-1} photon	0.28	0	0.5	1	8
K_c^{25}	Michaelis–Menten constant for carboxylation	μmol mol^{-1}	460	50	600	1	6
E_{Kc}	Activation energy of K_c^{25}	J mol^{-1}	59,356	30,000	150,000	1	6
E_{k0}	Activation energy of K_0^{25}	J mol^{-1}	35,948	10,000	60,000	1	7
K_0^{25}	Michaelis–Menten constant for oxygenation	mol mol^{-1}	0.33	0.2	0.5	1	7
E_{Vm}	Activation energy of V_m^{25}	J mol^{-1}	58,520	10,000	100,000	1	4
Γ_*^{25}	CO_2 compensation point without dark respiration	μmol mol^{-1}	42.5	10	200	1	5
r_{JmVm}	Ratio of J_m to V_m^{25} at 25 °C		1.79	1	5	1	9
R_{eco}^0	Whole ecosystem respiration at 0 °C	μmol CO_2 m^{-2} s^{-1}	2.5	1	5	3	15
Q_{10}	Temperature dependency of ecosystem respiration		2	1	3	3	15
V_m^{25}	Maximum carboxylation rate at 25 °C	μmol CO_2 m^{-2} s^{-1}	29	10	300	1	9
f_{Ci}	Ratio of internal CO_2 to air CO_2		0.87	0.5	0.9	1	3
k_n	Canopy extinction coefficient for light		0.8	0.7	0.9	1	14
$E_{\Gamma_*^{25}}$	Activation energy of CO_2 compensation point at 25 °C	J mol^{-1}	60,000	30,000	100,000	1	5
g_l	Empirical coefficient in Leuning model		1657	100	2000	2	11
D_0	Empirical coefficient in Leuning model	kPa	2.74	1	10	2	11

1) Knorr and Kattge (2005); 2) Van Wijk et al. (2000); 3) Novick et al. (2004).

In total, there were 16 parameters that govern the model's behavior (Table 2).

Ton site was distinct from other sites for high temperature and scarce precipitation during summer, while other sites had humid climate throughout the year (Table 1 and Fig. S1). Given the special pattern of climate at Ton, the response of respiration to soil moisture should be incorporated into the original model. Here we adopted the asymptote model (Gulledge and Schimel, 2000) and incorporated soil moisture and an additional parameter a_1 into the original model (Eq. (17)). Other processes and parameters remained the same as in the original model.

$$R_{eco} = \frac{\mathrm{M}}{\mathrm{M} + a_1} \times R_{eco}^0 \times Q_{10}^{T_a/10} \tag{17}$$

where M was soil moisture (g H_2O g^{-1} dry soil), a_1 represented the moisture at which the respiration rate was half the maximum. The prior range of a_1 was set from Gulledge and Schimel (2000). The asymptote model assumed that as moisture increased, respiration asymptotically approached some maximum rate (as allowed by other factors such as temperature), but that moisture did not directly alter the biota's temperature sensitivity.

2.4. Bayesian inversion with Markov Chain Monte Carlo technique

Data-model fusion technique can be used to inform model parameters by incorporating observations into a model. According to Bayes' theorem, posterior probability density functions (PDFs) of model parameters (p) can be calculated from prior knowledge and information generated by comparing modeled with observed values. The theorem can be described as (Xu et al., 2006; Mosegaard and Sambridge, 2002):

$$p\left(\theta|Z\right) = \frac{p\left(Z|\theta\right) p\left(\theta\right)}{p\left(Z\right)} \tag{18}$$

where $p\left(\theta|Z\right)$ is the posterior distribution of the parameters θ given the observations Z. $p\left(\theta\right)$ is a set of uniform distributions over the ranges specified in Table 2, and $p(Z)$ is the probability distribution function of observations. $p\left(Z|\theta\right)$ is a likelihood function.

In this study, we adopted those prior ranges of model parameters from Wu et al. (2009) and were set the same for each site (see Table 2). The prior ranges of parameters were defined according to values published in the literature, educated guesses and measurements (Knorr and Kattge, 2005). Then, posterior parameter space was selected by Metropolis-Hastings (M-H) algorithm. The main idea of M-H algorithm was to use MCMC technique to generate multi-dimensional PDFs of model parameters via a sampling procedure (Hastings, 1970; Metropolis et al., 1953). In this study, we ran the M-H algorithm by repeating two steps: a proposing step and a moving step (Xu et al., 2006). In each proposing step, the algorithm proceeded from the previously accepted parameter vector to produce a new parameter vector within the parameter space as a random walk (Δp). Δp was defined by a random number (r) between 0 and 1, the vector of parameter minima ($p_{\min}$) and maxima ($p_{\max}$) and a step length factor (s) as:

$$p^{i+1} = p^i + \Delta p = p^i + \frac{(r - 0.5)}{s} \times (p_{\max} - p_{\min}) \tag{19}$$

where p^i and p^{i+1} were previously accepted and new parameter vector, respectively.

In each moving step, p^{i+1} from the proposing step was tested against the Metropolis criterion (Metropolis et al., 1953; Xu et al., 2006) to determine whether it should be accepted or rejected. This criterion was a probability to accept the proposed parameters or not, which was generated from likelihood functions of proposed parameters relative to the parameters accepted at a previous step. Here we calculated likelihood function $p\left(Z|\theta\right)$ with the assumption that random flux error followed double-exponential (Laplace) distribution (Hollinger and Richardson, 2005; Richardson et al., 2006; Zhang et al., 2008):

$$p\left(Z|\theta\right) \propto \exp\left\{-\sum_{i=1}^{2}\sum_{t\,\in\, obs(Z_i)} -\left|\frac{Z_i(t) - X_i(t)}{\beta_i}\right|\right\} \tag{20}$$

where $Z_i(t)$ was observation and i represented ith dataset (in this study including GPP and *Reco*), X_i (t) was the simulated flux. β_i was the mean of the absolute deviations of the ith observations from their mean:

$$\beta_i = \frac{1}{N}\sum_{t=1}^{N}|Z_i(t) - \overline{Z_i}| \tag{21}$$

where $\overline{Z_i}$ was the mean of ith observations. The probability to accept the new parameters (moving to the next step) was calculated by:

$$P(p^i, p^{i+1}) = \min\left\{1, \frac{p_{i+1}\left(Z|\theta\right)}{p_i\left(Z|\theta\right)}\right\} \tag{22}$$

where $p_{i+1}\left(Z|\theta\right)$ was the likelihood function of generated parameters, $p_i\left(Z|\theta\right)$ was the likelihood function of accepted parameters in the last step (as in Eq. (18)). Once the Metropolis criterion was reached and the proposed parameters were accepted, the accepted

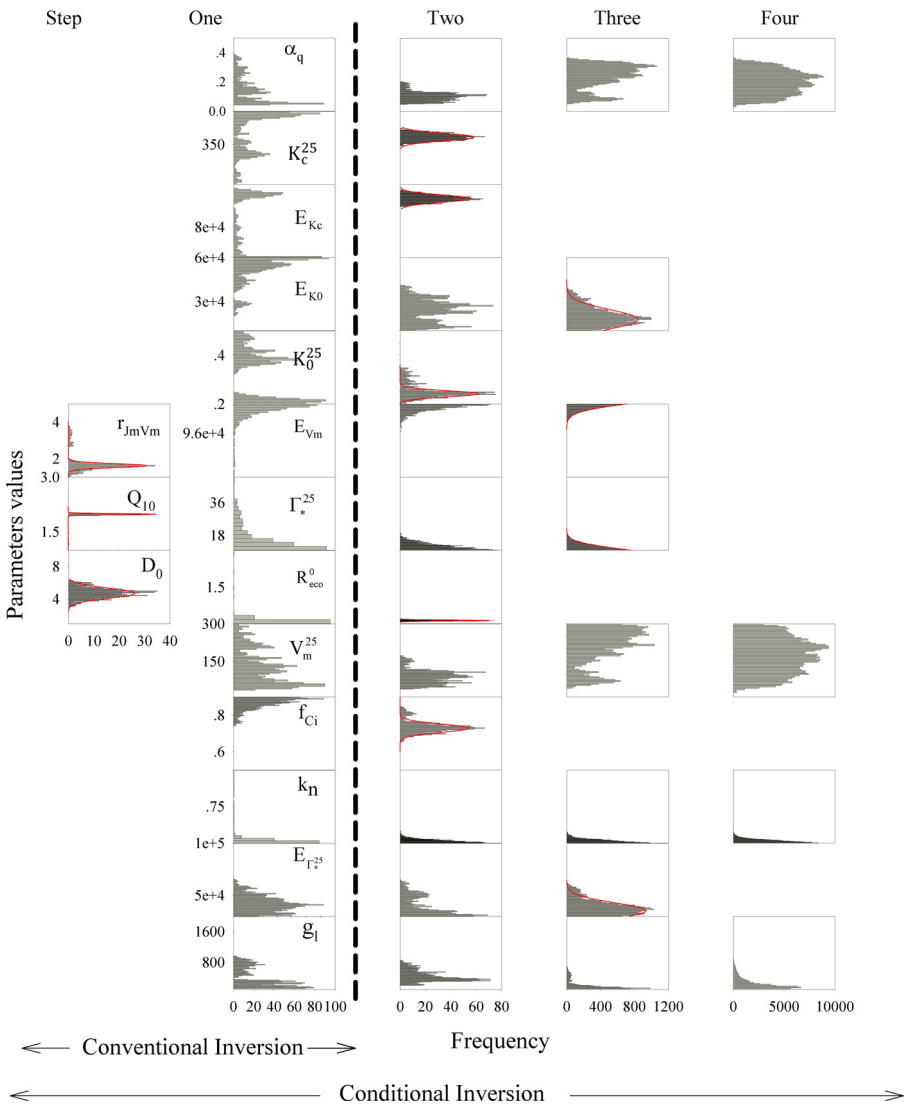

Fig. 3. An example about posterior frequency distribution of parameters by conditional inversion method at MMS. Panels on the left side of the dashed line showed the optimization result by one step inversion: only three parameters (e.g., r_{JmVm}, Q_{10} and D_0) were well-constrained. This one step was also called conventional inversion. Well-constrained parameters were evaluated by two criteria. Gaussian distribution curves were plotted to mark the well-constrained parameters. Panels on the right showed the step-by-step optimization results. After four steps, there were still four parameters (e.g., α_q, V_m^{25}, k_n and g_l) that were not constrained (please refer to Table 2 for abbreviation and unit of each parameter).

parameters were then used to generate a new set of parameters (repeating proposing step). Otherwise, the proposed parameters were abandoned and a new set of parameters was produced based on the accepted parameters from the previous iteration. After the model was run for enough times, the model parameters estimated by the M–H algorithm converged to stationary distributions.

In this study, we ran five parallel chains with different parameter initial values and set 1,000,000 iterations for each run. The first 10% of the accepted values were discarded as burn-in (Xu et al., 2006). Then the union of accepted parameter samples of five chains after burn-in were tested the convergence by the Gelman–Rubin (G–R) diagnostic method (Gelman and Rubin, 1992). Samples could only be used when statistical inferences G–R statistics approached to 1 (data not shown).

2.5. Conditional Bayesian inversion

The conditional Bayesian inversion approach was developed by Wu et al. (2009) based on the conventional Bayesian inversion method. After conducting conventional Bayesian inversion once, we obtained medians from posterior distribution as maximal likelihood estimates (MLEs) of parameters that were well constrained. Well-constrained parameters exhibit well-defined unimodal distributions and tend to have small retrieved standard deviations relative to the parameters' allowable ranges (Braswell et al., 2005). Here, we evaluate well-constrained parameters by two criteria: (1) whether or not the posterior distributions of parameters follow Gaussian distribution and (2) how well the modeled values are correlated with observed ones. If posterior distributions (i.e., histograms) of estimated parameters follow Gaussian distributions and the model fitting results are good, the estimated parameters are considered well-constrained (Du et al., 2015). Then, MLEs for those well-constrained parameters were set as prior values in the model for the next step of inversion. We repeated this method until no more parameters could be constrained. Since the information derived from the new constrained parameters was based on the previous step, this process was called conditional Bayesian inversion.

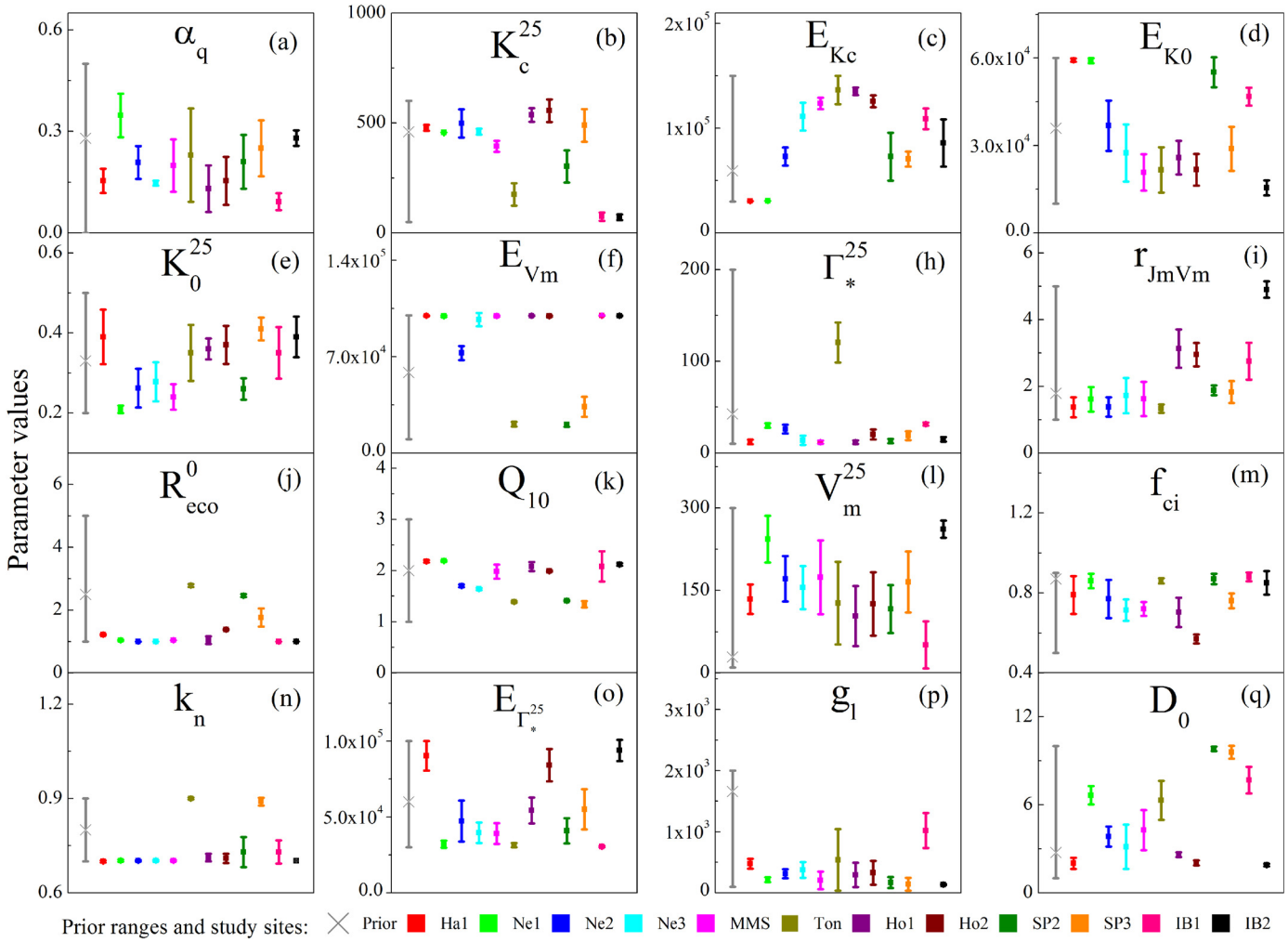

Fig. 4. Randomly selected initial values (gray crosses), prior ranges (gray lines), MLEs (points) and standard deviations (error bars) for all parameters across all sites. Points and error bars with different colors indicated different sites (please refer to Table 2 for abbreviation and unit of each parameter).

2.6. Statistical methods

We used SPSS 16.0 to analyze bivariate correlation for each pair of parameters in each inversion step. We conducted conditional inversion for all sites and obtained the MLEs of well-constrained parameters (sample medians for unconstrained parameters were also used when we could not acquire MLEs). Then the optimized parameter sets were used to generate GPP, *Reco* and NEE time series for model evaluation. For parameter variation analysis, we calculated means, standard deviations and coefficients of variation of parameter MLEs across all sites, among adjacent sites and based on PFTs where MLEs of parameters were available. Coefficient of variation (CV) is an index revealing the extent of variability across sites. Here we only considered the effects of climate conditions on the variation of parameters. We made stepwise multiple linear regression analysis between MLEs of parameters and three environmental variables in Statistical Analysis System (SAS). To determine which variable contributed most to the variation of parameters, we calculated coefficient of partial determination (i.e., partial R^2) at the significant level $P < 0.05$. Partial R^2 is the index indicating the improvement of overall model fit R^2 when each variable is added into the regression, thus could be used to represent the relative importance of each variable to the regression model.

3. Results

3.1. Parameter estimation

By applying the conditional inversion method to constrain parameteres at Morgan Monre State Forest (MMS) in 2006, we generated histograms to show posterior frequency distribution of parameters (Fig. 3). Panels on the left side of dashed line showed the parameter optimization results by one step inversion: only three parameters (i.e., r_{JmVm}, $Q_{10}$and D_0) were well-constrained. Gaussian distribution curves were plotted to mark the well-constrained parameters. Panels on the right showed the step-by-step optimization results. After four steps, there were still four parameters (i.e., α_q, V_m^{25}, k_n and g_l) that were not well-constrained.

Bivariate correlation analysis was made for each inversion step at MMS to represent how every two parameters were correlated (Table S1). In the first inversion step, parameters r_{JmVm}, Q_{10} and D_0 were constrained simultaneously and were highly correlated. Two parameters constrained in the second step were strongly correlated with parameters constrained in the first step: K_0^{25} was correlated with D_0, R_{eco}^0 was correlated with Q_{10} and r_{JmVm}. In the third inversion step, E_{k0}, E_{Vm} and Γ_*^{25} were constrained and they had strong correlations with at least one of the parameters constrained in the first inversion step. After four inversion steps, the parameters that

were still not well-constrained (i.e., α_q, V_m^{25} and g_l) had strong correlations among them.

After conducting conditional inversion for all sites, we obtained an optimized parameter set to match the observations in each site. The sequence of parameter optimization was shown in Table S2 and MLEs of well-constrained parameters were shown in Table S3. There were nine parameters (i.e., K_c^{25}, E_{k0}, K_0^{25}, E_{Vm}, Γ_*^{25}, r_{JmVm}, Q_{10}, f_{Ci} and D_0) well-constrained by datasets at all of the sites. Other parameters were only constrained by datasets at some sites (see Table S2). Prior ranges, MLEs and standard deviations for all parameters across all sites were shown in Fig. 4 (sample medians and standard deviations were also used for parameters that were not well-constrained).

3.2. Model evaluation results

The optimized parameter sets were used to generate the modeled GPP, *Reco* and NEE time series in the same year. The modeled flux values matched the observed ones well for all sites except for Ton (Table S4 and Fig. 5).

The poor optimization results at Ton was mainly due to the bad performance in simulating ecosystem respiration, as shown in Fig. S2a. *Reco* was significantly underestimated during the first half year while overestimated during the second half year by original model. Then we used the improved respiration model (Eq. (17)) to regenerate MLEs of parameters at Ton. After 8 steps, 14 out of 17 parameters were well-constrained, MLEs of parameters and standard deviations were shown in Table S3 and Fig. 4. Parameters related to respiration were well-constrained by ecosystem respiration flux data and soil moisture data, including the newly-added one a_1. The convergence of parameter samples was improved for R_{eco}^0 and Q_{10} in the improved model, and MLEs of these two parameters were higher than those in the original model (Fig. S3). The new simulations performed much better and captured the seasonal dynamic of *Reco* (Fig. S2b).

We carried out additional simulations in another year using the MLEs of parameters to prevent overfitting for each site. Model validation results were good at most of the sites except at Ton and IB1 (Table S5 and Fig. 4). For IB1, the model captured the NEE trend intra-annually but overestimated daily net CO_2 uptake in summer. This result might be due to the crop type change from corns to beans in 2007 in this cropland (Xin et al., 2015). But the MLEs of parameters at IB1 were still applicable for comparision with those at other sites in 2006 or 2004. For Ton, although optimization results were improved by the new simulations, the model did not capture the overall trend of daily NEE and model fit results were not satisfactory in another year. So the MLEs of parameters at Ton might not be acceptable.

3.3. Analysis of cross-site variability of parameters

The site of Ton was excluded from following analysis because of the unsatisfactory model validation results. Across the remaining sites, the CVs of g_l and D_0 were greater than 0.5, which indicated strong variability. Other parameters such as α_q, K_c^{25}, E_{Kc}, E_{k0}, K_0^{25}, E_{Vm}, Γ_*^{25}, r_{JmVm}, R_{eco}^0, Q_{10}, V_m^{25} and $E_{\Gamma_*^{25}}$ showed medium variability with CVs varying from 0.16 to 0.47. The CVs of f_{Ci} and k_n were relatively low (less than 0.15) and these parameters exhibited weak variability across sites (Table 3).

Within one cluster of the three sites in Nebraska, α_q, E_{k0}, and D_0 exhibited relatively strong variation compared with other parameters. Parameters optimized in Howland Forest varied little between the two towers Ho1 and Ho2. The third group SP1 and SP2 also had similar MLE for each parameter except for E_{k0}. Many parameters differed a lot between IB1 and IB2. For example, the CVs of α_q, E_{k0}, V_m^{25}, g_l and D_0 were greater than 0.7 (Table 4).

Table 3
Means, standard deviations and coefficients of variation of MLEs of parameters across sites (please refer to Table 2 for abbreviation and unit of each parameter).

Parameter	Mean value	Standard deviation	Coefficient of variation
α_q	0.21	0.10	0.47
K_c^{25}	392.2	172.1	0.44
E_{Kc}	100,711	25,444	0.25
E_{k0}	36,108	16,265	0.45
K_0^{25}	0.32	0.07	0.22
E_{Vm}	83,755	29,246	0.35
Γ_*^{25}	18.5	7.47	0.40
r_{JmVm}	2.29	1.07	0.47
R_{eco}^0	1.42	0.53	0.37
Q_{10}	1.88	0.31	0.16
V_m^{25}	169.3	76.6	0.45
f_{Ci}	0.77	0.09	0.12
k_n	0.75	0.08	0.10
$E_{\Gamma_*^{25}}$	57,636	23,116	0.40
g_l	422.5	315.7	0.75
D_0	4.87	3.04	0.62

3.4. Analysis of parameter variation within and across PFTs

We compared MLEs of parameters within and across PFTs after sites were grouped based on PFTs. There existed large variation within PFTs for some of the parameters with CVs greater than 0.5: E_{k0}, $E_{\Gamma_*^{25}}$, g_l and D_0 within DBF; α_q, K_c^{25}, V_m^{25} and g_l within CRO; E_{Vm} and D_0 within ENF. Some parameters varied little within the same PFTs with CVs smaller than 0.1: E_{Vm}, Γ_*^{25}, Q_{10} and f_{Ci} within DBF; f_{Ci} within CRO; k_n within ENF. Across various PFTs, CRO had the highest Γ_*^{25} value. GRA had relatively higher values for r_{JmVm} and V_m^{25} compared with other PFTs (Table S6 and Fig. 6).

3.5. Statistical correlations between MLEs of parameters and environmental variables

We chose three environmental variables: MAT, MAP and total PAR in a year (tPAR), which were representatives of a site's climate, to examine their influences on parameter variation. Ne1, Ne2 and IB1 were excluded from the regression analysis since they were irrigated croplands with substantial human disturbance.

As shown in Table 5, five out of 16 parameters had significant linear correlations with MAT, MAP or tPAR ($P < 0.05$). MAT explained most of the variation of E_{Kc}, E_{Vm}, Q_{10} and D_0, while MAP dominated the variation of R_{eco}^0. tPAR seemed not to be the factor related to the variation of parameters. E_{Kc}, E_{Vm} and Q_{10} were negatively correlated with MAT ($R^2 = 0.74$, 0.95 and 0.81 respectively, $P < 0.05$), while D_0 increased as MAT increased ($R^2 = 0.97$, $P < 0.05$). R_{eco}^0 had a positive linear correlation with MAP ($R^2 = 0.89$, $P < 0.05$).

4. Discussion

4.1. Variation of parameters across sites

The C_3 leaf photosynthesis model of Farquhar et al. (1980) and respiration model of van't Hoff (1899) have been widely used to project the biosphere responses of C cycles to global change in earth system models (Sellers et al., 1996; Sitch et al., 2003; Bernacchi et al., 2003; Rayner et al., 2005). Many of the involved parameters were calibrated based on site measurements. The variation of parameters has been well documented in field experiments, such as the maximum rate of electron transport (J_{max}) by Wullschleger (1993), maximum rate of carboxylation ($V_{c,max}$) by Wilson et al. (2000) and Q_{10} by Xu & Qi (2001). By synthesizing many A–C_i curves

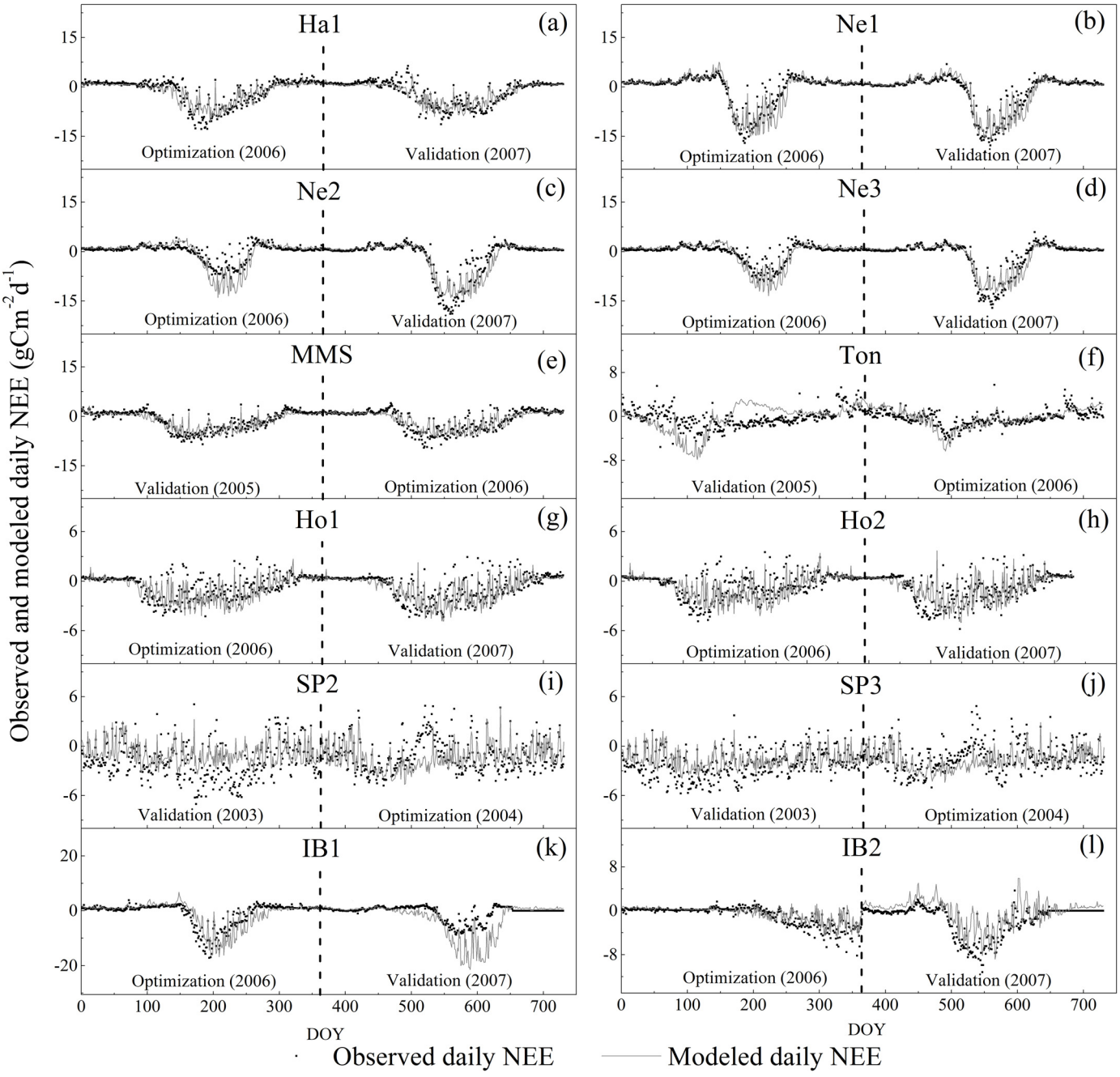

Fig. 5. Model optimization and validation results for all sites. In the optimization phases, the model captured the seasonal dynamics of NEE well for all sites (improved model was used at Ton). Forcing data and flux data in another year were used for model validation. The validation results were relatively good for almost all sites except at Ton and IB1.

in published studies, Medlyn et al. (2002) found herbaceous species and crops differed from tree species in several aspects, including activation energies of both J_{max} and $V_{c,max}$ and the ratio of J_{max}: $V_{c,max}$ at 25 °C (i.e., r_{JmVm}), suggesting that alternative parameter sets were required for modeling these two plant types. Our study was among the first studies of identifying the cross-site variation of all the involved parameters by data-model fusion technique from eddy-covariance data. Most of the site-averaged MLEs and standard deviations of parameters we obtained (Table 3) fell in ranges of the parameter values published before (see Table 2 in Wu et al., 2009). Many of the site-averaged MLEs of parameters were highly comparable with the commonly used values in literature such as α_q, r_{JmVm} and Q_{10} (e.g., Wang et al., 2001; Leuning, 1997; Friedlingstein et al., 2006).

The analysis of parameter variation across all sites with different vegetation and climate types shows an integral view of parameter variability. Based on the results from this study, f_{Ci} and k_n are the least variable parameters, indicating they could be used as constants for different sites in models. For other variable parameters (e.g., r_{JmVm}, g_l and D_0), it is not appropriate to set constant values for different sites. The different variability in parameters might be the reflection of the different internal properties of ecosystems and deserves further investigation about model structure.

Variation analysis of parameters at four groups of adjacent sites also gives us insights about how parameters vary under similar climate conditions. The large variation of many parameters optimized in IB1 and IB2 might be the result of the different vegetation covers at these two sites. Three sites in Nebraska were all

Table 4
Means and coefficients of variation of MLEs of parameters across adjacent sites. The most variable parameters within each group were highlighted (CVs greater than 0.5). Blank represented no available MLE for this parameter or there was only one site to calculate CV (please refer to Table 2 for abbreviation and unit of each parameter).

	α_q	K_c^{25}	E_{Kc}	E_{k0}	K_0^{25}	E_{Vm}	Γ_*^{25}	r_{JmVm}	R_{eco}^0	Q_{10}	V_m^{25}	f_{Ci}	k_n	$E_{\Gamma_*^{25}}$	g_l	D_0
Ne1, Ne2, Ne3																
Mean	0.23	471.5	91989	41120	0.25	89646	23.3	1.57	1.04	1.8	189.8	0.78		39654	301	4.5
CV	0.44	0.05	0.29	0.40	0.14	0.17	0.36	0.11		0.16	0.25	0.09		0.19	0.27	0.41
Ho1, Ho2																
Mean		546	130349	23757	0.37	99615	15.9	3.04	1.21	2		0.64	0.7	69223		2.32
CV		0.02	0.05	0.12	0.01	0.002	0.37	0.04	0.20	0.03		0.15	0.002	0.30		0.17
SP1, SP2																
Mean		395.4	71679	42011	0.34	27153	15.9	1.86	2.11	1.38		0.82	0.81	47912		9.7
CV		0.33	0.02	0.4	0.32	0.33	0.26	0.02	0.23	0.03		0.10	0.14	0.21		0.02
IB1, IB2																
Mean	0.19	72.7	97368	31114	0.37	99777	23.2	3.8		2.1	156.2	0.87	0.73	62131	577.1	4.8
CV	0.71	0.03	0.17	0.7	0.07	0.0	0.50	0.40		0.01	0.95	0.02			1.1	0.85

croplands but management strategies differed among them. Only some of the parameters showed relatively large variation among these three sites. SP2 and SP3 were two sites with major difference in site histories, and parameters exhibited small variation between them. Parameters optimized in Howland forest also varied little between the two towers Ho1 and Ho2, probably because of the similar environmental conditions, vegetation types and management strategies. From the analysis above, variance in vegetation types might be the major factor influencing the cross-site variation of parameters when climate conditions are quite the same.

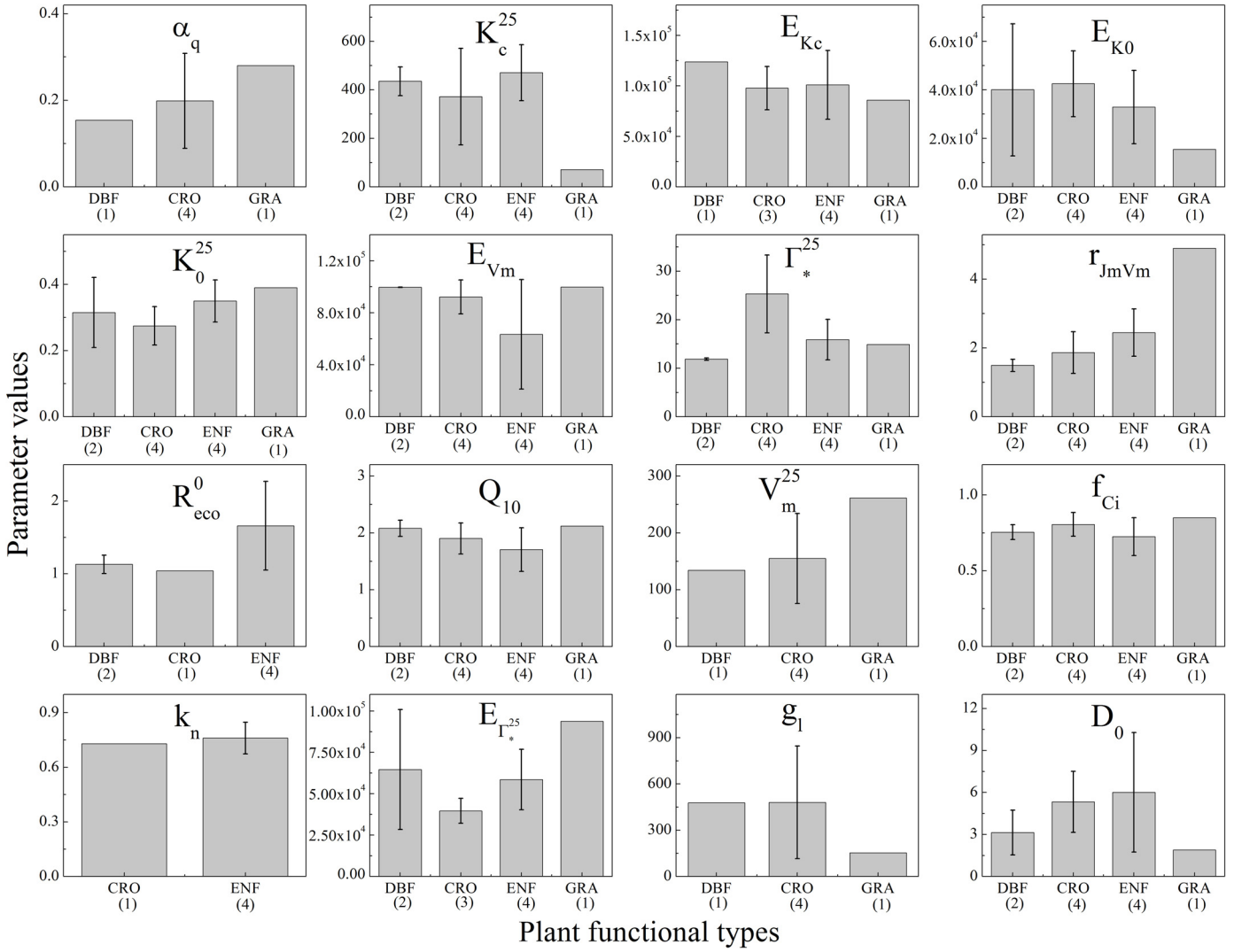

Fig. 6. Means and standard deviations of MLEs of parameters when sites were grouped based on PFTs. DBF, deciduous broadleaf forests; ENF, evergreen needleleaf forests; CRO, croplands; GRA, grasslands. The number at the bottom indicated the number of sites where MLEs of well-constrained parameters were available in each type (please refer to Table 2 for abbreviation and unit of each parameter).

Table 5
Stepwise multiple linear regression results between MLEs of parameters and environmental variables MAT, MAP and tPAR (only significant relationships were listed).

	Partial R^2[a]			Regression equation[a]
	MAT	MAP	tPAR	
E_{Kc}	0.74	0	0	E_{Kc} = −3654.7MAT + 145651 (R^2 = 0.74)
E_{Vm}	0.88	0	0.07	E_{Vm} = −9913MAT + 92138tPAR-75185 (R^2 = 0.95)
R^0_{eco}	0	0.89	0	R^0_{eco} = 0.004MAP-2.76 (R^2 = 0.89)
Q_{10}	0.81	0	0	Q_{10} = −0.05MAT + 2.38 (R^2 = 0.81)
D_0	0.93	0.04	0	D_0 = 0.43MAT + 0.005MAP-5.79 (R^2 = 0.97)

[a] All statistics involved are significant at 0.05 level.

In some land surface models, most parameters are given fixed constants based on PFTs (Bonan et al., 2003; Sitch et al., 2003; Krinner et al., 2005). Our results indicate that some parameters vary a lot within PFTs. The inadequate representation of site specialty might cause bias in simulating ecosystem dynamics (Groenendijk et al., 2011). In term of parameter variation across PFTs, Medlyn et al. (2002) found higher r_{JmVm} value for croplands than tree species, while we found grasslands had the highest r_{JmVm} value compared with forests and croplands. Groenendijk et al. (2011) estimated parameter values based on ecosystem types and found that cropland parameters had the highest values for V_m^{25} and α_q among seven ecosystem types. However, we found higher V_m^{25} value in the grassland sites than in the croplands or forest sites. In the future research, we need more sites to study the variation of parameters within and across PFTs.

4.2. Relationship between model parameters and environmental variables

In ecological models, parameters estimated by data-model fusion might vary spatially because all key environmental factors such as air and soil temperature, air and soil moisture, and soil organic matter are spatially heterogeneous (Zhou et al., 2009). In this study, the 12 sites covered a variety of climate types. Here we assumed that different climate types might exert a big influence on the variation of parameters either directly or indirectly through vegetation types.

Our correlation analysis showed that the negative correlation between Q_{10} with MAT, was probably due to temperature acclimation of respiratory processes. Previous experimental studies have shown that Q_{10} decreased with increasing temperature (Kirschbaum, 1995; Luo et al., 2001). We found that a 1 °C increase in MAT could reduce the Q_{10} value by 0.05, which was smaller than previous studies of 0.08 (Chen and Tian, 2005) and 0.06 (Peng et al., 2009). E_{Kc} and E_{Vm} are activation energy for carboxylation and related to enzyme kinetics. Enzyme kinetics is strongly regulated by temperature. Our results showed these two parameters had negative linear correlations with MAT. It is reasonable that the values of two parameters decrease with increasing temperature. R^0_{eco} is whole ecosystem respiration at 0 °C. In the original respiration model, the limiting effect of soil moisture on respiration was not considered. The strong positive correlation between R^0_{eco} and MAP might be the reflection of model deficiency. D_0 is the sensitivity of the stomata to VPD. Wang et al. (2001) found that estimated D_0 values were quite variable among six eddy-covariance sites. This study found a strong positive linear correlation between D_0 and MAT, which might explain the variation. In the future, with more study sites, we can further explore whether the correlations still exist within each PFT, and whether the correlations are different across PFTs.

In the past, we used to simulate the responses of ecosystems to elevated temperature or changed precipitation pattern with an assumption that parameters are constants. The use of constants rather than varying parameters to represent plants' instantaneous responses to environmental forcing is not able to reflect how plants acclimate to changes in the environment (Wythers et al., 2005; Smith and Dukes, 2013). Attempts to incorporate the acclimation of plants into ecosystem models have been done by some researchers. For example, King (2006) applied reduced sensitivity of plant respiration to temperature increase in a global terrestrial ecosystem model, GTEC 2.0, and found a reduced magnitude of the positive feedback between climate and the carbon cycle in a warming world. Kattge et al. (2009) quantified the relationship of maximum carboxylation rate at 25 °C (V_m^{25}) to leaf nitrogen content per unit area for each PFT and incorporated the relationship in a terrestrial biosphere model, BETHY. Their results showed that the accuracy of modeled GPP was improved. According to our results, the correlations between MLEs of parameters, MAT and MAP across sites also indicate that in order to accurately project plants' and microbes' acclimation to future environmental change, the possible adjustments of model parameters to global change scenarios are necessary.

4.3. Implications for model improvement

Model structure and parameters are important sources of uncertainties in model predictions (Luo et al., 2015). In this study, we applied the FBEM to 12 sites with different vegetation types and climate types. Results showed that this model worked well for almost all the sites except for Ton, which had Mediterranean climate with moist winter and dry summer. The significant underestimation of respiration during winter might result from missing processes of utilizing soil moisture when precipitation was sufficient but the temperature was low. And the overestimation during summer at Ton site might due to the lack of responses to water limitation. After soil moisture was introduced as an additional variable into the model, ecosystem respiration was much better simulated. But for other sites, only considering the effect of temperature fluctuations can simulate the respiration process relatively well. In addition, the overall simulations of ecosystem respiration in our study are not as good as those of photosynthesis, and more efforts should be paid to the improvement of respiration modeling in the future work, such as considering the effect from soil moisture.

Besides the model structure, parameters should also be modified for specific climate conditions. For example, Carlyle and Than (1988) found a moisture-dependent Q_{10} gave an improved match with soil respiration measured at sites than a constant Q_{10} under an attenuated Mediterranean climate. However, introducing a variable to represent a constant parameter will also add model complexity. When a complicated model integrates more knowledge, issues might also arise such as making more parameters less identifiable with certain datasets (Luo et al., 2009). How to strike a balance between model complexity and parameter identifiability is an important area to explore in the future.

Optimization method is also crucial for parameter identifiability. Most of the previous studies only assimilated NEE to optimize photosynthesis and respiration parameters simultaneously (Wang et al., 2001; Braswell et al., 2005; Wu et al., 2009; Kuppel et al.,

2012). This assimilation strategy only searches for the best match with the difference between GPP and *Reco*, rather than the respective process. In this way, separating the effects of changing the values of several parameters controlling photosynthesis from the effects of changing parameters that control respiration is difficult. Thus, this assimilation strategy might lead to some biases in estimated MLEs of parameters. Combining multiple observations has been proved to be able to reduce parameter uncertainty (Williams et al., 2005; Yuan et al., 2012; Du et al., 2015; Frasson et al., 2015). In this study, we also designed experiments that only used observed NEE for all sites. Results showed that MLEs of parameters were different from those using both GPP and *Reco* (Table S7). Site-averaged α_q, r_{JmVm} and Q_{10} MLEs were smaller than commonly used ones. Although GPP and *Reco* are modeled values from NEE, previous researches have shown that different gap-filling and partitioning methods have consistency in capturing different carbon dynamics across different sites (e.g., Moffat et al., 2007; Desai et al., 2008). Therefore joint assimilation of GPP and *Reco* is highly recommendable and can effectively constrain photosynthesis and respiration parameters in ecological models.

5. Conclusions

In order to extend the site-specifically calibrated ecological model to regional and global scales, identifying the possible variation of parameters at site scale is of fundamental importance. In this study, we show that different parameters have different variability across 12 sites. For example, g_l and D_0 exhibited high variation whereas f_{Ci} and k_n varied little among sites. Vegetation types exerted a big influence on variation of parameters at adjacent sites where climate conditions were quite similar. When sites were grouped based on PFTs, MLEs of parameters varied within and across PFTs. If effects from vegetation types were not considered, we found that E_{Kc}, E_{Vm}, Q_{10} and D_0 had statistical linear correlations with MAT. R^0_{eco} exhibited linear correlation with MAP. These correlations could be interpreted by the acclimation of plants and microbes to environment. Future studies should further examine parameter variation over space and time with more sites and investigate their correlations with environmental variables and various ecosystem attributes. Model structure (especially the respiration process) and optimization method itself should also be improved for a better projection of terrestrial ecosystem C cycle.

Acknowledgements

We thank Anlei Wei and Weiming Song for providing useful suggestion in the writing. We acknowledge the collection of AmeriFlux database by all people involved. Without the availability of data, this research would not have been possible. We also thank the two anonymous reviewers for their insightful comments and suggestions. This study was supported financially by National Basic Research Program (or 973 program) of China (2013CB556601).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ecolmodel.2016.05.016.

References

AmeriFlux, 2007. AmeriFlux Network, http://ameriflux.ornl.gov/.
Baldocchi, D.D., 2003. Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past present and future. Global Change Biol. 9, 479–492.
Bonan, G.B., Levis, S., Sitch, S., Vertenstein, M., Oleson, K.W., 2003. A dynamic global vegetation model for use with climate models: concepts and description of simulated vegetation dynamics. Global Change Biol. 9, 1543–1566.
Bernacchi, C.J., Singsaas, E.L., Pimentel, C., Portis Jr., A.R., Long, S.P., 2001. Improved temperature response functions for models of Rubisco-limited photosynthesis. Plant Cell Environ. 24, 253–259.
Bernacchi, C.J., Pimentel, C., Long, S.P., 2003. In vivo temperature response functions of parameters required to model RuBP-limited photosynthesis. Plant Cell Environ. 26, 1419–1430.
Braswell, B.H., Sacks, W.J., Linder, E., Schimel, D.S., 2005. Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations. Global Change Biol. 11, 335–355.
Carlyle, J.C., Than, U.B., 1988. Abiotic controls of soil respiration beneath an eighteen-year-old *Pinus radiata* stand in south-eastern Australia. J. Clim. 76, 654–662.
Chang, M., 2006. Forest Hydrology: An Introduction to Water and Forests. CRC Press.
Chen, H., Tian, H.Q., 2005. Does a general temperature-dependent Q_{10}. Model of soil respiration exist at biome and global scale? J. Integr. Plant Biol. 47, 1288–1302.
Curtis, P.S., Hanson, P.J., Bolstad, P., Barford, C., Randolph, J.C., Schmid, H.P., Wilson, K.B., 2002. Biometric and eddy-covariance based estimates of annual carbon storage in five eastern North American deciduous forests. Agric. For. Meteorol. 113, 3–19.
Desai, A.R., Richardson, A.D., Moffat, A.M., Kattge, J., Hollinger, D.Y., Barr, A., Stauch, V.J., 2008. Cross-site evaluation of eddy covariance GPP and RE decomposition techniques. Agric. For. Meteorol. 148, 821–838.
Du, Z., Nie, Y., He, Y., Yu, G., Wang, H., Zhou, X., 2015. Complementarity of flux-and biometric-based data to constrain parameters in a terrestrial carbon model. Tellus 67, 24102, http://dx.doi.org/10.3402/tellusb.v67.24102.
Farquhar, G.D., von Caemmerer, S.V., Berry, J.A., 1980. A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. Planta 149, 78–90.
Frasson, R.P.D.M., Bohrer, G., Medvigy, D., Matheny, A.M., Morin, T.H., Vogel, C.S., Curtis, P.S., 2015. Modeling forest carbon cycle response to tree mortality: Effects of plant functional type and disturbance intensity. J. Geophys. Res. Biogeosci. 120, 2178–2193.
Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., Zeng, N., 2006. Climate-carbon cycle feedback analysis: results from the CMIP4 model intercomparison. J. Clim. 19, 3337–3353.
Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Stat Sci. 7, 457–472.
Gholz, H.L., Clark, K.L., 2002. Energy exchange across a chronosequence of slash pine forests in Florida. Agric. For. Meteorol. 112, 87–102.
Gulledge, J., Schimel, J.P., 2000. Controls on soil carbon dioxide and methane fluxes in a variety of taiga forest stands in interior Alaska. Ecosystems 3, 269–282.
Goulden, M.L., Munger, J.W., Fan, S.M., Daube, B.C., Wofsy, S.C., 1996. Measurements of carbon sequestration by long-term eddy covariance: methods and a critical evaluation of accuracy. Global Change Biol. 2, 169–182.
Groenendijk, M., Dolman, A.J., Van der Molen, M.K., Leuning, R., Arneth, A., Delpierre, N., Wohlfahrt, G., 2011. Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data. Agric. For. Meteorol. 151, 22–38.
Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57, 97–109.
Hollinger, D.Y., Richardson, A.D., 2005. Uncertainty in eddy covariance measurements and its application to physiological models. Tree Physiol. 25, 873–885.
Jastrow, J.D., 1987. Changes in soil aggregation associated with tall grass prairie restoration. Am. J. Bot. 74, 1656–1664.
Kalfas, J.L., Xiao, X., Vanegas, D.X., Verma, S.B., Suyker, A.E., 2011. Modeling gross primary production of irrigated and rain-fed maize using MODIS imagery and CO_2 flux tower data. Agric. For. Meteorol. 151, 1514–1528.
Kattge, J., Knorr, W., Raddatz, T., Wirth, C., 2009. Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. Global Change Biol. 15, 976–991.
Kirschbaum, M.U., 1995. The temperature dependence of soil organic matter decomposition: and the effect of global warming on soil organic C storage. Soil Biol. Biochem. 27, 753–760.
King, A.W., 2006. ATMOSPHERE: plant respiration in a warmer world. Science 1114166, 312, http://dx.doi.org/10.1126/science.1114166.
Knorr, W., Kattge, J., 2005. Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling. Global Change Biol. 11, 1333–1351.
Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Prentice, I.C., 2005. A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. Global Biogeochem. Cycles, 19, http://dx.doi.org/10.1029/2003GB002199.
Kuppel, S., Peylin, P., Chevallier, F., Bacour, C., Maignan, F., Richardson, A.D., 2012. Constraining a global ecosystem model with multi-site eddy-covariance data. Biogeosciences 9, 3757–3776.
Leuning, R., 1995. A critical appraisal of a combined stomatal-photosynthesis model for C_3 plants. Plant Cell Environ. 18, 339–355.
Leuning, R., 1997. Scaling to a common temperature improves the correlation between the photosynthesis parameters J_{max} and $V_{c,max}$. J. Exp. Bot. 48, 345–347.

Lloyd, J., Taylor, J.A., 1994. On the temperature dependence of soil respiration. Funct. Ecol. 8, 315–323.
Long, S.P., 1991. Modification of the response of photosynthetic productivity to rising temperature by atmospheric CO_2 concentrations: has its importance been underestimated? Plant Cell Environ. 14, 729–739.
Luo, Y., Wan, S., Hui, D., Wallace, L.L., 2001. Acclimatization of soil respiration to warming in a tall grass prairie. Nature 413, 622–625.
Luo, Y., Weng, E., Wu, X., Gao, C., Zhou, X., Zhang, L., 2009. Parameter identifiability constraint, and equifinality in data-model fusion with ecosystem models. Ecol. Appl. 19, 571–574.
Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Schimel, D.S., 2011. Ecological forecasting and data-model fusion in a data-rich era. Ecol. Appl. 21, 1429–1442.
Luo, Y., Ahlström, A., Allison, S.D., Batjes, N.H., Brovkin, V., Carvalhais, N., Georgiou, K., 2015. Towards more realistic projections of soil carbon dynamics by earth system models. Global Biogeochem. Cycles 30, 40–56.
Matamala, R., Jastrow, J.D., Miller, R.M., Garten, C.T., 2008. Temporal changes in C and N stocks of restored prairie: implications for C sequestration strategies. Ecol. Appl. 18, 1470–1488.
Medlyn, B.E., Dreyer, E., Ellsworth, D., Forstreuter, M., Harley, P.C., Kirschbaum, M.U.F., Wang, K., 2002. Temperature response of parameters of a biochemically based model of photosynthesis. II. A review of experimental data. Plant Cell Environ. 25, 1167–1179.
Medvigy, D., Jeong, S.J., Clark, K.L., Skowronski, N.S., Schäfer, K.V., 2013. Effects of seasonal variation of photosynthetic capacity on the carbon fluxes of a temperate deciduous forest. J. Geophys. Res. G: Biogeosci. 118, 1703–1714.
Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21, 1087–1092.
Moffat, A.M., Papale, D., Reichstein, M., Hollinger, D.Y., Richardson, A.D., Barr, A.G., Falge, E., 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agric. For. Meteorol. 147, 209–232.
Mosegaard, K., Sambridge, M., 2002. Monte Carlo analysis of inverse problems. Inverse Prob. 18, R29.
Novick, K.A., Stoy, P.C., Katul, G.G., Ellsworth, D.S., Siqueira, M.B.S., Juang, J., Oren, R., 2004. Carbon dioxide and water vapor exchange in a warm temperate grassland. Oecologia 138, 259–274.
Owen, K.E., Tenhunen, J., Reichstein, M., Wang, Q., Falge, E., Geyer, R., Vogel, C., 2007. Linking flux network measurements to continental scale simulations: ecosystem carbon dioxide exchange capacity under non-water-stressed conditions. Global Change Biol. 13, 734–760.
Peng, S., Piao, S., Wang, T., Sun, J., Shen, Z., 2009. Temperature sensitivity of soil respiration in different ecosystems in China. Soil Biol. Biochem. 41, 1008–1014.
Rayner, P.J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., Widmann, H., 2005. Two decades of terrestrial carbon fluxes from a carbon cycle data-model fusion system (CCDAS). Global Biogeochem. Cycles, 19, http://dx.doi.org/10.1029/2004GB002254.
Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Grünwald, T., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Global Change Biol. 11, 1424–1439.
Richardson, A.D., Hollinger, D.Y., Burba, G.G., et al., 2006. A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes. Agric. For. Meteorol. 136, 1–18.
Rogers, A., 2014. The use and misuse of $V_{c,max}$ in Earth System Models. Photosynth. Res. 119, 15–29.
Schimel, D., Melillo, J., Tian, H., McGuire, A.D., Kicklighter, D., Kittel, T., Rizzo, B., 2000. Contribution of increasing CO_2 and climate to carbon storage by ecosystems in the United States. Science 287, 2004–2006.
Sellers, P.J., Berry, J.A., Collatz, G.J., Field, C.B., Hall, F.G., 1992. Canopy reflectance photosynthesis, and transpiration. III. A reanalysis using improved leaf models and a new canopy integration scheme. Remote Sens. Environ. 42, 187–216.
Sellers, P.J., Randall, D.A., Collatz, G.J., Berry, J.A., Field, C.B., Dazlich, D.A., Bounoua, L., 1996. A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: model formulation. J. Clim. 9, 676–705.
Sellers, P.J., Dickinson, R.E., Randall, D.A., Betts, A.K., Hall, F.G., Berry, J.A., Henderson-Sellers, A., 1997. Modeling the exchanges of energy, water, and carbon between continents and the atmosphere. Science 275, 502–509.
Shi, Z., Yang, Y., Zhou, X., Weng, E., Finzi, A.C., Luo, Y., 2015. Inverse analysis of coupled carbon–nitrogen cycles against multiple datasets at ambient and elevated CO_2. J. Plant Ecol., rtv059, http://dx.doi.org/10.1093/jpe/rtv059.
Sitch, S., Smith, B., Prentice, I.C., Arneth, A., Bondeau, A., Cramer, W., Thonicke, K., 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. Global Change Biol. 9, 161–185.
Smith, N.G., Dukes, J.S., 2013. Plant respiration and photosynthesis in global-scale models: incorporating acclimation to temperature and CO_2. Global Change Biol. 19, 45–63.
Tang, J., Baldocchi, D.D., 2005. Spatial–temporal variation in soil respiration in an oak-grass savanna ecosystem in California and its partitioning into autotrophic and heterotrophic components. Biogeochemistry 73, 183–207.
Tian, H., Melillo, J.M., Kicklighter, D.W., McGuire, A.D., Helfrich, J.V.K.I., 1999. The sensitivity of terrestrial carbon storage to historical climate variability and atmospheric CO_2 in the United States. Tellus 51, 414–452.
van't Hoff, J.H., 1899. Lectures on Theoretical and Physical Chemistry. Edward Arnold.
Van Wijk, M.T., Dekker, S.C., Bouten, W., Bosveld, F.C., Kohsiek, W., Kramer, K., Mohren, G.M.J., 2000. Modeling daily gas exchange of a Douglas-fir forest: comparison of three stomatal conductance models with and without a soil water stress function. Tree Physiol. 20, 115–122.
Wang, Y.P., Leuning, R., Cleugh, H.A., Coppin, P.A., 2001. Parameter estimation in surface exchange models using nonlinear inversion: how many parameters can we estimate and which measurements are most useful? Global Change Biol. 7, 495–510.
Wang, Y.P., Baldocchi, D., Leuning, R.A.Y., Falge, E.V.A., Vesala, T., 2007. Estimating parameters in a land-surface model by applying nonlinear inversion to eddy covariance flux measurements from eight Fluxnet sites. Global Change Biol. 13, 652–670.
Williams, M., Schwarz, P.A., Law, B.E., Irvine, J., Kurpius, M.R., 2005. An improved analysis of forest carbon dynamics using data assimilation. Global Change Biol. 11, 89–105.
Wilson, K.B., Baldocchi, D.D., Hanson, P.J., 2000. Spatial and seasonal variability of photosynthetic parameters and their relationship to leaf nitrogen in a deciduous forest. Tree Physiol. 20, 565–578.
Wullschleger, S.D., 1993. Biochemical limitations to carbon assimilation in C_3 plants—are retrospective analysis of the A/Ci curves from 109 species. J. Exp. Bot. 44, 907–920.
Wu, X., Luo, Y., Weng, E., White, L., Ma, Y., Zhou, X., 2009. Conditional inversion to estimate parameters from eddy-flux observations. J. Plant. Ecol.—UK 2, 55–68.
Wythers, K.R., Reich, P.B., Tjoelker, M.G., Bolstad, P.B., 2005. Foliar respiration acclimation to temperature and temperature variable Q_{10} alter ecosystem carbon balance. Global Change Biol. 11, 435–449.
Xiao, J., Davis, K.J., Urban, N.M., Keller, K., 2014. Uncertainty in model parameters and regional carbon fluxes: A model-data fusion approach. Agric. For. Meteorol. 189, 175–186.
Xin, Q., Broich, M., Suyker, A.E., Yu, L., Gong, P., 2015. Multi-scale evaluation of light use efficiency in MODIS gross primary productivity for croplands in the Midwestern United States. Agric. For. Meteorol. 201, 111–119.
Xu, M., Qi, Y., 2001. Spatial and seasonal variations of Q_{10} determined by soil respiration measurements at a Sierra Nevadan forest. Global Biogeochem. Cycles 15, 687–696.
Xu, T., White, L., Hui, D., Luo, Y., 2006. Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction. Global Biogeochem. Cycles, 20, http://dx.doi.org/10.1029/2005GB002468.
Yuan, W., Liang, S., Liu, S., Weng, E., Luo, Y., Hollinger, D., Zhang, H., 2012. Improving model parameter estimation using coupling relationships between vegetation production and ecosystem respiration. Ecol. Model. 240, 29–40.
Zhang, L., Yu, G.R., Luo, Y.Q., Gu, F.X., Zhang, L.M., 2008. Influences of error distributions of net ecosystem exchange on parameter estimation of a process-based terrestrial model A case of broad-leaved Korean pine mixed forest in Changbaishan, China. Stxben 28, 3017–3026.
Zhou, T., Luo, Y., 2008. Spatial patterns of ecosystem carbon residence time and NPP-driven carbon uptake in the conterminous United States. Global Biogeochem. Cycles, 22, http://dx.doi.org/10.1029/2007GB002939.
Zhou, T., Shi, P., Hui, D., Luo, Y., 2009. Global pattern of temperature sensitivity of soil heterotrophic respiration (Q_{10}) and its implications for carbon-climate feedback. J. Geophys. Res.: Biogeosci. 114, G2, http://dx.doi.org/10.1029/2008JG000850.