Received: 2 April 2018 | Accepted: 4 October 2018
DOI: 10.1002/hyp.13298



RESEARCH ARTICLE

WILEY

# A hierarchical Bayesian approach for multi-site optimization of a satellite-based evapotranspiration model

Yonghong Su[1] | Qi Feng[1] | Gaofeng Zhu[2] | Chunjie Gu[2] | Yunquan Wang[3] | Shasha Shang[2] | Kun Zhang[2] | Tuo Han[2] | Huiling Chen[2] | Jinzhu Ma[2]

[1] Key Laboratory of Eco-hydrology of Inland River Basin (CAS), Northwest Institute of Eco-Environment and Resources, CAS, Lanzhou, China

[2] Laboratory of Western China's Environmental Systems (Ministry of Education), Lanzhou University, Lanzhou, China

[3] School of Environmental Studies, China University of Geosciences, Wuhan, China

**Correspondence**
Gaofeng Zhu, Laboratory of Western China's Environmental Systems (Ministry of Education), Lanzhou University, Lanzhou 730000, China.
Email: zhugf@lzu.edu.cn

**Funding information**
National Key R & D Program of China, Grant/Award Number: 2016YFC0501002; National Natural Science Foundation of China, Grant/Award Numbers: 41571016 and 41871078

## Abstract

Modelling is an important tool in simulating and partitioning evapotranspiration (ET). To obtain realistic partitioning of ET, a hierarchical Bayesian (HB) method was used to fit the Priestly–Taylor Jet Propulsion Laboratory (PT-JPL) model against the multi-tower Flux Network (FLUXNET) datasets. Unique to the HB method is its ability to exchange of information between sites and simultaneously estimate the species- and PFT-level parameters. The results suggested that the sensitive parameters varied at the both species and PFT levels. The parameter $\beta$ (water control of soil evaporation) exhibited relatively wide species-and PFT-level posterior distributions, indicating that the original parameterization of soil moisture constraint may be problematic. Generally, the model with parameters determined by the HB approach showed better performance in predicting and partitioning ET than the original model, especially in evergreen needleleaf forests, open shrublands, closed shrublands, and woody savannas. To overcome the problem of parameter uncertainty (equifinality), direct observations of different components of ET are urgently needed in future studies, and assessments the extent to which the parameter uncertainties are reduced by the use of additional data.

**KEYWORDS**

differential evolution Markov chain algorithm, evaporation, evapotranspiration, hierarchical Bayesian method, process-based ET model, transpiration

## 1 | INTRODUCTION

Evapotranspiration (ET) is an important land surface process in climatology and a nexus for terrestrial water, carbon, and energy cycles (Jung et al., 2010). The individual components of ET include loss of intercepted water ($E_i$), soil evaporation ($E_s$), and canopy transpiration ($E_c$). Whereas $E_s$ and $E_i$ are abiotically controlled fluxes, $E_c$ is strongly influenced by plant physiology and environmental conditions. The ability to properly partition ET into its different components is critical for understanding ecosystem water use processes and its response to climate change (Newman et al., 2006). Efforts to accurately estimate ET and its components started in the 1970s with the development of new technologies (i.e., microlysimeters, sap flow, eddy covariance, and environmental stable isotope) and modelling (see Kool et al., 2014 for a comprehensive review). Among these methods, modelling provides a powerful tool and is becoming more and more popular (Shugart, 2000). The Priestly–Taylor Jet Propulsion Laboratory (PT-JPL) model (Fisher, Tu, & Baldocchi, 2008), which has a process-based structure to partition total ET into its different components, is physically sound and rigorous, and has been widely used in previous studies due to its minimal requirements for ground-based measurements and its good performance (Feng et al., 2015; Michel et al., 2016; Zhang et al., 2017; Zhu et al., 2016).

Despite these studies, there are still some insufficiencies in the application of the PT-JPL model. First, the PT-JPL model is highly complex with some ecophysiological parameters that may vary with the environmental conditions, plant functional types (PFTs), and other factors (Zhang et al., 2017). The widely used method to estimate

  

parameters is to calibrate the mechanistic process model independently for each site through simple Bayesian (SB) analysis (i.e., Samanta, Mackay, Clayton, Kruger, & Ewers, 2007; Zhu et al., 2014, 2016; Zhang et al., 2017). However, the SB method aggregated all uncertainties into a single term (residual error), which makes it be sensitive to measurement noise (Clark, 2005). Also, the observation data at some sites may not be sufficient to allow adequate identification of the parameter values. To overcome these shortages, we describe a statistically rigorous method to estimate the parameters of the PT-JPL model across a wider range of biomes and environmental conditions. Specifically, we implemented a hierarchical Bayesian (HB) framework that couples the PT-JPL model with multi-tower FLUXNET datasets in a single analysis, allowing exchange of information between sites and hierarchically quantifying different sources of uncertainty. Second, the parameters of a complex model can trade off and yield acceptable simulations of total ET but with great uncertainties in model's ET partitioning. This phenomenon that different parameters fit observations equally well without the ability to distinguish, which parameters are better than others is termed as parameter identifiability or equifinality within hydrologic community (Beven & Freer, 2001). However, few attempts have been made to quantify the uncertainty in model's ET partitioning caused by equifinality. This may be attributed to the fact that it is still a challenge for many current optimized algorithms (i.e., Metropolis–Hastings algorithm) to deal with the equifinality problem (Turner, Sederberg, Brown, & Steyvers, 2013; Zhu et al., 2018). Here, we used a population-based genetic algorithm, called the differential evolution Markov chain (DE-MC; Storn & Price, 1997; Ter Braak, 2006), to implement the HB estimation and investigate the influences of equifinality on model's ET partitioning. Finally, although many attention have been given to the quantification of the different components of ET in the global water cycle (Maxwell & Condon, 2016; Wei et al., 2017; Fatichi & Pappas, 2017; Russell and Biederman, 2017), studies based on long-term performance of the PT-JPL model over a wide type of ecosystems under various climatic conditions are still needed.

In this study, we applied the HB approach to calibrate the PT-JPL model against multi-tower FLUXNET datasets in a single analysis. The specific objectives were to: (a) simultaneously optimize the PT-JPL model parameters over different biomes using the HB method against multi-tower FLUXNET datasets, (b) quantify the uncertainty in model's ET partitioning caused by equifinality only based on measured latent heat flux data, and (c) reveal the variations of the ET partitioning across a wide variety of biomes based on long-term performance of the PT-JPL model.

## 2 | DATA AND METHODOLOGY

### 2.1 | Data source

The recently released FLUXNET2015 dataset (December 2015, http://fluxnet.fluxdata.org/data/fluxnet2015-dataset/; Pastorello et al., 2017) provides a continuous, high-quality dataset of surface heat fluxes and meteorological data across an extensive range of ecosystems (Baldocchi et al., 2001, 2001; Agarwal et al., 2010). Following the International Geosphere-Biosphere Programme (IGBP) classification, we selected the 65 eddy covariance towers from the FLUXNET2015 dataset across a wide range of climates and biomes, including croplands (CRO; 10 sites), deciduous broadleaf forests (DBF; 11 sites), evergreen broadleaf forests (EBF; six sites), evergreen needleleaf forests (ENF; 14 sites), grasslands (GRA; nine sites), mixed forests (MF; five sites), open shrublands (OSH; two sites), savannas (SAV; three sites), closed shrublands (CSH; two sites), and woody savannas (WAS; three sites). The inputs to the PT-JPL model are monthly values aggregated from half-hourly or hourly data from the towers. The monthly latent heat flux ($\lambda ET$, $W/m^2$) data were used to optimize parameters in the model. Thus, the sites were selected based on the requirement of a high fraction of original nongapfilled flux observations. Only monthly values of measured fluxes and meteorological variables with less than 20% of gapfilled data were used in our analysis. This yielded a total of 555 site-years of eddy covariance data, with the data coverage at each site ranging from at least 2 to 14 years. The observations span the period from 2001 to 2014. The spatial distribution and the information about each site can be found in Table S1 and Figure S1. The energy balance issue from eddy covariance data remains largely unexplained and the best way to correct the data are still under discussion (Barr, Morgenstern, Black, McCaughey, & Nesic, 2006; Ershadi, McCabe, Evans, Chaney, & Wood, 2014; Massman & Lee, 2002). The energy balance closure (the sum of sensible heat and latent heat against available energy) on daily and monthly basis across the selected sites ranged from 0.55 (CH-Dav) to 1.13 (DK-Sor) and 0.52 (CN-Din) to 1.10 (DK-Sor; Table S2), respectively. In addition, the energy balance for all sites combined was 0.83 on monthly basis (Figure S2), which fell in the range of 0.78 to 0.85 for the FLUXNET sites (Beer et al., 2010). Thus, the data quality at the selected sites was relatively high and suitable for the purposes of model performance evaluations. To account for the biased low latent heat flux measurements due to the energy imbalance at the selected sites, we used the Bowen ratio closure method to correct latent heat flux values (Twine et al., 2000).

Time series of the enhanced vegetation index (EVI) and the normalized difference vegetation index (NDVI), that are required for the model inputs, were extracted from moderate resolution imaging spectroradiometer (MODIS) products (MOD13Q1) that provide 250 m spatial and 16 day temporal resolution. The MOD13Q1 data are available online (https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod13q1_v006; Didan, 2015). We used an average of nine pixels covering and surrounding the flux tower to acquire the EVI and NDVI values. The leaf area index (LAI), a potential factor influencing ET segmentation, was extracted from MOD15A2H (https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod15a2h_v006; Myneni, Knyazikhin, & Park, 2015) at 500 m spatial and 8 day temporal resolution. We used linear interpolation to fill the day gaps between successive EVI, NDVI, and LAI records, and then integrated the daily data within a month into the monthly average.

### 2.2 | PT-JPL model

The PT-JPL (Fisher et al., 2008) is based on the Priestley–Taylor (PT) equation (Priestley & Taylor, 1972) to estimate actual ET through

using a number of ecophysiological constraint factors. In this model, total ET is partitioned into canopy transpiration ($\lambda E_c$), soil evaporation ($\lambda E_s$), and interception evaporation ($\lambda E_i$), which are defined as follows:

$$\lambda E_c = (1 - f_{\text{wet}}) f_g f_t f_m \alpha \frac{\Delta}{\Delta} + \gamma R_{\text{nc}}, \quad (1)$$

$$\lambda E_s = (f_{\text{wet}} + f_{\text{sm}}(1 - f_{\text{wet}})) \alpha \frac{\Delta}{\Delta} + \gamma (R_{\text{ns}} - G), \quad (2)$$

$$\lambda E_i = f_{\text{wet}} \alpha \frac{\Delta}{\Delta} + \gamma R_{\text{nc}}, \quad (3)$$

where $\lambda$ is the latent heat of vaporization (J $kg^{-1}$); $\alpha$ is known as the PT coefficient (1.26); $\Delta$ is the slope of the saturation water vapour pressure curve (Pa $K^{-1}$); $\gamma$ is the psychrometric constant (Pa $K^{-1}$); $f_{\text{wet}}$ is the relative surface wetness (unitless); $f_g$ is the green canopy fraction (unitless); $f_t$ is the plant temperature constraint (unitless); $f_m$ is the plant moisture constraint (unitless); $f_{\text{sm}}$ is the soil moisture constraint (unitless); $G$ is the soil heat flux (W $m^{-2}$); and $R_{\text{nc}}$ is the net radiation for the canopy (W $m^{-2}$) and is given by $R_{\text{nc}} = R_n - R_{\text{ns}}$, where $R_n$ is the net radiation (W $m^{-2}$) and $R_{\text{ns}}$ is the net radiation for surface soil, which is calculated as follows:

$$R_{\text{ns}} = R_n \exp(-k_{R_n} \times \text{LAI}), \quad (4)$$

where $k_{R_n}$ is the extinction coefficient (unitless) for net radiation; and LAI is the leaf area index ($m^2$ $m^{-2}$). Total ET is then $\lambda\text{ET} = \lambda E_c + \lambda E_s + \lambda E_i$.

The ecophysiological constraint factors used as a proxy for plant and soil water stress with values between 0 and 1 are given by the following:

$$f_{\text{wet}} = \text{RH}^4, \quad (5)$$

$$f_t = e^{-\left(\frac{T_a - T_{\text{opt}}}{T_{\text{opt}}}\right)^2}, \quad (6)$$

$$f_{\text{sm}} = \text{RH}^{\text{VPD}/\beta}, \quad (7)$$

$$f_g = \frac{f_{\text{APAR}}}{f_{\text{IPAR}}}, \quad (8)$$

$$f_m = \frac{f_{\text{APAR}}}{f_{\text{APAR}_{\text{max}}}}, \quad (9)$$

where RH is relative humidity (%); $T_a$ is the air temperature (°C); $T_{\text{opt}}$ is the optimum plant growth temperature (°C); VPD is vapour pressure deficit (kPa); $\beta$ is the sensitivity for $f_{\text{sm}}$ to VPD (kPa); $f_{\text{APAR}}$ ($f_{\text{IPAR}}$) is the fraction of absorbed (intercepted) photosynthetically active radiation by canopy; and $f_{\text{APAR}_{\text{max}}}$ is the maximum $f_{\text{APAR}}$. Note that as the input dataset does not include $f_{\text{APAR}}$, $f_{\text{IPAR}}$, and LAI, and they are calculated as follows (Fisher et al., 2008):

$$f_{\text{APAR}} = m_1 \text{EVI} + b_1, \quad (10)$$

$$f_{\text{IPAR}} = m_2 \text{NDVI} + b_2, \quad (11)$$

$$\text{LAI} = -\frac{\ln(1 - f_{\text{IPAR}})}{k_{\text{PAR}}}, \quad (12)$$

where $m_1$, $b_1$, $m_2$, and $b_2$ are parameters; $k_{\text{PAR}}$ is the extinction coefficient (unitless) for photosynthetically active radiation; EVI is the enhanced vegetation index; and NDVI is the normalized difference vegetation index. Based on our previous studies (Zhang et al., 2017), we found that three parameters ($m_1$,$\beta$,$T_{\text{opt}}$) were most sensitive to the model across different biomes. Thus, the parameter vector associated with the PT-JPL model has three components needed to be estimated, namely, $\theta = (m_1, \beta, T_{\text{opt}})$. The values of other parameters were taken the same as the original model (i.e., $k_{R_n} = 0.6$,$k_{\text{PAR}} = 0.5$, $b_1 = -0.048$, $m_2 = 1.0$, and $b_2 = -0.05$; Fisher et al., 2008).

## 2.3 | HB model

Traditional approaches (e.g., SB), fitted the model on a site-by-site basis, may tend to overestimate the species-level variability, while failing to obtain the proper PFT-level parameter values. To overcome these shortages, we present a HB approach for fitting the PT-JPL model to all available flux tower dataset in a single analysis. The advantages of this approach are that it can: (a) allow the information from one flux tower site to exchange with other sites (Carlin, Clark, & Gelfand, 2006); (b) simultaneously estimate the species- and/or PFT-specific parameters under varying environmental conditions; (c) partition the uncertainty and variability into multiple processes (i.e., species-to-species variability, PFT-to-PFT variability, and observation error), rather than lumping all variability into a single residual error term; and (d) incorporate prior information for parameters into the models, and improve model performance when data are limited.

The HB approach is composed of three primary components (Wikle, 2003; Clark, 2005): (a) the data model, which specifies the likelihood of observed ET data given a process model and data parameters; (b) the process model, which describes the biophysical mechanisms governing ET (say the PT-JPL model with process parameters), as well as process uncertainty associated with random effects; and (c) the parameter model, which accounts for the uncertainty in both the data and process parameters by assigning them prior distributions. Ultimately, these three components are combined to generate posterior distributions of all unknown species- and/or PFT-specific parameters and the variance/covariance terms (Clark, 2005). We describe each HB component in the context of the multi-tower flux dataset.

The data model: Let $O_{f,s}(t)$ correspond to the individual observation of ET at time $t$ made on plant species $s$ of plant functional type $f$. There are $m$ total PFTs, $n_f$ species in PFT $f$ (=1, 2, ..., $m$), $n = \sum_{f=1}^{m} n_f$ total species, and $t_s$ observations for species $s$ (=1, 2, ..., $n$). We assumed that the ET measurements could be described by a normal distribution, such that for observation at time $t$,

$$O_{f,s}(t) \sim N\left(S_{f,s}(t), \sigma_s^2\right), \quad (13)$$

where "~" is read "is distributed as"; $N$represents a normal distribution; $S_{f,s}(t)$ is the predicted ET at time $t$ (= 1, 2, ..., $t_s$) for plant species $s$ of

functional type $f$; and $\sigma_s^2$ is the variance parameter that describes the variability in the ET observation or measurement error for plant species $s$.

The process model: The process model describes the predicted ET, which was specified according to the PT-JPL model (Fisher et al., 2008):

$$S_{f,s}(t) = f_M\left(\mathbf{X}_{f,s}(t); \boldsymbol{\theta}_s\right), \tag{14}$$

where $f_M(\cdot)$ represents the mechanistic process model, that is, PT-JPL; $\mathbf{X}_{f,s}(t)$ is the model input data at time $t$ for species $s$ of functional type $f$; and $\boldsymbol{\theta}_s$ is the unknown species-specific parameter vector for species $s$, that is, $\boldsymbol{\theta}_s = (m_{1,s}, \beta_s, T_{\text{opt},s})$.

The parameter model: The final stage of the HB approach was the specification of the prior for the unknown parameters. Because the model parameters varied on a species and PFT levels, hierarchical and nested priors were chosen. Thus, under this framework, species-level parameters, which are directly related to single tower flux data, are assumed to be nested within PFTs. Thus, a prior uncertainty in the species-level parameters can be described as follows:

$$\boldsymbol{\theta}_s \sim N\left(\boldsymbol{\mu\theta}_f, \mathbf{V}_f\right), \tag{15}$$

where $\boldsymbol{\mu\theta}_f$ is the PFT-level mean vector for functional group $f$ (=1, 2, ⋯, $m$); and $\mathbf{V}_f$ is the covariance matrix that describes the species-to-species variability in parameter within the PFT $f$. The PFT-level parameters can differ between different functional group. Here, we used a common background distribution to describe the between-PFT variability, such that,

$$\boldsymbol{\mu\theta}_f \sim N\left(\overline{\boldsymbol{\mu\theta}}, \mathbf{R}\right), \tag{16}$$

where $\overline{\boldsymbol{\mu\theta}}$ is the overall mean parameter vector for all PFTs; $\mathbf{R}$ is the parameter covariance matrix that describes the PFT-to-PFT variability.

Finally, we must specify distributions for the hyperprior parameters (i.e., $\overline{\boldsymbol{\mu\theta}}$, $\mathbf{V}_f$, $\mathbf{R}$, and $\sigma_s^2$) to complete the model hierarchy. Because we do not have any prior information of the overall mean parameter values ($\overline{\boldsymbol{\mu\theta}}$), independent and relatively noninformative (diffuse) priors were employed for them; that is, we used normal densities with large variances, $\overline{\boldsymbol{\mu\theta}} \sim N(\mathbf{0}, \tau_0 \mathbf{I}_3)$ with $\mathbf{0}$ being three-dimensional zero vectors, $\tau_0$ = 1000 and $\mathbf{I}_3$ being the rank three identity matrix. For the residual variance of measurement error in Equation (13), because there is enough data available to estimate them, we can just assume independent uniform priors (Gelman, 2006), $p\left(\sigma_s^2\right) \propto 1$. For the covariance matrixes of Equations 14 and 15, we propose a multivariate Jeffreys prior density: $p(\mathbf{V}_f) \propto |\mathbf{V}_f|^{-(d+1)/2}$ and $p(\mathbf{R}) \propto |\mathbf{R}|^{-(d+1)/2}$ where $d$ = 3 is the number of dimensions. By combining the data, process, and parameter models as defined above in Equations 13–16, we derived the joint posterior as follows:

$$\begin{aligned} p\left(\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_n, \boldsymbol{\mu\theta}_1, \cdots, \boldsymbol{\mu\theta}_m, \overline{\boldsymbol{\mu\theta}}, \sigma_1^2, \cdots, \sigma_n^2, \mathbf{V}_1, \cdots, \mathbf{V}_m, \mathbf{R} | \mathbf{O}, \mathbf{X}\right) \propto \\ \left\{\prod_{f=1}^{m}\prod_{s=1}^{n_f}\prod_{t=1}^{t_s} N\left(O_{f,s}(t) | f_M\left(\mathbf{X}_{f,s}(t); \boldsymbol{\theta}_s\right)\sigma_s^2\right)\right\} \quad \text{(likelihood)} \\ \times \left\{\prod_{f=1}^{m}\prod_{s=1}^{n_f} N\left(\boldsymbol{\theta}_s | \boldsymbol{\mu\theta}_f, \mathbf{V}_f\right)\right\} \times \left\{\prod_{f=1}^{m} N\left(\boldsymbol{\mu\theta}_f | \overline{\boldsymbol{\mu\theta}}, \mathbf{R}\right)\right\} \quad \text{(prior)} \\ \times \left\{\prod_{f=1}^{m} |\mathbf{V}_f|^{-(d+1)/2}\right\} \times |\mathbf{R}|^{-(d+1)/2} \times N\left(\overline{\boldsymbol{\mu\theta}} | \mathbf{0}, \tau_0^2 \mathbf{I}_3\right) \quad \text{(hyperprior)} \end{aligned} \tag{17}$$

The first distribution of Equation (17) is the likelihood, and the middle two distributions are priors for model parameters. The remaining three are noninformative hyperprior distributions.

## 2.4 | Implementing the HB model

Usually, the Metropolis–Hastings Markov Chain Monte Carlo (MCMC) algorithms (Robert & Casella, 1999) are implemented to sample from the joint posterior, and from these samples, one can calculate measures of centrality (e.g., mean, median, and mode), spread (e.g., credible intervals, which are similar to confidence intervals), and correlations between parameters. However, the Metropolis–Hastings (MH)-based approaches often suffer from problems related to proper initialization and proposal density function, which may prevent the algorithm from efficiently reaching convergence (Haario, Laine, Mira, & Saksman, 2006). In this study, we used the DE-MC algorithm proposed by ter Braak (2006) to implement the HB model. Compared with the MH-based approaches, the DE-MC algorithm is more suitable to draw inference on high-dimensional models (Turner et al., 2013; Zhu et al., 2018). In the DE-MC method, $N$ chains are run in parallel and the proposals are generated on the basis of two randomly selected chains, the difference of which is multiplied by a scaling factor, and added to the current chain as follows:

$$\boldsymbol{\Theta}_p = \boldsymbol{\Theta}_i + \upsilon\left(\boldsymbol{\Theta}_{R_1} - \boldsymbol{\Theta}_{R_2}\right) + \boldsymbol{\varepsilon}, \tag{18}$$

where $\boldsymbol{\Theta}_p$ is the proposed set for the unknowns, including all parameters and variance terms (Clark & Bjønstad, 2004); $\boldsymbol{\Theta}_{R_1}$ and $\boldsymbol{\Theta}_{R_2}$ represent the randomly selected chains without replacement from the population $\boldsymbol{\Theta}_{-i}$ (the population without the current chain $\boldsymbol{\Theta}_i$); $\boldsymbol{\varepsilon}$ is drawn from a symmetrical distribution with a small variance compared with that of the target, but with unbounded support, for example, $\boldsymbol{\varepsilon} \sim N(0, \mathbf{b})$ with $\mathbf{b}$ small; and $\upsilon$ is the scaling factor that always take a positive value and can be set to vary between [0.4, 1] (Roberts & Rosenthal, 2001). The Metropolis ratio is then used to decide whether to accept or reject the proposals (ter Braak, 2006).

## 2.5 | Assessment of analysis method performance

Five statistical measures are used to evaluate model performance in this paper, including the coefficient of determination ($R^2$), bias, relative error (RE), root-mean-square error (RMSE), and the Nash–Sutcliffe efficiency coefficient (NSE). $R^2$ ranges between 0 and 1, with higher values indicating a good simulation result; the NSE values range from −∞ to 1, with NSE = 1 being the optimal value (Moriasi et al., 2007).

The calculation of the statistical measures can be found in Zhu et al. (2016).

A final characterization of model performance uses the Taylor diagram (Taylor, 2001), which is especially useful in testing multiple aspects of complex models (International Panel on Climate Change, 2001). Generally, the Taylor diagram characterizes a single point to indicate three different statistical relationships between the "test" field (simulation) and the "truth" field (observation). The statistics of each point can be scored using,

$$S = \frac{2(1+R)}{[(\sigma_m/\sigma_o) + 1/(\sigma_o/\sigma_m)]^2}, \quad (19)$$

where $S$ is the model skill metric bound by zero and unity (unity indicates agreement with observations), $R$ is the coefficient of determination, and $\sigma_m$ and $\sigma_o$ is the standard deviation of the simulation and observation, respectively.

## 3 | RESULTS

In this section, we conducted two case studies to evaluate the performance of the HB model in estimating the parameter values and quantifying the partitioning of ET. The first study serves as a benchmark experiment of the HB model based on synthetic ET datasets for which all-level parameter values are known. That is, we chose true values for the hyperprior parameters $\overline{\mu\theta}$ (mean) and $\mathbf{R}$ (variance) from their prior distributions, and drew the true mean parameter values for each PFT ($\mu\theta_f, f = 1, 2, \cdots, 10$) from the normal distribution $N(\overline{\mu\theta}, \mathbf{R})$. The true species-level parameters ($\theta_s$, $s = 1, 2, \cdots, 65$) can then be drawn from the normal distribution with mean $\mu\theta_f$ and known true covariance matrix $\mathbf{V}_f$ ($f = 1, 2, \cdots, 10$). Thus, the synthetic ET for each site were generated from the PT-JPL model using measured forcing data at the FLUXNET sites and true species-level parameter values. All the parameter values used in generating the ET datasets are given in Tables S3 and S4. The second study explores the performance of the HB model using the actual measured datasets at 65 FLUXNET sites (see details in the data source subsection). In all numerical experiments, the number of chains of the DE-MC algorithm is set to be 800 and iteration to be 10,000, where the initial parameter values were randomly sampled from the uniform distributions.

### 3.1 | Synthetic dataset

The posterior parameter distributions of the three sensitive parameters on species and PFT levels given by the HB model using the synthetic data are shown in Figures 1 and 2, respectively. The statistical results characterized by the posterior medians and 90% credible intervals (CIs) are summarized in Tables S3 and S4. Two main conclusions can be drawn from the results. First, the posterior median parameter values varied widely across different species and PFTs and the 90% CIs of the posterior distributions include true parameter values at both species and PFT levels (Figures 1 and 2), which indicates that the HB model is able to identify the true values in the parameter space. Second, the posterior distributions of the parameters, especially for $\beta$, are relatively widely spread on the prior bounds at both species and PFT levels (Figures 1 and 2). This is mainly caused by equifinality (Franks, Beven, Quinn, & Wright, 1997), where different combinations of parameters lead to the similar simulation accuracy in total $\lambda ET$ with potential uncertainties in its partitioning in different components.

To illustrate the influence of different combinations of parameters on the simulation results of ET and its components, we used the parameter sets obtained by the HB procedure to run the PT-JPL model. The monthly variation of water fluxes simulated with selected

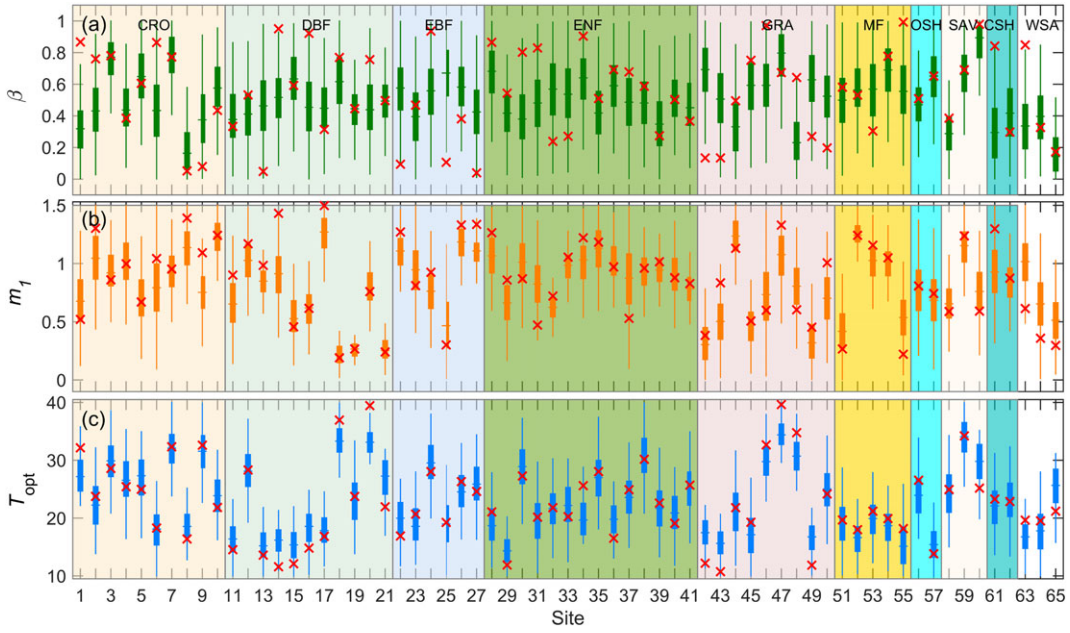


**FIGURE 1** Posterior distribution of the parameters at each site. Thin vertical lines indicate 90% confidence interval and thick vertical lines indicate interquartile range. The horizontal lines show the median parameter value at each site. The red cross symbol "×" indicates the actual parameter values used to generate the synthetic ET data

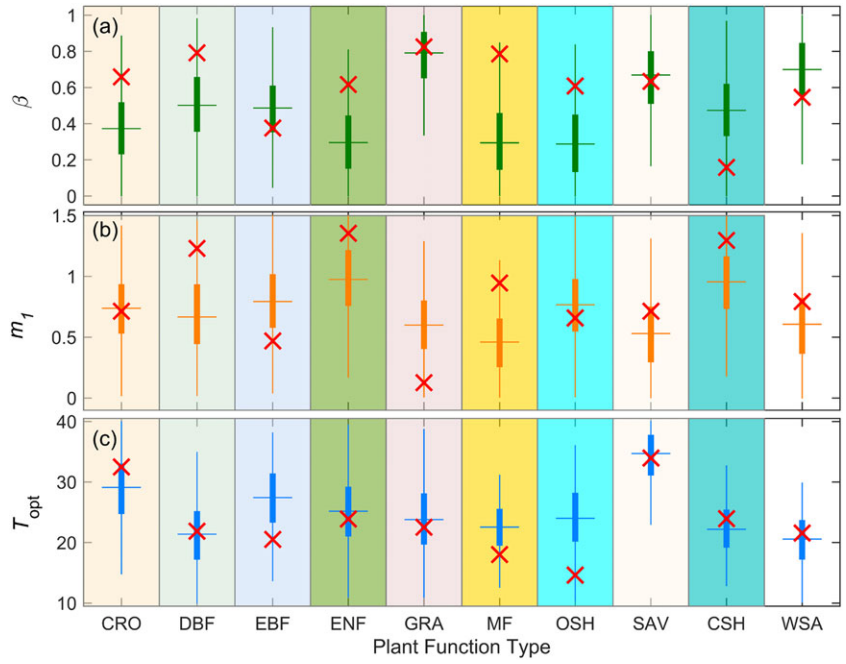


**FIGURE 2** Posterior distribution of the parameters at PFT level. Thin vertical lines indicate 90% confidence interval and thick vertical lines indicate interquartile range. The horizontal lines show the median parameter value at each site. The red cross symbol "×" indicates the actual parameter values used to generate the synthetic ET data

parameter sets (number = 60) from one crop ecosystem site (BE-Lon) are presented in Figure 3, as well as providing the synthetic observations. We can observe that the median values of simulated λ*ET* fluctuated tightly with the synthetic observations during the whole study period, and the 90% prediction intervals of simulated λ*ET* are relatively narrow (Figure 3a). These indicate that the PT-JPL model with the multiple parameter sets obtained by the HB procedure can yield quite similar satisfactory simulations in λ*ET*. However, the partitioning of λ*ET* between $\lambda E_c$ and $\lambda E_s$ tended to be different with the different parameter values, and showed relatively large uncertainties. The median values of simulated $\lambda E_c$ and $\lambda E_s$ may also tend to deviate from the true values (Figures 3b and 3c). Similar model performance is also observed in other sites (see details in Figure S3). Here, we also use two type synthetic observations (total ET and canopy transpiration) to constrain the model parameters. The results indicated that the parameters can be well updated compared with that constrained by only using ET data, and the medians of simulated ET partitioning were more comparable with the true values (Figure S4). However, direct field observations of different components of ET partitioning are still very sparse (Lawrence, Thornton, Oleson, & Bonan, 2007). Thus, special attention must be paid to the parameter uncertainties when using models, which were optimized only based on EC-measured data to partition ET into its different components.

To account for the impacts of parameter uncertainties on the partitioning of ET, we evaluated the model's ET partitioning (i.e., $\lambda E_c/\lambda ET$, $\lambda E_s/\lambda ET$, and $\lambda E_i/\lambda ET$) using the multiple combinations of parameters rather than a single parameter set (i.e., the median values of posterior parameter distributions, the parameter set that lead to the least

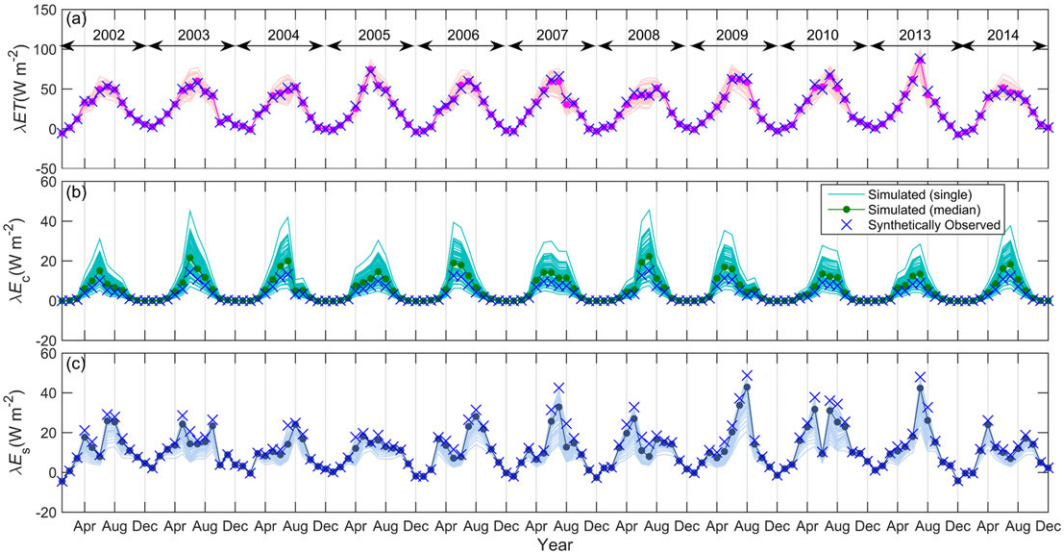


**FIGURE 3** Comparison between synthetically observed and simulated water fluxes for one cropland site (BE-Lon). (a) λET, (b) λEc, and (c) $\lambda E_s$. The cross symbol "×" indicates the synthetic water fluxes data, the solid circle '•' represents the median of set simulations, and one line corresponds to a simulation with one particular parameter set

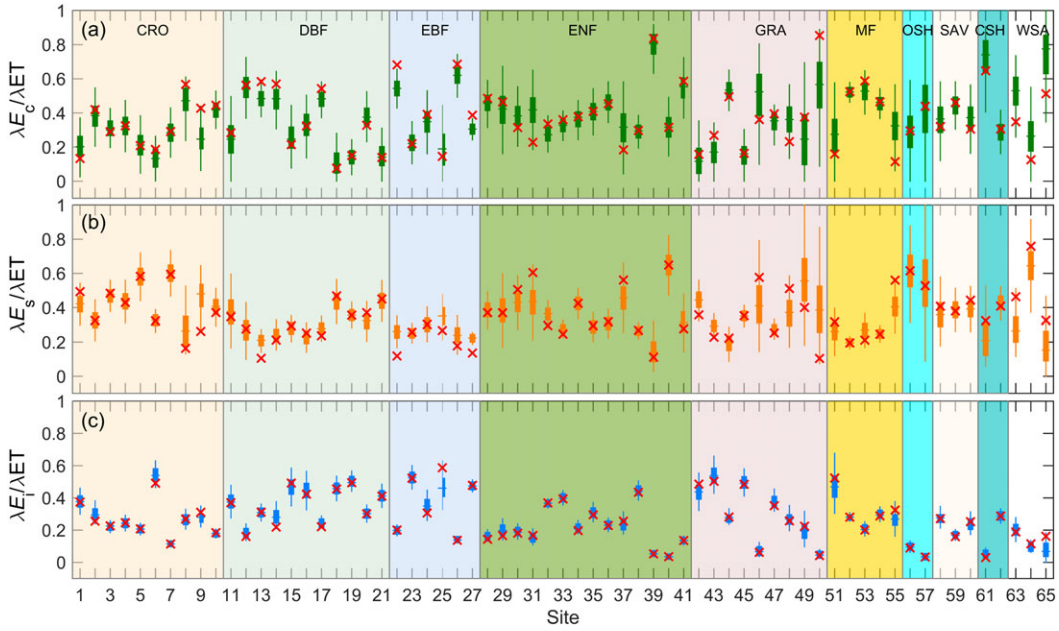


**FIGURE 4** The posterior medians and 90% CIs of the simulation sets of ET partitioning using the synthetic ET data. (a) $\lambda E_c/\lambda ET$, (b) $\lambda E_s/\lambda ET$, and (c) $\lambda E_i/\lambda ET$. The red cross symbol "×" indicates the observed true partitioning values across all site

normalized root mean square error or maximum likelihood). Figure 4 shows the posterior medians and 90% CIs of simulated partitioning using different parameter sets as well as the observed true values across all sites. It was found that the observed true partitioning of ET at most sites fell in the 50% CIs of simulated results, and exhibit good relationships with the medians of simulated partitioning values (linear regression slope = 0.77, 0.65, and 0.96 for observed vs simulated values of $\lambda E_c/\lambda ET$, $\lambda E_s/\lambda ET$, and $\lambda E_i/\lambda ET$ with $R^2$ = 0.74, 0.66, and 0.96, respectively; Figure S5). Thus, the HB approach presented here provided us the opportunity to get the multiple combinations of parameters that can fit the $\lambda ET$measurements satisfactorily. By taking the parameter uncertainties into accounts, we can get reasonable estimations about the posterior distributions of the ET partitioning for different ecosystems. The synthetic test gives us confidence in its estimations in ET partitioning for real-world hydrological problems. Hereafter, we called the simulations with multiple parameter sets obtained by the HB model as the set simulation approach.

## 3.2 | In situ FLUXNET dataset

For the in situ FLUXNET datasets, the posterior parameter distributions on the species and PFT levels are shown in Figures 5 and 6,

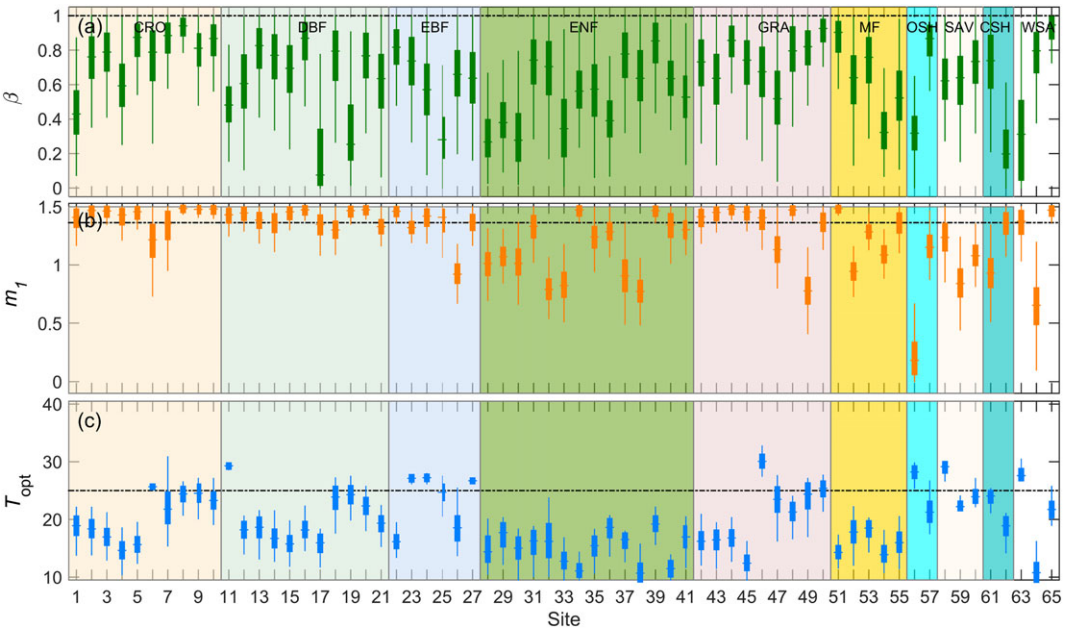


**FIGURE 5** Posterior distribution of the parameters at each FLUXNET site. The vertical bars indicate the 90% probability intervals; the dot line represents the original parameter value

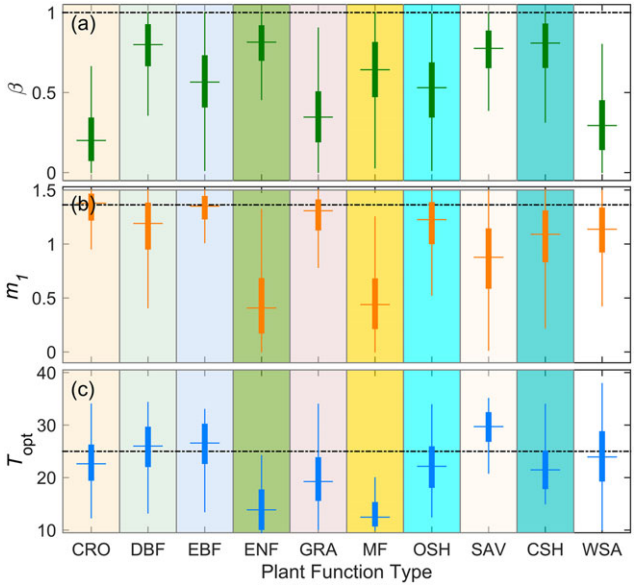


**FIGURE 6** Posterior distribution of the parameters at PFT level. The vertical bars indicate the 90% probability intervals; the dot line represents the original parameter value

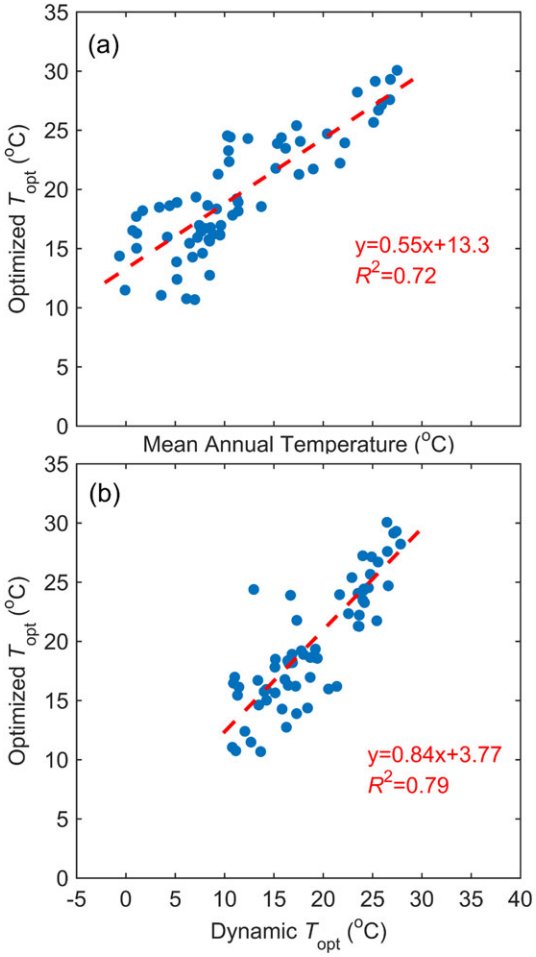


**FIGURE 7** Relationships between optimized $T_{opt}$ and (a) mean air temperature, and (b) dynamic $T_{opt}$ across all sites

respectively, as well as summarized in Tables S5 and S6 by posterior medians and 90% probability intervals. The HB model estimations of $T_{opt}$ and $m_1$ were well informed by the EC latent heat flux data. This was demonstrated by relatively narrow 90% CIs of posterior parameter distributions, posterior medians that were quite different from the corresponding prior means, and significant species- and PFT-level differences (Figures 5 and 6). Generally, the posterior median values of $m_1$ were high for crop and broad-leaf forests (about 1.35; Figure 6b) with a narrow species-level variation range (around 0.92–1.48; Figure 5b), whereas the posterior median values were slightly low for ENF, MF, and SAV (from 0.41 to 0.87; Figure 6b) and varied widely on species level (from 0.80 to 1.47; Figure 5b). This indicates that canopy types (i.e., needle-leaf, broad-leaf, leaf clumping, and orientations) have significant impacts on the estimations of $m_1$ (Gao, Huete, Ni, & Miura, 2000). The posterior median values of $T_{opt}$ were low for the MF and ENF (about 13°C; Figure 6c) with a narrow species-level variation range (from 11 to 19°C; Figure 5c), whereas the posterior median values were relative high for the DBF, EBF, and SAV (from 26 to 29°C; Figure 6c) and exhibited large species-level variation (from 16 to 29°C; Figure 5c). Moreover, the estimated species-level $T_{opt}$ exhibited good agreements with the respective growing mean annual temperature (slope = 0.55 and intercept = 13.3 with $R^2$ = 0.72; Figure 7a) and that estimated dynamically as the air temperature of the month with the highest NDVI, radiation and temperature and minimum VPD (hereafter named as dynamic method; Potter et al., 1993; Fisher et al., 2008) (slope = 0.84 and intercept = 3.77 with $R^2$ = 0.79; Figure 7b), indicating that our optimized values are reasonable. On the contrary, the parameter $\beta$ remained widely spread across their prior range of variation on both species and PFT levels at most sites (Figures 5a and 6a), indicating relatively larger uncertainties partly due to equifinality.

Figure 8 and Table 1 compare the estimates of monthly $\lambda ET$ using the original model and the medians of the simulation sets against the observed latent heat fluxes across all sites and PFTs. The temporal variation of observed and simulated $\lambda ET$ using the original model and the set simulation approach at selected typical sites was illustrated in Figures 9 and 10. We can observe that the model with the original parameters performed well for the CRO, DBF, EBF, GRA, and MF biomes (Figure 9), with $R^2$ ranging from 0.69 to 0.97 and NSE ranging from 0.58 to 0.95 (Table 1). However, the original model generally tended to overestimate the monthly $\lambda ET$at some ENF (CA-Man, CA-Qfo, CA-TP3 and IT-Lav), OSH (Mx-Lpa), and WSA (ES-LgS) sites, with $\sigma_{norm}$ greater than 1.3 and NSE less than 0.5 (Figures 8 and 10). Also, significant underestimations of the monthly $\lambda ET$by the original model was found at two sites of the ENF biomes (CA-Obs and FI-Sod) due to unrealistic estimations of parameter $T_{opt.}$ (Table 1 and Figure 10c). On the contrary, the set simulation approach performed better in matching the observations at most sites as indicated by lower bias and RMSE (Table 1), $\sigma_{norm}$ more closer to one, greater values of model skill metric and NSE (Figures 8, 9, and 10). Thus, the model parameter uncertainty seems to be the main source of the disagreements between the observed and simulated latent heat fluxes, and the set simulation approach is effective in improving the model performance at most sites.

Figures 11 and 12 showed the comparisons of partitioning of ET using the original model and the set simulation approach at species and PFT levels, respectively. Noticeably, the 90% CIs of the simulation sets of ET partitioning at the CRO, DBF, EBF, GRA, and MF sites are

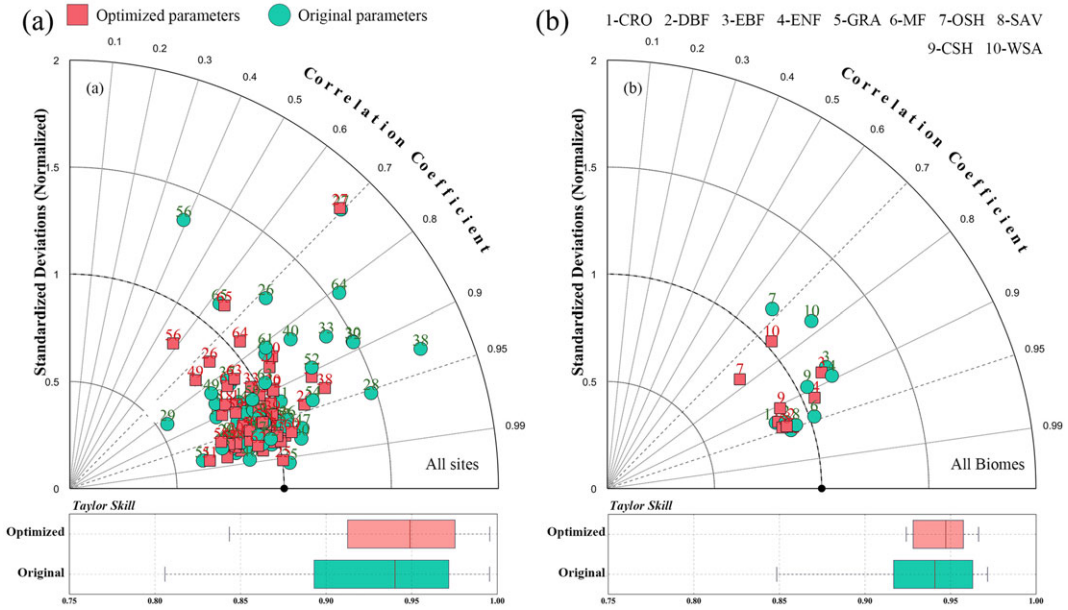


**FIGURE 8** Statistical comparison of model performance using the original and optimized parameters at both (a) the species level and (b) the PFT level

**TABLE 1** The statistics showing the algorithm performance by using different parameter sets over different sites during the study period

| | | | Original model | | | | | | | Medians of set simulations | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Site ID | PFT | Site name | Bias | $R^2$ | Slope | Intercept | RMSE | RE | NSE | Bias | $R^2$ | Slope | Intercept | RMSE | RE | NSE |
| 1 | CRO | BE-Lon | 5.97 | 0.79 | 0.79 | 3.04 | 17.3 | 0.41 | 0.76 | 6.84 | 0.79 | 0.77 | 2.83 | 17.6 | 0.41 | 0.75 |
| 2 | CRO | CH-Oe1 | 11.9 | 0.96 | 0.78 | −1.32 | 16.3 | 0.34 | 0.83 | 11.1 | 0.95 | 0.80 | −1.61 | 15.4 | 0.33 | 0.85 |
| 3 | CRO | DE-Geb | 5.76 | 0.85 | 0.80 | 0.89 | 13.7 | 0.41 | 0.82 | 5.16 | 0.86 | 0.82 | 0.72 | 13.0 | 0.39 | 0.83 |
| 4 | CRO | DE-Kli | 5.57 | 0.81 | 0.81 | 1.09 | 15.3 | 0.43 | 0.78 | 6.06 | 0.83 | 0.81 | 0.52 | 15.0 | 0.43 | 0.79 |
| 5 | CRO | FR-Gri | 11.7 | 0.94 | 0.76 | −0.32 | 16.9 | 0.36 | 0.82 | 11.0 | 0.94 | 0.77 | −0.46 | 16.2 | 0.35 | 0.83 |
| 6 | CRO | PA-SPn | −8.34 | 0.88 | 0.90 | 17.7 | 11.0 | 0.12 | 0.72 | −4.71 | 0.86 | 0.84 | 19.4 | 9.04 | 0.10 | 0.81 |
| 7 | CRO | US-ARM | 5.69 | 0.70 | 0.75 | 6.38 | 17.3 | 0.36 | 0.66 | 5.75 | 0.69 | 0.75 | 5.93 | 17.7 | 0.37 | 0.64 |
| 8 | CRO | US-Ne1 | 11.9 | 0.94 | 0.73 | 6.02 | 22.7 | 0.34 | 0.85 | 10.0 | 0.94 | 0.77 | 5.02 | 20.0 | 0.30 | 0.88 |
| 9 | CRO | US-Ne2 | 7.88 | 0.90 | 0.74 | 8.48 | 21.6 | 0.35 | 0.85 | 7.05 | 0.91 | 0.77 | 7.27 | 20.1 | 0.33 | 0.87 |
| 10 | CRO | US-Ne3 | 5.49 | 0.91 | 0.79 | 5.63 | 16.3 | 0.30 | 0.88 | 4.34 | 0.92 | 0.84 | 4.31 | 14.7 | 0.27 | 0.90 |
| 11 | DBF | AU-How | −3.43 | 0.84 | 0.85 | 18.51 | 13.4 | 0.14 | 0.83 | 2.32 | 0.86 | 0.92 | 5.98 | 12.4 | 0.13 | 0.85 |
| 12 | DBF | CA-Oas | 0.93 | 0.86 | 0.78 | 6.35 | 15.6 | 0.47 | 0.85 | 1.24 | 0.88 | 0.81 | 5.17 | 14.3 | 0.43 | 0.88 |
| 13 | DBF | DE-Hai | 2.13 | 0.89 | 0.93 | 0.52 | 10.9 | 0.30 | 0.89 | 1.92 | 0.90 | 0.94 | 0.42 | 10.9 | 0.30 | 0.89 |
| 14 | DBF | DE-Lnf | 0.27 | 0.93 | 0.89 | 3.34 | 8.86 | 0.26 | 0.93 | 0.08 | 0.95 | 0.93 | 2.22 | 7.22 | 0.21 | 0.95 |
| 15 | DBF | DK-Sor | 9.42 | 0.81 | 0.68 | 3.70 | 22.3 | 0.54 | 0.74 | 8.46 | 0.81 | 0.71 | 3.57 | 21.5 | 0.52 | 0.76 |
| 16 | DBF | FR-Fon | 9.55 | 0.90 | 0.86 | −2.54 | 16.0 | 0.31 | 0.85 | 7.77 | 0.91 | 0.92 | −3.40 | 14.4 | 0.28 | 0.87 |
| 17 | DBF | IT-Col | −8.39 | 0.91 | 0.99 | 8.81 | 15.6 | 0.34 | 0.87 | −0.10 | 0.94 | 0.91 | 4.29 | 10.8 | 0.23 | 0.94 |
| 18 | DBF | IT-Ro2 | −4.75 | 0.75 | 0.68 | 22.9 | 25.6 | 0.45 | 0.73 | −4.33 | 0.77 | 0.73 | 19.9 | 24.0 | 0.42 | 0.76 |
| 19 | DBF | US-MMS | −2.99 | 0.95 | 0.84 | 11.7 | 12.4 | 0.23 | 0.93 | −1.55 | 0.96 | 0.90 | 6.89 | 9.95 | 0.18 | 0.96 |
| 20 | DBF | US-Oho | 4.94 | 0.91 | 0.73 | 11.1 | 20.3 | 0.34 | 0.86 | 3.01 | 0.91 | 0.77 | 10.4 | 18.5 | 0.31 | 0.89 |
| 21 | DBF | US-UMB | −3.27 | 0.91 | 0.94 | 5.84 | 12.2 | 0.31 | 0.91 | −1.60 | 0.92 | 0.91 | 5.21 | 11.5 | 0.29 | 0.92 |
| 22 | EBF | AU-Tum | 8.18 | 0.84 | 0.81 | 3.58 | 17.2 | 0.27 | 0.79 | 4.40 | 0.90 | 0.98 | −2.89 | 12.9 | 0.21 | 0.88 |
| 23 | EBF | BR-Sa1 | −2.16 | 0.93 | 0.93 | 9.93 | 4.76 | 0.05 | 0.91 | −0.57 | 0.92 | 0.89 | 11.7 | 4.55 | 0.04 | 0.92 |
| 24 | EBF | BR-Sa3 | 5.70 | 0.91 | 0.99 | −4.86 | 6.99 | 0.07 | 0.72 | 6.15 | 0.89 | 1.09 | −15.9 | 8.16 | 0.08 | 0.62 |
| 25 | EBF | CN-Din | 0.05 | 0.99 | 1.03 | −1.64 | 3.74 | 0.06 | 0.98 | 2.06 | 0.98 | 1.00 | −1.78 | 4.51 | 0.07 | 0.98 |
| 26 | EBF | FR-Pue | −7.16 | 0.51 | 0.91 | 10.7 | 28.2 | 0.68 | 0.15 | 6.15 | 0.55 | 0.65 | 8.31 | 21.9 | 0.53 | 0.49 |
| 27 | EBF | GF-Guy | 2.97 | 0.49 | 1.26 | −33.5 | 14.5 | 0.13 | −0.84 | 3.67 | 0.48 | 1.26 | −33.6 | 14.7 | 0.13 | −0.90 |

(Continues)

**TABLE 1** (Continued)

| | | | Original model | | | | | | | Medians of set simulations | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Site ID | PFT | Site name | Bias | $R^2$ | Slope | Intercept | RMSE | RE | NSE | Bias | $R^2$ | Slope | Intercept | RMSE | RE | NSE |
| 28 | ENF | CA-Man | −10.3 | 0.91 | 1.41 | −0.83 | 18.1 | 0.66 | 0.46 | −0.08 | 0.92 | 0.95 | 1.34 | 7.22 | 0.26 | 0.91 |
| 29 | ENF | CA-Obs | 11.7 | 0.69 | 0.46 | 4.64 | 20.7 | 0.69 | 0.43 | 1.89 | 0.94 | 1.01 | −2.04 | 7.01 | 0.23 | 0.93 |
| 30 | ENF | CA-Qfo | −6.48 | 0.79 | 1.32 | −3.51 | 17.4 | 0.56 | 0.33 | 2.87 | 0.80 | 0.95 | −1.22 | 10.4 | 0.33 | 0.76 |
| 31 | ENF | CA-SF | −0.86 | 0.85 | 0.98 | 1.28 | 10.9 | 0.41 | 0.83 | 0.45 | 0.85 | 0.94 | 1.18 | 10.8 | 0.41 | 0.84 |
| 32 | ENF | CA-TP3 | −9.64 | 0.79 | 1.32 | −3.89 | 24.4 | 0.58 | 0.32 | 2.30 | 0.81 | 0.95 | −0.17 | 13.9 | 0.33 | 0.78 |
| 33 | ENF | CA-TP4 | −10.4 | 0.74 | 1.20 | 0.84 | 26.5 | 0.54 | 0.36 | 3.60 | 0.76 | 0.85 | 3.88 | 16.9 | 0.35 | 0.74 |
| 34 | ENF | CH-Dav | 14.4 | 0.84 | 0.86 | −6.58 | 20.5 | 0.38 | 0.70 | 14.1 | 0.85 | 0.88 | −7.63 | 20.1 | 0.37 | 0.71 |
| 35 | ENF | DE-Obe | 0.15 | 0.90 | 0.97 | 1.00 | 10.1 | 0.28 | 0.90 | 1.96 | 0.91 | 0.94 | 0.38 | 9.75 | 0.27 | 0.91 |
| 36 | ENF | FI-Hyy | −0.30 | 0.91 | 1.02 | −0.15 | 8.92 | 0.31 | 0.90 | 2.06 | 0.92 | 0.93 | −0.10 | 8.30 | 0.29 | 0.91 |
| 37 | ENF | FI-Sod | 6.87 | 0.68 | 0.73 | −2.30 | 11.6 | 0.68 | 0.50 | 2.42 | 0.82 | 1.13 | −4.63 | 9.14 | 0.53 | 0.69 |
| 38 | ENF | IT-Lav | −4.46 | 0.86 | 1.64 | −25.2 | 22.8 | 0.49 | 0.13 | 7.45 | 0.86 | 1.19 | −16.3 | 14.5 | 0.31 | 0.65 |
| 39 | ENF | US-Blo | 8.22 | 0.87 | 0.88 | −1.08 | 16.0 | 0.26 | 0.83 | 6.23 | 0.88 | 0.94 | −2.82 | 14.9 | 0.24 | 0.85 |
| 40 | ENF | US-GLE | 10.0 | 0.69 | 1.03 | −11.51 | 19.3 | 0.40 | 0.33 | 13.4 | 0.70 | 0.94 | −10.7 | 19.9 | 0.41 | 0.29 |
| 41 | ENF | US-Me2 | 6.16 | 0.68 | 0.91 | −2.10 | 18.8 | 0.41 | 0.55 | 7.41 | 0.73 | 0.93 | −4.23 | 17.6 | 0.38 | 0.60 |
| 42 | GRA | CH-Cha | 10.9 | 0.95 | 0.94 | −7.71 | 13.6 | 0.24 | 0.87 | 10.6 | 0.95 | 0.95 | −8.01 | 13.5 | 0.24 | 0.87 |
| 43 | GRA | CH-Fru | 11.6 | 0.95 | 0.90 | −6.67 | 14.3 | 0.29 | 0.84 | 9.79 | 0.95 | 0.98 | −9.04 | 13.0 | 0.25 | 0.88 |
| 44 | GRA | DE-Gri | 11.4 | 0.93 | 0.80 | −1.80 | 16.2 | 0.34 | 0.83 | 9.78 | 0.93 | 0.84 | −2.27 | 14.6 | 0.31 | 0.86 |
| 45 | GRA | IT-MBo | 11.7 | 0.92 | 0.90 | −7.51 | 16.2 | 0.37 | 0.84 | 10.8 | 0.92 | 0.93 | −7.73 | 15.4 | 0.35 | 0.85 |
| 46 | GRA | SD-Dem | −1.86 | 0.97 | 0.84 | 6.79 | 7.05 | 0.23 | 0.95 | 0.27 | 0.96 | 0.74 | 7.89 | 9.83 | 0.32 | 0.91 |
| 47 | GRA | US-Goo | −4.12 | 0.94 | 1.09 | −0.82 | 11.4 | 0.20 | 0.90 | 2.20 | 0.93 | 0.97 | −0.26 | 9.66 | 0.17 | 0.93 |
| 48 | GRA | US-IB2 | 12.0 | 0.92 | 0.74 | 5.15 | 20.5 | 0.31 | 0.83 | 10.1 | 0.93 | 0.78 | 4.37 | 18.2 | 0.28 | 0.86 |
| 49 | GRA | US-Var | −4.97 | 0.69 | 0.66 | 12.6 | 14.9 | 0.66 | 0.65 | 1.83 | 0.57 | 0.39 | 11.9 | 17.6 | 0.78 | 0.51 |
| 50 | GRA | US-Wkg | −2.47 | 0.91 | 0.86 | 5.67 | 7.71 | 0.34 | 0.90 | −1.47 | 0.91 | 0.84 | 5.09 | 7.63 | 0.34 | 0.90 |
| 51 | MF | AT-Neu | 18.3 | 0.96 | 0.62 | −0.70 | 26.6 | 0.58 | 0.70 | 16.6 | 0.96 | 0.65 | −0.58 | 24.4 | 0.53 | 0.74 |
| 52 | MF | BE-Bra | −7.92 | 0.80 | 1.13 | 3.94 | 17.6 | 0.57 | 0.58 | −0.35 | 0.80 | 0.89 | 3.90 | 12.2 | 0.39 | 0.80 |
| 53 | MF | CA-Gro | −0.99 | 0.93 | 1.01 | 0.50 | 9.45 | 0.25 | 0.92 | 1.27 | 0.94 | 0.97 | −0.03 | 8.18 | 0.22 | 0.94 |
| 54 | MF | RU-Fyo | −4.61 | 0.88 | 1.13 | 0.36 | 14.7 | 0.46 | 0.79 | 1.64 | 0.90 | 0.92 | 1.05 | 10.1 | 0.32 | 0.90 |
| 55 | MF | US-Syv | 1.53 | 0.90 | 0.91 | 1.79 | 10.4 | 0.29 | 0.89 | 1.92 | 0.89 | 0.90 | 1.63 | 10.6 | 0.30 | 0.89 |
| 56 | OSH | Mx-Lpa | −48.7 | 0.15 | 0.53 | 58.9 | 51.0 | 2.33 | −19.4 | −0.59 | 0.34 | 0.48 | 11.9 | 9.66 | 0.44 | 0.27 |
| 57 | OSH | US-Whs | −7.17 | 0.93 | 0.88 | 9.78 | 9.50 | 0.44 | 0.82 | −2.07 | 0.93 | 0.74 | 7.66 | 7.75 | 0.36 | 0.88 |
| 58 | SAV | AU-DaP | −13.8 | 0.93 | 0.71 | 31.0 | 20.6 | 0.35 | 0.78 | −3.14 | 0.91 | 0.71 | 20.6 | 16.4 | 0.28 | 0.86 |
| 59 | SAV | ZA-Kru | −24.1 | 0.81 | 0.85 | 31.0 | 28.9 | 0.62 | 0.36 | −3.38 | 0.83 | 0.68 | 18.2 | 16.5 | 0.35 | 0.79 |
| 60 | SAV | ZM-Mon | −12.7 | 0.96 | 1.08 | 7.39 | 15.0 | 0.23 | 0.78 | −0.07 | 0.94 | 1.03 | −2.23 | 8.53 | 0.13 | 0.93 |
| 61 | CSH | AU-Cpr | −14.0 | 0.66 | 0.91 | 16.4 | 16.8 | 0.59 | −0.43 | −0.36 | 0.70 | 0.57 | 12.6 | 7.98 | 0.28 | 0.68 |
| 62 | CSH | US-Los | −5.84 | 0.94 | 0.94 | 8.05 | 9.98 | 0.28 | 0.91 | −2.01 | 0.95 | 0.88 | 6.34 | 8.14 | 0.23 | 0.94 |
| 63 | WSA | AU-Dry | −8.95 | 0.77 | 0.91 | 15.3 | 19.4 | 0.27 | 0.68 | 2.95 | 0.69 | 0.77 | 13.7 | 19.5 | 0.27 | 0.68 |
| 64 | WSA | ES-LgS | −11.7 | 0.65 | 1.26 | 2.42 | 25.2 | 0.70 | −0.14 | 1.37 | 0.57 | 0.79 | 6.06 | 17.0 | 0.47 | 0.48 |
| 65 | WSA | US-SRM | 17.6 | 0.40 | 0.70 | -3.62 | 24.0 | 0.52 | 0.81 | 16.5 | 0.42 | 0.72 | -3.61 | 23.0 | 0.50 | -0.29 |

*Note*. CRO: croplands; CSH: closed shrublands; DBF: deciduous broadleaf forests; EBF: evergreen broadleaf forests; ENF: evergreen needleleaf forests; GRA: grasslands; MF: mixed forests; NSE: Nash–Sutcliffe efficiency; OSH open shrublands; PFT: plant functional type; RMSE: root-mean-square error; RE: relative error; SAV: savannas; WSA: woody savannas.

The better value is highlighted in bond.

relative narrow, indicating that information contained in observations was sufficiently used by our set simulation approach to constrain the model behaviour. Also, the medians of simulation sets are very similar to that estimated by the original model at these sites. Thus, it seems that the parameter uncertainties have less influence on the model's ET partitioning, and very satisfactory estimations of ET partitioning can be obtained by using the original model at these sites and PFTs (Figures 11 and 12). On the contrary, the 90% CIs of the simulation sets of ET partitioning showed relatively wider species- and PFT-variations for the ENF, OSH, SAV, CSH, and WSA biomes (Figures 11 and 12), indicating parameter uncertainties are large in these sites and biomes. The ET partitioning estimated by the original model at some sites of these biomes (i.e., CA-Obs, CA-TP3, CA-TP4, FI-Sod, IT-Lav, Mx-Lpa, ZA-Kru, AU-Dry, and ES-LgS) may be significantly different from that estimated by the set simulation approach (i.e., fall out 50% of CIs of simulation sets; Figure 11). Considering its relatively poor performance at these sites, the ET partitioning by the original model seemed to be unrealistic. Thus, special attention should be paid when

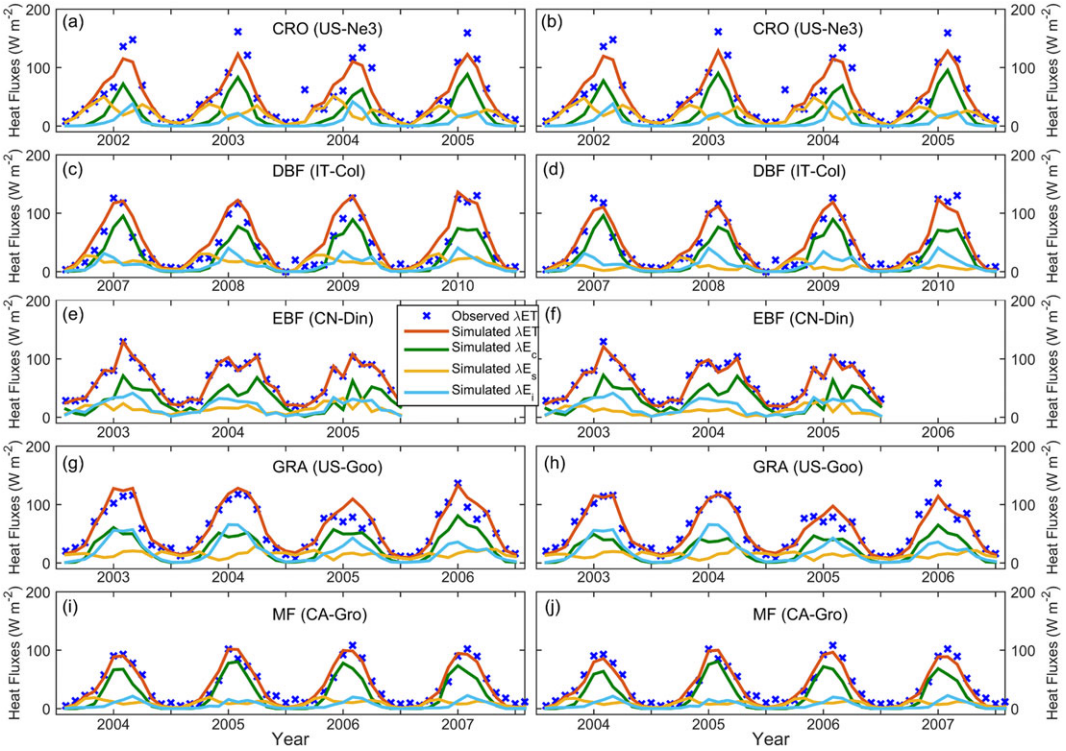


**FIGURE 9** Comparison water fluxes of observed, original model [left panel: (a), (c), (e), (g), and (i)] and optimized model [right panel: (b), (d), (f), (h), and (j)] in five different sites during the study period

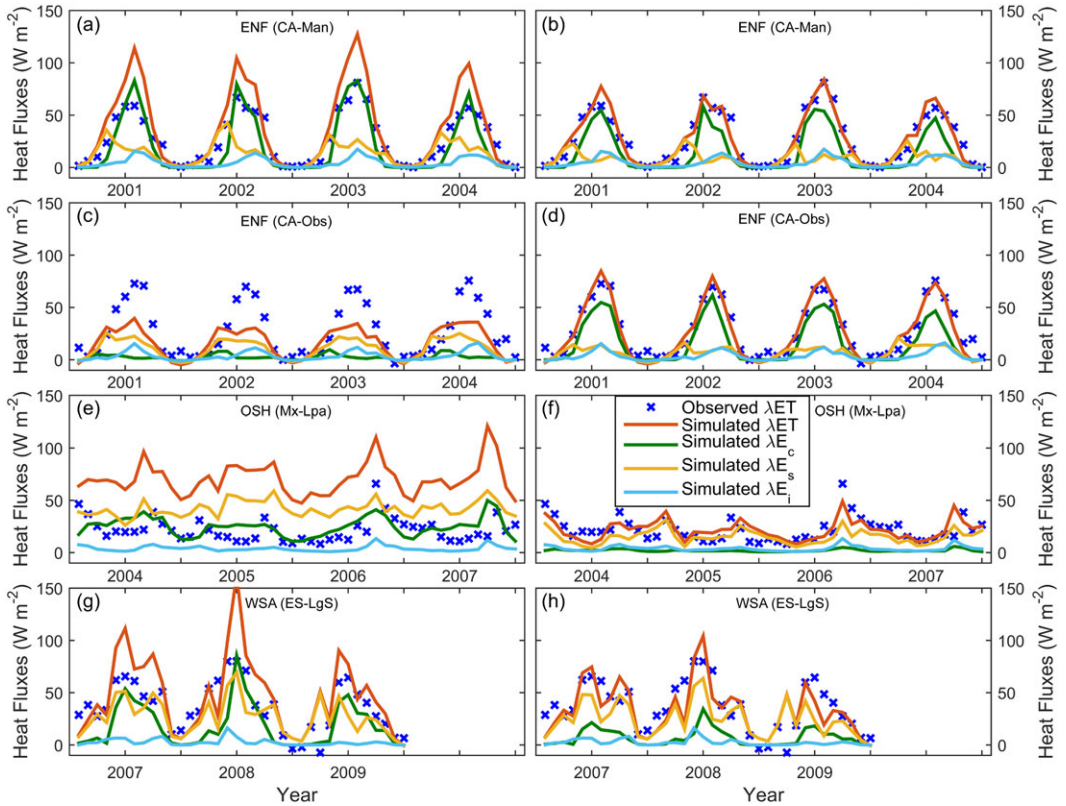


**FIGURE 10** Comparison water fluxes of observed, original model [left panel: (a), (c), (e), and (g)] and optimized model [right panel: (b), (d), (f), and (h)] in four different sites during the study period

using original model to partition ET over these biomes. Generally, the ET partitioning patterns calculated by the set simulation approach were in line with our intuitive expectation. For examples, the transpiration was the dominant component of ET over forest biomes with a PFT-level median value of $\lambda E_c/\lambda ET$ ranging from 0.49 (± 0.15 s. d.) for ENF to 0.59 ± 0.04 for DBF (Figure 12). Interestingly,

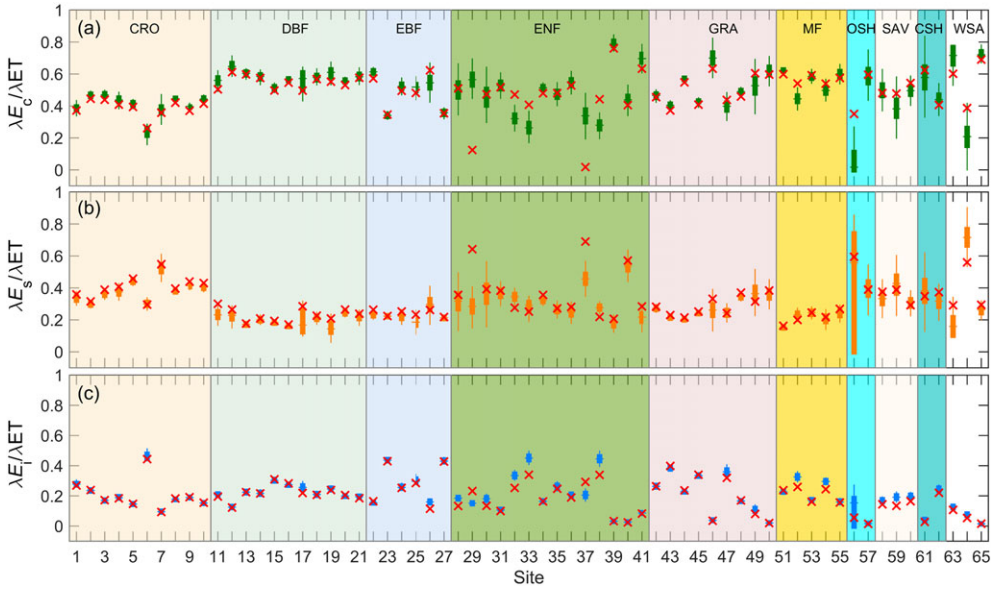


**FIGURE 11** The posterior medians and 90% CIs of the simulation sets of ET partitioning across all site: (a) $\lambda E_c/\lambda ET$, (b) $\lambda E_s/\lambda ET$, and (c) $\lambda E_i/\lambda ET$. The red cross symbol "×" indicates the partitioning results using the original model

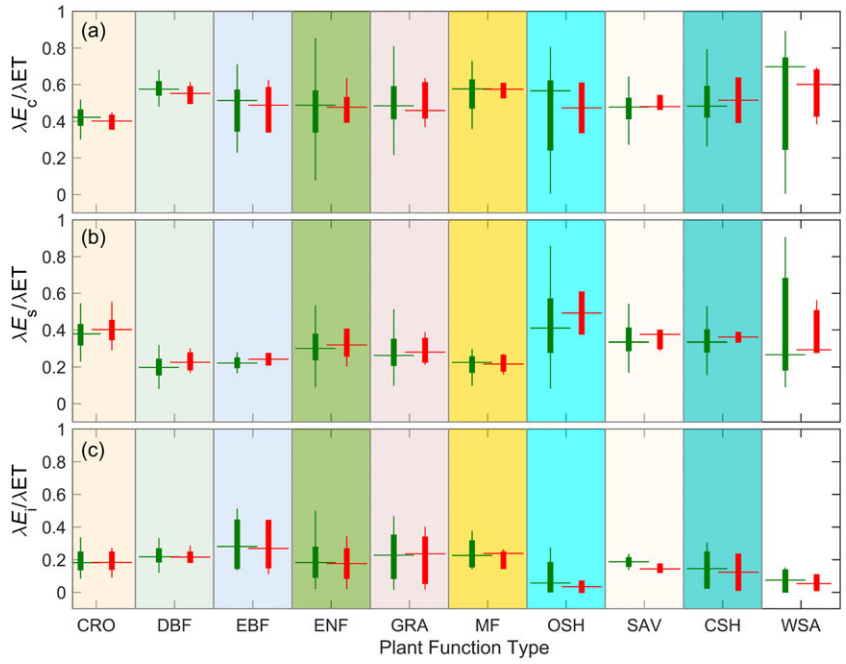


**FIGURE 12** Comparison of ET partitioning at PFT level: (a) $\lambda E_c/\lambda ET$, (b) $\lambda E_s/\lambda ET$, and (c) $\lambda E_i/\lambda ET$. The green bars represent the set simulation approach and the red bars represent the original model

large species-level variations in $\lambda E_c/\lambda ET$ were found within ENF (from 0.28 ± 0.03 at CA-TP4 to 0.79 ± 0.02 at IT-Lav), whereas relatively small species-level variations were observed within DBF (from 0.52 ± 0.02 to 0.65 ± 0.03) and MF (from 0.44 ± 0.03 to 0.62 ± 0.01) biomes (Figure 11). Soil evaporation is a significant ET component over CRO and OSH biomes with a PFT-level median value of 0.38 ± 0.07 and 0.41 ± 0.23, respectively (Figure 12). Canopy interception evaporation generally accounted for less than 30% of total ET with a maximum value of 0.28 ± 0.12 for the EBF and a minimum value of 0.03 ± 0.03 for OSH.

## 4 | DISCUSSION

### 4.1 | Simple vs hierarchical Bayesian approaches

Understanding the influence of model parameters on model response is significant elements in the development of robust regional and global ET products (Bastola, Murphy, & Sweeney, 2011; Brigode, Oudin, & Perrin, 2013; McCabe et al., 2016; Zhang et al., 2017; Zhu et al., 2014). Previous studies have shown that model parameters may vary by PFTs and species because of the genetic or local

environmental variation (Mu, Zhao, & Running, 2011; Wullschleger et al., 2014). Thus, there is a need to simultaneously incorporate species- and PFT-level parameter variability into the procedures of model calibration. The traditional SB approach fits the PT-JPL model on a site-by-site basis, and does not allow information to be exchanged between sites (Zhang et al., 2017). It is therefore difficult for the SB approach to simultaneously take the specie- and PFT-level variability into account. The failure to include a PFT-level constraint may also result in an overestimation of species-level variability. In addition, it attributes all uncertainties to the measurement errors (Feng & Dietze, 2013). Thus, even small amounts of measurement noise can cause significant errors in parameter estimation (Clark, 2005), especially for sites where the observation period is not long enough. In this study, we illustrate how to optimize the ET model parameters using the HB statistical methods that: (a) simultaneously utilize the multi-tower datasets. This ability enables the HB methods to exchange information across different sites and minimize the limitation of available data at some sites. Recently, Norros, Laine, Lignell, and Thingstad (2017) compared the performance of the SB and HB methods in estimating parameters of aquatic ecosystem models, and proved that the HB method can obtain the most precise parameter estimates due to its ability to combine information across different datasets; (b) explicitly incorporate different sources of uncertainty. In analyses across multiple levels, the uncertainty is partitioned and attributed to specific drivers (such as observations and parameters), which makes the HB methods are robust to measurement errors. In our prior tests using synthetic data, we found that the parameters estimated by the HB approach are very close to the actual values when the standard deviation of measurement error varies from 5 to 30% of the actual simulated values. Previous studies have illustrated that the overall uncertainty in latent heat flux measurements is about 13% (Wang, Zhuang, Wang, Liu, & Xu, 2014). Thus, the effect of measurement uncertainty plays a marginal role in identifying real-world ET model parameters for the HB method; (c) explicitly estimate parameter variations at PFT level by accommodating the species-level variability. Thus, the HB method offers an efficient way for estimating the PFT-specific parameters for the PFT-based models (discussed below).

## 4.2 | Parameter uncertainties and model performance

In the PT-JPL model, the three sensitive parameters (i.e., $T_{opt}$, $m_1$, $\beta$) are closely related to different processes of ET. In particular, $m_1$ determines the green canopy fraction. The larger $m_1$ is, the higher the green canopy fraction ($f_g$) is (Equation 8); $T_{opt}$ represents the optimum temperature for transpiration at which plant stomata is fully open. When air temperature is higher or lower than $T_{opt}$, plants close their stomata to prevent water loss through transpiration. $\beta$ is the unique parameter that influences the estimation of soil evaporation. The larger $\beta$ is, the higher soil evaporation rate is (Equation 7). Thus, parameters $T_{opt}$ and $m_1$ mainly influence the estimates of canopy transpiration, whereas $\beta$ influences the estimates of soil evaporation. To properly estimate the ET and its different components, it is important to get the robust values of the sensitive parameters over different biomes and climate conditions. Unique to the HB method is its ability to combine multi-tower FLUXNET datasets in a single analysis and simultaneously estimate species- and PFT-level parameters. Results for our HB approach illustrated that parameter $T_{opt}$ varied considerably on both a species and PFT level (Figures 4 and 5). Across PFT, the posterior estimations of $T_{opt}$ varied from 12.5 (MF) to 29.7°C (SAV; Figure 5 ; Table S3), which reflects different PFTs may have developed genotypic adjustments to their evolutionary temperature environment (Niinemets, Oja, & Kull, 1999). Also, this result provided good support for previous study that optimum air temperature for canopy carbon flux varied from 7.5 to 30°C for different PFTs (Baldocchi et al., 2001, 2001). Thus, it seems to be unreasonable to fix $T_{opt}$ to be 25°C across different PFTs for some global modelling studies (García et al., 2013; Yao et al., 2014; Yuan et al., 2010). Within a PFT, the posterior median estimations of $T_{opt}$ varied substantially from species to species (Figure 4), and were highly correlated with growing mean annual temperature (Figure 7a), indicating a species-specific adaptation to local environmental conditions. This also confirmed the non-trivial amount of site-to-site variation of $T_{opt}$ in our previous studies (Zhang et al., 2017), though as discussed above the SB approach likely overestimated the species-level variability. Indeed, the species-level variations within the ENF and MF biomes estimated by the HB approach were less than that obtained by the SB approach (Figure 4). Notably, the dynamic method used in the original model is prone to produce unrealistic estimations of $T_{opt.}$ at two sites (CA-Obs and FI-Sod) of the ENF biomes because the product of the two negative meteorological variables (i.e., radiation and temperature) is highest in winter at these sites. Improper estimates of $T_{opt}$ by the dynamic method were also observed in Mediterranean semiarid environments where the maximum peak for vegetation activity occurred in late winter (García et al., 2013). Thus, the established relationship between $T_{opt}$ and mean annual temperature provides a convenient way to generate reliable ET estimations by using the PT-JPL model.

The estimates of $m_1$ were well informed by the multi-tower datasets at both a species and PFT level, which was demonstrated by a narrow CIs and posterior means that were quite different from the corresponding prior means (Figures 4 and 5). Across PFT, the estimates of $m_1$ for ENF (0.41), MF (0.44), and SAV (0.88) were relatively low (Figure 5; Table S3), which might be caused by their canopy heterogeneity. In the remaining PFTs, the posterior median values of $m_1$ ranged from 1.09 to 1.38, and were close to the default value (1.36) of the original PT-JPL model. Thus, the empirical relationships between $f_{APAR}$ and vegetation index (i.e., EVI and NDVI) may be variable for different PFTs (Rautiainen et al., 2010). For example, Ogutu and Dash (2013) reported that $f_{APAR}$ was positively correlated with EVI for the Harvard deciduous broadleaf forest and the Mead irrigated cropland, and the value of $m_1$ for these two sites was 1.67 and 2.27, respectively. Within a PFT, the posterior median estimates of $m_1$ showed relatively less species-level variations for CRO, DBF and EBF. However, considerable species-level variations were observed for the remaining PFTs, indicating $f_{APAR}$ retrieval may be sensitive to local environmental conditions (i.e., soil reflectance, nonphotosynthetic plant components, and atmospheric conditions). Until now, accurately estimating $f_{APAR}$ is still a challenge (Xiao et al., 2004), and more studies are needed to better understand and quantify the relationship between $f_{APAR}$ and

satellite-derived vegetation indices (e.g., EVI, NDVI) across leaf, species, and PFT levels (Rautiainen et al., 2010).

The posterior median estimations of $\beta$ at the species and the PFT level ranged from 0.20 to 0.95 and 0.20 to 0.81, respectively, which were all lower than the default value (1.0) of the original model. In the PT-JPL model, $\beta$ is a unique parameter that influences the estimation of soil evaporation. It has been well documented that the default value of $\beta$ (= 1) tends to overestimate soil evaporation in arid regions (García et al., 2013; Zhu et al., 2016; Zhang et al., 2017). For a global application, Mu, Heinsch, Zhao, and Running (2007); Mu et al. (2011) set the value of $\beta$ to be 0.1 and 0.2 kPa in the old and revised MODIS16 algorithm, respectively. In addition, it should be noted that $\beta$ exhibited the most wide posterior distributions among the three sensitive parameters (Figures 4 and 5). This indicates that the parameterization of soil moisture constraint $f_{sm}$ (Equation 7), which is built based on the link between atmospheric water deficit (VPD and RH) and soil moisture, may be problematic. Recently, Novick et al. (2016) reported that the correlation coefficients between atmospheric water deficit and soil moisture are relative low at monthly scale across 38 Ameriflux sites, ranging from 0 to 0.54 with a mean of 0.20. Thus, it seems that the vertical adjacent atmosphere is not necessary in equilibrium with the underlying soil, and the link may tend to be decoupled. García et al. (2013) have proposed to parameterize $f_{sm}$ using apparent thermal inertia (ATI) to improve model performance in arid regions. Alternatively, a promising scheme for parameterizing soil moisture constraint is to formulate $f_{sm}$ directly based on surface soil moisture (Fisher et al., 2008), which now can be easily obtained by ground instruments, remote sensing techniques and models.

## 4.3 | ET partitioning

The parameters in the PT-JPL model are found to be highly correlated and can trade off. That is, very different combination of the parameter values ("trade-off" between the parameters) can make the model generate similar total ET predictions, but with great uncertainties in the model's ET partitioning. This phenomenon is well known as equifinality or parameter identifiability (Beven & Freer, 2001; Zhu et al., 2014, 2018). Thus, we cannot ensure the model's ET partitioning based on a single set of optimized parameters to be correct, even though the simulated total ET were in good agreement with measurements. The set simulation approach proposed here provides us the opportunity to estimate the posterior distributions of the model's ET partitioning.

Based on the set simulation approach, the values of $\lambda E_c/\lambda ET$ ranged from 0.43 (± 0.07 s. d.; CRO) to 0.72 ± 0.29 (WSA) with a mean of 0.53 ± 0.10 (Figure 12). These results are nearly within the scope of previous studies (Lawrence et al., 2007; Miralles, Gash, Holmes, De Jeu, & Dolman, 2010; Schlesinger & Jasechko, 2014; Coenders-Gerrits et al., 2014; Maxwell & Condon, 2016). For example, Schlesinger and Jasechko (2014) compiled 81 studies that partitioned ET into transpiration and evaporation (ignoring canopy interception), and indicated that transpiration accounted for 61% (± 15%) of ET with highest in tropical rainforests (70 ± 14%) and lowest in steppes, shrublands, and deserts (51 ± 15%). Maxwell and Condon (2016) found transpiration accounts from 47 ± 13% to 62 ± 12% of ET after considering the influence of lateral ground water flow on ET partitioning. However, relative higher transpiration ratios were obtained by using the isotope mass balance method. Jasechko et al. (2013) reported that transpiration could account for nearly 80–90% of the total ET from continents. This finding was criticized for overestimating the contribution of plant transpiration and underestimating data uncertainties (Coenders-Gerrits et al., 2014). After counting the input data uncertainties, Coenders-Gerrits et al. (2014) argued the transpiration portion of ET to be lower, at 35–80%. For other components, we found the values of $\lambda E_s/\lambda ET$ ranged from 0.20 ± 0.03 (EBF) to 0.41 ± 0.23 (OSH) with a mean of 0.28 ± 0.11 (Figure 12). Previous studies on contributions of soil evaporation to total ET mainly focused on croplands for better irrigation management practices, and results suggest that soil evaporation accounts for 20–40% of ET (Kool et al., 2014). This is generally consistent with our estimations for croplands ranging from 0.3 to 0.52 with a mean of 0.38 (Figure 11). Until now, direct observations on the contributions of canopy loss to total ET across different biomes are still relative sparse (Miralles et al., 2010), and estimations were mainly derived from model simulation. For example, Lawrence et al. (2007) used a modified Community Land Model version 3 (CLM3) to simulate the ET partitioning, and reported that intercept loss accounts for 17% of total ET over the global land surface. According to our results, the values for $\lambda E_i/\lambda ET$ranged from 0.08 ± 0.06 (WSA) to 0.28 ± 0.12 (EBF) with a mean of 0.18 ± 0.06, which is similar to that reported by Lawrence et al. (2007). Although the comparisons were relatively rough, the results indicated that the PT-JPL model can be used to partition ET at ecosystem scale after carefully considering the parameter uncertainties. Notably, our synthetic test also indicated that the partitioning results can be further improved by using two type data (Figure S2). Unfortunately, direct observations of different components of ET are still not accessible at most FLUXNET sites at present (Kool et al., 2014). Thus, a major step forwards should be made to extract direct observations of different components of ET at FLUXNET sites using new observation technologies, such as environmental stable isotope ($\delta^2H$ and $\delta^{18}O$), sap flow, and lysimeter (Kool et al., 2014), and systematic assessments the extent to which the uncertainty in model parameters and predictions is reduced by the use of additional data are needed in the future studies.

## 4.4 | Potential applications of the HB approach

Now, the PFT concept offers a tractable scheme to describe the landscape, and is well integrated into dynamic vegetation models [i.e., LPJ-DGVM (Sitch et al., 2003); BIOME4 (Kaplan et al., 2003); CLM4.5 (Oleson et al., 2013); ORCHIDEE (Krinner et al., 2005)] and some remote sensing based ET models [i.e., MODIS16 algorithm (Mu et al., 2011); P-LSH model (Zhang et al., 2009)]. Most current models use PFT-specific constant parameter values for regional or global applications (Wullschleger et al., 2014), and not accounting for species-level variation within PFTs. However, results from our HB approach demonstrated that variation of parameters at species level was also significant in comparison with the PFT-level variation (Figures. 5 and 6). This highlights the potential importance of accounting for species-level variability into models in order to obtain credible

predictions. Indeed, it has been well documented that the variation in plant traits related to carbon, water, and nutrient cycling is large within PFTs and often even greater than the difference in means among PFTs (Laughlin, Leppert, Moore, & Sieg, 2010; Wright et al., 2005). Also, this study indicates that the PFT-level parameter estimates and their associated uncertainties may be compromised if variability among species is ignored. Thus, it is important to account for species-level variation to obtain proper estimates of PFT-level parameters. The method presented here provides a consistent framework to obtain proper PFT-level parameter estimations by accommodating the species-level variability. In future studies, we will apply the HB approach on the PFT-based models to obtain a new set of PFT-specific parameters for the models, and evaluate to what extent the model performance can be improved by the new set of parameters.

## 5 | CONCLUSIONS

With the increased application of process-based ET model in multiple research areas, there is an urgent need for methods extracting empirically and theoretically sound parameter values that provide realistic ET predictions and partitioning estimates. The HB method presented herein allowed us to rigorously fit fairly complicated ET models to multi-tower FLUXNET data via a population-based Monte Carlo algorithm that: (a) simultaneously analysed the multi-tower data sources that informed the same underlying ET processes; (b) explicitly accounted for different sources of uncertainty; and (c) simultaneously estimated species- and PFT-level parameters. As such, the HB method can efficiently identify the parameters and obtain realistic estimating in model's ET partitioning. However, strict validation of the model's ET partitioning is not still possible because the required observations do not exist. Therefore, a major step forwards should be made to extract direct observations of different ET components at FLUXNET sites with the development of new observation technologies (i.e., sap flow, lysimeter, and environmental stable isotope). In addition, further analysis is needed to address the model structure uncertainty. This is beyond the scope of current paper, and we will consider using multiple different ET models to reduce the uncertainty caused by the model structure in future studies.

## ACKNOWLEDGMENTS

This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia, and USCCC. The ERA-Interim reanalysis data are provided by ECMWF and processed by LSCE. The FLUXNET eddy covariance data processing and harmonization was carried out by the European Fluxes Database Cluster, AmeriFlux Management Project, and Fluxdata project of FLUXNET, with the support of CDIAC and ICOS Ecosystem Thematic Center, and the OzFlux, ChinaFlux, and AsiaFlux offices. MODIS NDVI, EVI, LAI satellite products were obtained online (https://lpdaac.usgs.gov/).

The authors would thank Prof. A. Molini (Editor) for her continued help during the revisions of the paper. We also thank the anonymous reviews for their critical reviews and helpful comments. This research was supported by the National Natural Science Foundation of China (Nos. 41571016 and 31370467), and National Key R & D Program of China (2016YFC0501002). The data and code written in MATLAB can be obtained from the first author (zhugf@lzu.edu.cn) upon request.

## ORCID

*Gaofeng Zhu* http://orcid.org/0000-0001-6797-7965

## REFERENCES

Agarwal, D. A., Humphrey, M., Beekwilder, N. F., Jackson, K. R., Goode, M. M., & van Ingen, C. (2010). A data-centered collaboration portal to support global carbon-flux analysis. *Concurrency and Computation: Practice and Experience*, *22*(17), 2323–2334. https://doi.org/10.1002/cpe.1600

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., … Wofsy, S. (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, *82*, 2415–2434. https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2

Barr, A. G., Morgenstern, K., Black, T. A., McCaughey, J. H., & Nesic, Z. (2006). Surface energy balance closure by the eddy-covariance method above three boreal forest stands and implications for the measurement of the CO2 flux. *Agricultural and Forest Meteorology*, *140*(1–4), 322–337. https://doi.org/10.1016/j.agrformet.2006.08.007

Bastola, S., Murphy, C., & Sweeney, J. (2011). The sensitivity of fluvial flood risk in irish catchments to the range of IPCC AR4 climate change scenarios. *Science of the Total Environment*, *409*, 5403–5415. https://doi.org/10.1016/j.scitotenv.2011.08.042

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., … Papale, D. (2010). Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate. *Science*, *329*(5993), 834–838. https://doi.org/10.1126/science.1184984

Beven, K., & Freer, J. (2001). Equifinality, data assimilation, and data uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, *249*, 11–29. https://doi.org/10.1016/S0022-1694(01)00421-8

Brigode, P., Oudin, L., & Perrin, C. (2013). Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change? *Journal of Hydrology*, *476*, 410–425. https://doi.org/10.1016/j.jhydrol.2012.11.012

Carlin, B. P., Clark, J. S., & Gelfand, A. E. (2006). Elements of hierarchical Bayesian inference. In J. S. Clark, & A. E. Gelfand (Eds.), *Hierarchical modelling for the environmental sciences: Statistical methods and applications*. New York: Oxford University Press Inc.

Clark, J. S. (2005). Why environmental scientists are becoming Bayesians. *Ecology Letters*, *8*, 2–14.

Clark, J. S., & Bjønstad, O. N. (2004). Population time series: Process variability, observation errors, missing values, lags, and hidden states. *Ecology*, *85*, 3140–3150. https://doi.org/10.1890/03-0520

Coenders-Gerrits, A. M. J., van der Ent, R. J., Bogaard, T. A., Wang-Erlandsson, L., Hrachowitz, M., & Savenije, H. H. G. (2014). Uncertainties in transpiration estimates. *Nature*, *506*, E1–E2. https://doi.org/10.1038/nature12925

Didan, K. (2015). MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006. NASA EOSDIS LP DAAC. doi: https://doi.org/10.5067/MODIS/MOD13Q1.006.

Ershadi, A., McCabe, M. F., Evans, J. P., Chaney, N. W., & Wood, E. F. (2014). Multi-site evaluation of terrestrial evaporation models using FLUXNET data. *Agricultural and Forest Meteorology*, *187*, 46–61. https://doi.org/10.1016/j.agrformet.2013.11.008

Fatichi, S., & Pappas, C. (2017). Constrained variability of modeled T: ET ratio across biomes. *Geophysical Research Letters*, *44*, 6795–6803. https://doi.org/10.1002/2017GL074041

Feng, F., Chen, J. Q., Li, X. L., Yao, Y. J., Liang, S. L., Liu, M., … Sun, M. M. (2015). Validity of five satellite-based latent heat flux algorithms for semi-arid ecosystems. *Remote Sens-Basel*, *7*(12), 16733–16755. https://doi.org/10.3390/rs71215853

Feng, X., & Dietze, M. (2013). Scale dependence in the effects of leaf ecophysiological traits on photosynthesis: Bayesian parameterization of photosynthesis models. *New Phytologist*, *200*, 1132–1144. https://doi.org/10.1111/nph.12454

Fisher, J. B., Tu, K. P., & Baldocchi, D. D. (2008). Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites. *Remote Sensing of Environment*, *112*(3), 901–919. https://doi.org/10.1016/j.rse.2007.06.025

Franks, S., Beven, K. J., Quinn, P. F., & Wright, I. (1997). On the sensitivity of soil-vegetation-atmosphere transfer (SVAT) schemes: Equifinality and the problem of robust calibration. *Agricultural and Forest Meteorology*, *86*, 63–75. https://doi.org/10.1016/S0168-1923(96)02421-5

Gao, X., Huete, A. R., Ni, W. G., & Miura, T. (2000). Optical-biophysical relationships of vegetation spectra without background contamination. *Remote Sensing of Environment*, *74*, 609–620. https://doi.org/10.1016/S0034-4257(00)00150-4

García, M., Sandholt, I., Ceccato, P., Ridler, M., Mougin, E., Kergoat, L., … Domingo, M. (2013). Actual evapotranspiration in drylands derived from in-situ and satellite data: Assessing biophysical constraints. *Remote Sensing of Environment*, *131*, 103–118.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, *1*, 515–533. https://doi.org/10.1214/06-BA117A

Haario, H., Laine, M., Mira, A., & Saksman, E. (2006). DRAM: Efficient adaptive MCMC. *Statistics and Computing*, *16*(4), 339–354. https://doi.org/10.1007/s11222-006-9438-0

International Panel on Climate Change. (2001). *Climate change 2001: The scientific basis: Contribution of working group I to the third assessment report of the international panel on climate change*, edited by J. T. Houghton et al., Cambrige Univ. Press, Cambrige, U. K., and New York.

Jasechko, S., Sharp, Z. D., Gibson, J. J., Birks, S. J., Yi, Y., & Fawcett, P. J. (2013). Terrestrial water fluxes dominated by transpiration. *Nature*, *496*, 347–350.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., … Zhang, K. (2010). Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, *467*, 951–954. https://doi.org/10.1038/nature09396

Kaplan, J. O., Bigelow, N. H., Prentice, I. C., Harrison, S. P., Bartlein, P. J., Christensen, T. R., … Lozhkin, A. V. (2003). Climate change and arctic ecosystems: 2. Modeling, paleodata-model comparisons, and future projections. *Journal of Geophysical Research*, *108*(D19), 8171. https://doi.org/10.1029/2002JD002559

Kool, D., Agam, N., Lazarovitch, N., Heitman, J. L., Sauer, T. J., & Ben-Gal, A. (2014). A review of approaches for evapotranspiration partitioning. *Agricultural and Forest Meteorology*, *184*, 56–70. https://doi.org/10.1016/j.agrformet.2013.09.003

Krinner, G., Viovy, N., de Noblet-Ducoudre, 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 Biogeochemical Cycles, 19, GB1015, doi:https://doi.org/10.1029/2003GB002199.

Laughlin, D. C., Leppert, J. J., Moore, M. M., & Sieg, C. H. (2010). A multi-trait test of the leaf-height-seed plant strategy scheme with 133 species from a pine forest flora. *Functional Ecology*, *24*, 493–501. https://doi.org/10.1111/j.1365-2435.2009.01672.x

Lawrence, D. M., Thornton, P. E., Oleson, K. W., & Bonan, G. B. (2007). The partitioning of evapotranspiration into transpiration, soil evaporation, and canopy interception in a GCM: Impacts on land–atmosphere interaction. *Journal of Hydrometeorology*, *8*, 862–880. https://doi.org/10.1175/JHM596.1

Massman, W. J., & Lee, X. (2002). Eddy covariance flux corrections and uncertainties in long-term studies of carbon and energy exchanges. *Agricultural and Forest Meteorology*, *113*, 121–144. https://doi.org/10.1016/S0168-1923(02)00105-3

Maxwell, R. M., & Condon, L. E. (2016). Connections between groundwater flow and transpiration partitioning. *Science*, *353*(6297), 377–380. https://doi.org/10.1126/science.aaf7891

McCabe, M. F., Ershadi, A., Jimenez, C., Miralles, D. G., Michel, D., & Wood, E. F. (2016). The GEWEX LandFlux project: Evaluation of model evaporation using tower-based and globally gridded forcing data. *Geoscientific Model Development*, *9*(1), 283–305. https://doi.org/10.5194/gmd-9-283-2016

Michel, D., Jiménez, C., Miralles, D. G., Jung, M., Hirschi, M., Ershadi, A., … Fernández-Prieto, D. (2016). The WACMOS-ET project-Part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms. *Hydrology and Earth System Sciences*, *20*(2), 803–822. https://doi.org/10.5194/hess-20-803-2016

Miralles, D. G., Gash, J. H., Holmes, R. H., De Jeu, R. A. M., & Dolman, A. J. (2010). Global canopy interception from satellite observations. *Journal of Geophysical Research*, *115*, D16122. https://doi.org/10.1029/2009JD013530

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *T Asabe*, *50*(3), 885–900. https://doi.org/10.13031/2013.23153

Mu, Q., Heinsch, F. A., Zhao, M., & Running, S. W. (2007). Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sensing of Environment*, *111*(4), 519–536. https://doi.org/10.1016/j.rse.2007.04.015

Mu, Q. Z., Zhao, M., & Running, S. W. (2011). Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, *115*, 1781–1800. https://doi.org/10.1016/j.rse.2011.02.019

Myneni, R., Knyazikhin, Y., & Park, T. (2015). MOD15A2H MODIS/Terra Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V006. NASA EOSDIS Land Processes DAAC. http://doi.org/10.5067/MODIS/MOD15A2H.006.

Newman, B. D., Wilcox, B. P., Archer, S. R., Breshears, D. D., Clifford, N. D., Duffy, C. J., … Vivoni, E. R. (2006). Ecohydrology of water-limited environments: A scientific vision. *Water Resources Research*, *42*, W06302. https://doi.org/10.1029/2005WR004141

Niinemets, U., Oja, V., & Kull, O. (1999). Shape of leaf photosynthetic electron transport versus temperature response curve is not constant along canopy light gradients in temperate deciduous trees. *Plant, Cell and Environment*, *22*, 1497–1513. https://doi.org/10.1046/j.1365-3040.1999.00510.x

Norros, V., Laine, M., Lignell, R., & Thingstad, F. (2017). Parameterization of aquatic ecosystem functioning and its natural variation: Hierarchical Bayesian modelling of plankton food web dynamics. *Journal of Marine Systems*, *174*, 40–53. https://doi.org/10.1016/j.jmarsys.2017.05.004

Novick, K. A., Ficklin, D. L., Stoy, P. C., Williams, C. A., Bohrer, G., Oishi, A. C., … Phillips, R. P. (2016). The increasing importance of atmospheric demand for ecosystem water and carbon fluxes. *Nature Climate Change*, *6*, 1023–1027. https://doi.org/10.1038/nclimate3114

Ogutu, B. O., & Dash, J. (2013). An algorithm to derive the fraction of photosynthetically active radiation absorbed by photosynthetic elements of the canopy (FAPARps) from eddy covariance flux tower data. *New Phytologist*, *197*, 511–523.

Oleson, K., Lawrence, D. M., Bonan, G. B., Drewniak, B., Huang, M., Koven, C. D., … Yang, Z.-L. (2013). Technical description of version 4.5 of the Community Land Model (CLM). NCAR Technical Note NCAR/TN-503+STR. doi: https://doi.org/10.5065/D6RR1W7M.

Pastorello, G. Z., Papale, D., Chu, H., Trotta, C., Agarwal, D. A., Canfora, E., … Torn, M. S. (2017). A new data set to keep a sharper eye on land-air exchanges, Eos, 98, https://doi.org/10.1029/2017EO071597.

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., & Klooster, S. A. (1993). Terrestrial ecosystem

production-A process model-based on global satellite and surface data. *Global Biogeochemical Cycles*, *7*(4), 811–841. https://doi.org/10.1029/93GB02725

Priestley, C. H. B., & Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, *100*, 81–92. https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2

Rautiainen, M., Heiskanen, J., Eklundh, L., Ttus, M., Lukes, P., & Stenberg, P. (2010). Ecological applications of physically based remote sensing methods. *Scandinavian Journal of Forest Research*, *25*, 325–339. https://doi.org/10.1080/02827581.2010.497159

Robert, C. P., & Casella, G. (1999). *Monte Carlo statistical methods*. New York, NY, USA: Springer-Verlag. https://doi.org/10.1007/978-1-4757-3071-5

Roberts, G. O., & Rosenthal, J. S. (2001). Optimal scaling for various Metropolis–Hastings algorithms. *Statistical Science*, *16*(4), 351–367. https://doi.org/10.1214/ss/1015346320

Russell, L. S., & Biederman, J. A. (2017). Partitioning evapotranspiration using long-term carbon dioxide and water vapor fluxes. *Geophysical Research Letters*, *44*(13), 6833–6840.

Samanta, S., Mackay, D. S., Clayton, M. K., Kruger, E. L., & Ewers, B. E. (2007). Bayesian analysis for uncertainty estimation of a canopy transpiration model. *Water Resources Research*, *43*, W04424. http://dx.doi.org/10.1029/2006WR005028

Schlesinger, W. H., & Jasechko, S. (2014). Transpiration in the global water cycle. *Agricultural and Forest Meteorology*, *189-190*, 115–117. https://doi.org/10.1016/j.agrformet.2014.01.011

Shugart, H. H. (2000). Ecosystem modeling. In O. E. Sala, R. B. Jackson, H. A. Mooney, & R. W. Howarth (Eds.), *Methods in ecosystem science* (pp. 373–388). New York: Springer. https://doi.org/10.1007/978-1-4612-1224-9_26

Sitch, S., Smith, B., Prentice, I. C., Areneth, A., Bondeau, A., Cramer, W., … Venevsky, S. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, *9*, 161–185. https://doi.org/10.1046/j.1365-2486.2003.00569.x

Storn, R., & Price, K. (1997). Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, *11*, 341–359. https://doi.org/10.1023/A:1008202821328

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719

ter Braak, C. J. F. (2006). A Markov chain Monte Carlo version of the genetic algorithm differential evolution: Easy Bayesian computing for real parameter spaces. *Statistics and Computing*, *16*(3), 239–249. https://doi.org/10.1007/s11222-006-8769-1

Turner, B. M., Sederberg, P. B., Brown, S. D., & Steyvers, M. (2013). A method for efficiently sampling from distributions with correlated dimensions. *Psychological Methods*, *18*(3), 368–384. https://doi.org/10.1037/a0032222

Twine, T. E., Kustas, W. P., Norman, J. M., Cook, D. R., Houser, P. R., Meyers, T. P., … Wesley, M. L. (2000). Correcting eddy-covariance flux underestimates over a grassland. *Agricultural and Forest Meteorology*, *103*, 279–300. https://doi.org/10.1016/S0168-1923(00)00123-4

Wang, J. M., Zhuang, J. X., Wang, W. Z., Liu, S. M., & Xu, Z. W. (2014). Assessment of uncertainties in eddy covariance flux measurement based on intensive flux matrix of HiWATER-MUSOEXE. *IEEE Geoscience and Remote Sensing Letters*, *12*(2), 259–264.

Wei, Z., Yoshimura, K., Wang, L., Miralles, D. G., Jasechko, S., & Lee, X. (2017). Revisiting the contribution of transpiration to global terrestrial evapotranspiration. *Geophysical Research Letters*, *44*, 2792–2801. https://doi.org/10.1002/2016GL072235

Wikle, C. K. (2003). Hierarchical models in environmental science. *International Statistical Review*, *71*, 181–199.

Wright, I. J., Reich, P. B., Cornelissen, J. H. C., Falster, D. S., Garnier, E., Hikosaka, K., … Westoby, M. (2005). Assessing the generality of global leaf trait relationships. *The New Phytologist*, *166*, 485–496. https://doi.org/10.1111/j.1469-8137.2005.01349.x

Wullschleger, S. D., Epstein, H. E., Box, E. O., Euskirchen, E. S., Goswami, S., Iversen, C. M., … Xu, X. (2014). Plant functional types in Earth system models: Past experiences and future directions for application of dynamic vegetation models in high-latitude ecosystems. *Annals of Botany*, *114*, 1–16. https://doi.org/10.1093/aob/mcu077

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

Yao, Y. J., Liang, S. L., Zhao, S. H., Zhang, Y. H., Qin, Q. M., Cheng, J., … Liu, M. (2014). Validation and application of the modified satellite-based Priestley-Taylor algorithm for mapping terrestrial evapotranspiration. *Remote Sens-Basel*, *6*(1), 880–904. https://doi.org/10.3390/rs6010880

Yuan, W. P., Liu, S. G., Yu, G. R., Bonnefond, J. M., Chen, J. Q., Davis, K., … Verma, S. B. (2010). Global estimates of evapotranspiration and gross primary production based on MODIS and global meteorology data. *Remote Sensing of Environment*, *114*, 1416–1431. https://doi.org/10.1016/j.rse.2010.01.022

Zhang, K., Kimball, J. S., Mu, Q., Jones, L. A., Goetz, S. J., & Running, S. W. (2009). Satellite based analysis of northern ET trends and associated changes in the regional water balance from 1983 to 2005. *Journal of Hydrology*, *379*, 92–110. https://doi.org/10.1016/j.jhydrol.2009.09.047

Zhang, K., Ma, J., Zhu, G., Ma, T., Han, T., & Feng, L. L. (2017). Parameter sensitivity analysis and optimization for a satellite-based evapotranspiration model across multiple sites using Moderate Resolution Imaging Spectroradiometer and flux data. *Journal of Geophysical Research – Atmospheres*, *122*, 230–245. https://doi.org/10.1002/2016JD025768

Zhu, G. F., Li, X., Ma, J. Z., Wang, Y. Q., Liu, S. M., Huang, C. L., … Hu, X. L. (2018). A new moving strategy for the sequential Monte Carlo approach in optimizing the hydrological model parameters. *Advances in Water Resources*, *114*, 164–179.

Zhu, G. F., Li, X., Su, Y. H., Zhang, K., Bai, Y., Ma, J. Z., … He, J. H. (2014). Simultaneously assimilating multivariate data sets into the two-source evapotranspiration model by Bayesian approach: Application to spring maize in an arid region of northwestern China. *Geoscientific Model Development*, *7*(4), 1467–1482. https://doi.org/10.5194/gmd-7-1467-2014

Zhu, G. F., Li, X., Zhang, K., Ding, Z. Y., Han, T., Ma, J. Z., … Ma, T. (2016). Multi-model ensemble prediction of terrestrial evapotranspiration across north China using Bayesian model averaging. *Hydrological Processes*, *30*(16), 2861–2879. https://doi.org/10.1002/hyp.10832

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Su Y, Feng Q, Zhu G, et al. A hierarchical Bayesian approach for multi-site optimization of a satellite-based evapotranspiration model. *Hydrological Processes*. 2018;32:3907–3923. https://doi.org/10.1002/hyp.13298