Agricultural and Forest Meteorology 313 (2022) 108761

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

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

Comparison of big-leaf and two-leaf light use efficiency models for GPP simulation after considering a radiation scalar

Xiaobin Guan [a,b], Jing M. Chen [b,c,*], Huanfeng Shen [a,d], Xinyao Xie [e], Jianbo Tan [f]

[a] School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China
[b] Department of Geography and Planning, University of Toronto, Toronto ON M5S3G3, Canada
[c] School of Geographical Sciences, Fujian Normal University, Fuzhou 350117, China
[d] Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China
[e] Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610041, China
[f] School of Traffic and Transportation Engineering, Changsha University of Science & Technology, Changsha 410114, China

ARTICLE INFO

Keywords:
Gross primary productivity
Light use efficiency model
Big-leaf
Two-leaf
Radiation scalar

ABSTRACT

Light use efficiency (LUE) models, mainly including the big-leaf (BL) and two-leaf (TL) categories, are efficient approaches to simulate gross primary productivity (GPP). Recently, a TL LUE model considering a radiation scalar (RTL-LUE) was developed, which improved the GPP simulation and unified the same model structure with the BL models using only one maximum LUE. However, whether the radiation scalar is suitable for BL models is still unknown yet, and the core parameters of maximum LUE in the BL and TL models have not been fairly compared before. In this study, we step forward to modifying the MOD17 model using the radiation scalar (RMOD17), and compare the GPP simulations and maximum LUEs in BL (MOD17, RMOD17) and TL (RTL-LUE) models at global 169 FLUXNET eddy covariance (EC) sites. Results indicate that the GPP estimation from RMOD17 (R^2=0.72) matches better with EC GPP than those from the original MOD17 (R^2=0.65), because the radiation scalar effectively corrects the underestimations or overestimations in low or high photosynthetically active radiation (PAR) ranges caused by the radiation-independent LUE in MOD17. The RTL-LUE can further improve the accuracy (R^2=0.74) by alleviating the GPP underestimation of BL RMOD17 in high productivity ranges, which is mainly caused by the insufficient consideration of the shaded leaves contributions. The maximum LUE from RMOD17 and RTL-LUE both show more reasonable values and lower diurnal variations than MOD17, and further analysis proves the radiation scalar is the main reason for this. Besides, RTL-LUE presents a lower model parameter sensitivity compared to RMOD17, indicating that TL is a more robust strategy than BL to simulate GPP. This study highlights the importance of considering a radiation scalar and two-leaf strategy in GPP simulation in LUE models, to better describe the multi-order impacts of radiation on vegetation photosynthesis.

1. Introduction

The terrestrial carbon cycle is a critical component of the global carbon budget (Friedlingstein et al., 2019; Marcott et al., 2014). Gross primary productivity (GPP), describing the overall carbon fixation of plant photosynthesis at the ecosystem scale, is a principal indicator for monitoring the status and changes of the terrestrial carbon cycle (Chen et al., 2019; Monteith, 1972). Accurate simulation of GPP at regional or global scales is essential to assess the global carbon balance and understand the interactions between terrestrial ecosystems and global climate change (Running et al., 2004; Wang et al., 2020).

There are many approaches to estimate GPP at present, which can be mainly divided into *in-situ* measurement-based and model-based estimates (Xie et al., 2020; Yuan et al., 2014). Eddy covariance (EC) flux systems are the mainstream *in-situ* measurements for the net ecosystem exchange (NEE), which can be used to derive GPP with satisfactory accuracies (Baldocchi et al., 2001; Pastorello et al., 2020). Although regional and global EC networks (i.e., FLUXNET) have already been established and significantly contributed to advancing our understanding of vegetation photosynthesis processes at the canopy scale, they are still limited by the spatial footprint varying from tens of meters to several kilometers (Pastorello et al., 2020; Yu et al., 2006). Ecosystem models are efficient methods to estimate regional or global GPP, mainly including process-based models and light use efficiency (LUE) models

* Corresponding author.
E-mail address: jing.chen@utoronto.ca (J.M. Chen).

https://doi.org/10.1016/j.agrformet.2021.108761
Received 6 July 2021; Received in revised form 22 October 2021; Accepted 28 November 2021
Available online 8 December 2021

(He et al., 2018b; Xie et al., 2020). Compared to the massive parameters and data input requirements in process-based models, the simple model structure and fewer data inputs in LUE models make it much easier and more popular to be adopted in large-scale applications (Dong et al., 2015; Haxeltine and Prentice, 1996). LUE models can also be conveniently driven by the satellite vegetation index (VI) data, i.e., normalized difference vegetation index (NDVI) and leaf area index (LAI), ensuring its applicability in regional or global GPP simulations worldwide (Guan et al., 2019; Hilker et al., 2008; Yang et al., 2013).

In recent decades, numerous LUE models have been developed for GPP calculation, mainly on the basis of big-leaf (BL) or the two-leaf (TL) assumptions (Bai et al., 2018; Zhou et al., 2016). Although the BL and TL models both calculate GPP as the product of the absorbed photosynthetically active radiation (APAR) and actual LUE, significant differences still exist in the treatments of the vegetation canopy and APAR calculations (Xie and Li, 2020; Zhou et al., 2016). The BL models treat the entire vegetation canopy as a big extended leaf, assuming that all the leaves in the canopy are the same and absorb the direct and diffuse radiation without differences for photosynthesis (McCallum et al., 2013). The typical BL models include MOD17 (Running et al., 2004), EC-LUE (Yuan et al., 2007), Vegetation Photosynthesis Model (VPM) (Xiao et al., 2004), and Carnegie-Ames-Stanford approach (CASA) model (Potter et al., 1993). These models calculate APAR directly from the incoming photosynthetically active radiation (PAR) and the remote sensing VI-based fraction of photosynthetically active radiation (fPAR), and the major disparities of these models are the hydrothermal environmental scalars used to constraint the maximum LUE (Dong et al., 2015; Yuan et al., 2014). However, many studies indicated that the APAR and LUE of sunlit and shaded leaves in a canopy should be different due to their different exposure levels to sunlight, and the same treatment of sunlit and shaded leaves in BL models may induce large bias in GPP estimation (Alton et al., 2007; Braghiere et al., 2019; De Pury and Farquhar, 1997; He et al., 2013; Zhang et al., 2012). Sunlit leaves, which absorb direct and diffuse radiation simultaneously, are easily light-saturated with low LUE. In contrast, shaded leaves only absorb diffuse radiation, and thus their photosynthesis is usually limited by low APAR, and showing higher LUE than sunlit leaves (Chen et al., 1999; Farquhar et al., 1980; Liu et al., 1997). To address the above issue, TL LUE models were developed by separately calculating the LAI and absorbed PAR of sunlit and shaded leaves based on the approach in the Boreal Ecosystem Productivity Simulator (BEPS) (Liu et al., 1997), such as the TL-LUE model (He et al., 2013), MTL-LUE model (Xie and Li, 2020), and the revised EC-LUE model (Zheng et al., 2020). These TL models have been proved to outperform the corresponding BL models, suggesting that the separate treatment of sunlit and shaded leaves is essential in GPP modeling (Zan et al., 2018; Zhou et al., 2016). However, these TL models still neglect the integral impacts of radiation intensity on LUE, and differentiate the LUE of sunlit and shaded leaves only by the different two constants of maximum LUEs, which may also lead to uncertainties in GPP estimation.

Recently, a two-leaf LUE model considering a radiation scalar (RTL-LUE) is developed, in which the same maximum LUE is assigned to the sunlit and shaded leaves, and the differences in actual LUE between them are characterized and constrained by the respective radiation intensity they received (Guan et al., 2021). The RTL-LUE model unifies the same model structure with the BL models with only one parameter of maximum LUE for all the leaves in the canopy, rather than assigning two different values for sunlit and shaded leaves in traditional TL models. In fact, it is the actual LUEs that are different for sunlit and shaded leaves mainly caused by the disparate radiation intensity they received, but the maximum LUE between the two groups of leaves should be almost the same (Chen et al., 2012; De Pury and Farquhar, 1997; Ogren, 1993). The maximum LUE is only decided by the physiological traits of the leaf itself, which is nearly the same in a canopy, either for sunlit or shaded leaves (Chen, 1996; Hunt Jr and Running, 1992; Koyama and Kikuzawa, 2010; Liu et al., 1997). While a light response curve normally shows that the leaf photosynthesis rate increases rapidly in low radiation intensity but slowly when radiation intensity is high, suggesting that LUE in shaded leaves should be higher than that in sunlit leaves (Harbinson, 2012; Ogren, 1993). Therefore, it is more reasonable to assign the same maximum LUE for all the leaves in a canopy, and constrain it to the actual level respectively by the radiation scalars of sunlit and shaded leaves according to the light response curve (Leverenz, 1987; McCallum et al., 2013; Propastin et al., 2012). Validation from global EC sites proved the RTL-LUE model could significantly improve the GPP simulation by alleviating the overestimation in the original TL-LUE model under high incoming PAR conditions, highlighting the necessity of incorporating a radiation scalar into TL LUE models (Guan et al., 2021). However, it is still unknown whether the radiation scalar is suitable for BL LUE models, in which even the effects of radiation distribution within the canopy on GPP simulation are insufficiently considered. Furthermore, the core parameter of maximum LUE in BL and TL models has not yet been directly compared previously, because of the different number of parameters (i.e., 2 in TL models for sunlit and shaded leaves, and only 1 in BL models for all the leaves). Since the RTL-LUE model unifies the same model structure as the BL models, it provides the possibility to directly explore the difference in maximum LUE in the BL and TL LUE models.

In this context, it is necessary to compare the BL and TL LUE models for GPP simulation after both using a radiation scalar, and to clarify the contribution of the TL strategy and radiation scalar to describe the impacts of radiation on vegetation photosynthesis. Since the RTL-LUE model stemmed from the MOD17 model, the BL MOD17 model is selected for comparison, and a revised model is modified with the radiation scalar. The main purposes of this study are: (1) to prove whether the radiation scalar is suitable for the big-leaf models; (2) to discuss the contribution of two-leaf strategy in GPP simulation after using the radiation scalar; (3) to compare the difference of maximum LUE in big-leaf and two-leaf models.

2. Data and methods

2.1. Data

EC data from the FLUXNET2015 dataset (www.fluxdata.org) is applied to calibrate the parameters and validate the GPP simulations from the different LUE models (Baldocchi et al., 2001; Pastorello et al., 2017, 2020). Overall, 169 sites (Table S1) distributed across the globe are selected on the basis of more than one year of valid half-hourly GPP, incoming solar radiation, air temperature, and VPD during 2001–2015, with 1191 site-years data in total. These sites cover 12 vegetation types in the International Geosphere-Biosphere Program (IGBP) classification system, including: 19 deciduous broadleaf forest (DBF); 1 deciduous needleleaf forest (DNF); 12 evergreen needleleaf forest (ENF); 41 evergreen broadleaf forest (EBF); 8 mixed forest (MF); 32 grass (GRA); crop (CRO); 1 closed shrub (CSH); 11 open shrub (OSH); 18 wetlands (WET); 7 savannas (SAV); 6 woody savannas (WSA). Furthermore, the Global LAnd Surface Satellite (GLASS) LAI product during 2001 and 2015 is derived to drive all the LUE models, in order to ensure the fairness of the comparison between BL and TL models using the same key driven data. This LAI product is provided in an 8-days composition with a spatial resolution of 1 km, which has been shown to have a satisfactory performance in GPP simulation (Xiao et al., 2016; Xie et al., 2019). A modified Whittaker trend filtering method is used to further correct the residual noises that existed in the LAI time series (Chu et al., 2021; Li et al., 2020).

2.2. The MOD17 algorithm

The MOD17 algorithm is developed based on the radiation conversion efficiency concept of Monteith (1972). GPP is calculated as (Running et al., 2004):

$$GPP = \varepsilon_{MOD} \times PAR \times fPAR \times f(VPD) \times g(T_a) \quad (1)$$

where ε_{BL} is the maximum LUE varied with different vegetation types;*fPAR* is the fraction of PAR absorbed by the canopy; $f(VPD)$ and $g(T_a)$ are the scalars of *VPD* and minimum temperature.

$$fPAR = 1 - e^{-k \times LAI} \quad (2)$$

$$f(VPD) = \begin{cases} 0 & VPD \geq VPD_{\max} \\ \dfrac{VPD_{\max} - VPD}{VPD_{\max} - VPD_{\min}} & VPD_{\max} < VPD < VPD_{\max} \\ 1 & VPD \leq VPD_{min} \end{cases} \quad (3)$$

$$g(T_a) = \begin{cases} 1 & T_a \geq T_{\max} \\ \dfrac{T_a - T_{\min}}{T_{\max} - T_{\min}} & T_{\min} < T_a < T_{\max} \\ 0 & T_a \leq T_{\min} \end{cases} \quad (4)$$

where k is the light extinction coefficient and set as 0.5; *LAI* is the leaf area index of the whole canopy; *VPD* is the daily average VPD at daytime and T_a is the daily minimum temperature; $VPD_{max}, VPD_{min}, T_{max}$, and$T_{min}$ are parameters specific to vegetation types.

2.3. Modified MOD17 model considering scalar of radiation

After considering the scalar of radiation, the MOD17 algorithm is modified and named as the RMOD17 model. The modified calculation of GPP as follows:

$$GPP = \varepsilon_{RBL} \times PAR \times fPAR \times f(VPD) \times g(T_a) \times f(PPFD) \quad (5)$$

where ε_{RBL} is the maximum LUE for RMOD17 model; $f(PPFD)$ is the scalar of the radiation on vegetation photosynthesis that reduces LUE from its prescribed maximum to a realistic value in a way similar to the other environmental scalars, i.e., $f(VPD)$ and $g(T_a)$. A reciprocal function is used for *f(PPFD)* according to the format used in previous ecosystem models (e.g., BEPS and BIOME-BGC) to scale stomatal conductance (Liu et al., 1997; Running and Coughlan, 1988), as follow:

$$f(PPFD) = \frac{b}{a \times PPFD + b} \quad (6)$$

where a and b are the coefficients determining the relationship between light intensity and LUE. The b is set as a constant with the value of 1 mol m^{-2} hh^{-1} referring to the value set in previous models (Chen et al., 1999; Mäkelä et al., 2008), and only the parameter a is used to control the response of LUE to PPFD and to minimize the number of parameters that need to be parameterized. *PPFD* is the photosynthetic photon flux density (mol m^{-2} hh^{-1}) representing the light intensity in a specific area and percent time. In the big-leaf model, *PPFD* is approximated by the amount of PAR multiply the constant PAR-energy ratio of 4.55 mol/MJ (Chen et al., 1999).

2.4. The RTL-LUE model

The RTL-LUE model also stems from the MOD17 algorithm, and improves the calculation of canopy APAR by considering the difference between sunlit and shaded leaves according to the BEPS model (Chen et al., 1999; Liu et al., 1997). The RTL-LUE model assumes the same maximum LUE for all the leaves in the canopy, and the differences in real LUE between sunlit and shaded leaves are decided by the radiation scalars ranging from 0 to 1. According to the light response curve, shaded leaves only exposed to diffuse radiation would be nearer to the light compensation point and thus have a radiation scalar with higher values than sunlit leaves. In this way, higher LUE values are obtained for shaded leaves than sunlit leaves (Chen et al., 1999; Harbinson, 2012).

The calculation of GPP in RTL-LUE is modified as follows (Guan et al., 2021):

$$GPP = \varepsilon_{RTL} \times (f(PPFD_{su}) \times APAR_{su} + f(PPFD_{sh}) \times APAR_{sh}) \times f(VPD) \times g(T_a) \quad (7)$$

where ε_{RTL} is the maximum LUE of all the leaves within the canopy; $f(PPFD_{su})$ and $f(PPFD_{sh})$ are the scalar of the radiation on vegetation photosynthesis in sunlit and shaded leaves; $f(VPD)$ and $g(T_a)$ are calculated the same as it in the MOD17 algorithm with the same meaning; $APAR_{su}$ and $APAR_{sh}$ are the incoming PAR absorbed by sunlit and shaded leaves and calculated as follows:

$$APAR_{su} = (1-\alpha) \times \left[\frac{PAR_{dir} \times \cos(\beta)}{\cos(\theta)} + \frac{PAR_{dif} - PAR_{dif,u}}{LAI} + C \right] \times LAI_{su} \quad (8)$$

$$APAR_{sh} = (1-\alpha) \times \left[\frac{PAR_{dif} - PAR_{dif,u}}{LAI} + C \right] \times LAI_{sh} \quad (9)$$

where α is the canopy albedo that obtained related to vegetation types (Table. S2); β is the mean leaf-sun angle and the value for a canopy with spherical leaf angle distribution is set as 60°; θ is the solar zenith angle;PAR_{dif}, PAR_{dir} $PAR_{dif,u}$, and C are the diffuse, direct components of incoming PAR, the diffuse PAR under the canopy, and the multiple scattering of direct radiation, respectively, which are empirically partitioned according to the clearness index (*CI*) following Chen et al. (1999); LAI_{su} and LAI_{sh} are the LAI of sunlit and shaded leaves, separated according to the canopy LAI, clumping index (Ω) and solar zenith angle according to Chen et al. (1999). The LAI_{su} can be computed as follow, and the LAI_{sh} is the residual of *LAI* minus LAI_{su}.

$$\mathrm{LAI_{su}} = 2 \times \cos(\theta) \times \left(1 - exp\left(-0.5 \times \Omega \times \frac{LAI}{\cos(\theta)} \right) \right) \quad (10)$$

2.5. Model parameterization and evaluation

2.5.1. Model parameterization

Similar to previous studies, the parameters $VPD_{max}, VPD_{min}, T_{max}, T_{min}, \alpha, \Omega$and$a$ are empirically set both in the MOD17, RMOD17, and RTL-LUE models according to the previous literature (Guan et al., 2021; He et al., 2013; Zhou et al., 2016), as shown in Table S2. There is a respective parameter of maximum LUE that needs to be further optimized in three models, i.e., ε_{BL} in the MOD17 model, ε_{RBL} in the RMOD17 model, and ε_{RTL} in the RTL-LUE model. The parameters of the maximum LUE in the three models are all optimized site by site, so a set of parameter values can be obtained for each site within a given PFT, revealing the variance of the parameter (Huang et al., 2021). Similar to previous studies, the mean values for the sites within the same PFT are obtained as the parameter for this PFT (Xie and Li, 2020; Zhou et al., 2016). Only one site-year data are randomly selected for parameterization at each site, in order to compare the BL and TL models using as much data as possible. The shuffled complex evolution method of the University of Arizona was employed to implement the optimization (Duan et al., 1992; Zhou et al., 2016), which evaluates the model performance with the agreement index (d):

$$d = 1 - \frac{\sum_{i=1}^{N} (P_i - \mathrm{O}_i)^2}{\sum_{i=1}^{N} (|P_i - \overline{O}| + |O_i - \overline{O}|)^2} \quad (11)$$

where N is the total simulated experiment data point; P_i and O_i represent the predicted GPP and observed EC GPP, respectively; and $\overline{O}$ and $\overline{P}$ are the mean values of observations and predictions for all experimental data points.

2.5.2. Model evaluation

The accuracies of the GPP simulations from the three models were assessed against the EC GPP, both using the half-hourly data, the composited 8-day data, and the composited yearly data. The optimized maximum LUE for the three models were compared both in the spatial heterogeneity and the diurnal variation. In addition, four indexes were applied for quantitative accuracy evaluation, including the coefficient of determination (R^2), the root-mean-square error (RMSE), the mean predictive error (bias), and the Coefficient of Variation (CV):

$$R^2 = \left(\frac{\sum_{i=1}^{N}(P_i - \overline{O})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{N}(P_i - \overline{P})^2}\sqrt{\sum_{i=1}^{N}(O_i - \overline{O})^2}} \right)^2 \quad (12)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(P_i - O_i)^2}{N}} \quad (13)$$

$$bias = \frac{\sum_{i=1}^{N}(O_i - P_i)}{N} \quad (14)$$

$$CV = \frac{SD}{Mean} \times 100\% \quad (15)$$

where *SD* is the standard deviation, and *Mean* is the mean value.

3. Results

3.1. Improvements of RMOD17 after considering the scalar of radiation

The half-hourly GPP simulations from the MOD17, RMOD17, and RTL-LUE models are evaluated using the EC GPP from the global 169 sites. As shown in Table 1, the statistics of R^2, RMSE, and bias among 12 different vegetation types are listed, after calculating the mean value of all the sites within the same vegetation type. It can be observed that the accuracy of GPP simulations from the RMOD17 model is better than the results of the original MOD17 model in all of the 12 vegetation types, with higher R^2 or lower RMSE and bias. Overall, the mean R^2 and RMSE for MOD17 model results is 0.57 and 0.13 g C m^{-2} hh^{-1}, respectively, and the corresponding value is 0.60 and 0.11 g C m^{-2} hh^{-1} for the RMOD17 model, with an improvement of 0.03 in R^2 and a reduction of 0.02 g C m^{-2} hh^{-1} in RMSE. Considerable improvements can be found in all of the five forest vegetation types, especially the increase of R^2 reaches 0.05 and 0.04 in the deciduous and evergreen broadleaf forest sites, respectively. The value of R^2 ranges from 0.37 for open shrubs to 0.78 for closed shrubs in the results of the MOD17 model, and the range is from 0.40 to 0.81 for the same two vegetation types in the RMOD17 model. The R^2 and RMSE of half-hourly GPP simulations from the MOD17 and RMOD17 models at each site are further compared using the scatter plots in Fig. 1. Among all the 169 EC sites, there are in total 140 sites showing a reduced RMSE in the results of the RMOD17 model after considering the radiation scalar, and 119 sites have a higher R^2. Although the number of sites for an improved R^2 is not as many as it for a reduced RMSE, a significant difference of 0.06 in the mean value can be observed at these sites between the results of MOD17 (R^2=0.60) and RMOD17 (R^2=0.66).

In order to conduct an overall assessment of GPP simulations from the MOD17 and RMOD17 models, the half-hourly data were further composited into the yearly sum. The density scatterplots of the GPP simulations and EC GPP for all site years are shown in Fig. 2(a) and (b). After the yearly composition, more significant improvements can be observed in the RMOD17 results, with increased R^2 (+0.07) and reduced RMSE (−82.51 g C m^{-2} $year^{-1}$). Due to the accumulation error, many points from MOD17 results show a large discrepancy from EC GPP, while the points in the scatter plot of RMOD17 results are more concentrated near the 1:1 line, especially in the medium and high productivity ranges. The slope of the linear regress line improves from 0.69 to 0.73, which indicates minor systematic biases of the GPP simulations from the RMOD17 model after using the radiation scalar.

3.2. Comparison of the big-leaf RMOD17 model and two-leaf RTL-LUE model

The accuracies of GPP simulations from the big-leaf RMOD17 model and the two-leaf RTL-LUE model are also compared. The two models both use the radiation scalar but treat the vegetation canopy differently. As shown in Table 1, the RTL-LUE model generally shows a better performance than the RMOD17 model almost in all the 12 vegetation types, either with higher R^2 or reduced RMSE. Considerable improvements in R^2 can be found at DBF, DNF, crops, and wet sites. After being composited into yearly data, the GPP simulations from the RTL-LUE model also agree better with EC data than the RMOD17 results (Fig. 2 (b) and (c)), with improvements of R^2 (+0.02) and RMSE (20.73 g C m^{-2} $year^{-1}$). Although the gains are not as significant as those between MOD17 and RMOD17 model induced by the radiation scalar, an obviously better linear regression line can be found in the RTL-LUE results with a slope of 0.85, compared to the value of 0.69 and 0.73 in the MOD17 and RMOD17 results. This indicates that the distribution of GPP simulations from the RTL-LUE model is overall more similar to EC data with a smaller systematic error. The main reason for the improvements

Table 1
Statistic of the half-hour GPP simulation accuracy for different vegetation types.

Vegetation type [a]	MOD17			RMOD17			RTL-LUE		
	R^2	RMSE [b]	bias [b]	R^2	RMSE [b]	bias [b]	R^2	RMSE [b]	bias [b]
DBF	0.66	0.16	0.03	0.71	0.13	0.04	0.74	0.12	0.04
DNF	0.64	0.08	0.04	0.65	0.06	0.03	0.67	0.05	0.02
EBF	0.65	0.15	−0.01	0.69	0.11	−0.01	0.69	0.11	0.02
ENF	0.65	0.13	0.00	0.68	0.10	0.00	0.68	0.09	0.00
MF	0.71	0.19	0.02	0.73	0.13	0.01	0.73	0.15	0.01
GRA	0.50	0.13	0.03	0.53	0.11	0.03	0.53	0.10	0.02
CRO	0.39	0.19	0.05	0.40	0.17	0.05	0.43	0.15	0.03
CSH	0.78	0.09	−0.01	0.81	0.07	0.01	0.82	0.07	0.01
OSH	0.37	0.07	0.03	0.40	0.06	0.03	0.40	0.04	0.01
WET	0.58	0.15	0.08	0.59	0.10	0.03	0.62	0.09	0.03
SAV	0.49	0.15	0.04	0.49	0.13	0.01	0.49	0.11	0.01
WSA	0.52	0.12	0.03	0.52	0.10	0.04	0.54	0.09	0.03
All	0.57	0.13	0.03	0.60	0.11	0.02	0.61	0.10	0.02

[a] DBF: deciduous broadleaf forest; DNF: deciduous needleleaf forest; ENF: evergreen needleleaf forest; EBF: evergreen broadleaf forest; MF: mixed forest; GRA: grass; CRO: crop; CSH: closed shrub; OSH: open shrub; WET: wetlands; SAV: savannas; WSA: woody savannas

[b] the units of RMSE and bias are g C m^{-2} hh^{-1}.

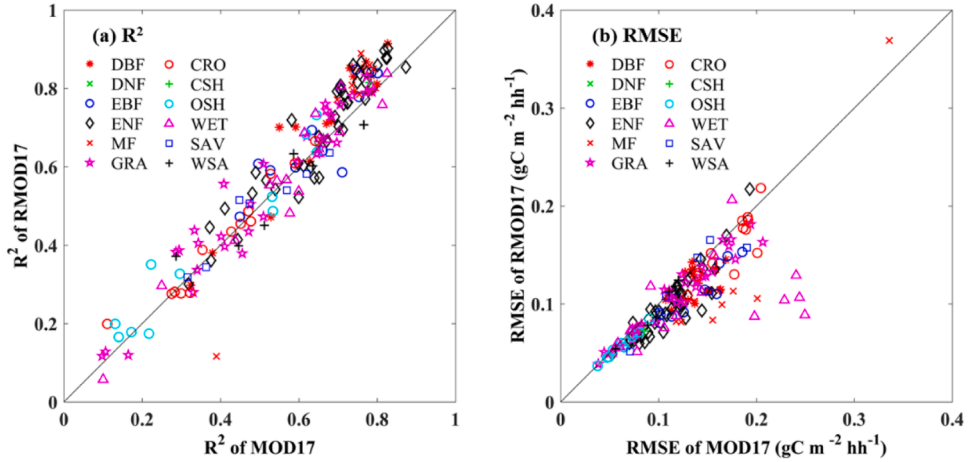

Fig. 1. Comparison of the R^2 and RMSE for GPP simulations from the MOD17 and RMOD17 models against tower measurements (EC GPP) at 169 EC sites.

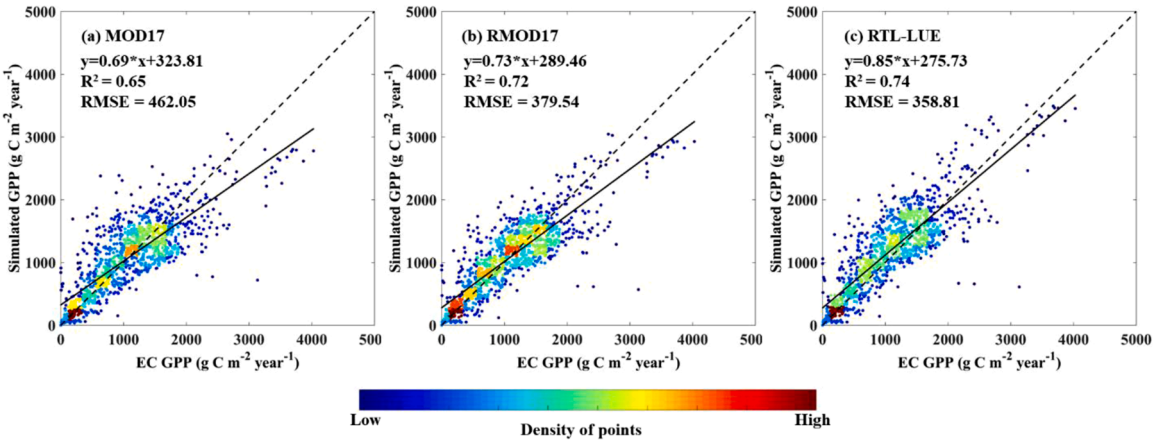

Fig. 2. The relation between the yearly EC GPP and the simulation results from (a) the MOD17 model, (2) the RMOD17 model, and (c) the RTL-LUE model.

is that RTL-LUE can better simulate GPP in high productivity ranges, where underestimation is typically found in the RMOD17 results (Fig. 2 (b)).

In order to further explore the improvements of slope in the RTL-LUE (RTL) model compared to the RMOD17 (RBL) model, the scatterplots of all the 8-days sum data points for the 12 vegetation types are provided in Fig. 3. The slopes of RTL-LUE simulations are much better than the RMOD17 results in all the 12 vegetation types, with the value much nearer to 1, which indicates a better agreement with EC data. Considerable improvements in slope can be found in EBF (+0.18), DBF (+0.16), ENF (+0.15), and wet (+0.23). The data points of results from the RMOD17 (gray) and RTL-LUE (black) models are displayed in different colors, so direct qualitative comparisons can be achieved. The underestimated points from the RMOD17 model in high productivity ranges are overall well raised by the RTL-LUE model, to be more concentrated near the 1:1 line. Especially for the sites covered by five forests and wet types, significant differences can be observed in the results from the two models.

3.3. Sensitivity of GPP simulations to PAR and CI in different models

In order to further investigate the reasons for the different accuracies of GPP simulations from the MOD17, RMOD17, and RTL-LUE models, simulated and EC GPP values are averaged over 0.01 (MJ m^{-2} hh^{-1}) bins of incoming PAR, as well as their differences ($\Delta GPP = GPP_{simulation} - GPP_{EC}$). R^2 between the average binned GPP simulations and the EC data in 12 different vegetation types are listed in Table 2. It can be observed that R^2 values in all three models are improved after binned into different PAR magnitudes in almost all the 12 vegetation types, because the averaging process can eliminate the outliers and abnormal values. GPP simulations from the RTL-LUE model still perform the best for all or individual vegetation types, with a higher R^2 than that of RMOD17 (+0.04) and MOD17 (+0.07). This indicates that the GPP simulations from the RTL-LUE model show more similar trends than those from the other two models with the EC GPP at different PAR intensities.

The variations of $\Delta GPP = GPP_{simulation} - GPP_{EC}$ for 12 typical sites with different vegetation types are selected to show the variation of GPP estimation errors with radiation intensity, as shown in Fig. 4. GPP simulated by the MOD17 model can only approach the EC data under the average incoming PAR conditions, but showing apparent underestimation in low PAR ranges and overestimation in high PAR ranges. It is because the radiation-independent LUE used in the MOD17 model is only suitable for the average incoming PAR conditions, and the actual LUE should decrease with the increase of incoming PAR. As a result, the applied constant LUE will certainly be lower than the actual LUE when the incoming PAR is low, resulting in the underestimation of GPP. The conditions will be opposite when the incoming PAR is high, leading to overestimation. After using the radiation scalar, this problem is mostly overcome in the RMOD17 model, with reduced bias in GPP simulation under different PAR conditions. However, the RMOD17 model tends to underestimate GPP under the high incoming PAR conditions, which is consistent with the result in Figs. 2 and 3, i.e., the underestimation in high productivity ranges. It is mainly caused by the insufficient

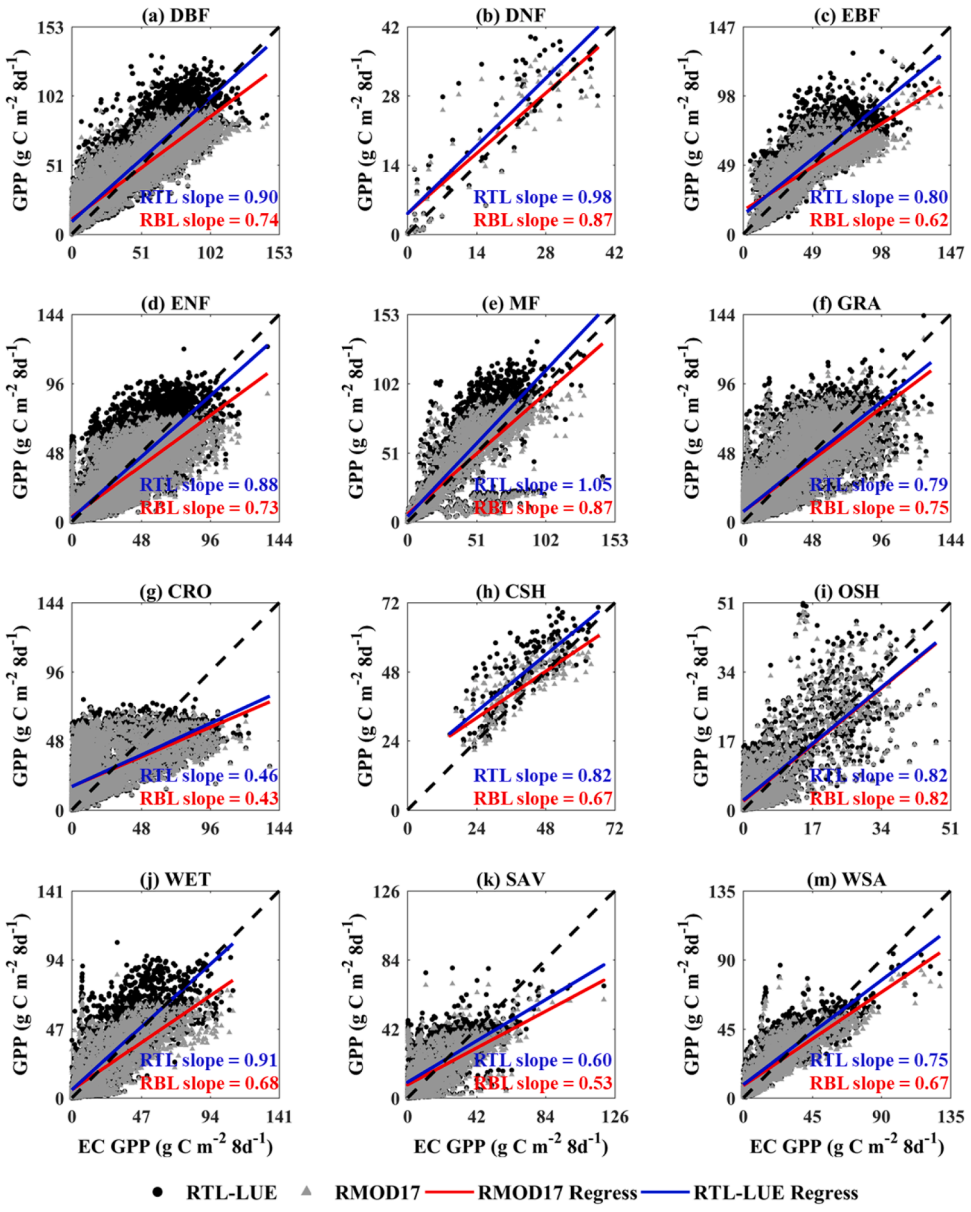

Fig. 3. Scatter plot of GPP simulations from the RMOD17 and RTL-LUE models after composited into 8-day totals. The "RTL slope" denotes the slope of the linear regress line for RTL-LUE model results, and the "RBL slope" is for the results of the big-leaf RMOD17 model.

Table 2
Statistic of the R^2 between GPP simulations from the three models and EC GPP after averaged over 0.01 bins of incoming PAR in different vegetation types.

Vegetationtype [a]	DBF	DNF	EBF	ENF	MF	GRA	CRO	CSH	OSH	WET	SAV	WSA	All
MOD17	0.82	0.78	0.78	0.80	0.86	0.64	0.66	0.91	0.54	0.70	0.69	0.64	0.74
RMOD17	0.82	0.76	0.92	0.86	0.88	0.71	0.70	0.92	0.59	0.71	0.68	0.64	0.77
RTL-LUE	0.89	0.83	0.90	0.87	0.90	0.72	0.74	0.96	0.59	0.79	0.73	0.74	0.81

[a] The abbreviations of different vegetation types are the same as it in Table 1.

consideration of the contribution of shaded leave in the big-leaf RMOD17 model, which has been described in many previous studies (He et al., 2018a; Zhang et al., 2012). Due to the shaded leaves only expose to diffuse radiation, they can still maintain relatively high LUE even with high incident radiation on the canopy. Thus, the big-leaf models using the same average LUE for all leaves in the canopy will certainly lead to a lower LUE and cause an underestimation of GPP. Furthermore, the unreasonably high level of PAR used to represent the average radiation intensity of the canopy would be another reason because the large fraction of shaded leaves in the canopy only receive diffuse radiation. As a result, the RMOD17 model will use a relatively low *f(PPFD),* inducing a lower LUE to underestimate GPP under the high incoming PAR conditions. This problem is alleviated in the RTL-LUE model by separately considering the photosynthesis of sunlit and shaded leaves, so the LUE values of different groups of leaves are individually determined according to their own received radiation intensity. As a result, the RTL-LUE model does not show any significant tendency to overestimate or underestimate GPP under any levels of incoming

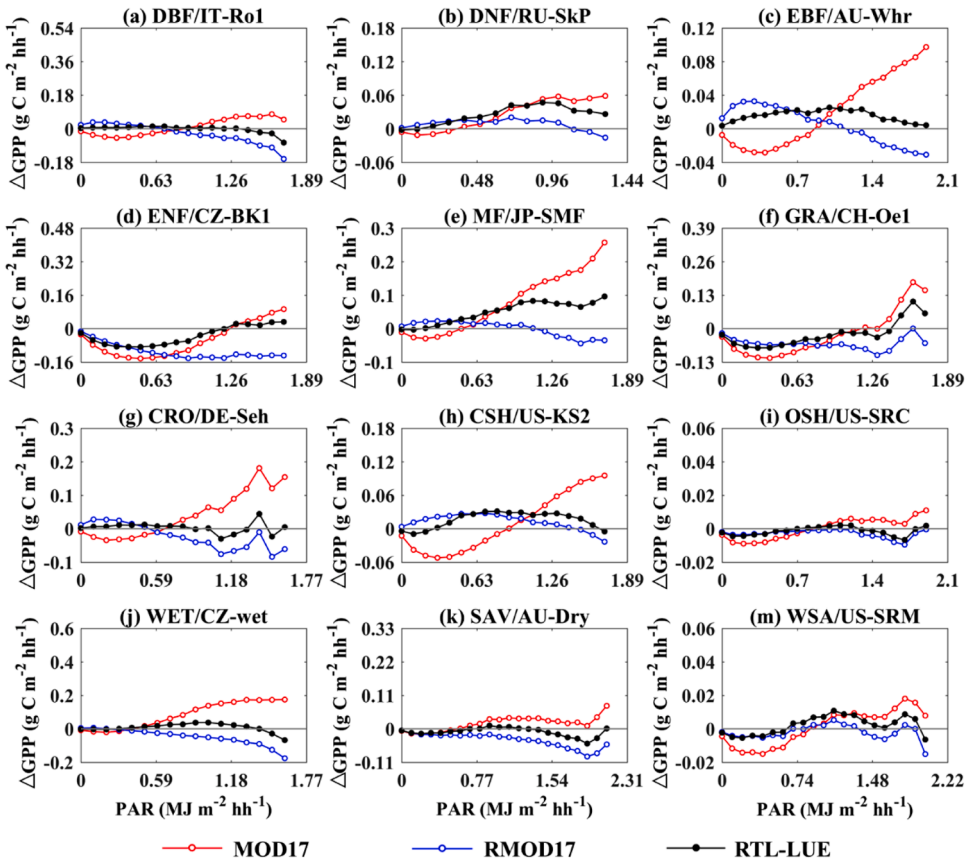

Fig. 4. The sensitivity of GPP simulations to PAR for 12 specific sites with different vegetation types ($\Delta GPP = GPP_{simulation} - GPP_{EC}$). The title of each subplot is named as "vegetation types/site name".

PAR, and the simulated GPP can fit better with the EC data than the other two big-leaf models.

Furthermore, the canopy LUEs (LUEc=GPP/PAR) (He et al., 2013) calculated by the models and EC data are also averaged over 0.05 bins of sky clearness index (CI), as shown in Fig. 5. At all sites, the EC LUEc generally decreases with the increase of CI. This is because the amount of diffuse radiation, which can be used by both sunlit and shaded leaves with high light use efficiency, is negatively correlated with the R. As a result, the higher CI is, the less diffuse radiation is, and the lower LUE of a canopy is. However, the LUEc simulated by the MOD17 model shows weak responses to CI only with similar values to EC LUEc at high CI ranges. This is because the MOD17 model considered neither the impacts of radiation on LUE nor the differences between sunlit and shaded leaves, so the LUE of the model results does not vary with the amount of radiation or the percentage of the diffuse component. As a result, the MOD17 model usually underestimates LUE under low CI conditions, leading to GPP underestimation in low PAR ranges (Fig. 4). After using the radiation scalar, the RMOD17 and RTL-LUE models can generally reproduce the decreasing trends of LUEc with CI shown in EC data. Especially under the high CI conditions (CI> 0.4) with more direct radiation, results from both RMOD17 and RTL-LUE are almost the same as the EC data.

3.4. Differences of parameterized maximum LUE across biomes

The maximum LUE is a crucial parameter in LUE models. Since the differences between maximum LUE in the big-leaf and two-leaf models have not yet been previously investigated, it is necessary to conduct a comparison for the three models. As shown in Table 3, the optimized ε_{BL} in the MOD17 model is the lowest in all the 12 vegetation types, and ε_{RBL} in the RMOD17 model shows the highest value with approximately 2.5 times higher than ε_{BL}. As a result, ε_{RBL} in the RMOD17 model indicates that all vegetation types should have a much higher maximum LUE after considering the radiation scalar. It is because that the ε_{BL} in the original MOD17 model is not the actual maximum LUE but only is an average value constrained by radiation scalar at an intermediate radiation level. The maximum LUE should be the maximum photosynthetic rate for vegetation under the completely ideal environmental conditions, which definitely includes the radiation environment. As a result, ε_{BL} after the constraint by the average radiation scalar will certainly be much lower than ε_{RBL}.

Furthermore, ε_{RBL} in RMOD17 is also higher than ε_{RTL} in RMOD17 in all of the 12 vegetation types. It is mainly caused by the insufficient consideration of the contribution of shaded leaves in the big-leaf RMOD17 model by treating the whole canopy as a big leaf, so a higher maximum LUE is needed to avoid underestimation and approach the data used to optimize it. Besides, the direct use of PAR as the average radiation intensity for the canopy in the big-leaf RMOD17 model would be another reason. Since only sunlit leaves in the canopy can receive the total incoming radiation while shaded leaves can only receive diffuse radiation, the average radiation intensity for a canopy should be much lower than the value directly calculated by PAR. As a result, the RMOD17 model tends to use a generally more strong radiation constraint to LUE with a much lower *f(PPFD)* value, and finally results in the overestimated ε_{RBL}. In contrast, the separate consideration of the radiation intensity incident on sunlit and shaded leaves in the RTL-LUE model can avoid these problems, and shows a more reasonable ε_{RTL} to approximate actual values.

For all of the three models, the greatest values of maximum LUE are found in the crop sites, and the lowest values are found in different vegetation types, which are CSH, SAV, and OSH respectively for ε_{BL}, ε_{RBL}, and ε_{RTL}. The rank of ε_{RBL} and ε_{RTL} among the 12 vegetation types

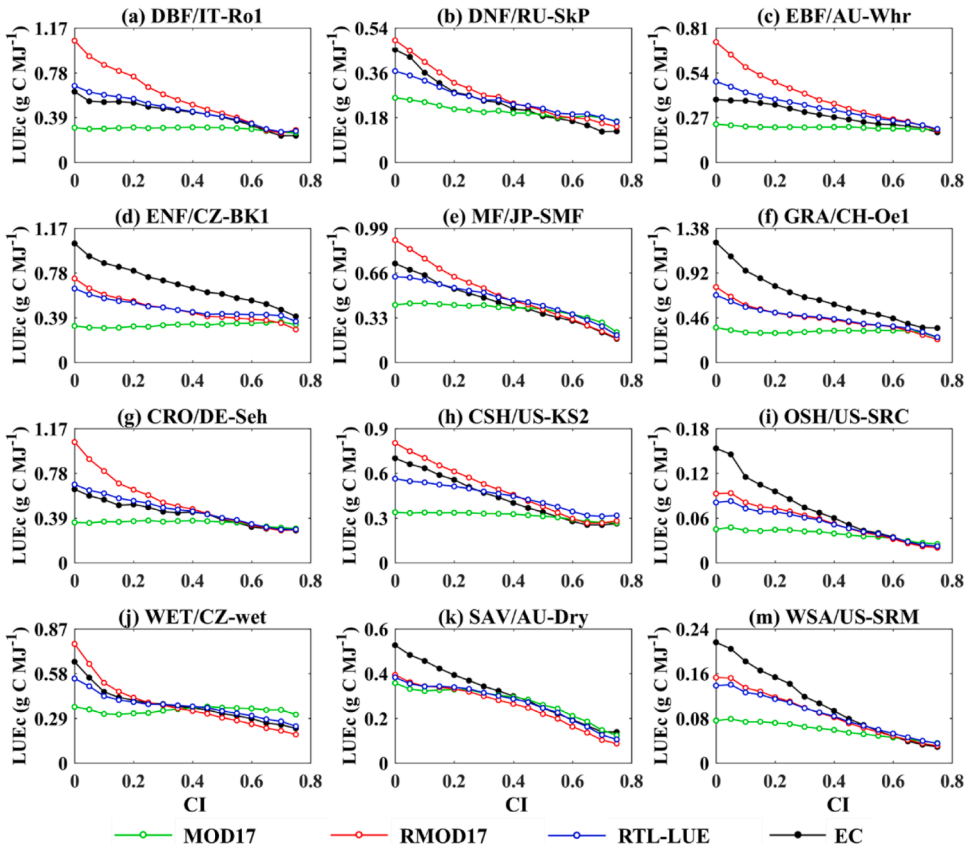

Fig. 5. Comparison of the dependence of canopy light use efficiency (LUEc = GPP/PAR) on clear sky index (CI) averaged over 0.05 bins at 12 specific sites with different vegetation types. The title of each subplot is named as "vegetation types/site name".

Table 3
Mean value, Standard Deviations (SD), and CV of the optimized different maximum LUE for the three models, i.e. ε_{BL}, ε_{RBL}, and ε_{RTL}, for different vegetation types.

Vegetation type [a]	Mean (g C MJ^{-1})			SD (g C MJ^{-1})			CV (%)		
	ε_{BL}	ε_{RBL}	ε_{RTL}	ε_{BL}	ε_{RBL}	ε_{RTL}	ε_{BL}	ε_{RBL}	ε_{RTL}
DBF	1.22	4.75	2.85	0.34	2.28	0.91	28.01	48.06	32.08
DNF	1.01	1.95	1.68	/	/	/	/	/	/
EBF	1.01	3.32	2.16	0.16	1.96	0.83	15.67	59.05	38.46
ENF	1.11	2.67	2.14	0.34	0.87	0.58	30.24	32.63	27.07
MF	1.50	3.34	2.54	0.97	0.93	0.76	64.67	27.84	30.05
GRA	1.52	3.39	2.39	0.74	2.25	1.39	48.31	66.29	58.01
CRO	1.63	5.06	3.32	0.46	4.30	2.04	28.49	85.12	61.33
CSH	0.91	2.21	1.68	/	/	/	/	/	/
OSH	0.93	1.95	1.58	0.56	1.41	1.01	59.87	72.06	63.97
WET	1.71	3.82	2.82	1.04	3.09	1.89	60.66	80.88	67.08
SAV	1.55	1.72	1.60	1.16	1.23	1.07	74.45	71.62	66.85
WSA	1.21	2.47	1.96	0.32	0.97	0.63	26.78	39.15	31.95

[a] The abbreviations of different vegetation types are the same as it in Table 1.

are very similar with an R^2 reaching 0.94, while the rank of ε_{BL} is quite different from ε_{RBL} (R^2=0.25) and ε_{RTL} (R^2=0.41). The CV of ε_{BL} is the lowest among the three maximum LUE except for the MF and SAV sites, and the CV of ε_{RTL} is lower than ε_{RBL} in 11 vegetation types except for the MF sites. It indicates that the spatial variation of ε_{RTL} in the RTL-LUE model would be much lower than ε_{RBL} in the RMOD17 model.

3.5. *Diurnal variation of* ε_{BL}, ε_{RBL}, *and* ε_{RTL}

The LUE models are usually driven by biome specific maximum LUE, so it is essential to clarify the suitability of different model structures using a constant maximum LUE without any temporal changes. Because the maximum LUE is only determined by the physiological traits of the leaf itself, which should not change considerably within a day, there should be no significant diurnal variation for the ε_{BL}, ε_{RBL}, and ε_{RTL}. As a result, it can be a standard to verify the reasonability of the maximum LUE in the three models, and half-hourly ε_{BL}, ε_{RBL}, and ε_{RTL} time series at each site are all calculated using the simple ratio of the EC GPP and the GPP simulations without the maximum LUE (i.e., the maximum LUE sets to 1 in all the three models, denoted as $GPP_{sim\_\varepsilon=1}$). Afterward, the mean value and CV for ε_{BL}, ε_{RBL}, and ε_{RTL}, as well as R^2 between EC GPP and $GPP_{sim\_\varepsilon=1}$ from the three models are calculated for all the 333,919 site-days. Since the different values of maximum LUE in the three models will not influence R^2 between GPP simulations and EC GPP, so R^2 between EC GPP and $GPP_{sim\_\varepsilon=1}$ in each site-days can represent the accuracy of the model simulations. The direct comparisons among the three models are plotted in Fig. 6, and the statistics for 12 vegetation types are listed in Table. 4. As shown in Fig. 6(a) and (d), it can be observed that

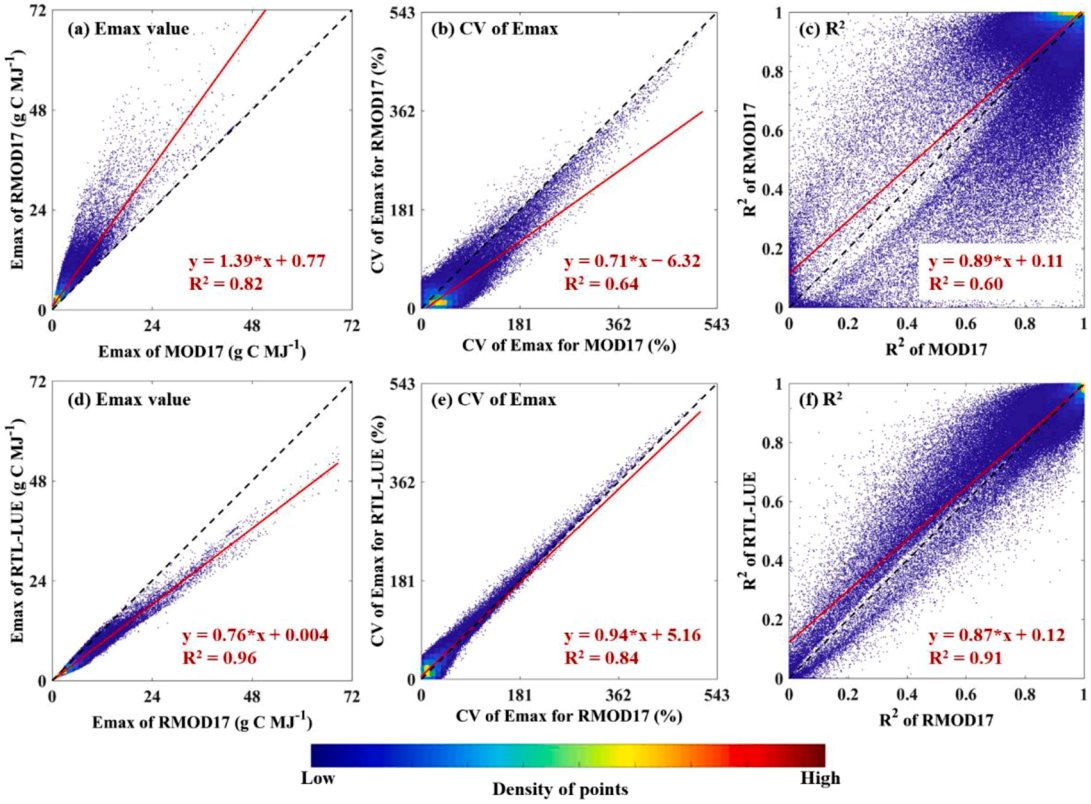

Fig. 6. Comparison of the daily mean maximum LUE value, the CV of maximum LUE diurnal variation, R^2 between $GPP_{sim\_\varepsilon=1}$ (GPP simulation results with maximum LUE set to 1) and EC GPP in all the 333,919 site-days: (a), (b), and (c) for MOD17 and RMOD17 models; (d), (e), and (f) for RMOD17 and RTL-LUE models.

ε_{RBL} is generally much higher than ε_{BL} and ε_{RTL}, and the relationship between ε_{RBL} and ε_{RTL} (R^2=0.96) is obviously much better than it between ε_{RBL} and ε_{BL} (R^2=0.82), which is consistent with the results in Section 3.3. The RMOD17 model can significantly reduce the diurnal variation of maximum LUE with lower CV (Fig. 6(b)) after considering the radiation scalar, and improve R^2 between the diurnal variations of GPP simulations and EC data (Fig. 6(c)). The results in Fig. 6(e) and (f) also indicate the overall improvements of CV and R^2 in the results of the two-leaf RTL-LUE model compared to the big-leaf RMOD17 model. However, different from the scatter plot in Fig. 6(e), the RTL-LUE model shows a slightly higher mean CV value for all site-days (24.1%) than the RMOD17 model (23.2%) in Table 4. It is mainly caused by the higher value of the RTL-LUE model in the abnormal points with very high CV ranges (such as >100%), as demonstrated in Fig. 6(e). The R^2 of RTL-LUE model (0.89) results is higher than it for the RMOD17 model (0.87) both for all or individual vegetation types. The percentage of site-days showing a higher R^2 and lower CV than the results of the MOD17 model are also calculated for both RMOD17 and RTL-LUE models. Obviously higher ratios can be found for both improved R^2 and reduced CV in the results of the RTL-LUE model compared to the RMOD17 model. The percentage for a higher R^2 is 82.0% in the results RTL-LUE model, and it is 71.0% for the RMOD17 model; the percentage for a

Table 4
Statistic of the R^2 for GPP diurnal variation and CV for maximum LUE diurnal variation in different vegetation types.

Vegetation type [a]	R^2 [b]			CV [c] (%)			Percentage [d] (%)		Percentage [d] (%)	
	BL [e]	RBL [e]	RTL [e]	BL [e]	RBL [e]	RTL [e]	R^2-RBL	R^2-RTL	CV-RBL	CV-RTL
DBF	0.91	0.91	0.95	35.36	19.21	15.77	60.31	79.46	75.60	87.69
DNF	0.90	0.90	0.92	37.15	22.91	25.05	59.44	72.90	80.93	91.78
EBF	0.91	0.95	0.96	33.05	13.19	15.77	75.24	92.42	86.33	96.79
ENF	0.90	0.94	0.94	35.60	15.32	21.37	81.03	90.61	88.73	94.86
MF	0.90	0.90	0.91	31.23	14.03	18.94	73.62	83.09	82.50	88.32
GRA	0.84	0.86	0.86	44.90	26.86	30.20	74.23	79.09	85.86	89.76
CRO	0.89	0.90	0.93	37.52	21.86	17.92	60.46	78.49	72.72	87.08
CSH	0.94	0.96	0.97	27.07	11.01	13.03	71.78	90.48	83.63	97.12
OSH	0.76	0.81	0.81	60.10	38.83	39.36	75.56	80.32	89.13	92.36
WET	0.91	0.91	0.94	32.53	18.51	15.54	62.24	76.77	73.24	85.39
SAV	0.75	0.73	0.75	45.38	41.17	40.98	44.43	47.94	71.93	74.91
WSA	0.76	0.73	0.76	46.88	35.35	34.76	45.32	54.13	76.11	87.14
ALL	0.86	0.87	0.89	38.90	23.19	24.06	70.99	81.96	83.10	90.96

[a] The abbreviations of different vegetation types are the same as it in Table 1.
[b] R^2 is the mean for the diurnal variation of EC GPP and GPP simulations with the maximum LUE set to 1.
[c] CV is the mean for the diurnal variation of maximum LUE in each site-days.
[d] percentage is the percentage of site-days that the diurnal variation of GPP simulated by RTL-LUE (RTL) or RMOD17 (RBL) model is better than MOD17 model, i.e., with a higher R^2 and lower CV.
[e] BL, RBL, RTL denotes the results for the MOD17, RMOD17, RTL-LUE model, respectively.

lower CV is 91.0% and 83.1%, respectively. As a result, it can be concluded that the RMOD17 (R^2=0.87, CV=23.2%) and RTL-LUE (R^2=0.89, CV=24.1%) models can significantly reduce the diurnal variation of maximum LUE compared to MOD17 model (R^2=0.86, CV=38.9%), with much lower CV and higher R^2 in most of the site-days.

We also select 12 representative diurnal variation curves from 12 different months and 12 different vegetation types to further qualitatively compare the diurnal variation of maximum LUE in the three models. As shown in Fig. 7, monthly mean values for each half-hour were calculated and plotted to reflect the overall diurnal variation at monthly time steps. It can be observed that ε_{BL} in MOD17 shows evident diurnal variation in all the months and vegetation types, generally with a much lower value at noon and a higher value in the morning and dusk. This is because that the MOD17 model does not consider the impact of radiation on LUE, i.e., *f(FFPD)*, so these impacts are manifested in ε_{BL}. Due to much higher radiation intensity at noon, the LUE at noon should be lower than those in the morning and dusk. The variations of ε_{RBL} and ε_{RTL} in RMOD17 and RTL-LUE are generally much smaller after considering the radiation scalar, with only slight or even no diurnal variations. Due to the constant parameter *a* used for all the sites with the same vegetation types, ε_{RBL} and ε_{RTL} cannot be the ideal constants over the diurnal cycle, but are generally much less variable than ε_{BL}. The ε_{RBL} tends to show an inverse diurnal variation to ε_{BL} with higher values at noon and lower values in the morning and dusk. This is because PAR used to calculate the radiation scalar in the RMOD17 model will overweight the radiation intensity at noon and the contribution of shaded leaves to GPP that is not accounted for in the model, resulting in lower *f(FFPD)* and higher ε_{RBL} to compensate for the contribution of shaded leaves.

3.6. Sensitivity of RMOD17 and RTL-LUE models to parameter *a*

The parameter *a* in Eq. (6) is the only parameter to control the relationship between LUE and light intensity. Since the big-leaf RMOD17 and two-leaf RTL-LUE model treat the canopy differently and use different light intensities (i.e. *PAR* in RMOD17 model and PAR_{su} and PAR_{sh} in RTL-LUE model) to calculate the radiation scalar, it is necessary to discuss the sensitivity of these two models to the parameter *a*. Fig. 8 demonstrates the variation of R^2 between the GPP simulations ($GPP_{sim\_\varepsilon=1}$) and EC GPP with different *a* values ranged from 0 to 4. Due to the RMSE and bias of the simulation results are not only impacted by *a* but also related to the maximum LUE (i.e. ε_{RBL}, and ε_{RTL}), so only the R^2 is demonstrated to show the impact of the parameter *a* on GPP simulation. It can be observed that R^2 generally firstly increases and then decreases with increasing *a*, and an optimal *a* can be found with the highest R^2. For the RMOD17 model, when $a = 0$, *f(PPFD)* is a constant equal to 1, and it becomes the original MOD17 model. In all of the 169 sites, the optimal *a* (with the highest R^2) is found to be higher than 0 for the RMOD17 model, and significant improvements in R^2 can be observed. This analysis further proves the necessity of considering the radiation scalar when simulating GPP. Huge differences can be found between the optimal *a* values (with the highest R^2) for the two models, with 162 out of 169 sites showing a higher value in the RTL-LUE model than it in the RMOD17 model. The mean optimal *a* for different vegetation types are listed in Table. 5, the value of the RTL-LUE model is almost three times higher than that of the RMOD17 model. The optimal *a* here is different from the value used in this study (Table S2) because only R^2 is considered here and all site-years data are involved, while the *a* value in Table S2 is optimized using the agreement index with the random selected one-year data for each site. The optimal *a* values for the EC LUEc and PAR are also calculated based on Eq. (6), which are similar to

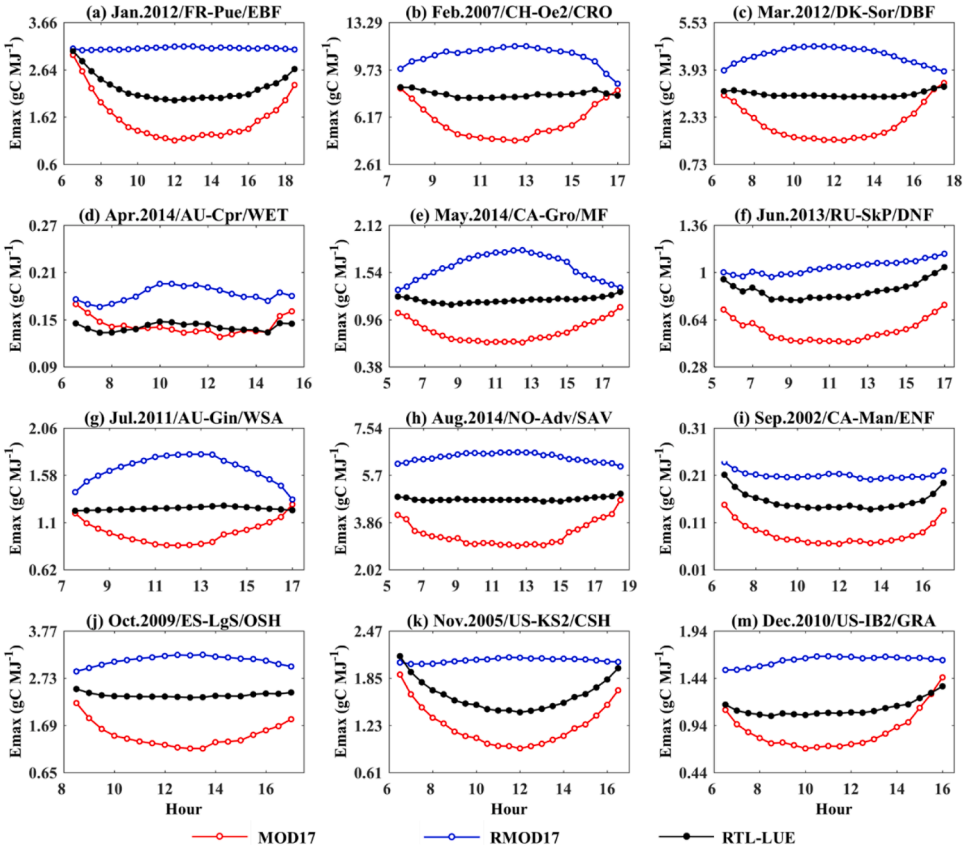

Fig. 7. Diurnal variations of the optimized different maximum LUE for the three models in different months and different vegetation types. The title of each subplot is named as "month. year/site name/vegetation types".

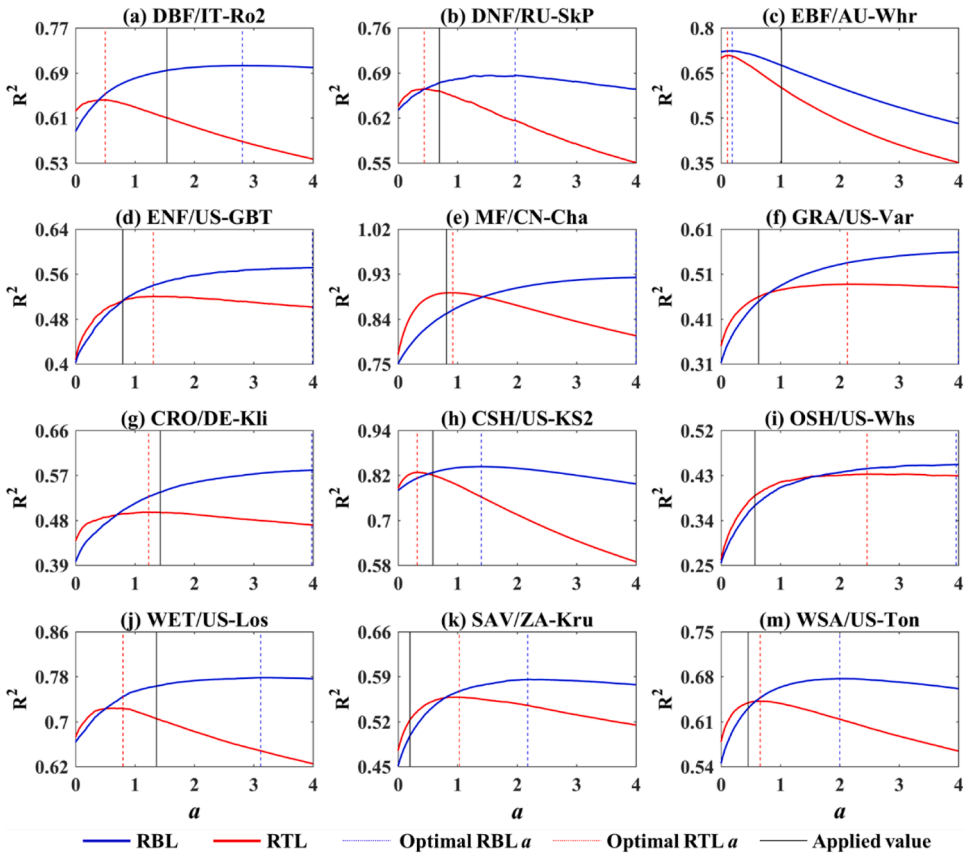

Fig. 8. The variation of R^2 between GPP simulations and EC GPP with different parameter *a* for the RMOD17 (RBL) and RTL-LUE (RTL) model.

Table 5
Statistics of the GPP simulation results for the RMOD17 (RBL) and RTL-LUE (RTL) model with respective optimal parameter *a*.

Vegetation types [a]	Optimal *a* [b]			R^2			CV (%)		
	RBL [c]	RTL [c]	EC [c]	RBL [c]	RTL [c]	Diff [d]	RBL [c]	RTL [c]	Diff [d]
DBF	0.77	3.06	1.29	0.75	0.78	0.03	15.95	15.72	−0.23
DNF	0.44	1.97	2.87	0.67	0.69	0.02	24.35	22.43	−1.91
EBF	0.59	1.79	1.90	0.72	0.72	0.00	11.28	13.77	2.49
ENF	0.69	2.02	2.12	0.71	0.72	0.01	17.13	17.98	0.84
MF	0.60	2.42	1.41	0.77	0.79	0.02	13.30	13.69	0.39
GRA	0.77	1.56	2.57	0.55	0.56	0.01	27.75	28.74	0.99
CRO	0.76	2.73	1.12	0.44	0.47	0.03	22.14	20.99	−1.16
CSH	0.32	1.40	1.50	0.83	0.85	0.02	10.73	9.85	−0.89
OSH	1.16	2.13	3.58	0.43	0.44	0.01	36.12	37.71	1.59
WET	0.75	1.83	2.12	0.63	0.65	0.02	17.55	17.72	0.16
SAV	0.48	1.27	1.83	0.53	0.55	0.02	39.57	39.23	−0.34
WSA	0.43	0.97	2.24	0.56	0.58	0.02	37.71	35.48	−2.23
ALL	0.72	2.02	2.07	0.63	0.65	0.02	22.80	22.77	−0.03

[a] The abbreviations of different vegetation types are the same as it in Table 1.
[b] the optimal a is for the results with the highest R^2.
[c] RBL, RTL, EC denotes the results for the RMOD17 model, RTL-LUE model, and EC data, respectively.
[d] Diff means the differences between the results of the RMOD17 and RTL-LUE model.

those of the RTL-LUE model, suggesting that the optimal *a* for RTL-LUE is more reasonable and can approach the regression parameter using EC data. Besides, R^2 between EC GPP and RTL-LUE GPP simulations using the optimal *a* is also higher than that between EC GPP and RMOD17 GPP simulations in all 12 vegetation types, demonstrating the advantage of RTL-LUE in GPP simulation. The lower optimal *a* in the big-leaf RMOD17 model is associated with the use of the total incoming PAR to calculate radiation scalar, causing overestimation of the light intensity on average leaves in the canopy because shaded leaves only receive the diffuse radiation. With overestimated radiation intensities, a lower value of parameter *a* is needed to calculate the proper radiation scalar and achieve the optimal results.

Furthermore, the variation of mean CV for diurnal variation of maximum LUE (i.e., ε_{RBL} and ε_{RTL}) with parameter *a* is also discussed, in order to further assess the model stability. As shown in Fig. 9, the CV values for maximum LUE diurnal variations in the two models both first decrease and then increase with increasing *a*. Another optimal *a* can also be found for the lowest CV, which is also higher than 0 in all the 169 sites for the RMOD17 model and proves the necessity of using radiation scalar. The optimal *a* here generally show different values from the optimal *a* for the highest R^2. Consistently, the optimal *a* with the lowest CV for the RTL-LUE model is also much higher than that for the RMOD17

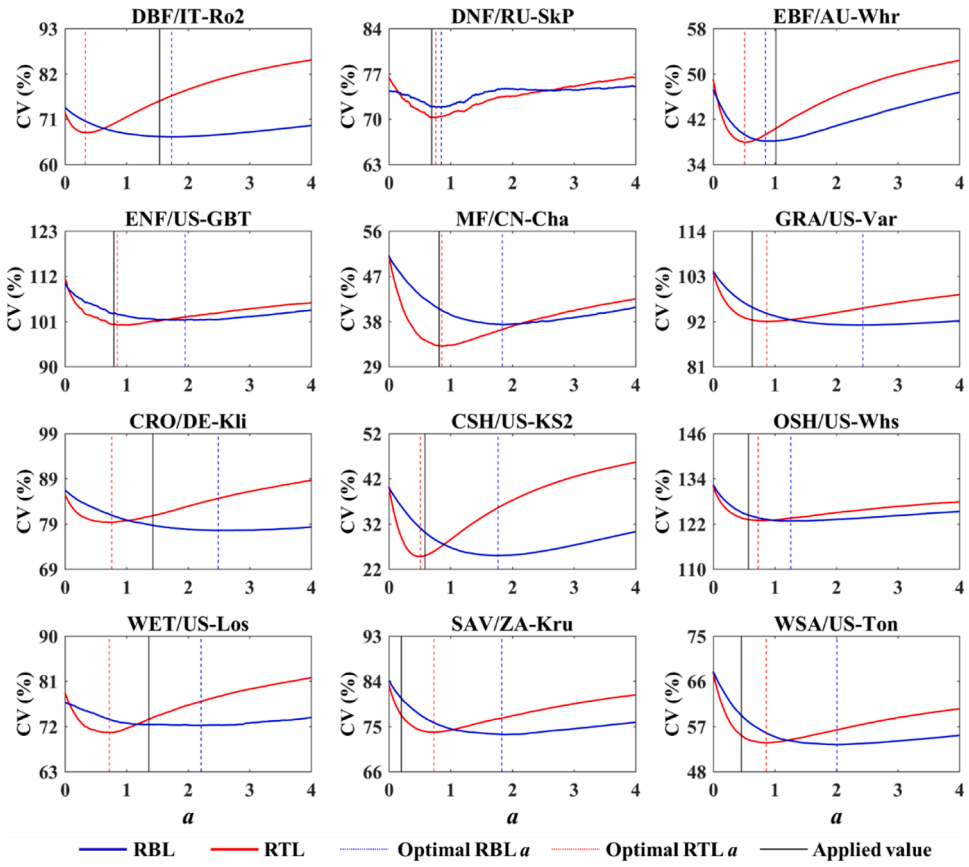

Fig. 9. The variation of mean CV for maximum LUE diurnal variation with different parameter *a* in the RMOD17 (RBL) and RTL-LUE (RTL) model.

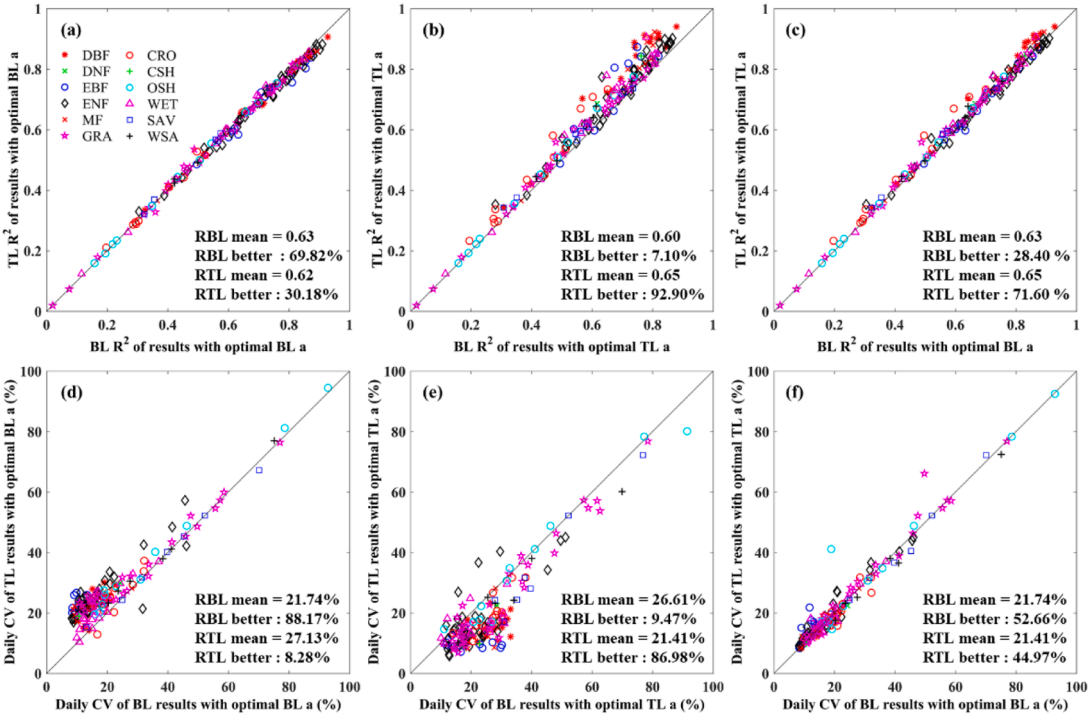

Fig. 10. Comparison of R^2 and daily CV between the RMOD17 (BL) and RTL-LUE (TL) model with different *a*: (a) and (d) are the results both using the optimal *a* for BL; (b) and (e) are the results both using the optimal *a* for TL; (a) and (d) are the results for the two models using their respective optimal *a*.

model. However, only minor differences can be found in the CV values using the respective optimal *a* for the RTL-LUE and RMOD17 model, as listed in Table 6. Overall, the optimal CV for the RTL-LUE model (22.77%) is slightly lower than that for the RMOD17 model (22.80%), but discrepancies are observed among different vegetation types. The performances of these two models are much better than the original MOD17 model, suggesting that it is the radiation scalar that can significantly reduce the diurnal variation of maximum LUE, and the two-leaf separation strategy contributes little to it but can improve the R^2 between the GPP simulations and EC data.

In order to compare the parameter sensitivity of the big-leaf (RMOD17) and two-leaf (RTL-LUE) model, the R^2 and CV at all the 169 sites are further calculated using different optimal *a* values. Fig. 10 (a) and (d) are the results for the two models both using the optimal *a* for the RMOD17 model, Fig. 10(b) and (e) are the results using the optimal *a* for the RTL-LUE model, and Fig. 10(c) and (f) are the results for the two models with their respective optimal *a*. It can be observed that there are still 30.18% of the 169 sites showing a higher R^2 in the results of the RTL-LUE model than it in the RMOD17 model, even using the optimal *a* for the RMOD17 model, and the mean difference between the two R^2 is only 0.01. On the contrary, when the optimal *a* for the RTL-LUE model is used, only 7.10% of the sites do not show a higher R^2 in the RTL-LUE mode results, and the mean value for the RTL-LUE model (R^2=0.65) is notably higher than the RMDO17 model (R^2=0.60). When the two models using their respective optimal *a*, 71.60% of the sites show a higher R^2 in the RTL-LUE model than the RMDO17 model, with a mean difference of 0.02. As a result, it can be concluded that the two-leaf RTL-LUE model shows much lower sensitivity to parameter *a*, and is a more robust strategy to estimate GPP compared to the big-leaf RMOD17 model. However, the results of the CV for maximum LUE diurnal variations are different, with almost no differences can be observed when the two models using their respective optimal parameters. Both the two models performed much poorer when using the optimal *a* for the other model (Fig. 10(d) and (e)).

4. Discussion

4.1. Improvements of the radiation scalar

Consideration of the radiation scalar can considerably improve the accuracy of GPP estimation, which cooperatively constrain LUE from optimal to reality with environmental scalars of water and temperature. In previous studies, it has been proved that GPP simulations from the RTL-LUE model show better accuracy than the original TL-LUE model after using the radiation scalar, by alleviating the overestimation under the conditions of high incoming PAR (Guan et al., 2021). This study also proves that GPP simulations from RMOD17 model is more converged to the EC GPP with many fewer outliers after using the radiation scalar, because the underestimation of GPP in low PAR ranges and overestimation in high PAR ranges are well corrected. The flat variation of LUEc with the clear sky index in the original MOD17 model is also altered by the radiation scalar to be more similar to the decreasing trends of EC GPP. As a result, it can be concluded that the radiation scalar is necessary and helpful to improve the accuracy of GPP simulation for both big-leaf and two-leaf models.

Furthermore, the radiation scalar is the main reason for the improvements in the CV for the maximum LUE diurnal variation. Since the LUE models usually assign a constant maximum LUE without any temporal variations to unique vegetation types, a stable maximum LUE is important to capture the actual diurnal GPP variation and improve the estimation accuracy (Chen, 1999; Huang et al., 2021; Madani et al., 2017). The diurnal variation of maximum LUE is significantly reduced by the radiation scalar both in the big-leaf RMOD17 model and the two-leaf RTL-LUE model, and this reduction is of great significance to model the diurnal variation of GPP. Besides, the similar diurnal behavior of these two models indicates that the radiation scalar is the main reason for the reduction of the diurnal CV in the maximum LUE, and the two-leaf strategy contributes little to this issue. The seasonal variation of maximum LUE is not analyzed in this study, because previous studies have proved that seasonal changes in maximum LUE is related to the vegetation physiological traits and it is difficult to evaluate the usefulness of the radiation scalar in different models in simulating seasonal variations in GPP (Chen et al., 2020b; Lin et al., 2017).

4.2. Improvements of the two-leaf model

The improvements of the two-leaf model from the big-leaf model in GPP simulation have already been shown in many studies (Braghiere et al., 2019; He et al., 2018a; Zheng et al., 2020). The main reason is that the two-leaf model can correct the underestimations at high productivity ranges (Chen et al., 2020a; He et al., 2013; Zhang et al., 2012). A similar conclusion is also obtained in this study, and we further find that the underestimation in big-leaf models is not only caused by the insufficient consideration of the shaded leaf contributions but also related to the unreasonable PAR used to represent the average radiation intensity on leaves in the canopy. Although R^2, RMSE, and bias of GPP simulations from the RMOD17 model have been considerably improved after using the radiation scalar, the apparent underestimation in high productivity ranges can still be observed both in the yearly and 8-days results. This problem can be well adjusted in the two-leaf RTL-LUE model, showing a much better linear regression line in all of the 12 vegetation types. Furthermore, the two-leaf RTL-LUE model also shows a lower sensitivity to parameter *a* than the big-leaf RMOD17 model. The RTL-LUE model can obtain overall more stable results using different optimal parameter *a*, thus indicates it has better fault tolerance when simulating regional or global GPP. Since the LUE models are usually driven by the biomes-specific parameters (Chen et al., 2020b; Huang et al., 2021; Madani et al., 2017), this is an important advantage for the two-leaf RTL-LUE model because it can be more robust to simulate GPP with less impact by the spatial and seasonal heterogeneity of parameters. For BL models, previous studies improve their GPP simulations simply using CI to quantify the effects of diffuse radiation on LUE (Wang et al., 2015; Zhang et al., 2016), further efforts may be directed towards combining the radiation scalar and CI to better describe the impacts of radiation on vegetation photosynthesis in BL models.

4.3. Uncertainties

This study highlights the importance of the radiation scalar and two-leaf strategies in simulating GPP in LUE models. Although some conclusions are found, there are still some limitations and uncertainties that should be declared. First of all, only the key parameter maximum LUE is optimized for all the three models in this study, the parameters for environmental scalars (i.e., VPD_{max}, VPD_{min}, T_{min}, T_{max}) are all inherited from the original MOD17 model (He et al., 2013; Running et al., 2004). Some studies indicated the optimization of these parameters could further improve the accuracy of GPP simulation (Huang et al., 2021; Zheng et al., 2018). In order to enhance the reliability of the comparison between BL and TL models and provide the full set of parameters for the future applications of the three models (i.e., MOD17, RMOD17, and RTL-LUE), we also optimized these parameters based on the joint optimization methods using observations from all sites across each PFTs. Only the selected one site year data in each site (Table S1) are used for validation, and all the other data are used for parameterization, in order to obtain as reliable parameters as possible. The optimized parameters can be found in Table S3, and the comparison for the three models using these parameters is shown in Table S4. It can be observed that although the overall accuracy of GPP simulation can be slightly improved in all three models with optimized parameters, similar conclusions can still be drawn that RTL-LUE outperformed MOD17 and RMOD17. These results further supported the importance of the radiation scalar and the two-leaf strategy for LUE models. Whereas, the uncertainty of parameters from

calibration processes is not well addressed here, and deserves further exploration to enhance the reliability of the comparison of parameters among different models (Huang et al., 2021).

Furthermore, the uncertainties induced by the remote sensing LAI data are ignored in this study, which may impact both the parameter calibration and model evaluation. The sensitivities of GPP performance to different LAI products should be further investigated among LUE models, and analysis using LAI products with high spatial resolutions and high accuracy in conjunction with flux footprints is also needed. Besides, the EC sites considered in this study are selected by low criteria in order to assess the model performance as broadly as possible. The neglect of heterogeneity in land cover may also induce uncertainties in model parameterization and validation, which needs further assessments only using the EC data from sites covered by pure vegetation (Dou et al., 2021; Zhou et al., 2016). Finally, both the RMOD17 and RTL-LUE models stem from the MOD17 model, which excludes the impacts of soil water content, CO2 fertilization, and nutrient conditions. Further efforts may be directed towards improving the environmental scalars to constrain LUE and differentiating them between sunlit and shaded leaves.

5. Conclusion

In this study, the performances of two big-leaf models (MOD17 and RMOD17) and a two-leaf model (RTL-LUE) in GPP simulations are compared. The accuracies of GPP simulations are assessed based on global 169 EC sites covering 12 vegetation types, and the differences between the optimized maximum LUE in the three models are also analyzed. The importance of the radiation scalar (differences between MOD17 and RMOD17) and the improvements of the two-leaf separation strategy (differences between RMDO17 and RTL-LUE) are both discussed. The main conclusions can be drawn as follows:

(1) An RMOD17 model simply modified with a radiation scalar can significantly improve the GPP simulation by the original MOD17 model. RMOD17 (R^2=0.72, RMSE=379.54 g C m^{-2} $year^{-1}$) agrees much better with the EC data than MOD17 (R^2=0.65, RMSE=462.05 g C m^{-2} $year^{-1}$) both for all or individual vegetation types, and the outliers can be widely corrected by the RMOD17 model to approach to the EC data. There are underestimations or overestimations in low or high photosynthetically active radiation (PAR) ranges in the MOD17 model results caused by the radiation-independent LUE, which are well corrected by the radiation scalar in the RMOD17 model.

(2) GPP simulations can be further improved by the two-leaf LUE model (RTL-LUE) (R^2=0.74, RMSE=358.81 g C m^{-2} $year^{-1}$) compared to the RMOD17 model, even though the radiation scalar has already been considered. After the separate treatment of sunlit and shaded leaves, the underestimation of the RMOD17 model results in high productivity ranges can be well alleviated, allowing the simulated results to approach EC data. The two-leaf model is also a more robust strategy to simulate GPP with much lower sensitivity to the parameter a, which can minimize the errors caused by biomes-specific parameterization in large-scale applications.

(3) The maximum LUE in the original MOD17 model is generally lower than that in the RMOD17 and RTL-LUE models, because it is not the actual maximum LUE but an average value constrained by radiation scalar at an intermediate radiation level. The diurnal variations of maximum LUE in the RMOD17 and RTL-LUE models are also significantly reduced compared to the original MOD17 model. The radiation scalar is the main reason for this reduction, which further proves the importance of considering the radiation scalar on LUE.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work used the EC data acquired and shared by the FLUXNET community. We are grateful to all the PIs who contributed to the FLUXNET2015 dataset (https://fluxnet.org/data/fluxnet2015-dataset/) and provided valuable data for the research community. The corresponding literature for the EC sites was listed in the supporting information. We also thank the GLASS group for providing their LAI product. This research was supported by the National Key Research and Development Program of China (2019YFB2102903), the National Natural Science Foundation of China (42001371), and the China Scholarship Council.

References

Alton, P., Ellis, R., Los, S., North, P., 2007. Improved global simulations of gross primary product based on a separate and explicit treatment of diffuse and direct sunlight. J. Geophys. Res.: Atmospheres 112 (D7).
Bai, Y., Zhang, J., Zhang, S., Yao, F., Magliulo, V., 2018. A remote sensing-based two-leaf canopy conductance model: global optimization and applications in modeling gross primary productivity and evapotranspiration of crops. Remote Sens. Environ. 215, 411–437.
Baldocchi, D., et al., 2001. FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bull. Am. Meteorol. Soc. 82 (11), 2415–2434.
Braghiere, R.K., Quaife, T., Black, E., He, L., Chen, J., 2019. Underestimation of global photosynthesis in Earth system models due to representation of vegetation structure. Glob. Biogeochem. Cycles 33 (11), 1358–1369.
Chen, B., et al., 2020a. Importance of shaded leaf contribution to the total GPP of Canadian terrestrial ecosystems: evaluation of MODIS GPP. J. Geophys. Res.: Biogeosciences 125 (10), e2020JG005917.
Chen, J., Liu, J., Cihlar, J., Goulden, M., 1999. Daily canopy photosynthesis model through temporal and spatial scaling for remote sensing applications. Ecol. Modell. 124 (2–3), 99–119.
Chen, J.M., 1996. Canopy architecture and remote sensing of the fraction of photosynthetically active radiation absorbed by boreal conifer forests. IEEE Trans. Geosci. Remote Sens. 34 (6), 1353–1368.
Chen, J.M., 1999. Spatial scaling of a remotely sensed surface parameter by contexture. Remote Sens. Environ. 69 (1), 30–42.
Chen, J.M., et al., 2019. Vegetation structural change since 1981 significantly enhanced the terrestrial carbon sink. Nat. Commun. 10 (1), 1–7.
Chen, J.M., et al., 2012. Effects of foliage clumping on the estimation of global terrestrial gross primary productivity. Glob. Biogeochem. Cycles 26 (1).
Chen, Y., Feng, X., Fu, B., Wu, X., Gao, Z., 2020b. Improved global maps of the optimum growth temperature, maximum light use efficiency and gross primary production for vegetation. J. Geophys. Res.: Biogeosciences, e2020JG005651.
Chu, D., et al., 2021. Long time-series NDVI reconstruction in cloud-prone regions via spatio-temporal tensor completion. Remote Sens. Environ. 264 (112632).
De Pury, D., Farquhar, G., 1997. Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models. Plant Cell Environ. 20 (5), 537–557.
Dong, J., et al., 2015. Comparison of four EVI-based models for estimating gross primary production of maize and soybean croplands and tallgrass prairie under severe drought. Remote Sens. Environ. 162, 154–168.
Dou, P., Shen, H., Li, Z., Guan, X., 2021. Time series remote sensing image classification framework using combination of deep learning and multiple classifiers system. Int. J. Appl. Earth Obs. Geoinf. 103, 102477.
Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28 (4), 1015–1031.
Farquhar, G.D., von Caemmerer, S.v., Berry, J.A., 1980. A biochemical model of photosynthetic CO 2 assimilation in leaves of C 3 species. Planta 149 (1), 78–90.
Friedlingstein, P., et al., 2019. Global carbon budget 2019. Earth Syst. Sci. Data 11 (4), 1783–1838.
Guan, X., Chen, J.M., Shen, H., Xie, X., 2021. A modified two-leaf light use efficiency model for improving the simulation of GPP using a radiation scalar. Agric. For. Meteorol. 307 (108546).
Guan, X., Shen, H., Li, X., Gan, W., Zhang, L., 2019. A long-term and comprehensive assessment of the urbanization-induced impacts on vegetation net primary productivity. Sci. Total Environ. 669, 342–352.
Harbinson, J., 2012. Modeling the protection of photosynthesis. Proc. Natl. Acad. Sci. 109 (39), 15533–15534.
Haxeltine, A., Prentice, I., 1996. A general model for the light-use efficiency of primary production. Funct. Ecol. 551–561.

He, L., et al., 2018a. Changes in the shadow: the shifting role of shaded leaves in global carbon and water cycles under climate change. Geophys. Res. Lett. 45 (10), 5052–5061.
He, M., et al., 2013. Development of a two-leaf light use efficiency model for improving the calculation of terrestrial gross primary productivity. Agric. For. Meteorol. 173, 28–39.
He, Y., Piao, S., Li, X., Chen, A., Qin, D., 2018b. Global patterns of vegetation carbon use efficiency and their climate drivers deduced from MODIS satellite data and process-based models. Agric. For. Meteorol. 256, 150–158.
Hilker, T., Coops, N.C., Wulder, M.A., Black, T.A., Guy, R.D., 2008. The use of remote sensing in light use efficiency based models of gross primary production: a review of current status and future requirements. Sci. Total Environ. 404 (2–3), 411–423.
Huang, X., Xiao, J., Wang, X., Ma, M., 2021. Improving the global MODIS GPP model by optimizing parameters with FLUXNET data. Agric. For. Meteorol. 300, 108314.
Hunt Jr., E.R., Running, S.W., 1992. Simulated dry matter yields for aspen and spruce stands in the North American boreal forest. Canadian J. Remote Sens. 18 (3), 126–133.
Koyama, K., Kikuzawa, K., 2010. Geometrical similarity analysis of photosynthetic light response curves, light saturation and light use efficiency. Oecologia 164 (1), 53–63.
Leverenz, J.W., 1987. Chlorophyll content and the light response curve of shade-adapted conifer needles. Physiol. Plant. 71 (1), 20–29.
Li, L., Cai, M., Guan, X., Chu, D., 2020. Piecewise Adaptive-Norm Trend Filtering Method for ICESat/GLAS Waveform Data Denoising. IEEE Access 8, 168965–168979.
Lin, X., et al., 2017. Seasonal fluctuations of photosynthetic parameters for light use efficiency models and the impacts on gross primary production estimation. Agric. For. Meteorol. 236, 22–35.
Liu, J., Chen, J., Cihlar, J., Park, W., 1997. A process-based boreal ecosystem productivity simulator using remote sensing inputs. Remote Sens. Environ. 62 (2), 158–175.
Madani, N., Kimball, J.S., Running, S.W., 2017. Improving global gross primary productivity estimates by computing optimum light use efficiencies using flux tower data. J. Geophys. Res.: Biogeosciences 122 (11), 2939–2951.
Mäkelä, A., et al., 2008. Developing an empirical model of stand GPP with the LUE approach: analysis of eddy covariance data at five contrasting conifer sites in Europe. Glob. Chang. Biol. 14 (1), 92–108.
Marcott, S.A., et al., 2014. Centennial-scale changes in the global carbon cycle during the last deglaciation. Nature 514 (7524), 616–619.
McCallum, I., et al., 2013. Improved light and temperature responses for light-use-efficiency-based GPP models. Biogeosciences 10 (10), 6577–6590.
Monteith, J., 1972. Solar radiation and productivity in tropical ecosystems. J. Appl. Ecol. 9 (3), 747–766.
Ogren, E., 1993. Convexity of the photosynthetic light-response curve in relation to intensity and direction of light during growth. Plant Physiol. 101 (3), 1013–1019.
Pastorello, G., et al., 2017. A new data set to keep a sharper eye on land-air exchanges. Eos, Trans. Am. Geophys. Union 98 (8) (Online).
Pastorello, G., et al., 2020. The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. Sci. Data 7 (1), 1–27.
Potter, C.S., et al., 1993. Terrestrial ecosystem production: a process model based on global satellite and surface data. Glob. Biogeochem. Cycles 7 (4), 811–841.
Propastin, P., Ibrom, A., Knohl, A., Erasmi, S., 2012. Effects of canopy photosynthesis saturation on the estimation of gross primary productivity from MODIS data in a tropical forest. Remote Sens. Environ. 121, 252–260.
Running, S.W., Coughlan, J.C., 1988. A general model of forest ecosystem processes for regional applications I. Hydrologic balance, canopy gas exchange and primary production processes. Ecol. Modell. 42 (2), 125–154.
Running, S.W., et al., 2004. A continuous satellite-derived measure of global terrestrial primary production. Bioscience 54 (6), 547–560.
Wang, S., et al., 2015. Improving the light use efficiency model for simulating terrestrial vegetation gross primary production by the inclusion of diffuse radiation across ecosystems in China. Ecol. Complex. 23, 1–13.
Wang, S., et al., 2020. Recent global decline of CO2 fertilization effects on vegetation photosynthesis. Science 370 (6522), 1295–1300.
Xiao, X., et al., 2004. Satellite-based modeling of gross primary production in an evergreen needleleaf forest. Remote Sens. Environ. 89 (4), 519–534.
Xiao, Z., et al., 2016. Long-time-series global land surface satellite leaf area index product derived from MODIS and AVHRR surface reflectance. IEEE Trans. Geosci. Remote Sens. 54 (9), 5301–5318.
Xie, X., Li, A., 2020. An Adjusted Two-Leaf Light use efficiency model for improving GPP simulations over mountainous areas. J. Geophys. Res.: Atmospheres 125 (13), e2019JD031702.
Xie, X., et al., 2019. Assessment of five satellite-derived LAI datasets for GPP estimations through ecosystem models. Sci. Total Environ. 690, 1120–1130.
Xie, X., et al., 2020. Assessments of gross primary productivity estimations with satellite data-driven models using eddy covariance observation sites over the northern hemisphere. Agric. For. Meteorol. 280, 107771.
Yang, J., et al., 2013. The role of satellite remote sensing in climate change studies. Nat. Clim. Chang. 3 (10), 875–883.
Yu, G.-R., et al., 2006. Overview of ChinaFLUX and evaluation of its eddy covariance measurement. Agric. For. Meteorol. 137 (3–4), 125–137.
Yuan, W., et al., 2014. Global comparison of light use efficiency models for simulating terrestrial vegetation gross primary production based on the LaThuile database. Agric. For. Meteorol. 192-193, 108–120.
Yuan, W., et al., 2007. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. Agric. For. Meteorol. 143 (3–4), 189–207.
Zan, M., et al., 2018. Performance of a two-leaf light use efficiency model for mapping gross primary productivity against remotely sensed sun-induced chlorophyll fluorescence data. Sci. Total Environ. 613, 977–989.
Zhang, F., et al., 2012. Evaluating spatial and temporal patterns of MODIS GPP over the conterminous US against flux measurements and a process model. Remote Sens. Environ. 124, 717–729.
Zhang, Y., et al., 2016. Development of a coupled carbon and water model for estimating global gross primary productivity and evapotranspiration based on eddy flux and remote sensing data. Agric. For. Meteorol. 223, 116–131.
Zheng, Y., et al., 2020. Improved estimate of global gross primary production for reproducing its long-term variation, 1982–2017. Earth Syst. Sci. Data 12 (4), 2725–2746.
Zheng, Y., et al., 2018. Sources of uncertainty in gross primary productivity simulated by light use efficiency models: model structure, parameters, input data, and spatial resolution. Agric. For. Meteorol. 263, 242–257.
Zhou, Y., et al., 2016. Global parameterization and validation of a two-leaf light use efficiency model for predicting gross primary production across FLUXNET sites. J. Geophys. Res.: Biogeosciences 121 (4), 1045–1072.