Towards Robust Parameterizations in Ecosystem-level Photosynthesis Models

Shanning Bao[1,1], Lazaro Alonso[2,2], Siyuan Wang[3,3], Johannes Gensheimer[2,2], Ranit De[2,2], and Nuno Carvalhais[3,3]

[1]National Space Science Center Chinese Academy of Sciences
[2]Max-Planck Institute for Biogeochemistry
[3]Max-Planck-Institute for Bio-geochemistry

December 7, 2022

Abstract

In a model simulating the dynamics of a system, parameters can represent system properties and unresolved processes, therefore affecting the model accuracy and uncertainties. For a light use efficiency (LUE) model, which is a typical tool to estimate gross primary productivity (GPP), the plant-functional-type (PFT)-dependent parameterization method was widely used to extrapolate parameters to larger spatial scales. However, the method cannot capture the spatial variability within PFT and introduces misclassification errors. To overcome the shortage, here we proposed an ecosystem-property-based parameterization method (mNN-GPP) for an LUE model. This method refers to predicting model parameters using the multi-output artificial neural network based on collected variables including PFT, climate types, bioclimatic variables, vegetation features, atmospheric deposition and soil properties at 196 FLLUXNET eddy covariance flux sites. The neural network was optimized according to GPP errors and constraints on sensitivity functions of the LUE model. We compared mNN-GPP with eleven other typical parameter extrapolating methods, including PFT-, climate-specific parameterization, global and PFT-based parameter optimization, site-similarity-based, and regression methods. These twelve methods were assessed using Nash-Sutcliffe model efficiency (NSE), determination coefficient and normalized root mean squared error of the simulated GPP. The simulated results were also contrasted with those of site-specific calibration based on full-time-series GPP estimated from observational net ecosystem exchange. The N-fold cross-validated results showed that mNN-GPP had the best performance across various temporal and spatial scales (e.g., NSE=0.62 at the daily scale). No extrapolated parameters reached the same performance as the calibrated parameters (NSE=0.82), but the ranges of predicted parameters were constrained. Furthermore, the Shapley values, layer-wise relevance and partial dependence of the input features showed that bioclimatic variables, PFT, and vegetation features are the key variables determining parameters. We recommend using the parameterization method considering both ecosystem properties and prediction errors to other GPP models and across spatio-temporal scales.

A Robust Light Use Efficiency Model Parameterization Method Based on Ecosystem Properties

Shanning Bao[1,2*], Lazaro Alonso[1], Siyuan Wang[1], Johannes Gensheimer[1], Ranit De[1], Nuno Carvalhais[1,3*]

1. Department for Biogeochemical Integration, Max-Planck-Institute for Biogeochemistry, Jena, Germany

2. National Space Science Center, Chinese Academy of Sciences, Beijing, China

3. Departamento de Ciências e Engenharia do Ambiente, DCEA, Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

([*]Corresponding authors: Shanning Bao, baoshanning@nssc.ac.cn, Nuno Carvalhais, ncarvalhais@bgc-jena.mpg.de)

Abstract

In a model simulating the dynamics of a system, parameters can represent system properties and unresolved processes, therefore affecting the model accuracy and uncertainties. For a light use efficiency (LUE) model, which is a typical tool to estimate gross primary productivity (GPP), the plant-functional-type (PFT)-dependent parameterization method was widely used to extrapolate parameters to larger spatial scales. However, the method cannot capture the spatial variability within PFT and introduces misclassification errors. Here we proposed an ecosystem-property-based parameterization method (mNN-GPP) for an LUE model to overcome the issues. This method refers to predicting model parameters using the multi-output artificial neural network based on collected variables including PFT, climate types, bioclimatic variables, vegetation features, atmospheric deposition and soil properties at 196 FLLUXNET eddy covariance flux sites. The neural network was optimized according to GPP errors and constraints on sensitivity functions of the LUE model. We compared mNN-GPP with eleven other typical parameter extrapolating methods, including PFT-, climate-specific parameterization, global and PFT-based parameter optimization, site-similarity-based, and regression methods. These twelve methods were assessed using Nash-Sutcliffe model efficiency (NSE), determination coefficient and normalized root mean squared error of the simulated GPP. The simulated results were also contrasted with those of site-specific calibration based on full-time-series GPP estimated from observational net ecosystem exchange. The N-fold cross-validated results showed that mNN-GPP had the best performance across various temporal and spatial scales (e.g., NSE=0.62 at the daily scale). No extrapolated parameters reached the same performance as the calibrated parameters (NSE=0.82), but the ranges of predicted parameters were constrained. Furthermore, the Shapley values, layer-wise relevance and partial dependence of the input features showed that bioclimatic variables, PFT, and vegetation features are the key variables determining parameters. We recommend using the parameterization method considering both ecosystem properties and prediction errors to other GPP models and across spatio-temporal scales.

Keywords: parameter extrapolation, gross primary productivity, site-specific parameterization, vegetation and climate features, feature importance, spatio-temporal scales

1. Introduction

Growing multi-model-based studies reveal that large uncertainties resulted from various model structures, driver data and parameters (Huntzinger et al., 2017; Medlyn, Robinson, Clement, & McMurtrie, 2005; Zheng et al., 2018) are remained in modeling global carbon cycle and system responses to environmental changes (Baldocchi, Ryu, & Keenan, 2016; Bloom, Exbrayat, Van Der Velde, Feng, & Williams, 2016; Piao et al., 2020). Although model parameters contribute to considerable uncertainties, most global vegetation models were parameterized using fixed, biome- or plant functional type (PFT)- based values, which cannot capture the spatial variability of carbon process (Bloom et al., 2016). The fixed and PFT-based parameterization were also widely used in and introduced uncertainties to gross primary productivity (GPP) models (Groenendijk et al., 2011; Ryu, Berry, & Baldocchi, 2019), including light use efficiency (LUE), leaf-scale-process-based, machine-learning and sun-induced fluorescence models (Frankenberg et al., 2011; Jung et al., 2011; Running et al., 2004; Tian et al., 2020; Zhang et al., 2012). A more robust and physically intuitive parameterization method is desired for constraining the global GPP estimation.

LUE models are typical tools for the estimation of GPP at a large scale and the global scale (Mahadevan et al., 2008; Potter et al., 1993; Running et al., 2004; Tian et al., 2020; Yuan et al., 2007). The kind of models incorporates the knowledge of environmental constraints to the originally proposed empirical LUE model, Monteith et al.'s model (Monteith, 1972), having advantages of high efficiency and algorithmic transparency compared to leaf-scale-process-based models and machine-learning-based models, respectively.

The first global GPP product based on MODIS LUE algorithm (Running et al., 2004) proposed a set of PFT-dependent parameters. Other later published global LUE models inherit the PFT-based parameterization method or parameters directly from papers (He et al., 2013; Mahadevan et al., 2008; Xiao et al., 2004; Xie & Li, 2020). However, applying parameters in regions which it was not previously used for or evaluated against might easily lead to the erroneous conclusion that a specific model structure and parameter set falls short in locally describing ecosystem GPP. To overcome this limitation, LUE modelers usually calibrate parameters in their physical ranges according to the distance to the observational GPP (Nuno Carvalhais et al., 2008; Horn & Schulz, 2011a; Mäkelä et al., 2008; Yan et al., 2017; Zhou et al., 2016), i.e., the GPP estimated from eddy covariance (EC) carbon flux. This method has to be supported by the availability of EC flux towers, was therefore unable to be used globally. To extrapolate parameters to regions without observations, Carvalhais et al parameterized the CASA model based on the inversed parameter vector at the EC site which has the same PFT and the most similar climate and vegetation features to the target region (N Carvalhais et al., 2010). Other studies select to use the average site-level optimized parameters per PFT (Guan, Chen, Shen, Xie, & Tan, 2022; Yuan, Cai, Xia, et al., 2014; Zhou et

al., 2016), PFT-specific optimized parameters (Tian et al., 2020; Zheng et al., 2020), globally optimized parameters (Stocker et al., 2020; Yuan et al., 2019) or globally fixed parameters (Mengoli et al., 2022; H. Wang et al., 2017). Yuan et al showed using six different LUE models that the modeled GPP using globally optimized parameters was not significantly different from that using PFT-specific optimized parameters (Yuan, Cai, Liu, et al., 2014). A study afterwards based on PRELES model, nevertheless, illustrated that at least the variance in parameters across PFT needs to be considered to reach a promise model performance level (Tian et al., 2020). In general, most studies did not account variances of parameters within PFT, but all assumed that the LUE model parameters are related to detailed characteristics of the vegetation and the growing environment.

In some studies, the drivers for spatial changes of model parameters were analyzed based on the site-level calibrated parameters. For example, Horn et al (2011b) found that the parameters of a LUE model, which represent the maximum light use efficiency, the sensitivity of GPP to temperature and soil moisture, varied across climate zones and biomes and can be predicted using vegetation and environmental properties. The relationship between parameters and plant traits also existed in process-based GPP models (Peaucelle et al., 2019). Moreover, Luo et al (2020) claimed that model parameters, which can represent both the evolving ecosystem properties and the unresolved ecosystem processes, should be determined according to our knowledge about the changing ecosystem properties. These studies all confirmed the control of vegetation and environmental attributes on model parameters, which represents GPP sensitivities. These findings inspire the possible next generation of parameterization method based on the physical connection between model parameters and ecosystem properties.

In this study we aim to propose a new LUE model parameterization method (or parameter extrapolation method) and explore the drivers for the model parameters. Our hypothesis is that the PFT-based parameterization cannot perform as well as the site-specific parameterization. We assume that the spatial variations of parameters, indicating the GPP sensitivities to environmental forcings, can be predicted by the ecosystem properties. To test the hypotheses, we compared 12 different parameterization methods based on a LUE model with appropriate environmental drivers and sensitivity functional forms selected from an ensemble of 5600 LUE models (Bao et al., 2022). These parameterization methods were assessed according to the accuracy of the simulated GPP (GPP_{sim}) across different time scales at the site-level, per PFT, per climate type and globally. The neural-network-based method taking GPP as the fitting target (see the method in section 2.4.6) reach the highest performance and is recommended by our study to be applied at other spatio-temporal scales and for other kinds of vegetation models.

2. Method

2.1 Light use efficiency model

LUE models define GPP as a product of the photosynthetically active radiation (PAR), the fraction of photosynthetically active radiation (FAPAR) and the maximum light use efficiency ($\boldsymbol{\varepsilon_{max}}$), regulated by environmental sensitivity functions. The environmental drivers and sensitivity functional forms differ across LUE models. To minimize the effect of the selection of environmental drivers and sensitivity functions, we select a LUE model based on a model evaluation study (Bao et al., 2022) which considers the impacts of temperature (T), vapor pressure deficit (VPD), atmospheric CO_2 concentration (c_a), soil moisture (W), light intensity (L) and cloudiness index (CI) on GPP dynamics (see equations 1-8).

$$GPP=PAR\cdot FAPAR\cdot\boldsymbol{\varepsilon_{max}}\cdot f\mathrm{T}\cdot f\mathrm{VPD}\cdot f\mathrm{W}\cdot f\mathrm{L}\cdot f\mathrm{CI} \quad (1)$$

$$f\mathrm{T}=\frac{2e^{-\left(T_f-\mathbf{T_{opt}}\right)/\mathbf{k_T}}}{1+e^{\left(-\left(T_f-\mathbf{T_{opt}}\right)/\mathbf{k_T}\right)^2}} \quad (2)$$

$$f\mathrm{VPD}=e^{\boldsymbol{\kappa}\left(\frac{\mathbf{C_{a0}}}{c_a}\right)^{\mathbf{C_\kappa}}\mathrm{VPD}}\left(1+\frac{c_a-\mathbf{C_{a0}}}{c_a-\mathbf{C_{a0}}+\mathbf{C_m}}\right) \quad (3)$$

$$f\mathrm{W}=\frac{1}{1+e^{\mathbf{k_W}\left(\mathbf{k_W}-\mathbf{W_I}\right)}} \quad (4)$$

$$f\mathrm{L}=\frac{1}{1+\gamma\cdot\mathrm{APAR}} \quad (5)$$

$$f\mathrm{CI}=\mathrm{CI}^{\boldsymbol{\mu}} \quad (6)$$

$$T_f(t)=(1-\boldsymbol{\alpha_T})\cdot T(t)+\boldsymbol{\alpha_T}\cdot T_f(t-1) \quad (7)$$

$$T_f(t)=(1-\boldsymbol{\alpha_W})\cdot W(t)+\boldsymbol{\alpha_W}\cdot W_f(t-1) \quad (8)$$

The LUE model includes thirteen parameters (in **bold**) in total. All sensitivity functions (*f*T, *f*VPD, *f*W, *f*L and *f*CI) are scaled from zero to one, representing from strong to no constraints. The physical meanings and units of the parameters and references of these sensitivity functions are summarized in Table 1. The sensitivity function of VPD, *f*VPD, includes the effect of both VPD and c_a, which jointly control the leaf internal CO_2 concentration. The pure CO_2 fertilization effect is described only by the right part of *f*VPD (i.e., the summary of one and c_a function). The product of PAR and FAPAR, the absorbed photosynthetically active radiation (APAR), is the estimate of the light energy intercepted by the vegetation canopy, thus was used as the light intensity input of the sensitivity function of L (*f*L). The T and W were temporally filtered using the lag parameters, $\boldsymbol{\alpha_T}$ and $\boldsymbol{\alpha_W}$, at boreal and arid regions, respectively.

Table 1. List of LUE model parameters

Parameters	Meanings	Range	Units	References
$\boldsymbol{\varepsilon_{max}}$	Maximum light use efficiency	0 - 10	$gC\cdot MJ^{-1}$	(Running et al., 2004)
$\mathbf{T_{opt}}$	Optimal temperature	5 - 35	°C	(Horn & Schulz, 2011a)

$\mathbf{k_T}$	Sensitivity to temperature changes	1 - 20	-	As above
$\boldsymbol{\kappa}$	Sensitivity to vapor pressure deficit changes	-10^{-1} - -10^{-4}	kPa^{-1}	(Mäkelä et al., 2008)
$\mathbf{C_{a0}}$	Minimum optimal atmospheric CO_2 concentration	340 – 390	ppm	(Kalliokoski, Makela, Fronzek, Minunno, & Peltoniemi, 2018)
$\mathbf{C_\kappa}$	Sensitivity to atmospheric CO_2 concentration changes	0 – 10	-	As above
$\mathbf{C_m}$	CO_2 fertilization intensity indicator	100 – 4000	ppm	As above
$\mathbf{k_W}$	Sensitivity to soil moisture changes	-30 - -5	-	(Horn & Schulz, 2011a)
$\mathbf{W_I}$	Optimal soil moisture	0.01 – 0.99	$cm \cdot cm^{-1}$	As above
$\boldsymbol{\gamma}$	Light saturation curvature indicator	0 - 1	$MJ^{-1} \cdot m^2 \cdot d$	(Mäkelä et al., 2008)
$\boldsymbol{\mu}$	Sensitivity to cloudiness index changes	10^{-3} - 1	-	(Bao et al., 2022)
$\boldsymbol{\alpha_T}$	Lag parameter for temperature effect	0.0 - 0.9	-	(Horn & Schulz, 2011a)
$\boldsymbol{\alpha_W}$	Lag parameter for soil moisture effect	0.0 - 0.9	-	As above

2.2 Light use efficiency model forcings and calibrated parameters

The forcing data for the LUE model were collected at 196 EC sites (listed in Table S1) from FLUXNET (www.fluxnet.org). The detailed sources and algorithms of the forcing data are summarized in Table S2. The GPP estimated from the observed net ecosystem exchange (NEE) at EC sites (GPP_{obs}) were also collected to calibrate parameters (see the details about parameter calibration in section S1) and evaluate parameterization methods. We considered the simulated GPP using calibrated parameters (GPP_{calib}) as a reference for its good fitness to GPP_{obs}. As a derivative-free global searching algorithm, CMAES (Hansen & Kern, 2004), was used to calibrate parameters in its physical range according to the full time series of GPP_{obs}, we assumed that GPP_{calib} can reach the model potential (i.e., highest model performance).

2.3 Input features for predicting parameters

To extrapolate the parameters to the global scale, we collect mainly the variables that can represent the ecosystem properties available at both local (i.e., site-level) and global scales. These variables include the PFT, climate classification types, nineteen bioclimatic variables (BIO1-19), two aridity features (AI1-2), eleven vegetation features (VIF1-11), atmospheric Nitrogen and Phosphorus deposition ($Ndep_{NHX}$, $Ndep_{NOY}$, Pdep) and seventeen soil properties (Table 2).

Table 2. List of the input features for predicting parameters

Class name	Short names	Definitions	References
PFT	PFT	Plant functional types	See Table S1, eleven types in total
Clim	Clim	Koeppen-Geiger climate classification types	See Table S1, five main climate types and fourteen specific classification types in total
BioClim	BIO1	Annual Mean Temperature	Calculated based on the ANUCLIM algorithm (Xu & Hutchinson, 2011) using CRUNCEP dataset (Viovy, 2018) from 1986-2015.
	BIO2	Mean Diurnal Range (Mean of monthly maximum temperature minus minimum temperature)	
	BIO3	Isothermality (BIO2 divided by BIO7 and 100)	
	BIO4	Temperature Seasonality (standard deviation of temperature multiply with 100)	
	BIO5	Max Temperature of Warmest Month	
	BIO6	Min Temperature of Coldest Month	
	BIO7	Temperature Annual Range (BIO5 minus BIO6)	
	BIO8	Mean Temperature of Wettest Quarter	
	BIO9	Mean Temperature of Driest Quarter	
	BIO10	Mean Temperature of Warmest Quarter	
	BIO11	Mean Temperature of Coldest Quarter	
	BIO12	Annual Precipitation	
	BIO13	Precipitation of Wettest Month	
	BIO14	Precipitation of Driest Month	
	BIO15	Precipitation Seasonality (Coefficient of Variation)	
	BIO16	Precipitation of Wettest Quarter	
	BIO17	Precipitation of Driest Quarter	
	BIO18	Precipitation of Warmest Quarter	
	BIO19	Precipitation of Coldest Quarter	
	AI1	Mean annual aridity index (ratio between mean annual precipitation and potential evapotranspiration)	Calculated using the CRUNCEP dataset from 1986-2015
	AI2	Seasonality of aridity index (standard deviation of mean monthly aridity index)	
VIF	VIF1	Annual mean EVI (enhanced vegetation index)	Calculated based on the bioclimatic variables (BIO1-BIO11) algorithm using the gap-filled Landsat-based EVI (Walther et al., 2022) from 1986-2015
	VIF2	Mean monthly EVI range	
	VIF3	Mean EVI variability (VIF2 divided by VIF7)	
	VIF4	EVI seasonality (standard deviation of EVI)	
	VIF5	Max EVI of Warmest Month	
	VIF6	Min EVI of Coldest Month	

	VIF7	Annual EVI Range (BIO5 minus BIO6)	
	VIF8	Mean EVI of Wettest Quarter	
	VIF9	Mean EVI of Driest Quarter	
	VIF10	Mean EVI of Warmest Quarter	
	VIF11	Mean EVI of Coldest Quarter	
NPdep	$Ndep_{NHX}$	Average atmospheric nitrogen deposition (NH_3 and NH_4)	Extracted from the product of the atmospheric chemistry transport model TM3 (R. Wang et al., 2017)
	$Ndep_{NOY}$	Average atmospheric nitrogen deposition (NO and NO_2)	
	Pdep	Average atmospheric phosphorus deposition	
Soil	BDRICM	Depth to bedrock (R horizon) up to 200 cm	Extracted from the Soil Grids product (de Sousa et al., 2020)
	BDRLOG	Probability of occurrence (0-100%) of R horizon	
	BDTICM	Absolute depth to bedrock (in cm)	
	BLDFIE	Bulk density (fine earth) in kg/m3 at depth 0.00 m	
	CECSOL	Cation exchange capacity of soil in cmol/kg at depth 0.00 m	
	CLYPPT	Clay content (0-2 micro meter) mass fraction in % at depth 0.00 m	
	CRFVOL	Coarse fragments volumetric in % at depth 0.00 m	
	ORCDRC	Soil organic carbon content (fine earth fraction) in g/kg at depth 0.00 m	
	PHIHOX	Soil pH*10 in H2O at depth 0.00 m	
	PHIKCL	Soil PH (mulity with 10) in KCl at depth 0.00 m	
	SLTPPT	Silt content (2-50 micro meter) mass fraction in % at depth 0.00 m	
	SNDPPT	Sand content (50-2000 micro meter) mass fraction in % at depth 0.00 m	
	AWCh1	Derived available soil water capacity (volumetric fraction) with FC = pF 2.0 for depth 0 cm	
	AWCh2	Derived available soil water capacity (volumetric fraction) with FC = pF 2.3 for depth 0 cm	
	AWCh3	Derived available soil water capacity (volumetric fraction) with FC = pF 2.5 for depth 0 cm	
	WWP	Derived available soil water capacity (volumetric fraction) until wilting point for depth 0 cm	
	AWCtS	Saturated water content (volumetric fraction) teta-S for depth 0 cm	

The categorical variables (PFT and climate types) were converted to one or zero to indicate whether the target location belongs to a specific type or not. All non-categorical variables were normalized by subtracting the mean of each feature and a division by the standard deviation (equation 9). The normalized feature and original features are represented by *var* and var_{nor}. The mean and standard deviation per feature are represented by *mean* and *std*.

$$var_{nor} = \frac{var\text{-}mean}{std} \quad (9)$$

2.4 Parameterization methods

We extrapolate the parameters based on N-fold cross-validation strategy using the collected ecosystem property variables. In other word, the samples, here refers to the EC sites, were divided into ten groups randomly (see the group number of each site in Table S1). We trained every time the parameterization models using nine of ten groups and validate the result using the left one group, and repeated ten times until getting validated results of all sites. All PFT and climate classification types (eleven PFT and fourteen climate classification types in total, see Table S1) were included in each training dataset.

The twelve parameterization methods can be divided into six groups (see details in section 2.4.1-2.4.5).

2.4.1 Arithmetic methods ('PFT_{mean}', '$Clim_{mean}$', 'PFT_{med}', and '$Clim_{med}$')

In the regions without observational data, the parameters were decided by the arithmetic mean of the calibrated parameters at the sites with the same PFT (Guan et al., 2022; Yuan, Cai, Xia, et al., 2014; Zhou et al., 2016). Here we tested the methods of using the mean and median parameters per PFT and climate type.

'PFT_{mean}': the mean of the calibrated parameter vectors per PFT;

'$Clim_{mean}$': the mean of the calibrated parameter vectors per main climate type;

'PFT_{med}': the median of the calibrated parameter vectors per PFT;

'$Clim_{med}$': the median of the calibrated parameter vectors per main climate type.

2.4.2 Similarity-based method ('PFT_{sim}')

The site similarity is defined by Carvalhais et al.(2010) which measures the similarity (D) of the ecosystem properties between site *i* and site *j* as equation 10:

$$D_{i,j} = 1 - \frac{\sum_{n=1}^{N} (\boldsymbol{V}_{i,n} - \boldsymbol{V}_{j,n})^2}{\sum_{n=1}^{N} (\boldsymbol{V}_{i,n} - \bar{\boldsymbol{V}}_i)^2} \quad (10)$$

Here, $\boldsymbol{V}$ is a vector including the normalized daily mean of the air temperature, precipitation (in logarithm), global radiation and LANDSAT-based normalized difference vegetation index (NDVI, see data source and processing method in Table S2) between 1986 and 2015.

To determine the parameters of a target location, we calculated the D to each training site within the same PFT as the target location. The parameter vector at the site with the maximum D was used.

'PFT_{sim}': parameter vectors for each site from the most similar site.

2.4.3 Optimization-based methods ('OPT-All' and 'OPT-PFT')

The parameters can be optimized across all sites or at sites per PFT (Yuan, Cai, Liu, et al., 2014). Here we adopted the same algorithm, CMAES, and the same cost functions as the site-specific calibration method (see section S1).

'OPT-All': a parameter vector optimized using all sites in the training dataset.

'OPT-PFT': parameter vectors per PFT optimized using the sites within the same PFT in the training dataset.

2.4.4 Regression-based methods ('sRF', 'mRF', 'mNN-Par')

To test the assumption that the calibrated parameters are determined by the ecosystem properties, we here predict the calibrated parameters using the normalized features based on different regression methods.

'sRF': parameter vectors per site of which each parameter was predicted sequentially based on the single-output random forest (trees number=100; Breiman, 2001).

'mRF': parameter vectors per site predicted simultaneously based on the multi-output random forest (trees number=100; Pedregosa et al., 2011).

'mNN-Par': parameter vectors per site predicted simultaneously based on the multi-layer perceptron neural network (hidden layers number=2, neurons number=16; Gardner & Dorling, 1998; McCulloch & Pitts, 1943).

2.4.5 GPP-targeting method ('mNN-GPP')

Due to the model equifinality problem, which means different parameter vectors might result in the same model performance, the calibrated parameters might not represent the true parameters which reflect the GPP sensitivities controlled by the environmental properties. Here we additionally test the assumption that the ecosystem-properties-predicted parameter vector which might differ from the calibrated parameter vector can simulate GPP with a good accuracy. Instead of directly predicting parameters, we applied the neural network to predict GPP based on the LUE model using parameters predicted by the input features. The flowchart of this method is as Figure 1.

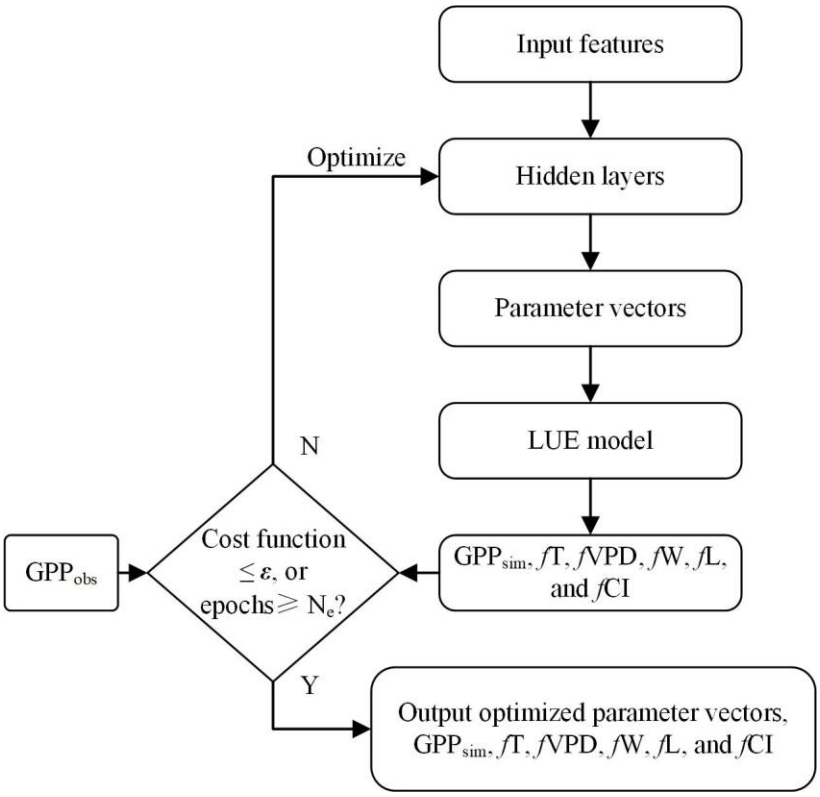


Figure 1. Flowchart of the GPP-targeting method. The parameter vectors per site are optimized until the cost function (see the definition in section S2) is lower than the threshold (ε=10^{-2}) or epochs is more than the maximum epochs (=2×10^3).

At the first step, the neural network predicts the parameter vectors based on the normalized features. The GPP_{sim} is then simulated using the predicted parameter vectors and compared with GPP_{obs} to measure the model error (see the definition in section S2). The neural network back-propagates the error to each hidden layer and optimizes the weight and bias of each neuron based on ADAM algorithm (Kingma & Ba, 2014). We repeated the optimization process until the epochs reach 2×10^3. To overcome the overfitting problem, we set the learning rate (=10^{-3}), L_2 regularization coefficient (=10^{-4}), mini batch size (=32), neurons per layer (=16) and hidden layers (=2) according to a grid-searching experiment (not shown here). We further applied the drop-out strategy (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014) on the input and hidden layers. The outputs of the network include the simulated GPP, sensitivity functions and predicted parameter vectors per site.

'mNN-GPP': parameter vectors per site predicted using the multi-layer perceptron neural network on the target of minimizing GPP errors.

2.4.6 Globally-fixed method ('P-model')

P-model (Stocker et al., 2020; H. Wang et al., 2017) derived based on Farquhar et al (1980) and Fick's law together with an optimality theory (Prentice, Dong, Gleason, Maire, & Wright, 2014) adopts a globally-fixed parameter vector upscaled from the leaf-scale process. Mengoli et al (2022) improved the model by adding an

acclimation process for the photosynthetic parameters. Here we ran Mengoli model based on the daily data, which is the same as the inputs for the other parameterization methods (section 2.2), using the initial parameters given in the paper and compared the model outputs, GPP, with other methods.

'P-model': a globally-fixed parameter vector from paper.

2.5 Statistical analysis for parameterized results

All the parameterization methods were assessed according to the GPP accuracy measured by Nash-Sutcliffe model efficiency (NSE, -∞-1; NSE=1 indicates a perfect model), determination coefficient (R^2, 0-1; R^2=1 indicates a perfect model) and normalized root mean squared error (NRMSE, 0-∞; NRMSE=0 indicates a perfect model) which is equal to the root mean squared error divided by the mean observational variable. Only good-quality data were used to calculate NSE, R^2 and NRMSE. Here the good-quality data refers to the input vector that the relevant quality flags (see 'QA' in Table S2) of all forcing data, GPP_{obs}, represented by the quality of NEE, are higher than 0.8 at the daily scale. When aggregated to longer time scales, the good quality data means the average quality flags are all higher than 0.7 at the weekly and monthly scales, and 0.5 at the yearly scale. Besides, the predicted parameters were compared to the calibrated parameters to test if the model equifinality problem exists.

2.5.1 Site-level temporal GPP assessment

We forced the LUE model at the daily scale and got the daily GPP_{sim} as a result. The weekly, monthly and yearly GPP_{sim} and GPP_{obs} were calculated based on the mean daily GPP_{sim} and GPP_{obs}, respectively. These time series of site-level GPP at different time scales were evaluated using NSE, R^2 and NRMSE. The vectors of NSE, R^2 and NRMSE were compared per PFT and climate types.

2.5.2 Spatial variability of GPP assessment

The site-mean GPP_{obs} across sites represent the spatial variance of GPP. We used NSE, R^2 and NRMSE to measure the accuracy of the site-mean GPP_{sim} compared with GPP_{obs} to evaluate the ability of the parameterization methods to capture the spatial variability of GPP.

2.5.3 Comprehensive assessment across spatio-temporal scales based on model likelihood

The likelihood of each parameterization method, ***P***, was calculated according to Bao et al (2022). To avoid selecting a method falling shortly in locally describing ecosystem GPP, ***P*** represents an overall performance at 200 different site groups. In every group, 100 sites were selected randomly from all sites and two site-years were then randomly extracted from each of these 100 sites. The 200 site-years GPP_{sim} were compared to GPP_{obs} based on NSE, R^2 and NRMSE at each site-year. The differences between the daily, weekly, and monthly NSE, R^2 and NRSE vectors (with 200 elements) and the yearly NRMSE vectors per parameterization method were tested using Kolmogorov-Smirnov statistical and t tests. The method with statistically higher NSE, R^2 or lower NRNSE than others has the largest score (=1, otherwise =0) at each site group. In case that two or more methods were

statistically equal and better than others, the NSE, R^2 or NRMSE across all site-years was additionally computed to sort the methods independently. ***P*** is equal to the average score across all site groups. The average ***P*** across different statistical metrics (NSE, R^2 and NRMSE) and time scales (daily, weekly, monthly and yearly) was used to detect the best parameterization method.

2.5.4 Comparison between predicted parameters and calibrated parameters

Since the thirteen parameters have different meanings and ranges, they were compared independently. The similarity between the predicted parameters using methods introduced in section 2.4 and the calibrated parameters based on the observational data (see section S1) was assessed using NSE, R^2 and NRMSE.

2.6 Feature importance estimation

We evaluated the importance of input features using three methods and select the most importance features based on the average normalized feature importance values.

2.6.1 Shapley-based feature importance (SHAP)

The Shapley value of a feature is calculated based on the deviation of the predicted parameter at a certain input from the average prediction (Lundberg & Lee, 2017), which represent the contribution of a feature to the output. Here SHAP is equal to the average absolute Shapley value across all inputs. The average SHAP across all cross-validation groups(Friedman, 2001) was used to assess the contribution of features for each parameter and all parameters.

2.6.2 Layer-wise-relevance-propagation-based feature importance (LRP)

The layer-wise relevance propagation refers to a strategy which allows to decompose the prediction of neural network over an input feature (Montavon, Binder, Lapuschkin, Samek, & Müller, 2019). It is usually used in deep classification neural network, here we applied LRP to assess a shallow regression neural network yet. We calculated the relevance vector according to Bach et al. (2015) and measured the feature importance according to the average relevance across different cross-validation groups.

2.6.3 Partial-dependence-based feature importance (PD)

We estimated the partial dependence of the prediction to each input feature based on Friedman's (2001) algorithm. The PD was measured according to the partial dependence, which is equal to the standard deviation of the partial dependence if the input feature is non-categorical variables, otherwise is equal to the one fourth of the absolute partial dependence range (Greenwell, Boehmke, & McCarthy, 2018).

3. Results

3.1 Temporal and spatial assessment

The parameterization method based on neural network aiming at minimizing GPP errors, mNN-GPP, had the best performance compared with other typical parameterization methods. All the assessing metrics at daily, weekly, monthly and yearly scales, NSE (Figure 2), R^2 (Figure S1) and NRMSE (Figure S2), showed that mNN-GPP was better at more sites (i.e., more bright color blocks in Figure 2). The spatial variability of GPP can be also better captured by mNN-GPP, which had higher NSE, R^2 and lower NRMSE measured by site-mean GPP_{obs} and GPP_{sim} (Figure 3). The accuracy of time series and site-mean GPP_{sim} using other methods were all significantly worse than mNN-GPP. Although mNN-GPP cannot perform as well as the site-specific calibration, it is the best parameter extrapolation method globally.

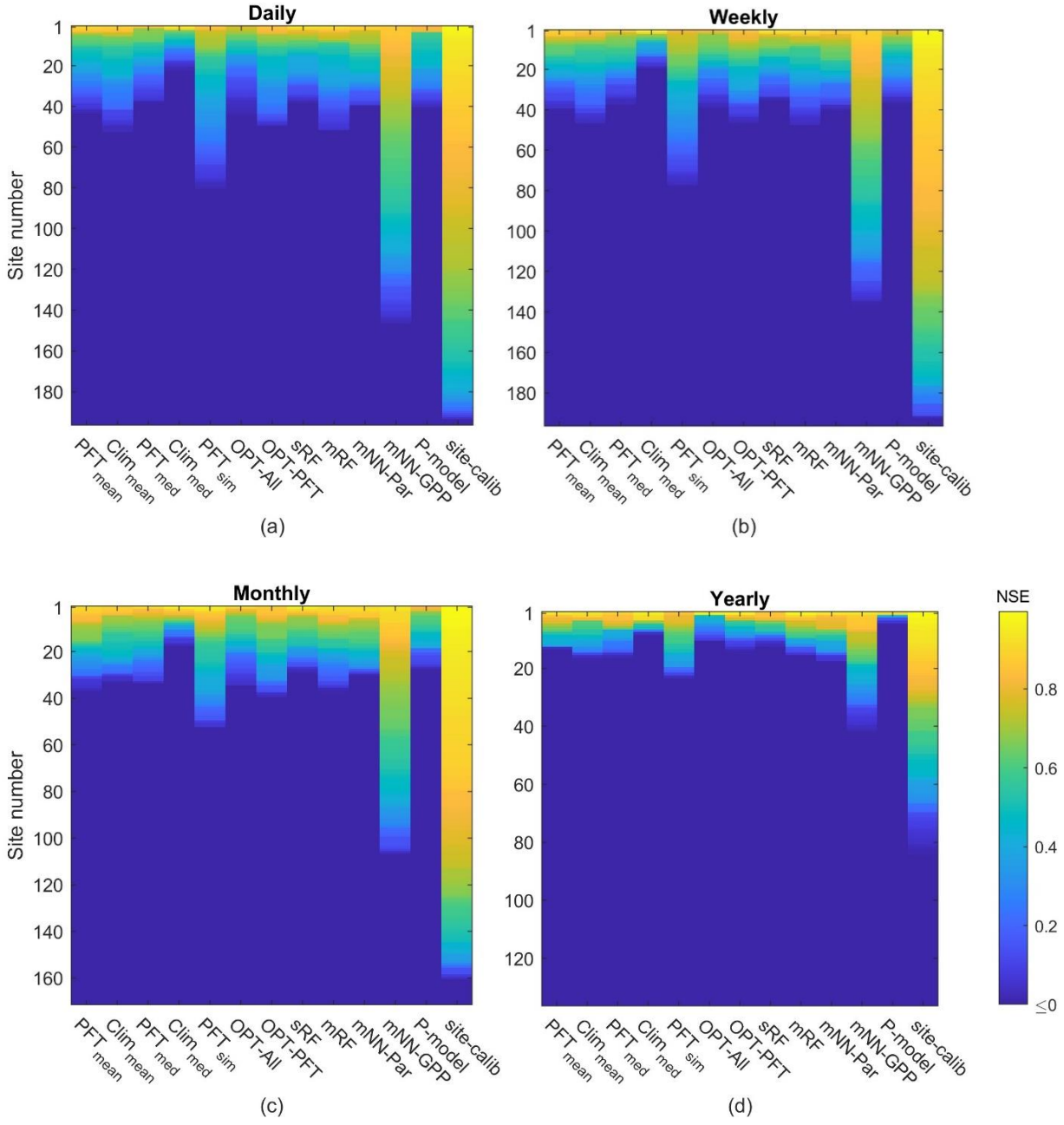


Figure 2. Comparison of NSE between GPP_{obs} and GPP_{sim} based on twelve different parameterization methods (see definitions of PFT_{mean}, $Clim_{mean}$, PFT_{med}, $Clim_{med}$, PFT_{sim}, OPT-All, OPT-PFT, sRF, mRF, mNN-Par, mNN-GPP, and P-model in section 2.4), and between GPP_{obs} and GPP_{calib} (site-calib, see the calibration process in section S1) at daily (a), weekly (b), monthly (c) and yearly (d) scales. The sites with less than four good-quality (defined in section 2.5) months or years were removed from penal c and d, respectively.

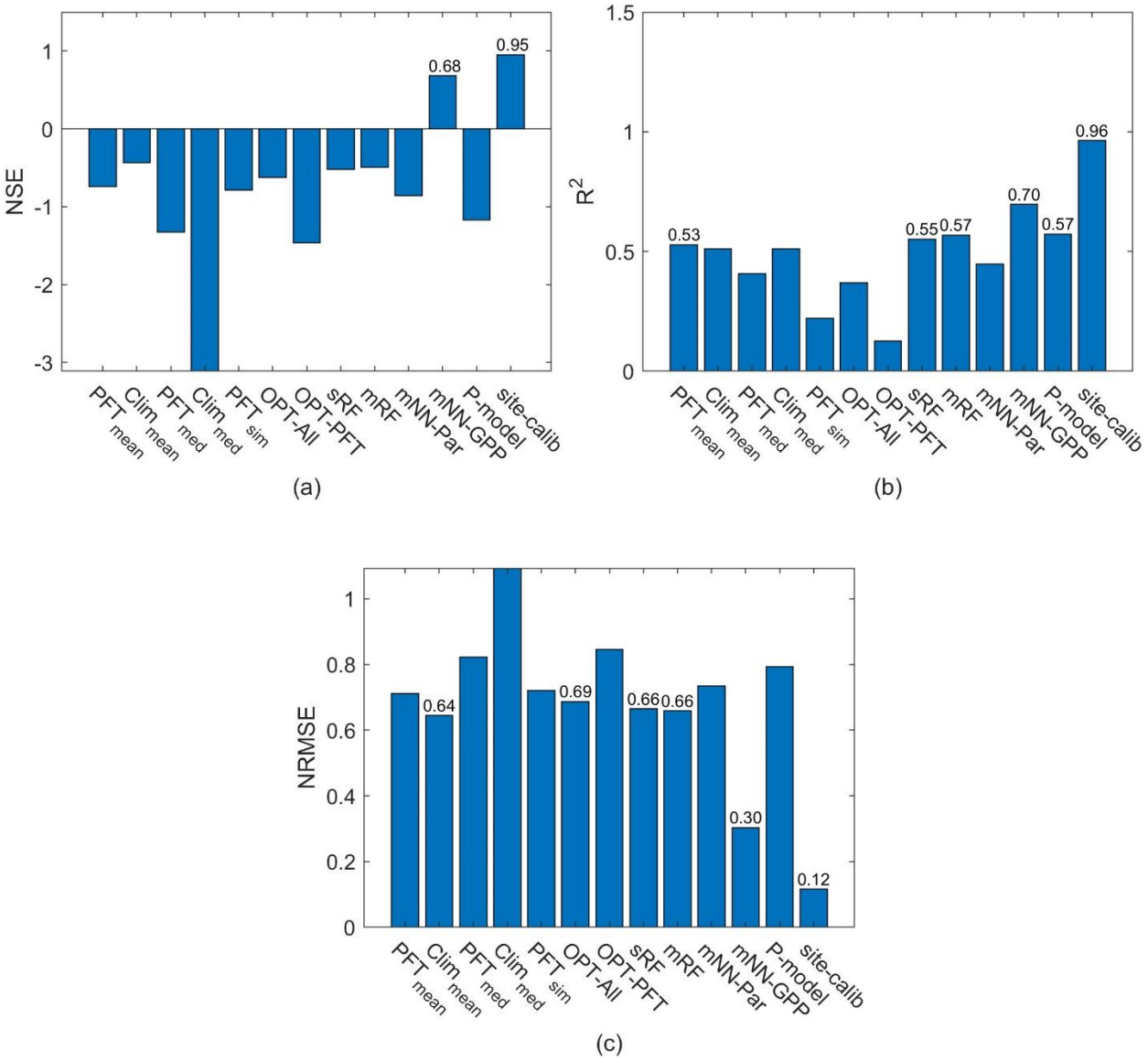


Figure 3. Comparison of NSE (a), R^2 (b), and NRMSE (c) between the site-mean GPP_{obs} and GPP_{sim}. The positive NSE, highest six R^2 and lowest six NRMSE values are displayed on the top of bars.

The global best parameterization method, mNN-GPP, outperformed across various PFT and climate types. It had the highest daily NSE quantiles for each PFT and climate type considered in this study (Figure 4). While mNN-GPP was relative better than other methods, no extrapolated parameters can provide accurate GPP dynamics (NSE>0.4) at closed shrubland (CSH in Figure 4a), tropical (A in Figure 4b) and polar (E in Figure

4b) climate types given that the model using calibrated parameters was good (grey colors in Figure 4). It demonstrated that the variance of current extrapolated parameters was still insufficient. Using R^2 or NRMSE as the assessing metric (Figure S3-4), the parameterization methods showed smaller but robust relative differences, i.e., the mNN-GPP was still the best method. In general, no extrapolated parameters can reach the highest model performance, mNN-GPP was the best option at areas without observational data.

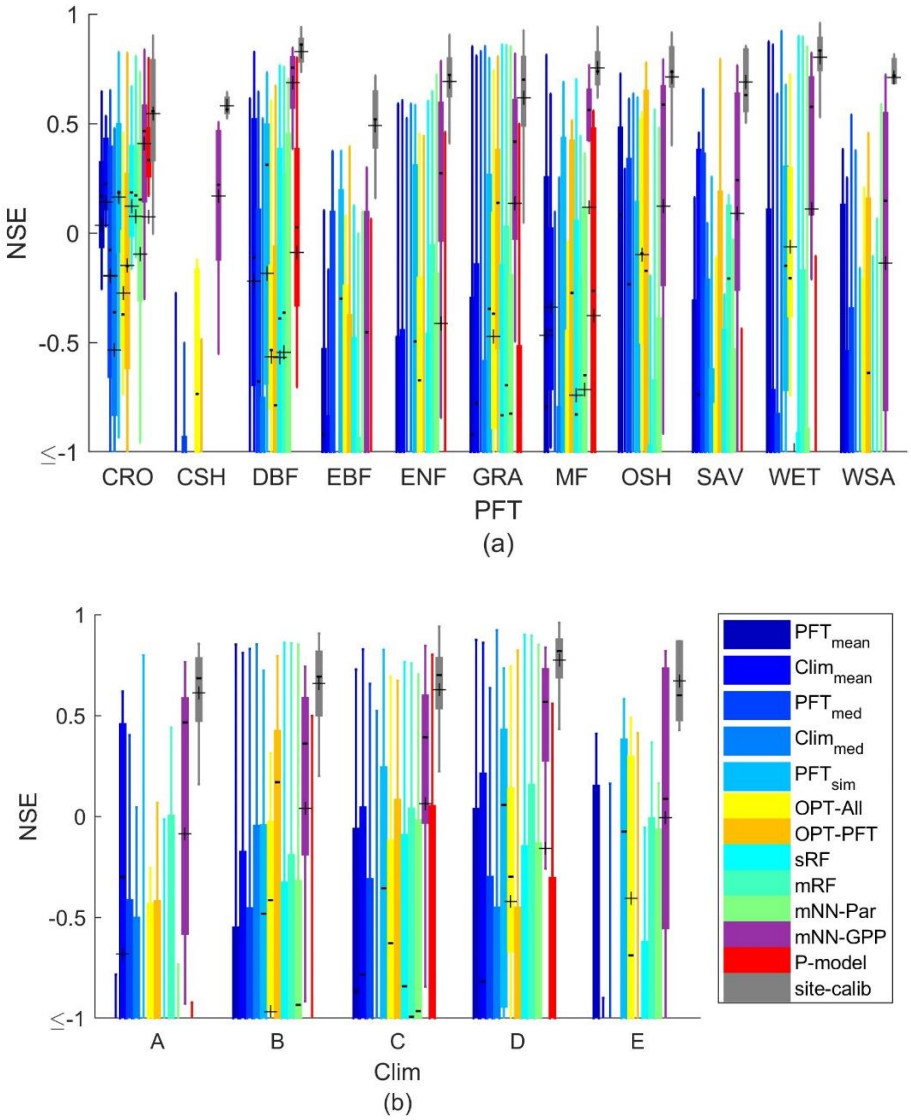


Figure 4. Site-level daily NSE comparison per plant functional type (a, PFT) and climate type (b, Clim). The mean and median per type are represented by the black cross and line, respectively. CRO=crop, CSH=closed shrubland, DBF=deciduous broadleaf forest, EBF=evergreen broadleaf forest, ENF=evergreen needleleaf forest, GRA=grass, MF=mixed forest, OSH=open shrubland, SAV=savanna, WET=wetland, WSA=woody savanna. A=tropic climate, B=arid climate, C=temperate climate, D=cold climate, E=polar climate

The model likelihood, ***P***, which represents the likelihood of a model statistically better than others across various site groups, illustrated that mNN-GPP was the best method to extrapolate parameters, following by OPT-All and $Clim_{med}$ with likelihoods lower than 0.06 (i.e., at less than 6% groups of sites the two methods can outperform). The average ***P*** of mNN-GPP across daily, weekly, monthly, and yearly scales, and across various assessing metrics was also significantly higher than the other methods. It represented that the method is robust across various temporal and spatial scales.

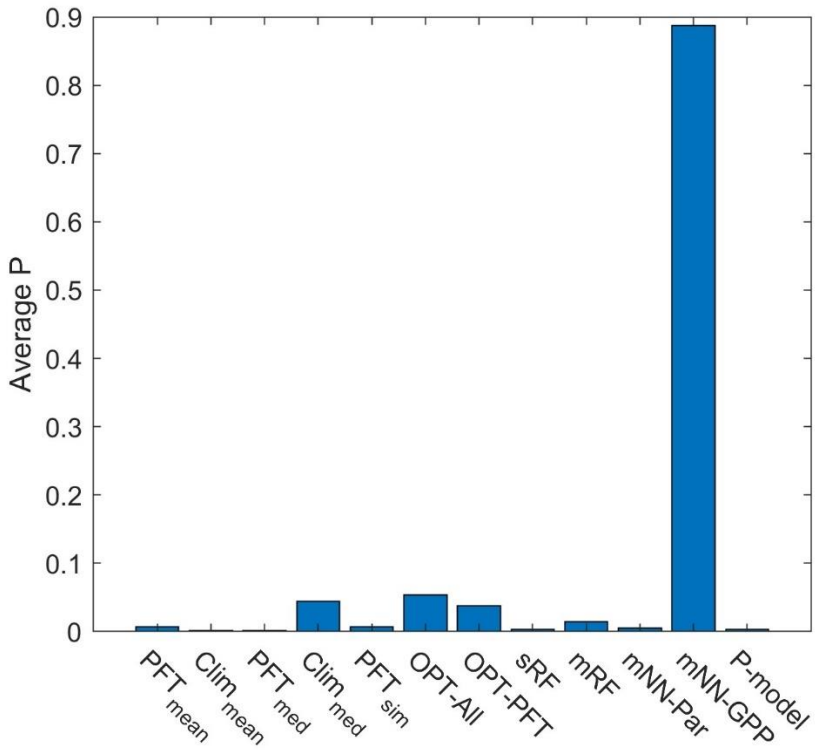


Figure 5. The average model likelihood (*P*) of parameterization methods

3.2 Difference between calibrated parameters and predicted parameters

The predicted parameters displayed different distribution pattern from the calibrated parameters. Taking the best method, mNN-GPP, as an example (Figure 6), the ranges of the predicted parameters were narrower than the calibrated parameters given the same predefined range. Further, the predicted parameters had no 'edge-hitting' problem, which means that the parameter frequently reaches its maximum or minimum values, e.g., the calibrated parameters T_{opt}, k_T, C_κ, C_{a0}, C_m and k_W (Figure 6b-c, f-h, j). The other parameterization methods also showed narrower ranges but no edge-hitting (e.g., mRF, $Clim_{med}$ and OPT-PFT in Figure S5-7). NSE between the predicted parameters using mNN-GPP and calibrated parameters across sites were all negative. The maximum R^2 was 0.08 and the lowest NRMSE was 0.08. Thus, the predicted parameters were not comparable to the calibrated parameters while they can produce similar GPP, highlighting the parameter equifinality in the LUE model.

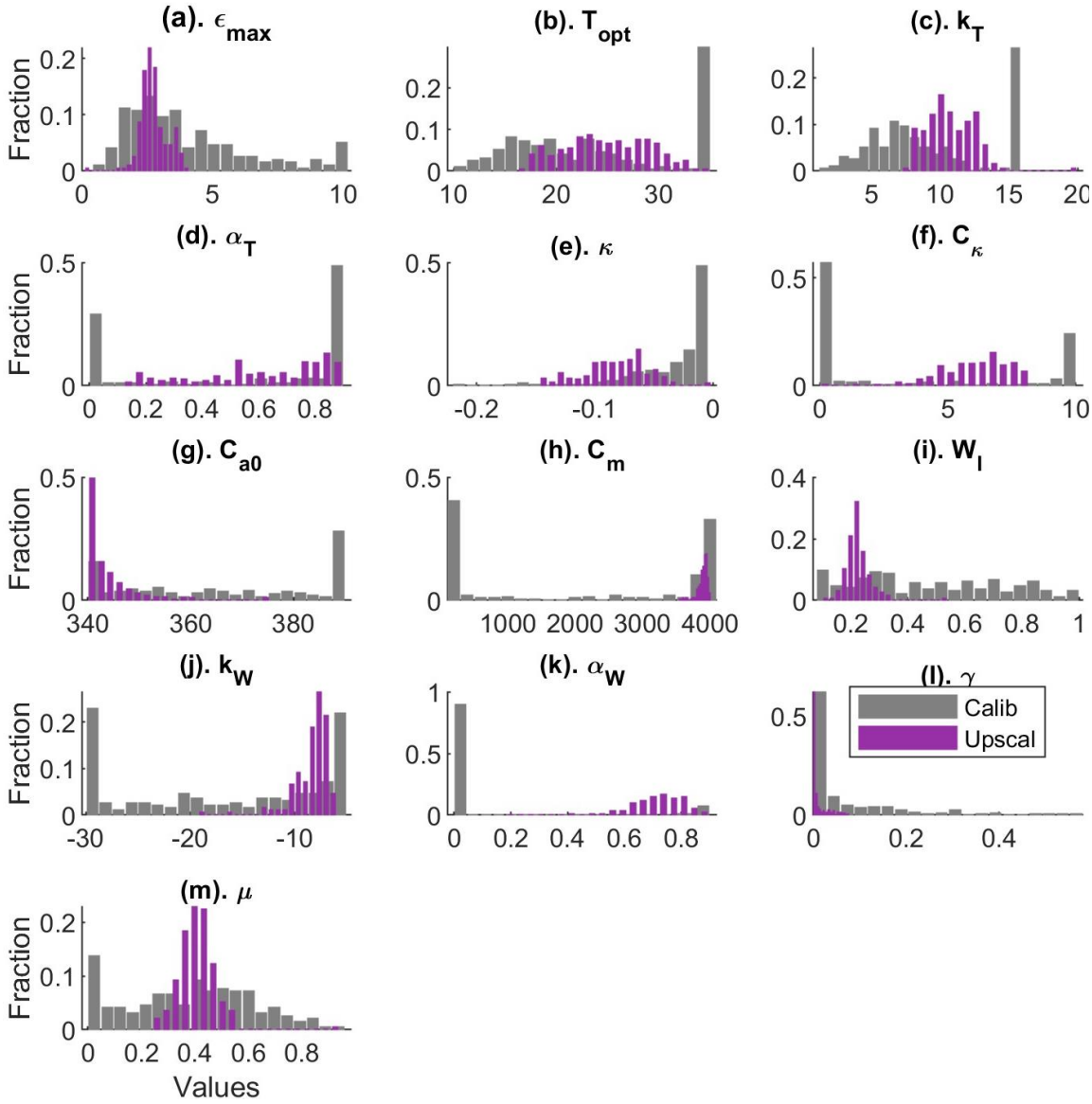
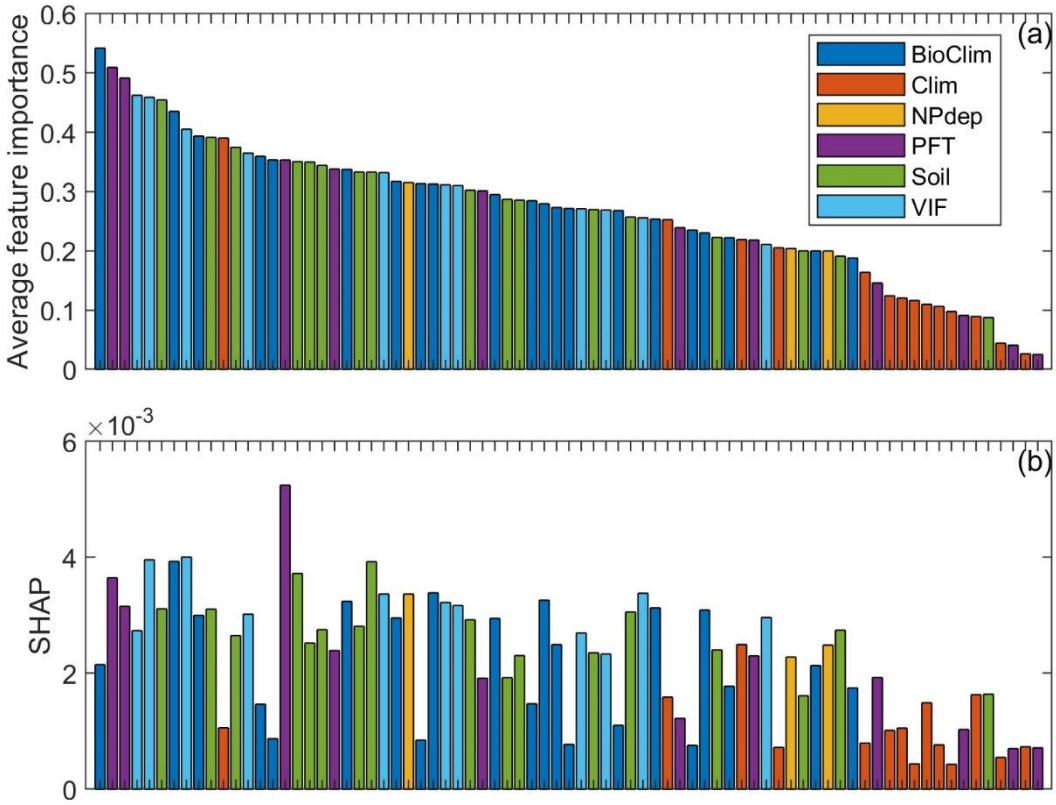


Figure 6. Distribution histogram of the calibrated parameters and the predicted parameters by mNN-GPP.

3.3 Important features for controlling model parameters

The average values of SHAP, LRP and PD illustrated that the bioclimate variables, PFT, and vegetation features were important features controlling spatial variability of parameters (Figure 7). The importance values differed across three different methods, but all of them showed that most climate types were not important for determining model parameters. Most bioclimate variables were shown to have higher importance than other features. The SHAP and LRP of most PFT were higher than other features. For a specific parameter, the most importance feature was not the same as the one for all parameters. For example, ε_{max} was controlled mainly by AI1 (mean annual aridity index, see figure S8) , nonetheless W_I was controlled primarily by VIF7 (the range of mean annual

EVI, see figure S9). In general, the bioclimate, PFT and vegetation features are determining the parameters which represent the GPP sensitivities to environment changes.



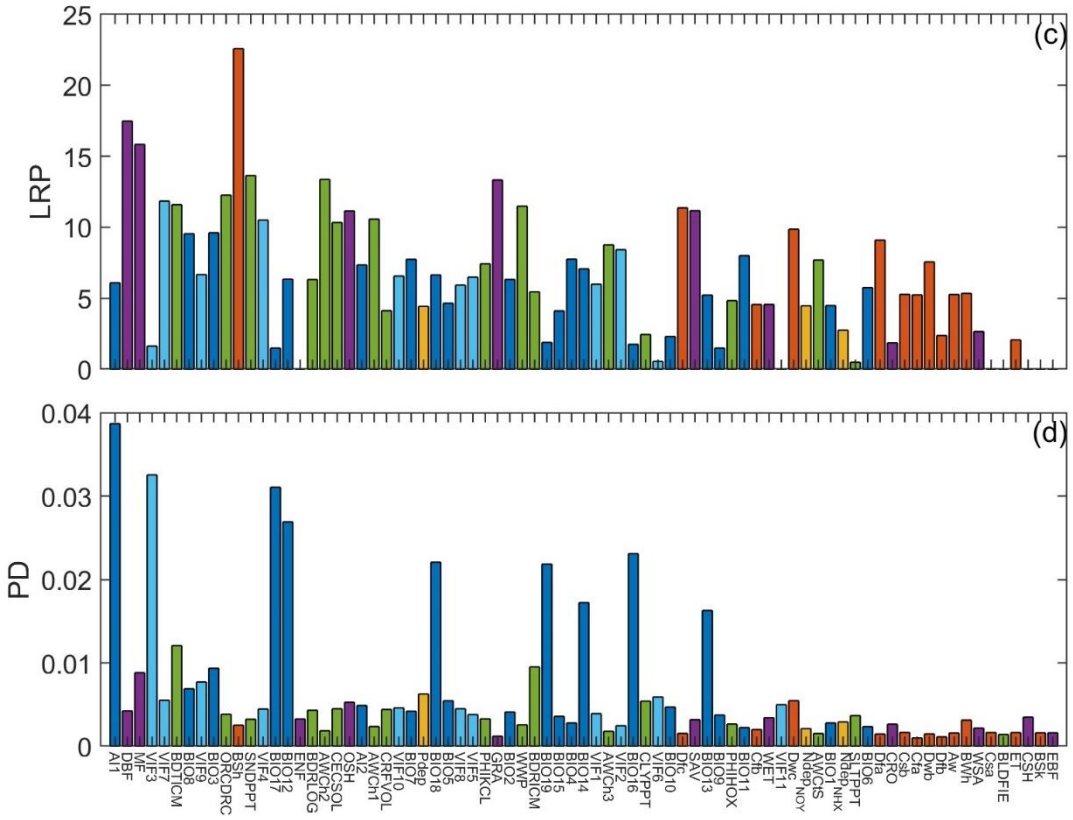


Figure 7. The input features sorted by the average normalized SHAP, LRP and PD for all parameters (a, see definitions in section 2.6). The detailed average SHAP, LRP and PD values for all parameters are displayed in b-d. The features (see definitions in **Table 2**) are classifed into plant functional types (PFT, purple), Koeppen-Geiger climate classification types (Clim, orange), bioclimate variables (BioClim, dark blue), vegetation features (VIF, shallow blue), atmospheric deposition (NPdep, yellow) and soil properties (Soil, green).

4. Discussion

4.1 Well-constrained site-specific parameterization is better than PFT-dependent parameterization

The PFT-based parameterization has been applied for a long time (Running et al., 2004), which was shown cannot capture the variance in parameters within PFT (Bloom et al., 2016) and can introduce the misclassification errors. The method to directly use parameters from papers without local or global evaluation can be also risky. P-model which adopted the globally fixed parameters upscaled from the leaf-scale might not include the PFT errors (Mengoli et al., 2022; Stocker et al., 2020), but had limited accuracy across the temporal (Figure 2) and spatial scales (Figure 3). Actually, our results showed that the globally fixed parameterization methods (e.g., P-model) were worse than PFT-based method (e.g., OPT-PFT, and PFT_{mean}). The globally optimization method (OPT-All) had slightly better performance than PFT-based optimization at the global scale (Figure 5) due to higher spatial generalizability (e.g., Figure 3), the same as Yuan et al (2014), but had accurate prediction at less sites (i.e., less bright blocks in Figure 2). This agrees with a study using PRELES model (Tian et al., 2020),

which demonstrated that globally optimized parameters are not sufficient to reflect the variability of the GPP sensitivities. Luo et al (2020) also confirmed that model parameters should vary with the spatial and temporal changes of ecosystem properties. While the site-specific method (PFT_{sim}, sRF, mRF, and mNN-Par) include wider spatial variability than PFT- and climate-type-based method (PFT_{mean}, $Clim_{mean}$, PFT_{med} and $Clim_{med}$), it did not show robust advantage due to uncertainties remained in the calibrated parameters, which were used to constrain the predicted parameters. However, the site-specific parameterization which considers the GPP prediction error, mNN-GPP, reaches the highest performance, highlighting that the well-constrained site-specific parameterization method can provide more reliable outputs than PFT-based method. This is opposite from the conclusion of Tian et al. (2020) which tested only the site-specific optimization method showing higher uncertainties than PFT-based optimization method.

4.2 Reduce parameter uncertainty by considering the relationship between parameters and ecosystem properties

Our results reveal the equifinality of model parameters, which consequently increases the model uncertainty. While no extrapolated parameter vectors outperformed calibrated ones, the parameter ranges were constrained in all methods based on site-specific input features (e.g., sRF, mRF, mNN-Par and mNN-GPP) compared with calibration and optimization methods (see Figure S5-7). It demonstrated that considering the physical links between GPP sensitivities and the ecosystem properties can reduce the parameter uncertainties. This is also true in other LUE models (Horn & Schulz, 2011b). Furthermore, mNN-GPP, considering only the GPP errors but not the distance to calibrated parameters, avoids inheriting uncertainties from model calibration. In general, the model parameterization relying on ecosystem properties can reduce the parameter uncertainty resulted from model equifinality.

4.3 Drivers of the spatial variability of GPP sensitivities

We assume that GPP sensitivities to environmental changes are reflected by the model parameters. The feature importance assessment showed that bioclimate variables, PFT, and vegetation features are controlling the spatial variability of the GPP sensitivities. This is similar to the results of other independent studies using a different LUE model (Horn & Schulz, 2011b) and terrestrial biosphere models (Peaucelle et al., 2019). However, we here used only the variables that represent the climate, vegetation, atmospheric deposition and soil properties. Other unknown important features such as illumination features, which was shown important in Horn et al's study (2011b), were not tested. This might result in insufficient spatial variance in the predicted parameters. Further, we found that the detected important features differed across measuring methods (i.e., SHAP, LRP and PD) and input features (see the result of using a subset of input features which achieve similar accuracy, Figure S10). Several input features are correlated (Figure S11) but not refined. This additionally increase the difficulty to identify the key features determining variability of GPP sensitivities. Besides, the temporal changes in the

ecosystem properties can affect the parameters (Luo & Schuur, 2020). To extrapolate parameters to longer time scales (e.g., decades), the temporal variation in parameters needs to be considered.

5. Conclusion

In this study we find a method to parameterize a LUE model based on the link between parameters and ecosystem properties and the distance to observed GPP. This method enables the parameter extrapolation in regions without observational data with a significantly higher accuracy than the widely-used PFT-based and globally fixed parameterization methods. This method can reduce the parameter uncertainty by predicting them using the ecosystem properties without reference of calibrated parameters which usually have high uncertainties. bioclimate variables, PFT, and vegetation features are the most important ecosystem properties controlling the spatial variability of LUE model parameters. The temporal variation in the parameters and the relationship between them and ecosystem properties need to be further explored. Since our parameterization method has high robustness across various temporal and spatial scales, we encourage the application to other GPP models and spatio-temporal scales.

Acknowledgement

This paper is funded by the European Space Agency (ESA) via the Earth System Data Lab (ESDL) project.

References

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.-R., & Samek, W. (2015). On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *Plos One, 10*(7), e0130140.

Baldocchi, D., Ryu, Y., & Keenan, T. (2016). Terrestrial carbon cycle variability. *F1000Research, 5*.

Bao, S., Wutzler, T., Koirala, S., Cuntz, M., Ibrom, A., Besnard, S., . . . Weber, U. (2022). Environment-sensitivity functions for gross primary productivity in light use efficiency models. *Agricultural and Forest Meteorology, 312*, 108708.

Bloom, A. A., Exbrayat, J.-F., Van Der Velde, I. R., Feng, L., & Williams, M. (2016). The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times. *Proceedings of the National Academy of Sciences, 113*(5), 1285-1290.

Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5-32.

Carvalhais, N., Reichstein, M., Collatz, G., Mahecha, M., Migliavacca, M., Neigh, C., . . . Seixas, J. (2010). Deciphering the components of regional net ecosystem fluxes following a bottom-up approach for the Iberian Peninsula. *Biogeosciences, 7*(11), 3707-3729.

Carvalhais, N., Reichstein, M., Seixas, J., Collatz, G. J., Pereira, J. S., Berbigier, P., . . . Papale, D. (2008). Implications of the carbon cycle steady state assumption for biogeochemical modeling performance and inverse parameter retrieval. *Global Biogeochemical Cycles, 22*(2).

de Sousa, L. M., Poggio, L., Batjes, N. H., Heuvelink, G. B., Kempen, B., Riberio, E., & Rossiter, D. (2020). SoilGrids 2.0: producing quality-assessed soil information for the globe. *Soil Discuss, 1*(10.5194).

Farquhar, G. D., von Caemmerer, S. v., & Berry, J. A. (1980). A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. *planta, 149*(1), 78-90.

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J. E., . . . Kuze, A. (2011). New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters, 38*(17).

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

Gardner, M. W., & Dorling, S. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric environment, 32*(14-15), 2627-2636.

Greenwell, B. M., Boehmke, B. C., & McCarthy, A. J. (2018). A simple and effective model-based variable importance measure. *arXiv preprint arXiv:1805.04755*.

Groenendijk, M., Dolman, A., Van der Molen, M., Leuning, R., Arneth, A., Delpierre, N., . . . Verbeeck, H. (2011). Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data. *Agricultural and Forest Meteorology, 151*(1), 22-38.

Guan, X., Chen, J. M., Shen, H., Xie, X., & Tan, J. (2022). Comparison of big-leaf and two-leaf light use efficiency models for GPP simulation after considering a radiation scalar. *Agricultural and Forest Meteorology, 313*, 108761.

Hansen, N., & Kern, S. (2004). *Evaluating the CMA evolution strategy on multimodal test functions.* Paper presented at the International conference on parallel problem solving from nature.

He, M., Ju, W., Zhou, Y., Chen, J., He, H., Wang, S., . . . Li, Y. (2013). Development of a two-leaf light use efficiency model for improving the calculation of terrestrial gross primary productivity. *Agricultural and Forest Meteorology, 173*, 28-39.

Horn, J., & Schulz, K. (2011a). Identification of a general light use efficiency model for gross primary production. *Biogeosciences, 8*(4), 999-1021.

Horn, J., & Schulz, K. (2011b). Spatial extrapolation of light use efficiency model parameters to predict gross primary production. *Journal of Advances in Modeling Earth Systems, 3*(4).

Huntzinger, D. N., Michalak, A. M., Schwalm, C., Ciais, P., King, A. W., Fang, Y., . . . Zhao, F. (2017). Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions. *Scientific Reports, 7*(1), 4765. Retrieved from https://doi.org/10.1038/s41598-017-03818-2. doi:10.1038/s41598-017-03818-2

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., . . . Chen, J. (2011). Global patterns of land‐atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research: Biogeosciences, 116*(G3).

Kalliokoski, T., Makela, A., Fronzek, S., Minunno, F., & Peltoniemi, M. (2018). Decomposing sources of uncertainty in climate change projections of boreal forest primary production. *Agricultural and Forest Meteorology, 262*, 192-205. Retrieved from <Go to ISI>://WOS:000444659300018. doi:10.1016/j.agrformet.2018.06.030

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Lundberg, S. M., & Lee, S.-I. (2017, 2017

Dec 04-09). *A Unified Approach to Interpreting Model Predictions.* Paper presented at the 31st Annual Conference on Neural Information Processing Systems (NIPS), Long Beach, CA.

Luo, Y., & Schuur, E. A. G. (2020). Model parameterization to represent processes at unresolved scales and changing properties of evolving systems. *Global change biology, 26*(3), 1109-1117. Retrieved from <Go to ISI>://WOS:000506663700001. doi:10.1111/gcb.14939

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., . . . Gottlieb, E. W. (2008). A satellite-based biosphere parameterization for net ecosystem CO2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM). *Global Biogeochemical Cycles, 22*(2).

Mäkelä, A., Pulkkinen, M., Kolari, P., Lagergren, F., Berbigier, P., Lindroth, A., . . . Hari, P. (2008). Developing an empirical model of stand GPP with the LUE approach: analysis of eddy covariance data at five contrasting conifer sites in Europe. *Global change biology, 14*(1), 92-108.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics, 5*(4), 115-133.

Medlyn, B. E., Robinson, A. P., Clement, R., & McMurtrie, R. E. (2005). On the validation of models of forest CO2 exchange using eddy covariance data: some perils and pitfalls. *Tree Physiology, 25*(7), 839-857. Retrieved from https://doi.org/10.1093/treephys/25.7.839. doi:10.1093/treephys/25.7.839

Mengoli, G., Agusti-Panareda, A., Boussetta, S., Harrison, S. P., Trotta, C., & Prentice, I. C. (2022). Ecosystem Photosynthesis in Land-Surface Models: A First-Principles Approach Incorporating Acclimation. *Journal of Advances in Modeling Earth Systems, 14*(1). Retrieved from <Go to ISI>://WOS:000751889900010. doi:10.1029/2021ms002767

Montavon, G., Binder, A., Lapuschkin, S., Samek, W., & Müller, K.-R. (2019). Layer-wise relevance propagation: an overview. *Explainable AI: interpreting, explaining and visualizing deep learning*, 193-209.

Monteith, J. (1972). Solar radiation and productivity in tropical ecosystems. *Journal of applied ecology, 9*(3), 747-766.

Peaucelle, M., Bacour, C., Ciais, P., Vuichard, N., Kuppel, S., Peñuelas, J., . . . Chen, J. (2019). Covariations between plant functional traits emerge from constraining parameterization of a terrestrial biosphere model. *Global ecology and biogeography, 28*(9), 1351-1365.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research, 12*, 2825-2830.

Piao, S., Wang, X., Wang, K., Li, X., Bastos, A., Canadell, J. G., . . . Sitch, S. (2020). Interannual variation of terrestrial carbon cycle: Issues and perspectives. *Global change biology, 26*(1), 300-318.

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.

Prentice, I. C., Dong, N., Gleason, S. M., Maire, V., & Wright, I. J. (2014). Balancing the costs of carbon gain and water transport: testing a new theoretical framework for plant functional ecology. *Ecology letters, 17*(1), 82-91.

Running, S. W., Nemani, R. R., Heinsch, F. A., Zhao, M., Reeves, M., & Hashimoto, H. (2004). A continuous satellite-derived measure of global terrestrial primary production. *Bioscience, 54*(6), 547-560.

Ryu, Y., Berry, J. A., & Baldocchi, D. D. (2019). What is global photosynthesis? History, uncertainties and opportunities. *Remote Sensing of Environment, 223*, 95-114. Retrieved from https://www.sciencedirect.com/science/article/pii/S0034425719300161. doi:https://doi.org/10.1016/j.rse.2019.01.016

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research, 15*, 1929-1958. Retrieved from <Go to ISI>://WOS:000344638300002.

Stocker, B. D., Wang, H., Smith, N. G., Harrison, S. P., Keenan, T. F., Sandoval, D., . . . Prentice, I. C. (2020). P-model v1. 0: an optimality-based light use efficiency model for simulating ecosystem gross primary production. *Geoscientific Model Development, 13*(3), 1545-1581.

Tian, X., Minunno, F., Cao, T., Peltoniemi, M., Kalliokoski, T., & Mäkelä, A. (2020). Extending the range of applicability of the semi-empirical ecosystem flux model PRELES for varying forest types and climate. *Global change biology, 26*(5), 2923-2943.

Viovy, N. (2018). CRUNCEP version 7-atmospheric forcing data for the community land model. *Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, 10*.

Walther, S., Besnard, S., Nelson, J. A., El-Madany, T. S., Migliavacca, M., Weber, U., . . . Jung, M. (2022). Technical note: A view from space on global flux towers by MODIS and Landsat: the FluxnetEO data set. *Biogeosciences, 19*(11), 2805-2840. Retrieved from <Go to ISI>://WOS:000807411900001. doi:10.5194/bg-19-2805-2022

Wang, H., Prentice, I. C., Keenan, T. F., Davis, T. W., Wright, I. J., Cornwell, W. K., . . . Peng, C. (2017). Towards a universal model for carbon dioxide uptake by plants. *Nature plants, 3*(9), 734-741.

Wang, R., Goll, D., Balkanski, Y., Hauglustaine, D., Boucher, O., Ciais, P., . . . Sardans, J. (2017). Global forest carbon uptake due to nitrogen and phosphorus deposition from 1850 to 2100. *Global change biology, 23*(11), 4854-4872.

Xiao, X., Hollinger, D., Aber, J., Goltz, M., Davidson, E. A., Zhang, Q., & Moore III, B. (2004). Satellite-based modeling of gross primary production in an evergreen needleleaf forest. *Remote Sensing of Environment, 89*(4), 519-534.

Xie, X., & Li, A. (2020). An adjusted two-leaf light use efficiency model for improving GPP simulations over mountainous areas. *Journal of Geophysical Research: Atmospheres, 125*(13), e2019JD031702.

Xu, T., & Hutchinson, M. (2011). ANUCLIM version 6.1 user guide. *The Australian National University, Fenner School of Environment and Society, Canberra, 90*.

Yan, H., Wang, S. Q., Yu, K. L., Wang, B., Yu, Q., Bohrer, G., . . . Shugart, H. H. (2017). A novel diffuse fraction-based two-leaf light use efficiency model: An application quantifying photosynthetic seasonality across 20 AmeriFlux flux tower sites. *Journal of Advances in Modeling Earth Systems, 9*(6), 2317-2332.

Yuan, W., Cai, W., Liu, S., Dong, W., Chen, J., Arain, M. A., . . . Georgiadis, T. (2014). Vegetation-specific model parameters are not required for estimating gross primary production. *Ecological modelling, 292*, 1-10.

Yuan, W., Cai, W., Xia, J., Chen, J., Liu, S., Dong, W., . . . Beringer, J. (2014). Global comparison of light use efficiency models for simulating terrestrial vegetation gross primary production based on the LaThuile database. *Agricultural and Forest Meteorology, 192*, 108-120.

Yuan, W., Liu, S., Zhou, G., Zhou, G., Tieszen, L. L., Baldocchi, D., . . . Goulden, M. L. (2007). Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. *Agricultural and Forest Meteorology, 143*(3-4), 189-207.

Yuan, W., Zheng, Y., Piao, S., Ciais, P., Lombardozzi, D., Wang, Y., . . . Hu, Z. (2019). Increased atmospheric vapor pressure deficit reduces global vegetation growth. *Science advances, 5*(8), eaax1396.

Zhang, F., Chen, J. M., Chen, J., Gough, C. M., Martin, T. A., & Dragoni, D. (2012). Evaluating spatial and temporal patterns of MODIS GPP over the conterminous US against flux measurements and a process model. *Remote Sensing of Environment, 124*, 717-729.

Zheng, Y., Shen, R., Wang, Y., Li, X., Liu, S., Liang, S., . . . Yuan, W. (2020). Improved estimate of global gross primary production for reproducing its long-term variation, 1982–2017. *Earth System Science Data, 12*(4), 2725-2746.

Zheng, Y., Zhang, L., Xiao, J., Yuan, W., Yan, M., Li, T., & Zhang, Z. (2018). Sources of uncertainty in gross primary productivity simulated by light use efficiency models: Model structure, parameters, input data, and spatial resolution. *Agricultural and Forest Meteorology, 263*, 242-257.

Zhou, Y., Wu, X., Ju, W., Chen, J. M., Wang, S., Wang, H., . . . Ibrom, A. (2016). Global parameterization and validation of a two-leaf light use efficiency model for predicting gross primary production across FLUXNET sites. *Journal of Geophysical Research: Biogeosciences, 121*(4), 1045-1072.

Journal of Advances in Modeling Earth Systems

Supporting Information for

Towards Robust Parameterizations in Ecosystem-level Photosynthesis Models

Shanning Bao[1,2*], Lazaro Alonso[1], Siyuan Wang[1], Johannes Gensheimer[1], Ranit De[1], Nuno Carvalhais[1,3,4*]

[1]Department for Biogeochemical Integration, Max-Planck-Institute for Biogeochemistry, 07745, Jena, Germany.

[2]National Space Science Center, Chinese Academy of Sciences, 100190, Beijing, China.

[3]Departamento de Ciências e Engenharia do Ambiente, DCEA, Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.

[4]ELLIS Unit Jena, 07745, Jena, Germany.

Contents of this file

Introduction

The supporting information includes the descriptions of model parameters calibration (text S1), cost functions used in SPIE (text S2), and figures and tables to supplement the information about the data and results appeared in the main text.

Text S1. Model parameters calibration

To get the highest model performance, we here calibrated the LUE model at each site using the full time series of GPP_{obs}. The purpose of model calibration is to find the parameter vector that can minimize the cost function, a metric to measure the model error, and to reduce

the model uncertainties associated with model parameters. In the calibration process, the parameters were optimized in their physical ranges (Table 1) using a stochastic and derivative-free evolutionary algorithm, CMAES(Hansen & Kern, 2004). CMAES, which is a reliable tool for global optimization (Trautmann et al., 2018).

We define the cost function (*cf*) as the sum of the GPP errors (cf_1, equation S- 2), the ET errors (cf_2, equation S- 3), and the environmental sensitivity functions (*f*X) constraints (cf_3 and cf_4).

$$cf = cf_1 + cf_2 + (cf_3 + cf_4) \quad S\text{-}1$$

$$cf_1 = \sum_{t=1}^{N_t} \sqrt{\left(GPP_t - \widehat{GPP}_t\right)^2 \cdot \sigma_{NEE_t}^{-2}} \quad S\text{-}2$$

$$cf_2 = \sum_{t=1}^{N_t} \sqrt{\left(ET_t - \widehat{ET}_t\right)^2 \cdot \sigma_{LE_t}^{-2}} \quad S\text{-}3$$

The cf_1 and cf_2 are to measure the sum of squares for errors of simulated GPP and ET, which is used to optimize the parameters of water availability index (WAI, see Table S2), at each time step *t*. The simulated GPP using the calibrated LUE parameters ($\widehat{GPP}$) and simulated ET using the calibrated WAI parameters ($\widehat{ET}$) were compared to GPP_{obs} (*GPP*) and ET_{obs} (*ET*), respectively. N_t denotes the total number of time steps. Due to the uncertainties in observation and the different units of GPP and ET, we weighted the model errors using the estimated uncertainty of *GPP* (σ_{NEE}) and *ET* (σ_{LE}), respectively. We assume that the parameter vector that minimizes the sum of cf_1 and cf_2 is the best for the LUE model and WAI, respectively.

We follow the concept of ecological and dynamic constraints (EDC, by (Bloom & Williams, 2015)) to regularize the inversion approach via two additional constraints: cf_3 (equation S- 4) and cf_4 (equation S- 5).

$$cf_3 = \left(\left(1\text{-}\max(f\mathrm{T}_r)\right) + \left(1\text{-}\max(f\mathrm{VPD}_r)\right) + \left(1\text{-}\max(f\mathrm{W}_r)\right) + \left(1\text{-}\max(f\mathrm{L}_r)\right)\right) \cdot c \quad S\text{-}4$$

$$cf_4 = \left(\sum_r \left(fT_r(T<0\,°C) > \theta_{fT}\right) + \sum_r \left(fVPD_r(VPD>2kPa) > \theta_{fVPD}\right) + \sum_r \left(fW_r(W<0.01) > \theta_{fW}\right)\right) \cdot c \quad S\text{-}5$$

These impose constraints on the simulated *f*X (i.e., *f*T, *f*VPD, *f*W, *f*L and *f*CI) based on two assumptions: the instantaneous ε ($=\varepsilon_{max} \cdot f\mathrm{T} \cdot f\mathrm{VPD} \cdot f\mathrm{W} \cdot f\mathrm{L} \cdot f\mathrm{CI}$) of vegetation can reach its potential, ε_{max}, under some specific environmental condition (cf_3) and is inhibited under a non-ideal growing condition (cf_4). Here cf_3 and cf_4 were calculated independently from cf_1 and cf_2, using analog inputs (PAR=0-20 $MJ \cdot m^{-2} \cdot day^{-1}$, FAPAR=0-1, T=-10-40 °C, VPD=0-2 kPa, W=0-1 and CI=0-1). *cf*3 is to set the maximum of *f*T, *f*VPD, excluding the CO_2 fertilization effect (the right part of equation 3), *f*W, and *f*L to one, which implies that the corresponding environmental factor does not limit ε at a certain point within the given ranges of PAR$\in$ [0,20] (in $MJ \cdot m^{-2} \cdot day^{-1}$), FAPAR$\in$[0,1], T$\in$[-10,40] (in °C), VPD$\in$[0,2] (in kPa), W$\in$[0,1] and CI$\in$[0,1], represented by the subscript, *r*, in equations S- 4-S- 5 (e.g., $\max(f\mathrm{T}_r)$ represents the maximum *f*T when the temperature is ranging between -10 and 40 °C).

Another constraint, cf_4, is to guarantee the *f*T, *f*VPD, excluding the CO_2 fertilization part, and *f*W lower than the threshold (θ_{fT}, θ_{fVPD}, and θ_{fW}) under the non-ideal conditions (T<0 °C, VPD>2 kPa, or W<0.01). Here the thresholds (θ_{fT}=0.2, θ_{fVPD}=0.9, and θ_{fW}=0.2) were estimated according to the normalized ratio of GPP to APAR at all sites. The other non-ideal conditions were not included since they vary across sites. The *c* in equations S- 4-S- 5 denotes a penalty term ($=10^4$, an empirical value) to coordinate the scales of cf_1, cf_2, cf_3, and cf_4.

Since the WAI parameters were not predicted in this study, the calibrated WAI parameters were used in the parameterization experiments and the cf_2 was not considered in the optimization-based parameterization methods, i.e., 'OPT-All' and 'OPT-PFT'.

Text S2. Cost functions used in SPIE

The cost function (cf_{NN}, equation S- 6) for SPIE was similar to the sum of cf_1, cf_3, and cf_4. Since normalizing the cost function can significantly improve the training efficiency of neural network, we used normalized NSE (Nossent & Bauwens, 2012), ranging from 0-1, rather than the sum of squares (S- 7).

$$cf_{NN}=cf_{NN1}+cf_3+cf_4 \qquad S\text{- }6$$

$$cf_{NN1}=\frac{\sum_{t=1}^{N_t}\left(GPP_t-\widehat{GPP}_t\right)^2\cdot\sigma_{NEE_t}{}^{-2}}{\sum_{t=1}^{N_t}\left(GPP_t-\widehat{GPP}_t\right)^2\cdot\sigma_{NEE_t}{}^{-2}+\sum_{t=1}^{N_t}\left(GPP_t-\overline{GPP}\right)^2\cdot\sigma_{NEE_t}{}^{-2}} \qquad S\text{- }7$$

GPP_t and $\widehat{GPP}_t$ are the observed and simulated GPP at time step, t. The normalized NSE is the ratio between the sum of the GPP errors across all time steps (N_t) to the sum of GPP errors and the sum of GPP changes to the average GPP ($\overline{GPP}$). To consider the EDC, we added cf_{NN1} to cf_3 and cf_4 as defined in section S1. The only difference was that the empirical coefficient, c, was changed to 0.2 here due to the small range of cf_{NN1}.

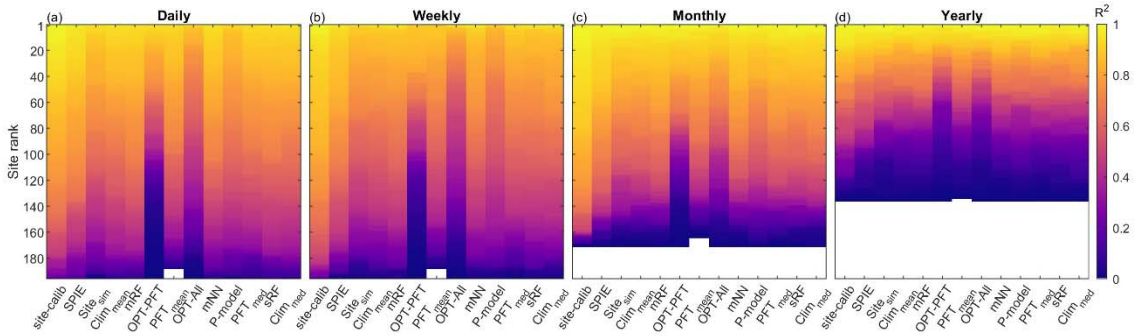


Figure S1. Comparison of R^2 between GPP_{obs} and GPP_{sim} based on twelve different parameterization methods, and between GPP_{obs} and GPP_{calib} (site-calib) at daily (a), weekly (b), monthly (c) and yearly (d) scales. The sites with less than four good-quality months or years are removed from panel c and d, respectively. The sites with p-value larger than 0.05 are shown in white.

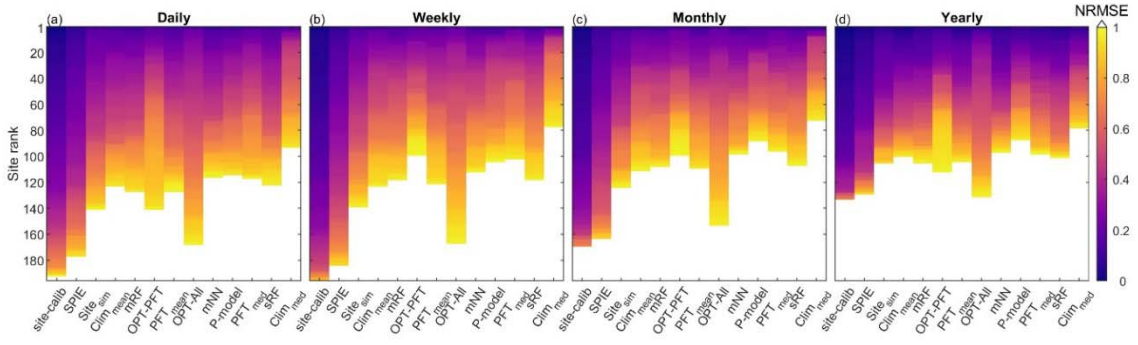


Figure S2. Comparison of NRMSE between GPP_{obs} and GPP_{sim} based on twelve different parameterization methods, and between GPP_{obs} and GPP_{calib} (site-calib) at daily (a), weekly (b), monthly (c) and yearly (d) scales.

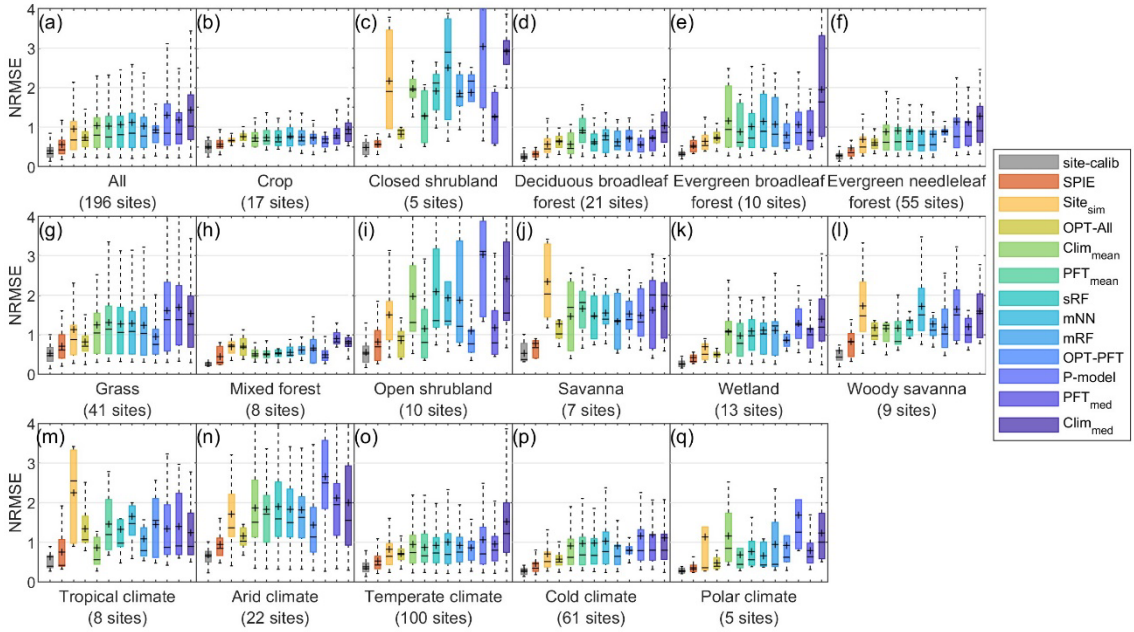


Figure S3. Site-level daily NRMSE comparison across all sites (a), per PFT (b-l) and per Clim (m-q). The mean and median per type are represented by the black cross and line, respectively

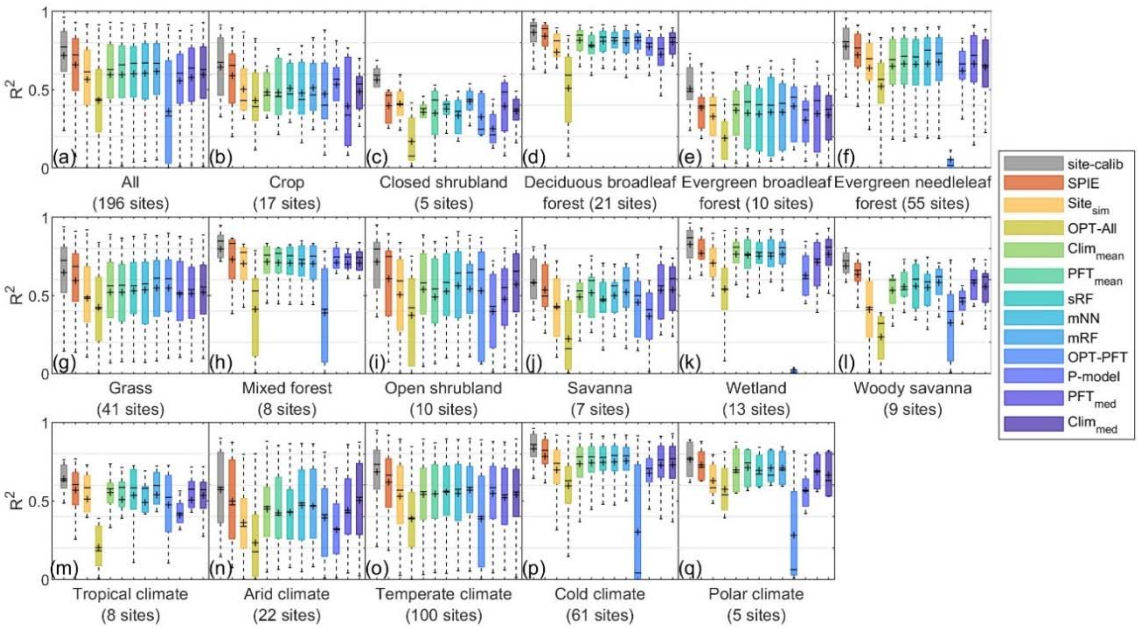


Figure S4. Site-level daily R^2 comparison across all sites (a), per PFT (b-l) and per Clim (m-q). The mean and median per type are represented by the black cross and line, respectively

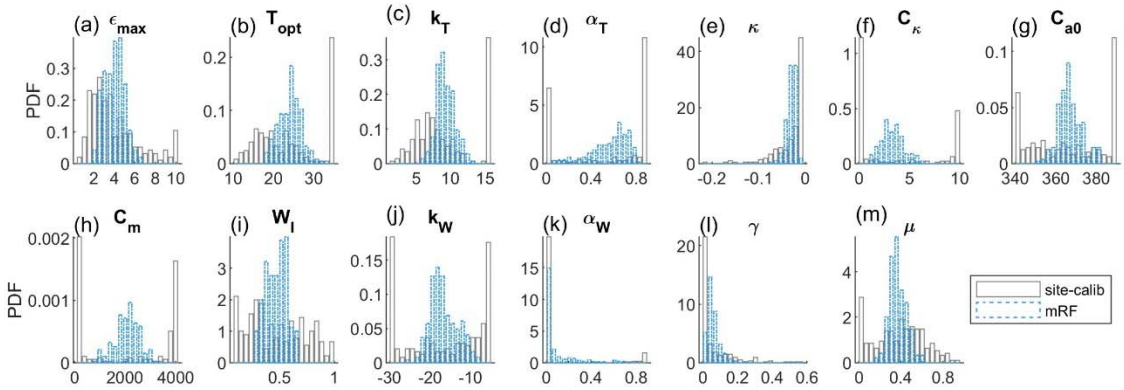


Figure S5. The probability distribution function (PDF) of the predicted parameters by mRF (Upscal) and the calibrated parameters (Calib)

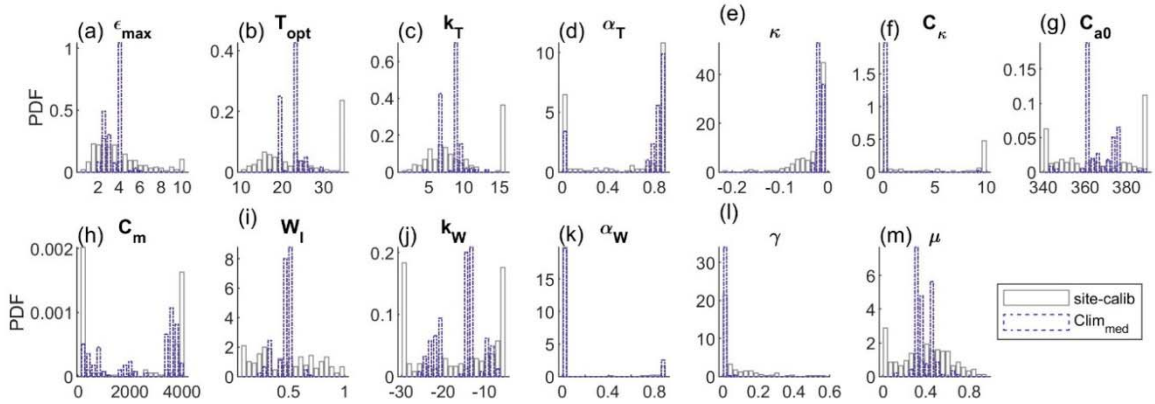


Figure S6. The probability distribution function (PDF) of the predicted parameters by $Clim_{med}$ and the calibrated parameters

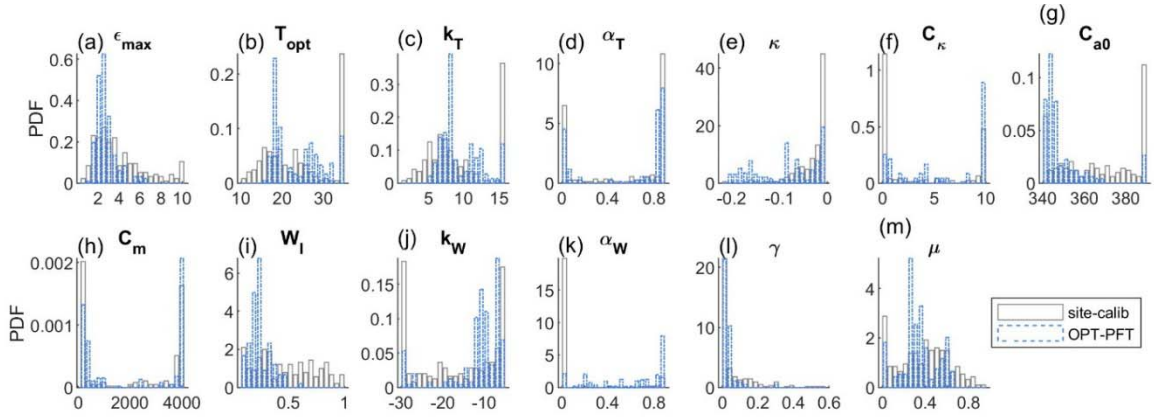


Figure S7. The probability distribution function (PDF) of the predicted parameters by OPT-PFT and the calibrated parameters

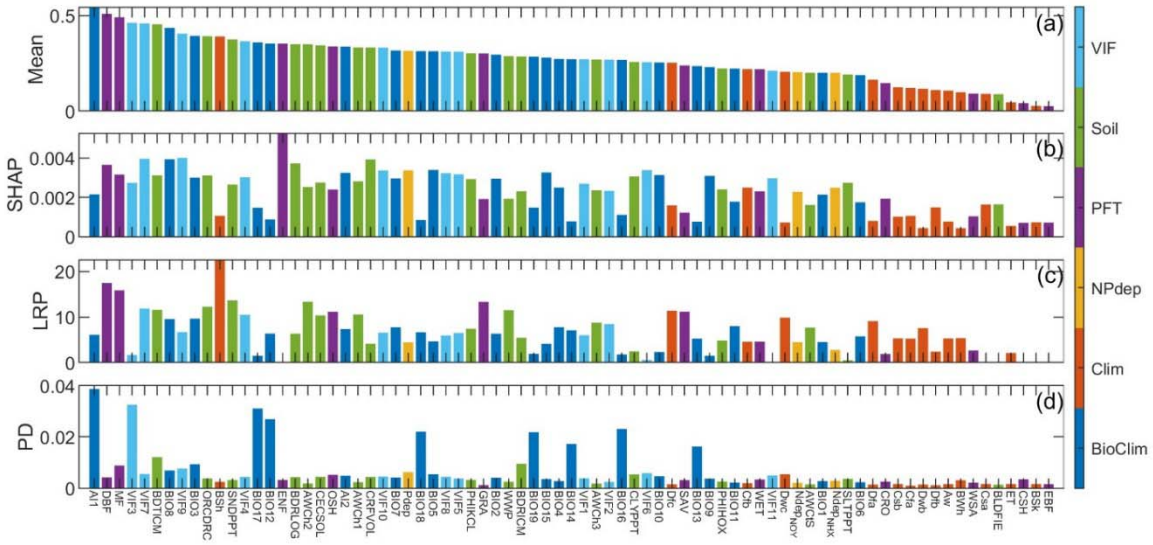


Figure S8. The input features for predicting LUE model parameters sorted according to the average normalized SHAP, LRP and PD values (a). The absolute SHAP, LRP and PD of each feature are shown in b-d in the same order.

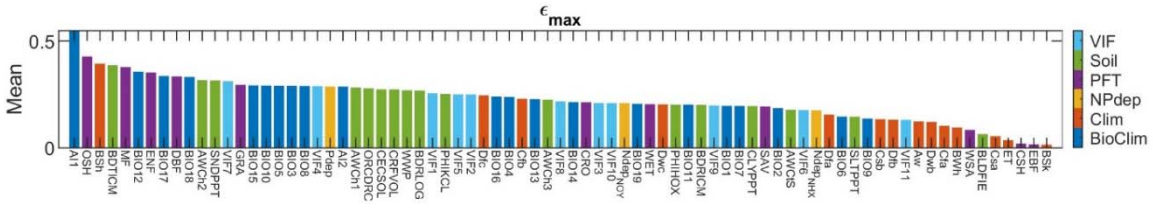


Figure S9. The features sorted by the average normalized SHAP, LRP and PD for ε_{max}

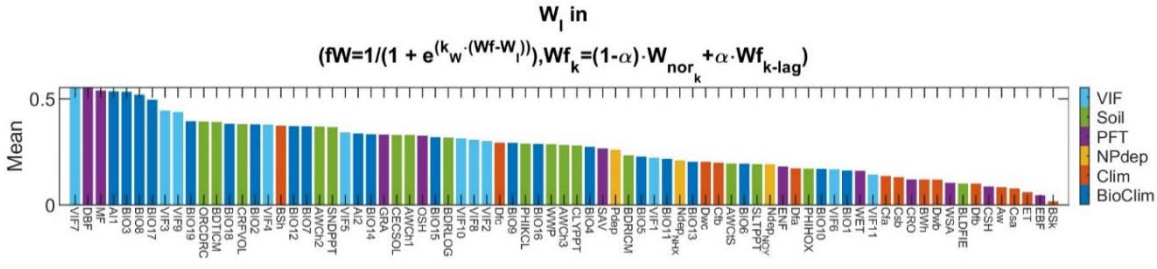


Figure S10. The features sorted by the average normalized SHAP, LRP and PD for W_I

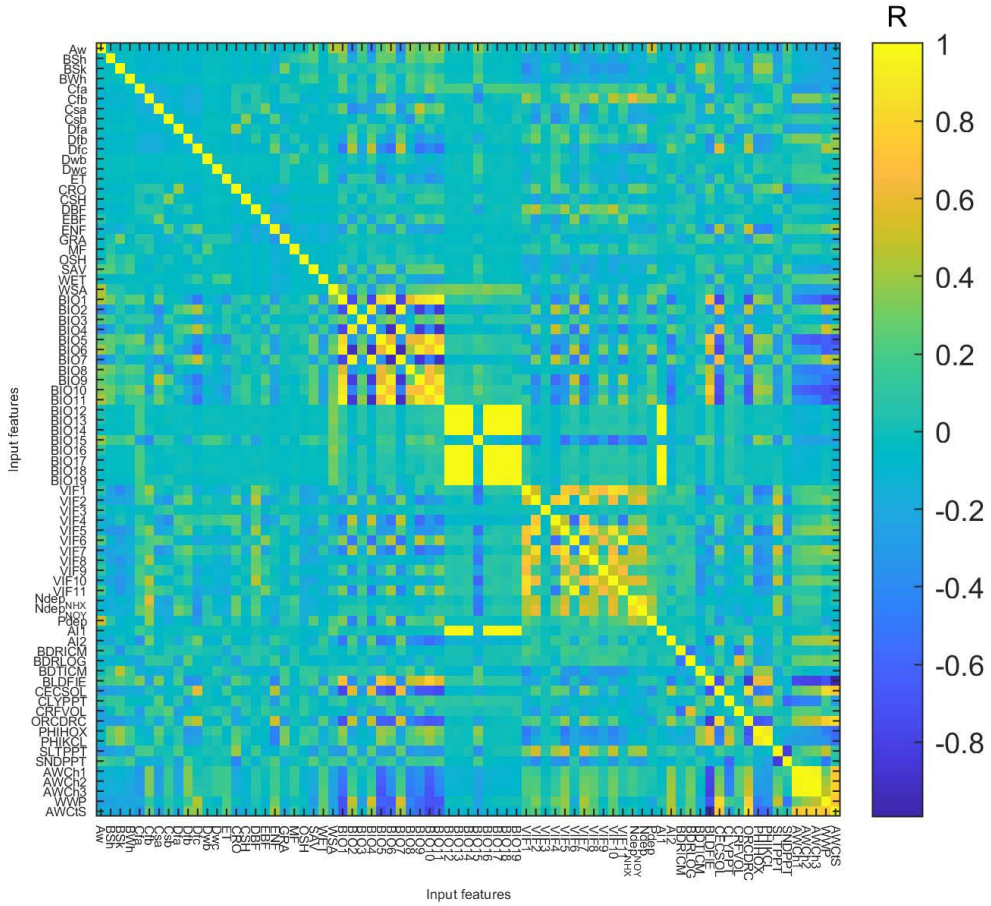


Figure S11. Correlation coefficient (R) matrix between input features.

Table S1. Eddy covariance flux site list used in this study. The latitude (Lat), longitude (Lon) and plant functional types (PFT) are collected from FLUXNET website. The data length differs across site and is determined by the years between 'data start' and 'data end'. The climate type is extracted from the Koeppen-Geiger climate classification map (at 5 arc min; Rubel et al., 2017). The elevation is collected from the site ancillary information, papers and satellite images (see the footnote below the table). The site group refers to the group number of each site used to validate the training result.

SiteID	Lat	Lon	Data start (year)	Data end (year)	PFT	Climate type	Elevation (m)	Site group	Reference
AR-SLu	-33.5	-66.5	2010	2011	MF	BSh	506*e	4	(Ulke et al., 2015)
AT-Neu	47.1	11.3	2002	2012	GRA	Dfc	970	10	(Wohlfahrt et al., 2008)
AU-ASM	-22.3	133.3	2010	2014	ENF	BWh	606*b	5	(Cleverly et al., 2013)

AU-Cpr	-34.0	140.6	2010	2014	SAV	BSk	62[*e]	3	(Bloomfield et al., 2018; Meyer et al., 2015)
AU-Cum	-33.6	150.7	2012	2014	EBF	Cfa	20	1	(Renchon et al., 2018)
AU-DaP	-14.1	131.3	2009	2013	GRA	Aw	71[*e]	3	(Hutley et al., 2011)
AU-DaS	-14.2	131.4	2010	2014	SAV	Aw	110	6	(Hutley et al., 2011)
AU-Dry	-15.3	132.4	2008	2014	SAV	Aw	175	5	(Hutley et al., 2011)
AU-Emr	-23.9	148.5	2011	2013	GRA	BSh	170	3	(Schroder, 2014)
AU-Gin	-31.4	115.7	2013	2014	WSA	Csa	51	3	(Beringer et al., 2016)
AU-GWW	-30.2	120.7	2011	2014	SAV	BSh	450	1	(Beringer et al., 2016)
AU-How	-12.5	131.2	2001	2014	WSA	Aw	64	2	(Beringer et al., 2003)
AU-RDF	-14.6	132.5	2011	2013	WSA	Aw	188[*e]	7	(Bristow et al., 2016)
AU-Rig	-36.7	145.6	2011	2014	GRA	Cfa	152	10	(Beringer et al., 2016)
AU-Stp	-17.2	133.4	2008	2014	GRA	BSh	250[*b]	1	(Hutley et al., 2011)
AU-TTE	-22.3	133.6	2012	2014	OSH	BWh	553	3	(Cleverly et al., 2016)
AU-Tum	-35.7	148.2	2001	2014	EBF	Cfb	1200	8	(Leuning et al., 2005)
AU-Wom	-37.4	144.1	2010	2014	EBF	Cfb	705	6	(Griebel et al., 2016)
AU-Ync	-35.0	146.3	2012	2014	GRA	BSk	126[*e]	6	(Yee et al., 2015)
BE-Bra	51.3	4.5	2000	2014	MF	Cfb	16[*a]	7	(Carrara et al., 2004)
BE-Lon	50.6	4.8	2004	2014	CRO	Cfb	167	2	(Aubinet et al., 2009)
BE-Vie	50.3	6.0	2000	2014	MF	Cfb	450[*a]	1	As above
BR-Ban	-9.8	-50.2	2003	2006	EBF	Aw	120	6	(Da Rocha et al., 2009)
BR-Sp1	-21.6	-47.7	2001	2002	WSA	Aw	690	9	As above
BW-Ma1	-19.9	23.6	2000	2001	WSA	BSh	950	4	(Veenendaal et al., 2004)

CA-Ca1	49.9	-125.3	2000	2005	ENF	Cfb	300	7	(Humphreys et al., 2006)
CA-Ca2	49.9	-125.3	2000	2005	ENF	Csb	300	7	As above
CA-Ca3	49.5	-124.9	2001	2005	ENF	Csb	300	7	As above
CA-Gro	48.2	-82.2	2003	2014	MF	Dfb	340	9	(Pejam et al., 2006)
CA-Let	49.7	-112.9	2000	2005	GRA	BSk	960	7	(Flanagan et al., 2002)
CA-Mer	45.4	-75.5	2000	2005	WET	Dfb	70	1	(Lafleur et al., 2003)
CA-NS2	55.9	-98.5	2002	2005	ENF	Dfc	260	6	(Beringer et al., 2011)
CA-NS3	55.9	-98.4	2001	2005	ENF	Dfc	260	10	As above
CA-NS4	55.9	-98.4	2002	2005	ENF	Dfc	260	6	As above
CA-NS5	55.9	-98.5	2002	2005	ENF	Dfc	260	2	As above
CA-NS6	55.9	-99.0	2001	2005	OSH	Dfc	244	2	As above
CA-NS7	56.6	-100.0	2002	2005	OSH	Dfc	297	5	As above
CA-Oas	53.6	-106.2	2000	2010	DBF	Dfc	530	6	(Black et al., 1996)
CA-Obs	54.0	-105.1	2000	2010	ENF	Dfc	628.94	7	(Jarvis et al., 1997)
CA-Ojp	53.9	-104.7	2000	2005	ENF	Dfb	579	1	(Baldocchi et al., 1997)
CA-Qcu	49.3	-74.0	2001	2006	ENF	Dfb	392.3	4	(Giasson et al., 2006)
CA-Qfo	49.7	-74.3	2004	2010	ENF	Dfc	382	7	(Bergeron et al., 2007)
CA-SF1	54.5	-105.8	2003	2006	ENF	Dfc	536	10	(Mkhabela et al., 2009)
CA-SF2	54.3	-105.9	2001	2005	ENF	Dfc	520	6	As above
CA-SF3	54.1	-106.0	2001	2005	OSH	Dfc	540	5	As above
CA-SJ1	53.9	-104.7	2001	2005	ENF	Dfb	580	8	(Howard et al., 2004)
CA-SJ2	53.9	-104.7	2003	2005	ENF	Dfc	580	1	(Coursolle et al., 2012)
CA-TP1	42.7	-80.6	2008	2014	ENF	Dfb	265	4	(Peichl & Arain, 2007)
CA-TP3	42.7	-80.4	2008	2014	ENF	Dfb	184	2	As above

CA-TP4	42.7	-80.4	2008	2014	ENF	Dfb	184	8	(Arain & Restrepo-Coupe, 2005)
CA-TPD	42.6	-80.6	2012	2014	DBF	Dfb	260	1	As above
CA-WP1	55.0	-112.5	2003	2005	WET	Dfc	540	4	(Syed et al., 2006)
CH-Cha	47.2	8.4	2005	2014	GRA	Cfb	393	5	(Lutz Merbold et al., 2014)
CH-Dav	46.8	9.9	2000	2014	ENF	ET	1639	9	(Wolf et al., 2013; Zielis et al., 2014)
CH-Fru	47.1	8.5	2005	2014	GRA	Cfb	982	3	(Imer et al., 2013)
CH-Oe1	47.3	7.7	2002	2008	GRA	Cfb	450	3	(Ammann et al., 2009)
CN-Cha	42.4	128.1	2003	2005	MF	Dwb	738	1	(Zhang et al., 2006)
CN-Cng	44.6	123.5	2007	2010	GRA	BSk	171[*d]	10	(Pastorello et al., 2020)
CN-Dan	30.5	91.1	2004	2005	GRA	Dwc	4286	8	(Shi et al., 2006)
CN-Du2	42.1	116.3	2006	2008	GRA	Dwb	1350[*b]	5	(Chen et al., 2009)
CN-Ha2	37.6	101.3	2003	2005	WET	Dwc	3357	2	(Pastorello et al., 2020)
CN-Xfs	44.1	116.3	2004	2006	GRA	BSk	1250	4	(Chen et al., 2009)
CZ-BK1	49.5	18.5	2004	2014	ENF	Dfb	908[*a]	8	(Krupková et al., 2017)
CZ-BK2	49.5	18.5	2009	2012	GRA	Dfb	855	7	(Acosta et al., 2013)
CZ-wet	49.0	14.8	2007	2014	WET	Cfb	426	3	(Dušek et al., 2012)
DE-Geb	51.1	10.9	2001	2014	CRO	Cfb	161.5	7	(Anthoni et al., 2004b)
DE-Gri	51.0	13.5	2004	2014	GRA	Cfb	385	4	(Prescher et al., 2010)
DE-Hai	51.1	10.5	2000	2009	DBF	Cfb	430[*a]	9	(Knohl et al., 2003)
DE-Har	47.9	7.6	2005	2006	ENF	Cfb	201	1	(Pastorello et al., 2020)

DE-Kli	50.9	13.5	2004	2014	CRO	Cfb	478	8	(Prescher et al., 2010)
DE-Lnf	51.3	10.4	2002	2012	DBF	Cfb	451	9	(Anthoni et al., 2004a)
DE-Meh	51.3	10.7	2003	2006	MF	Cfb	293[*a]	7	(DON et al., 2009)
DE-Obe	50.8	13.7	2008	2014	ENF	Cfb	734	4	(Pastorello et al., 2020)
DE-SfN	47.8	11.3	2013	2014	WET	Cfb	590	2	(Hommeltenberg et al., 2014)
DE-Tha	51.0	13.6	2000	2014	ENF	Cfb	380[*a]	3	(Bernhofer et al., 2003)
DE-Wet	50.5	11.5	2002	2006	ENF	Cfb	785[*a]	8	(Rebmann et al., 2010)
DK-Ris	55.5	12.1	2004	2005	CRO	Cfb	10	10	(Pastorello et al., 2020)
DK-Sor	55.5	11.6	2000	2014	DBF	Cfb	40[*a]	10	(Pilegaard & Ibrom, 2020)
ES-Amo	36.8	-2.3	2000	2014	OSH	BSh	58	3	(López-Ballesteros et al., 2017)
ES-ES1	39.4	-0.3	2007	2012	ENF	Csa	5[*a]	3	(Sanz M J, 2004)
ES-ES2	39.3	-0.3	2000	2006	CRO	Csa	10	9	As above
ES-LgS	37.1	-3.0	2004	2006	OSH	Csb	2267	9	(Reverter et al., 2010)
ES-LJu	36.9	-2.8	2005	2011	OSH	Csa	1600	6	(Serrano-Ortiz et al., 2009)
ES-LMa	39.9	-5.8	2004	2006	SAV	Csa	258[*a]	5	(Perez-Priego et al., 2017)
ES-VDA	42.2	1.5	2007	2009	GRA	Cfb	1765[*a]	2	(Pastorello et al., 2020)
FI-Hyy	61.9	24.3	2004	2006	ENF	Dfc	181[*a]	8	(Suni et al., 2003)
FI-Kaa	69.1	27.3	2000	2014	WET	Dfc	155	2	(MIKA AURELA et al., 2007)
FI-Let	60.6	24.0	2000	2006	ENF	Dfb	111	5	(Koskinen et al., 2014)
FI-Lom	68.0	24.2	2009	2012	WET	Dfc	269[*a]	2	(M. Aurela et al., 2015)

FI-Sod	67.4	26.6	2007	2009	ENF	Dfc	180[*a]	1	(Thum et al., 2007)
FR-Fon	48.5	2.8	2008	2014	DBF	Cfb	92[*a]	8	(Michelot et al., 2011)
FR-Gri	48.8	2.0	2005	2013	CRO	Cfb	125	6	(Loubet et al., 2011)
FR-Hes	48.7	7.1	2004	2014	DBF	Cfb	300[*a]	2	(Granier et al., 2000)
FR-LBr	44.7	-0.8	2000	2006	ENF	Cfb	61[*a]	1	(Berbigier et al., 2001)
FR-Lq1	45.6	2.7	2000	2008	GRA	Cfb	1040	9	(Pastorello et al., 2020)
FR-Lq2	45.6	2.7	2004	2006	GRA	Cfb	1040	1	(Pastorello et al., 2020)
FR-Pue	43.7	3.6	2004	2006	EBF	Csa	270[*a]	5	(Rambal et al., 2004)
GL-ZaH	74.5	-20.6	2000	2014	GRA	ET	48	4	(Lund et al., 2012)
HU-Bug	46.7	19.6	2002	2006	GRA	Cfb	111[*a]	7	(Pastorello et al., 2020)
IL-Yat	31.3	35.1	2001	2006	ENF	Csa	650	7	(Tatarinov et al., 2016)
IT-Amp	41.9	13.6	2002	2006	GRA	Cfb	884[*a]	6	(Papale et al., 2015)
IT-BCi	40.5	15.0	2004	2012	CRO	Csa	20	7	(Vitale et al., 2016)
IT-CA1	42.4	12.0	2011	2014	DBF	Csa	200	4	(Sabbatini et al., 2016)
IT-CA2	42.4	12.0	2011	2014	CRO	Csa	200	5	(Sabbatini et al., 2016)
IT-CA3	42.4	12.0	2011	2014	DBF	Csa	197	5	(Sabbatini et al., 2016)
IT-Col	41.9	13.6	2004	2014	DBF	Cfb	1560[*a]	4	(VALENTINI et al., 1996)
IT-Cpz	41.7	12.4	2000	2008	EBF	Csa	68	8	(Tirone et al., 2003)
IT-Isp	45.8	8.6	2013	2014	DBF	Cfa	210	9	(Ferréa et al., 2012)
IT-Lav	46.0	11.3	2004	2014	ENF	Cfb	1353	2	(B. Marcolla et al., 2003)
IT-Lec	43.3	11.3	2005	2006	EBF	Csa	314	8	(Pastorello et al., 2020)

IT-MBo	46.0	11.1	2003	2013	GRA	Dfb	1550[*a]	5	(Barbara Marcolla et al., 2005)
IT-Noe	40.6	8.2	2004	2014	CSH	Csa	25	10	(Papale et al., 2015)
IT-Non	44.7	11.1	2001	2006	MF	Cfa	25[*c]	1	(Nardino, 2002)
IT-PT1	45.2	9.1	2002	2004	DBF	Cfa	60	8	(Migliavacca et al., 2009)
IT-Ren	46.6	11.4	2002	2013	ENF	Dfc	1730[*a]	10	(Barbara Marcolla et al., 2005)
IT-Ro1	42.4	11.9	2000	2008	DBF	Csa	235	9	(Rey et al., 2002)
IT-Ro2	42.4	11.9	2002	2012	DBF	Csa	224[*a]	4	(TEDESCHI et al., 2006)
IT-SR2	43.7	10.3	2013	2014	ENF	Csa	4	10	(Pastorello et al., 2020)
IT-SRo	43.7	10.3	2000	2012	ENF	Csa	4[*a]	7	(Chiesi et al., 2005)
IT-Tor	45.8	7.6	2008	2012	GRA	ET	2160	10	(Galvagno et al., 2013)
NL-Ca1	52.0	4.9	2003	2006	GRA	Cfb	0.7	3	(Jacobs et al., 2007)
NL-Loo	52.2	5.7	2000	2014	ENF	Cfb	25[*a]	9	(Dolman et al., 2002)
PT-Cor	39.1	-8.3	2010	2017	EBF	Csa	170[*c]	6	(Pastorello et al., 2020)
PT-Esp	38.6	-8.6	2002	2006	EBF	Csa	95[*a]	9	(Rodrigues et al., 2011)
PT-Mi1	38.5	-8.0	2003	2005	EBF	Csa	264[*a]	10	(Pereira et al., 2007)
PT-Mi2	38.5	-8.0	2004	2006	GRA	Csa	190	8	(Pereira et al., 2007)
RU-Fyo	56.5	32.9	2002	2014	ENF	Dfb	265[*a]	9	(Kurbatova et al., 2008)
RU-Ha1	54.7	90.0	2002	2004	GRA	Dfb	446	4	(Belelli Marchesini et al., 2007)
RU-Zot	60.8	89.4	2002	2004	ENF	Dfc	90	10	(Arneth et al., 2002)
SD-Dem	13.3	30.5	2007	2009	SAV	BWh	500	2	(Ardö et al., 2008)

SE-Deg	64.2	19.6	2001	2005	WET	Dfc	270	6	(Sagerfors et al., 2008)
SE-Fla	64.1	19.5	2000	2002	ENF	Dfc	226[*c]	3	(Valentini et al., 2000)
US-AR1	36.4	-99.4	2009	2012	GRA	Cfa	611	6	(Billesbach D, 2016)
US-AR2	36.6	-99.6	2009	2012	GRA	Cfa	646	10	(Billesbach D, 2016)
US-ARb	35.6	-98.0	2003	2012	GRA	Cfa	424	8	(Pastorello et al., 2020)
US-ARc	35.6	-98.0	2005	2006	GRA	Cfa	424	5	(Pastorello et al., 2020)
US-ARM	36.6	-97.5	2005	2006	CRO	Cfa	314	2	(Pastorello et al., 2020)
US-Atq	70.5	-157.4	2003	2008	WET	ET	15	10	(Pastorello et al., 2020)
US-Aud	31.6	-110.5	2002	2006	GRA	BSk	1469	9	(Pastorello et al., 2020)
US-Bar	44.1	-71.3	2004	2005	DBF	Dfb	272	8	(Ouimette et al., 2018)
US-Bkg	44.4	-96.8	2004	2006	GRA	Dfa	510	4	(Gilmanov et al., 2005)
US-Blo	38.9	-120.6	2000	2007	ENF	Csb	1315	1	(Goldstein et al., 2000)
US-Bo1	40.0	-88.3	2000	2007	CRO	Cfa	219	2	(Pastorello et al., 2020)
US-Bo2	40.0	-88.3	2004	2006	CRO	Cfa	219	9	(Pastorello et al., 2020)
US-Cop	38.1	-109.4	2011	2013	GRA	BSk	1520	1	(D., 2016)
US-CRT	41.6	-83.4	2001	2007	CRO	Dfa	180	8	(Pastorello et al., 2020)
US-Dk1	36.0	-79.1	2001	2005	GRA	Cfa	168	7	(Pastorello et al., 2020)
US-Dk3	36.0	-79.1	2001	2005	ENF	Cfa	163	4	(Pastorello et al., 2020)

US-Fmf	35.1	-111.7	2000	2006	ENF	Csb	2160	5	(Pastorello et al., 2020)
US-FPe	48.3	-105.1	2004	2006	GRA	BSk	634	3	(Pastorello et al., 2020)
US-FR2	30.0	-98.0	2005	2006	WSA	Cfa	271.9	5	(Heinsch et al., 2004)
US-Goo	34.3	-89.9	2002	2006	GRA	Cfa	87	10	(T., 2016)
US-Ha1	42.5	-72.2	2000	2012	DBF	Dfb	340	2	(Urbanski et al., 2007)
US-Ho1	45.2	-68.7	2000	2004	ENF	Dfb	60	10	(Hollinger et al., 1999)
US-IB1	41.9	-88.2	2005	2007	CRO	Dfa	226.5	3	(Pastorello et al., 2020)
US-IB2	41.8	-88.2	2004	2011	GRA	Dfa	226.5	5	(Pastorello et al., 2020)
US-Ivo	68.5	-155.8	2004	2007	WET	ET	568	1	(Epstein et al., 2004)
US-KS2	28.6	-80.7	2003	2006	CSH	Cfa	3	6	(Powell et al., 2006)
US-Los	46.1	-90.0	2000	2014	WET	Dfb	480	4	(Sulman et al., 2009)
US-Me2	44.5	-121.6	2000	2014	ENF	Csb	1253	2	(Kwon et al., 2018; Thomas et al., 2009)
US-Me3	44.3	-121.6	2004	2006	ENF	Csb	1005	6	(Vickers et al., 2012)
US-Me5	44.4	-121.6	2002	2014	ENF	Csb	1188	4	(Law et al., 2001; Williams et al., 2001)
US-Me6	44.3	-121.6	2004	2009	ENF	Csb	998	7	(Ruehr et al., 2014)
US-MMS	39.3	-86.4	2000	2002	DBF	Cfa	275	7	(Roman et al., 2015)
US-MOz	38.7	-92.2	2010	2014	DBF	Cfa	219.4	3	(Gu et al., 2016)
US-Myb	38.1	-121.8	2011	2014	WET	Csa	-1	2	(Pastorello et al., 2020)
US-NC1	35.8	-76.7	2005	2006	OSH	Cfa	5	4	(Noormets et al., 2012)

US-NC2	35.8	-76.7	2005	2006	ENF	Cfa	5	2	(Pastorello et al., 2020)
US-Ne1	41.2	-96.5	2000	2014	CRO	Dfa	361	7	(Pastorello et al., 2020)
US-Ne2	41.2	-96.5	2001	2013	CRO	Dfa	362	8	(Pastorello et al., 2020)
US-Ne3	41.2	-96.4	2001	2013	CRO	Dfa	363	6	(Pastorello et al., 2020)
US-NR1	40.0	-105.6	2001	2013	ENF	Dfc	3050	9	(Monson et al., 2002)
US-Oho	41.6	-83.8	2004	2013	DBF	Dfa	230	4	(DeForest et al., 2006)
US-Prr	65.1	-147.5	2011	2014	ENF	Dfc	210	3	(Ikawa et al., 2015; Nakai et al., 2013)
US-SO2	33.4	-116.6	2004	2006	CSH	Csb	1394	3	(Lipson et al., 2005)
US-SO3	33.4	-116.6	2001	2006	CSH	Csb	1429	5	(Lipson et al., 2005)
US-SO4	33.4	-116.6	2004	2006	CSH	Csb	1429	5	(Lipson et al., 2005)
US-SP2	29.8	-82.2	2000	2004	ENF	Cfa	50	9	(Clark et al., 1999)
US-SP3	29.8	-82.2	2000	2004	ENF	Cfa	50	6	(Clark et al., 1999)
US-SRC	31.9	-110.8	2008	2014	OSH	BSh	991	6	(Pastorello et al., 2020)
US-SRG	31.8	-110.8	2008	2014	GRA	Csa	1291	2	(Scott et al., 2015)
US-SRM	31.8	-110.9	2004	2014	WSA	BSk	1120	5	(Scott et al., 2009)
US-Syv	46.2	-89.4	2001	2014	MF	Dfb	540	6	(Desai et al., 2005)
US-Ton	38.4	-121.0	2001	2014	WSA	Csa	177	8	(Ma et al., 2016)
US-Twt	38.1	-121.7	2009	2014	CRO	Csa	-7	5	(Pastorello et al., 2020)
US-UMB	45.6	-84.7	2000	2014	DBF	Dfb	234	10	(Gough et al., 2008)
US-Var	38.4	-121.0	2000	2014	GRA	Csa	129	9	(Ma et al., 2011)

US-WCr	45.8	-90.1	2000	2014	DBF	Dfb	520	4	(Cook et al., 2004)
US-Whs	31.7	-110.1	2011	2013	OSH	BSk	1370	10	(Pastorello et al., 2020)
US-Wi4	46.7	-91.2	2007	2014	ENF	Dfb	352	7	(Noormets et al., 2007)
US-Wi9	46.6	-91.1	2002	2005	ENF	Dfb	350	1	(Noormets et al., 2007)
US-Wkg	31.7	-109.9	2004	2005	GRA	BSk	1531	8	(Scott, 2010)
US-WPT	41.5	-83.0	2004	2014	WET	Cfa	175	3	(Pastorello et al., 2020)
US-Wrc	45.8	-122.0	2000	2006	ENF	Csb	371	1	(Wharton et al., 2012)
ZA-Kru	-25.0	31.5	2000	2012	SAV	BSh	359	3	(Archibald et al., 2009)
ZM-Mon	-15.4	23.3	2007	2009	WSA	Aw	1053	9	(L. Merbold et al., 2009)

Note.

[*a]: collected from (Flechard et al., 2020).

[*b]: collected from (Hao et al., 2019).

[*c]: collected from (B. Tang et al., 2018).

[*d]: collected from (X. Tang et al., 2020).

[*e]: extracted from google earth.

Other elevation data were collected from https://fluxnet.org/, http://www.europe-fluxdata.eu/, http://www.ozflux.org.au/, https://ameriflux.lbl.gov/, http://www.asiaflux.net/, http://www.chinaflux.org/, and ancillary information of LaThuile dataset (https://fluxnet.org/data/la-thuile-dataset/).

131 **Table S2.** List of the forcing variables for the LUE model. The variables in bold are used to
132 calibrate the model parameters.

Abbreviation	Definition	Unit	Equation or source	Reference
LE	Latent heat flux, 'LE_F_MDS' in FLUXNET2015 dataset or 'LE_f' in LaThuile dataset	$MJ·m^{-2}·d^{-1}$	EC observations	See Table S1
NEE	Net ecosystem exchange, 'NEE_VUT_REF' or 'NEE_f'	$gC·m^{-2}·d^{-1}$	EC observations	
Precip	Precipitation, 'P_F' or 'precip'	mm	EC observations	

QA	Quality flag of the variable from EC measurement, e.g., 'SW_IN_F_QC' is the QA of global radiation in FLUXNET2015 dataset, and 'Rg_fqcOK' is QA of that in LaThuile dataset.	Unitless (0-1)	FLUXNET dataset	(Pastorello et al., 2020)
QC	Quality flags for all the reflectance of MCD43A4 product	Unitless	MCD43A2 quality assessment product	(Schaaf & Wang, 2015)
R_g[a]	Global radiation, 'SW_IN_F' or 'Rg_f'	MJ·m^{-2}·d^{-1}	EC observations	
R_p[a]	Potential radiation, 'SW_IN_POT' or 'Rg_pot'	MJ·m^{-2}·d^{-1}	EC observations	See Table S1
R_n[a]	Net radiation, 'NETRAD' or 'Rn_f'	MJ·m^{-2}·d^{-1}	EC observations	
r_{red}	Reflectance at red band	Unitless (0-1)	MCD43A4 version 6 Nadir BRDF-Adjusted Reflectance product	(Schaaf & Wang, 2015)
r_{nir}	Reflectance at near-infrared band	Unitless (0-1)	As above	
T[a]	Air temperature, 'TA_F' or 'Tair_f'	°C	EC observations	See Table S1
VPD[a]	Vapor pressure deficit, 'VPD_F' or 'VPD_f'	kPa	EC observations	See Table S1
CI	Cloudiness index	Unitless (0-1)	$1-R_g/R_p$	(Fu & Rich, 1999; Turner et al., 2006)
CO_2	Atmospheric CO_2 concentration	ppm	Observations by NOAA/ESRL. The global annual mean atmospheric CO_2 concentration was converted to daily time steps using a linear interpolation function	www.esrl.noaa.gov/gmd/ccgg/trends/
ET_{obs}[d]	Evapotranspiration	mm	converted from LE using a latent heat of vaporization changing with T	(Henderson - Sellers, 1984)
PET	Potential ET	mm	Estimated using Rn and T	(Priestley & Taylor, 1972)

GPP_{obs}[d]	Gross primary productivity, 'GPP_NT_VUT_REF' or 'GPP_f'	$gC·m^{-2}·d^{-1}$	Estimated from NEE using the night-time partitioning method	(Reichstein et al., 2005)
NDVI[c]	MODIS-based Normalized differential vegetation index	Unitless (-1-1)	$\frac{r_{nir}-r_{red}}{r_{nir}+r_{red}}$	(Rouse et al., 1974)
PAR	Photosynthetically active radiation	$MJ·m^{-2}·d^{-1}$	$R_g \times 0.45$	(Running & Zhao, 2015; Weiss & Norman, 1985)
WAI	Water availability index	mm	Estimated using Precip and PET, with two site-level calibrated parameters	See the algorithm of WAI in (Boese et al., 2019; Tramontana et al., 2016) and detailed calibration process in section S1 in (Bao et al., 2022)
W	Soil water supply	Unitless (0-1)	$W = \min(1, WAI/AWC)$	(Bao et al., 2022)
σ_{LE}	Random uncertainty of ET, 'LE_RANDUNC' or 'LE_fsd_UncNew_fullDay_m1'	$MJ·m^{-2}·d^{-1}$	Standard deviation of LE	(Pastorello et al., 2020)
σ_{NEE}	Random uncertainty of GPP, 'NEE_VUT_REF_RANDUNC' or 'NEE_fsd_UncNew_fullDay_m1'	$gC·m^{-2}·d^{-1}$	Standard deviation of NEE	
FAPAR[b]	Fraction of absorbed PAR	Unitless (0-1)	$\begin{cases} = NDVI\ (NDVI>0) \\ = 0\ (NDVI\leq 0) \end{cases}$	(Myneni et al., 1997)

Note. All the above variables are at the daily scale;

[a]The gaps in the R_g, R_p, R_n, T, and VPD were filled using machine-learning-based downscaling (Besnard et al., 2019) of gridded product from CRUNCEP (Viovy, 2018);

[b]The linear relationship between FAPAR and NDVI was assumed according to (Myneni et al., 1997).

[c]The gaps in NDVI was filled using FluxnetEO dataset (Walther et al., 2022). The time-series NDVI were filtered by Savitzky-Golay filter (window size was eleven and polynomial order was three) (Savitzky & Golay, 1964).

[d]Since GPP_{obs} is estimated from NEE and ET_{obs} is estimated from LE, the QA of GPPobs and ETobs are represented by QA of NEE ('NEE_VUT_REF_QC' or 'NEE_fsd_UncNew_fullDay_m1') and LE ('LE_F_MDS_QC' or 'LE_fsd_UncNew_fullDay_m1'), respectively.

133 **Table S3.** Eddy covariance flux site list used in this study. The latitude (Lat), longitude (Lon)
134 and plant functional types (PFT) are collected from FLUXNET website. The data length differs

across site and is determined by the years between 'data start' and 'data end'. The climate type is extracted from the Koeppen-Geiger climate classification map (at 5 arc min; Rubel et al., 2017). The elevation is collected from the site ancillary information, papers and satellite images (see the footnote below the table). The site group refers to the group number of each site used to validate the training result.

Class name	Short names	Definitions	References
PFT	PFT	Plant functional types	See Table S1, eleven types in total
Clim	Clim	Koeppen-Geiger climate classification types	See Table S1, five main climate types and fourteen specific classification types in total
BioClim	BIO1	Annual Mean Temperature	Calculated based on the ANUCLIM algorithm (Xu & Hutchinson, 2011) using CRUNCEP dataset (Viovy, 2018) from 1986-2015.
	BIO2	Mean Diurnal Range (Mean of monthly maximum temperature minus minimum temperature)	
	BIO3	Isothermality (BIO2 divided by BIO7 and 100)	
	BIO4	Temperature Seasonality (standard deviation of temperature multiply with 100)	
	BIO5	Max Temperature of Warmest Month	
	BIO6	Min Temperature of Coldest Month	
	BIO7	Temperature Annual Range (BIO5 minus BIO6)	
	BIO8	Mean Temperature of Wettest Quarter	
	BIO9	Mean Temperature of Driest Quarter	
	BIO10	Mean Temperature of Warmest Quarter	
	BIO11	Mean Temperature of Coldest Quarter	
	BIO12	Annual Precipitation	
	BIO13	Precipitation of Wettest Month	
	BIO14	Precipitation of Driest Month	
	BIO15	Precipitation Seasonality (Coefficient of Variation)	
	BIO16	Precipitation of Wettest Quarter	
	BIO17	Precipitation of Driest Quarter	
	BIO18	Precipitation of Warmest Quarter	
	BIO19	Precipitation of Coldest Quarter	
	AI1	Mean annual aridity index (ratio between mean annual precipitation and potential evapotranspiration)	Calculated using the CRUNCEP dataset from 1986-2015
	AI2	Seasonality of aridity index (standard deviation of mean monthly aridity index)	

VIF	VIF1	Annual mean EVI (enhanced vegetation index)	Calculated based on the bioclimatic variables (BIO1-BIO11) algorithm using the gap-filled Landsat-based EVI (Walther et al., 2022) from 1986-2015
	VIF2	Mean monthly EVI range	
	VIF3	Mean EVI variability (VIF2 divided by VIF7)	
	VIF4	EVI seasonality (standard deviation of EVI)	
	VIF5	Max EVI of Warmest Month	
	VIF6	Min EVI of Coldest Month	
	VIF7	Annual EVI Range (BIO5 minus BIO6)	
	VIF8	Mean EVI of Wettest Quarter	
	VIF9	Mean EVI of Driest Quarter	
	VIF10	Mean EVI of Warmest Quarter	
	VIF11	Mean EVI of Coldest Quarter	
NPdep	$Ndep_{NHX}$	Average atmospheric nitrogen deposition (NH_3 and NH_4)	Extracted from the product of the atmospheric chemistry transport model TM3 (Wang et al., 2017)
	$Ndep_{NOY}$	Average atmospheric nitrogen deposition (NO and NO_2)	
	Pdep	Average atmospheric phosphorus deposition	
Soil	BDRICM	Depth to bedrock (R horizon) up to 200 cm	Extracted from the Soil Grids product (Poggio et al., 2021)
	BDRLOG	Probability of occurrence (0-100%) of R horizon	
	BDTICM	Absolute depth to bedrock (in cm)	
	BLDFIE	Bulk density (fine earth) in kg/m3 at depth 0.00 m	
	CECSOL	Cation exchange capacity of soil in cmol/kg at depth 0.00 m	
	CLYPPT	Clay content (0-2 micro meter) mass fraction in % at depth 0.00 m	
	CRFVOL	Coarse fragments volumetric in % at depth 0.00 m	
	ORCDRC	Soil organic carbon content (fine earth fraction) in g/kg at depth 0.00 m	
	PHIHOX	Soil pH*10 in H2O at depth 0.00 m	
	PHIKCL	Soil PH*10 in KCl at depth 0.00 m	
	SLTPPT	Silt content (2-50 micro meter) mass fraction in % at depth 0.00 m	
	SNDPPT	Sand content (50-2000 micro meter) mass fraction in % at depth 0.00 m	
	AWCh1	Derived available soil water capacity (volumetric fraction) with FC = pF 2.0 for	

	depth 0 cm
AWCh2	Derived available soil water capacity (volumetric fraction) with FC = pF 2.3 for depth 0 cm
AWCh3	Derived available soil water capacity (volumetric fraction) with FC = pF 2.5 for depth 0 cm
WWP	Derived available soil water capacity (volumetric fraction) until wilting point for depth 0 cm
AWCtS	Saturated water content (volumetric fraction) teta-S for depth 0 cm

140