

Contents lists available at ScienceDirect

# Agricultural and Forest Meteorology





# An improved light use efficiency model by considering canopy nitrogen concentrations and multiple environmental factors

Helin Zhang <sup>a, c</sup>, Jia Bai <sup>a, c</sup>, Rui Sun <sup>a, b, \*</sup>, Yan Wang <sup>d</sup>, Zhiqiang Xiao <sup>a</sup>, Bowen Song <sup>e, f</sup>

<sup>a</sup> State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

<sup>b</sup> Faculty of Arts and Sciences, Beijing Normal University, Zhuhai 519087, China

<sup>c</sup> Beijing Engineering Research Center for Global Land Remote Sensing Products, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

<sup>d</sup> Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China

e Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100094, China

<sup>f</sup> University of Chinese Academy of Sciences, Beijing 100049, China

#### ARTICLE INFO

Keywords: Gross primary production (GPP) Nitrogen (N) Carbon dioxide (CO<sub>2</sub>) Environmental factors Light use efficiency (LUE)

# ABSTRACT

Although the light use efficiency (LUE) models are widely employed to estimate ecosystem gross primary production (GPP), the majority of these models inadequately consider the effects of environmental and biological factors on GPP, resulting in considerable uncertainty. In addition, most developed LUE models have assumed that the maximum LUE ( $\varepsilon_{max}$ ) is a fixed value for different vegetation types, while  $\varepsilon_{max}$  should be dynamic under environmental changes. The canopy nitrogen (N) concentrations were considered to have a significant linear relationship with  $\varepsilon_{max}$  and could be estimated using various vegetation indices. In this study, we selected a vegetation index to characterize the canopy N concentrations and further simulate the dynamic  $\varepsilon_{max}$ . We then developed an improved LUE model that simultaneously integrated the effects of canopy N concentrations, temperature, water, atmospheric carbon dioxide (CO<sub>2</sub>) and radiation components on the GPP estimates. Different forms of LUE models that partially integrate the above factors were also constructed for comparison. Our results showed that (1) the green chlorophyll index (CIgreen) correlated well with measured canopy N concentrations (R<sup>2</sup> = 0.68), and the model using the CI<sub>green</sub> to characterize canopy N concentrations performed the best; (2) the GPP estimated using the improved model gave the best accuracy ( $R^2 = 0.69$ , RMSE = 2.13 gC/m<sup>2</sup>/d, MAE=1.36  $m^2/d$ , IOA = 0.915) and performed well for different vegetation types when validated against the FLUXNET GPP; and (3) the estimated GPP had the best accuracy compared with MOD17 GPP and the revised EC-LUE GPP on a both daily and yearly scale. Overall, this study was an attempt to integrate N into the LUE model to obtain the spatiotemporally dynamic  $\varepsilon_{max}$  while simultaneously taking into account the impacts of multiple environmental variables on the GPP estimates. The proposed model has the potential for satisfactory GPP simulations on a global or regional scale.

# 1. Introduction

The total carbon dioxide (CO<sub>2</sub>) absorbed from the atmosphere via photosynthesis by vegetation is defined as the gross primary production (GPP), which has an important role in the global carbon budget and climate system (Beer et al., 2010; Bonan et al., 2019; Chen et al., 2012; Fang et al., 2018). In the context of rapid global change, accurate quantification of GPP on a regional or global scale is critical to strengthening our comprehension of how the terrestrial ecosystems and environmental changes interact (Guan et al., 2021).

Although the flux observation network based on the eddy covariance

technique supplies much observational data (such as FLUXNET, Ameriflux, and Chinaflux), due to the finite number and nonuniform distribution of sites, it is still unable to meet the needs of vegetation productivity monitoring and evaluation on a regional or global scale (Xiao et al., 2008). In recent decades, satellite observations and digital earth technologies have become a meaningful way to capture global change and the carbon cycle (Wang et al., 2021a). Using models based on satellite data to estimate GPP has become an essential and widely accepted research method. Over the last few decades, GPP has been quantified by many different models, including meteorological-based models (Sharpe 1975; Uchijima and Seino 1985), process-based

https://doi.org/10.1016/j.agrformet.2023.109359

Received 5 July 2022; Received in revised form 31 January 2023; Accepted 4 February 2023 Available online 16 February 2023 0168-1923/© 2023 Elsevier B.V. All rights reserved.

<sup>\*</sup> Corresponding author at: State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China *E-mail address:* sunrui@bnu.edu.cn (R. Sun).

models (Hunt et al., 1996; Liu et al., 1997; Melillo et al., 1993; Parton et al., 1993; Running and Coughlan 1988), light use efficiency (LUE) models (Guan et al., 2021; He et al., 2013; Heinsch et al., 2003; Monteith 1972; Monteith et al., 1977; Potter et al., 1993; Wang et al., 2021b; Xiao et al., 2004) and data-driven models (Alemohammad et al., 2017; Dou and Yang 2018; Ichii et al., 2017; Xiao et al., 2014). Among these models, LUE models, which have the advantages of high accuracy, simple model structure, and easy access to input requirements, are favored by many applications (Bao et al., 2022; Guan et al., 2021; Zheng et al., 2020).

In past years, many studies have tried to improve the estimation of terrestrial GPP to minimize the uncertainty in the global carbon cycle (Ryu et al., 2019). However, the interannual variations in GPP estimated using the LUE models still have considerable deviations from in situ measurements (Yuan et al., 2014; Zheng et al., 2020). In addition, there are significant differences in the global annual GPP estimates using different LUE models (Wang et al., 2021b). One of the major reasons for the uncertainty in the GPP estimates is that the influence of environmental factors on photosynthesis, such as soil moisture, atmospheric  $CO_2$ , and nitrogen. was not entirely taken into account in the LUE models (Stocker et al., 2019).

In general, the LUE is determined by the maximum LUE ( $\varepsilon_{max}$ ) with some adjusted environmental stress factors. Among all environmental factors, temperature has an important influence on enzyme activity and the electron transport rate, and a sufficient water supply ensures stomatal openness and physiochemical reactions in plants. These are the most common environmental stress factors in LUE models (Bao et al., 2022; Medlyn et al., 2002). In addition, atmospheric CO<sub>2</sub> is another critical driving factor of the vegetation photosynthesis; it can diffuse into leaves through stomata (Farquhar et al., 1980; Leuning 1995). Many studies have shown that increasing  $\mathrm{CO}_2$  concentrations promote an increase in vegetation productivity (Ainsworth and Long 2005; Norby et al., 2005; Norby et al., 1999; Wang et al., 2020). According to NOAA's CarbonTracker, version CT2019B (Jacobson et al., 2020), the global average atmospheric CO2 concentrations have increased by almost 23% since the 1980s, which increased global annual photosynthesis by approximately 12% (Keenan et al., 2021). However, with the exception of the CFix (Veroustraete et al., 2002), P-model (Stocker et al., 2020), revised EC-LUE (Zheng et al., 2020), PRELES models (Kalliokoski et al., 2018), and the algorithm of dry matter productivity (DMP) products (Swinnen et al., 2021), the CO<sub>2</sub> fertilization effect has not been included in most LUE models, which results in traditional LUE models being insensitive to the effect of increasing CO<sub>2</sub> concentrations on vegetation productivity. Moreover, many current studies have revealed that direct and diffuse radiation also have a particular impact on the LUE and suggested that the effect of radiation conditions should be considered in LUE estimation (Guan et al., 2021; He et al., 2013; Wang et al., 2015; Zhang et al., 2011; Zheng et al., 2020). Generally, sunlit leaves tend to reach light saturation due to their ability to simultaneously receive direct radiation and diffuse radiation, which results in their low LUE. In contrast, shaded leaves that only absorb diffuse radiation receive less incoming radiation and have a higher LUE than sunlit leaves (Chen et al., 1999; Guan et al., 2021; He et al., 2013). Alton et al. (2007) indicated that for northern forests, temperate forests, and tropical forests, when the sky emits mainly diffuse radiation rather than direct radiation, the LUE increases by 6%–33%. Therefore, when the total income radiation is constant, the LUE would increase with the proportion of diffuse radiation. In addition, many studies have used hyperspectral data to calculate the photochemical reflectance index (PRI) and found that it is related to LUE. This relationship may help us obtain LUE directly from reflectance data (Barton and North 2001; Guo and Trotter 2004).

Nitrogen (N) is one of the most essential components of enzymes and pigments, and its availability is an important constraint for photosynthesis (He et al., 2020; Hikosaka 2004; Lepine et al., 2016; Loozen et al., 2020; Reich 2012; Tang et al., 2018). Balzarolo et al. (2019) discovered that the  $\varepsilon_{max}$  of cold forests increases with N deposition. Nevertheless,

almost all LUE models disregard the influence of N on GPP estimation. Peltoniemi et al. (2012) suggested that canopy N and  $\varepsilon_{max}$  are linearly correlated. In most LUE models, the  $\varepsilon_{max}$  of each vegetation type is assumed to be a constant value, but it should be dynamic under various environmental conditions, especially in human activities (Chen et al., 2021; Lin et al., 2017; Madani et al., 2014; Zhou and Xin 2019). Therefore, the  $\varepsilon_{max}$  estimated by the canopy N, which seasonally and spatially varies, makes the LUE models more suitable for the actual situation of leaves adapting to changes in environmental conditions (Houborg et al., 2009). Unfortunately, although many studies have mapped the canopy N distribution using random forests or state-of-the-art Bayesian model on a large scale (Butler et al., 2017; Loozen et al., 2020; Moreno-Martínez et al., 2018), canopy N variation was compressed to a static value (Butler et al., 2017). These existing canopy N products or leaf N content products only have spatial distributions and lack temporal variation that cannot reflect changes in  $\varepsilon_{max}$ . A previous study also used the PROSAIL model, which is a physically based radiative transfer model, to estimate canopy N content (Clevers and Kooistra 2012). However, the complex parameters and calculation process of the radiative transfer model increase the uncertainty and error of the simulation (Marie et al., 2000; Wen et al., 2019). In addition, vegetation indices based on satellite or field-measured reflectance have been commonly utilized to estimate canopy N concentrations on a local scale due to their simple calculations and high accuracy (Chen et al., 2010; He et al., 2020; He et al., 2016; Loozen et al., 2018; Ollinger et al., 2008; Verrelst et al., 2021; Wen et al., 2019). Therefore, calculating  $\varepsilon_{max}$ using a vegetation index that characterizes canopy N concentrations may have implications for GPP simulations and facilitate the evaluation of the long-term impacts of N deposition on GPP (Dong et al., 2017).

Overall, it is necessary to integrate the temperature, water, radiation components, CO<sub>2</sub> fertilization and N into the LUE model to improve the quantification of GPP. The primary objectives are listed as follows: 1) to select a vegetation index that can characterize the vegetation canopy N concentrations and to develop a LUE model considering multiple environmental factors and canopy N concentrations; 2) to evaluate the effectiveness and accuracy of this model using in situ observations and to compare it with multiple forms of models that do not fully consider environmental factors and N; and 3) to compare our model with other LUE models on a daily and annual scale.

# 2. Data and methods

# 2.1. Data

#### 2.1.1. Flux and measured data

The FLUXNET2015 dataset (https://FLUXNET.fluxdata.org/) includes carbon flux and other meteorological variables over 200 sites globally. In this study, we obtained daily GPP based on the nighttime partitioning method (GPP NT VUT REF) and meteorological variables, including incident shortwave radiation (SW), air temperature  $(T_a)$ , vapor pressure deficit (VPD), latent heat flux (LE) and sensible heat flux (H), for model development and validation. First, we chose high-quality data with a quality flag above 0.8 and deleted the negative GPP values to ensure the quality of the data. Second, we aggregated the daily values to an 8-day time step to match the 8-day temporal resolution of the leaf area index (LAI) and the fraction of absorbed photosynthetically active radiation (FPAR) products. This step ensured that high-quality valid values exceeded 5 days per the 8-day time step. To verify the homogeneity of the surrounding landscape of sites, we selected FLUXNET sites with land cover types consistent with moderate-resolution imaging spectroradiometer (MODIS) land cover products and filtered these sites with high-spatial-resolution Google Earth imagery to ensure that the land cover of these sites was the same as the dominant type within a 5  $km\,\times\,5$  km window around the site. Last, a total of 104 sites were selected (Fig. 1, Table S5). The land cover type of each site is consistent with the MODIS land cover product (MCD12Q1), including 22 evergreen



Fig. 1. The MODIS land cover product (MCD12Q1) with the International Geosphere-Biosphere Program (IGBP) classification schemes and the distribution of FLUXNET sites used in this study. The vegetation types at these sites include evergreen needleleaf forest (ENF), evergreen broadleaf forest (EBF), deciduous needleleaf forest (DNF), deciduous broadleaf forest (DBF), mixed forest (MF), closed shrubland (CSH), open shrubland (OSH), woody savanna (WSAV), savanna (SAV), grassland (GRA), permanent wetland (WET), cropland (CRO), urban areas (UA), snow and ice (SI), barren or sparsely vegetated (BSV), water body (WB), and cropland/nature vegetation mosaic (C/NV).

needleleaf forest (ENF), 5 evergreen broadleaf forest, 1 deciduous needleleaf forest (DNF), 10 deciduous broadleaf forest (DBF), 6 mixed forest (MF), 3 closed shrubland (CSH), 10 open shrubland (OSH), 4 woody savanna (WSAV), 8 savanna (SAV), 14 grassland (GRA), 8 permanent wetland (WET), and 13 cropland (CRO) sites.

To corroborate the ability of the selected vegetation index to characterize canopy N concentrations, we collected canopy mean N concentration measurements from previous studies that had time stamps (Table S1). The canopy N concentrations were determined by the mean of dry-mass-based foliar N concentrations for all species in multiple field plots within each site (Ollinger et al., 2008).

# 2.1.2. Satellite and reanalysis input data

The Global Change Data Processing and Analysis Center of Beijing Normal University generated and published the Global land surface satellite (GLASS) product set (Liang et al., 2013), and the LAI is a product in the GLASS series products. The latest version of the GLASS AVHRR LAI product from 1981 to 2018, which was derived from the Long-Term Data Record (LTDR) of the Advanced Very High Resolution Radiometer (AVHRR) reflectance data (GLASS AVHRR) with a spatiotemporal resolution of 0.05° and 8 days (http://www.glass.umd.edu/L AI/AVHRR/), is more continuous in spatiotemporal distribution (Xiao et al., 2017). Here, we obtained the GLASS AVHRR LAI product from 2001 to 2014 as vegetation structure parameters to drive the models and used it to decompose the APAR into the APAR of shaded leaves (APAR<sub>sh</sub>) and that of sunlit leaves (APAR<sub>su</sub>). The GLASS AVHRR FPAR product was derived from the GLASS AVHRR LAI products (http://www.glass. umd.edu/FAPAR/AVHRR/) with the same spatiotemporal resolution as the LAI product (Xiao et al., 2018).

The NOAA Earth System Research Laboratory (ESRL) provided global three-hourly distributions of the CO<sub>2</sub> mole fraction with a  $3^{\circ} \times 2^{\circ}$  spatial resolution (Jacobson et al., 2020) (https://gml.noaa.gov/aftp/pr oducts/carbontracker/co2/molefractions/co2\_total/). We obtained CO<sub>2</sub>

mole fraction data files for CT2019B from 2001 to 2014 and aggregated them into daily average  $CO_2$  concentrations, which were used to calculate the  $CO_2$  fertilization effect in the models.

Since MODIS reflectance data can provide long time series data and support global applications, we collected MODIS reflectance products MCD43A4 with a 500 m spatial resolution and MYDOCGA with a 1 km spatial resolution, which contain MODIS reflectance bands 1 to 7 and bands 8 to 16, respectively. To select vegetation indices that characterize canopy N concentrations for the  $\varepsilon_{max}$  estimation, we extracted reflectance data for each band at FLUXNET sites from 2001 to 2014. The reflectance data were controlled by quality data, and the original time series vegetation indices were smoothed using the Savitzky–Golay (S-G) filtering method (Savitzky and Golay 1964).

The input remote sensing data of the model were resampled to a  $0.05^\circ \times 0.05^\circ$  spatial resolution by bilinear interpolation to match the LAI and FPAR data to satisfy the application of the model on a global scale.

# 2.2. Model description

Form 1:

To investigate the simulation effect of the model developed in this study, we constructed four forms of LUE models for comparison.

First, we constructed the most commonly employed LUE model as Form 1 (Eq. (1)). The GPP of the vegetation canopy was calculated as follows:

$$GPP = \varepsilon_{\max} \times f(W) \times f(T) \times PAR \times FPAR$$
<sup>(1)</sup>

where  $\varepsilon_{\text{max}}$  is the maximum LUE of the whole canopy; f(W) and f(T) are the scalars of water and temperature, respectively; and PAR was calculated from shortwave radiation as follows:

$$PAR = 0.48 \times SW \tag{2}$$

where *SW* represents the shortwave radiation variable obtained from FLUXNET site data.

f(T) was calculated based on the equation developed in the Terrestrial Ecosystem Model (TEM) (Raich et al., 1991; Yuan et al., 2014; Yuan et al., 2007; Zheng et al., 2020). f(T) was calculated as follows:

$$f(T) = \frac{(T_a - T_{\min})(T_a - T_{\max})}{(T_a - T_{\min})(T_a - T_{\max}) - (T_a - T_{opt})^2}$$
(3)

where  $T_a$  represents the air temperature (°C);  $T_{opt}$ ,  $T_{min}$  and  $T_{max}$  are the optimum, minimum, and maximum air temperatures, respectively, for vegetation growth;  $T_{max}$  was set to 40 °C in this study; and  $T_{opt}$  and  $T_{min}$  for various vegetation types were determined according to Huang et al. (2019) and Yang et al. (2021) (Table S4).

For f(W), we used the evaporative fraction (EF) to characterize the effect of water on photosynthesis. Previous studies have suggested that EF can better reflect soil moisture conditions and indicate regional heterogeneity of soil moisture compared to VPD (Kurc and Small 2004; Suleiman and Crago 2004; Yuan et al., 2007). The LUE is more sensitive to EF than VPD (Wang et al., 2021c; Zhang et al., 2015); it was calculated as follows:

$$f(W) = EF = \frac{LE}{LE + H}$$
(4)

where LE and H denote the latent heat flux  $(W/m^2)$  and sensible heat flux  $(W/m^2)$ , respectively.

Second, the  $CO_2$  fertilization effect was integrated into the LUE model described in Form 1, and GPP was calculated as follows: Form 2:

$$GPP = \varepsilon_{\max} \times f(W) \times f(T) \times f(CO_2) \times PAR \times FPAR$$
(5)

where  $f(CO_2)$  represents the effect of atmospheric CO<sub>2</sub> on GPP and was calculated with reference to the revised EC-LUE model (Zheng et al., 2020) as follows:

$$f(CO_2) = \frac{C_i - \varphi}{C_i + 2\varphi} \tag{6}$$

$$C_i = [CO_2] \times \chi \tag{7}$$

where  $\varphi$  is the CO<sub>2</sub> compensation point in the absence of dark respiration and was set for different vegetation types according to Zheng et al. (2020) (Table S4);  $C_i$  indicates the internal leaf CO<sub>2</sub> concentration and was calculated using the product of the atmospheric CO<sub>2</sub> concentration ([ $CO_2$ ]) and  $\chi$ , which represents the ratio of  $C_i$  and [ $CO_2$ ] and can be calculated as follows:

$$\chi = \frac{\xi}{\xi + \sqrt{VPD}} \tag{8}$$

$$\xi = \sqrt{\frac{356.51K}{1.6\eta^*}}$$
(9)

where parameter  $\xi$  represents the sensitivity of  $\chi$  to VPD, *K* is the Michaelis–Menten coefficient of Rubisco, and  $\eta^*$  is the viscosity of water relative to its value at 25 °C.

$$K = K_c \left( 1 + \frac{P_o}{K_o} \right) \tag{10}$$

$$K_c = 39.97 \times e^{\frac{79.43 \times (T_a - 298.15)}{298.15 \times R \times T_a}}$$
(11)

$$K_o = 27,480 \times e^{\frac{36.38 \times (T_a - 298.15)}{298.15 \times R \times T_a}}$$
(12)

where  $P_0$  is the partial pressure of O<sub>2</sub>, approximated as 21,278.25 Pa;  $K_c$  and  $K_o$  are the Michaelis–Menten constants of CO<sub>2</sub> and O<sub>2</sub>, respectively;

 $T_a$  is the air temperature with unit K; and R is the molar gas constant and is set to 8.314 J/mol/K.

Third, we further considered the effect of sunlit and shaded leaves. Previous studies have indicated that the differences between the LUE of sunlit leaves and that of shaded leaves are mainly influenced by light intensity and that their  $\varepsilon_{max}$  should be similar (Guan et al., 2021; Koyama and Kikuzawa 2010; Leverenz 1987). Liu et al. (2021) found that the relationship between PAR and LUE follows a hyperbolic relationship and that the LUE significantly decreases with increasing PAR. To account for the nonlinear response of the LUE to APAR, MÄKelÄ et al. (2008) defined a rectangular hyperbolic light modifier to correct for the LUE. Guan et al. (2021) used the scalar of photosynthetic photon flux density (PPFD) of shaded and sunlit leaves as the radiation constraint for the LUE model. Based on the above studies, we applied the PAR of the shaded ( $PAR_{sh}$ ) and sunlit leaves ( $PAR_{su}$ ) to calculate the radiation constraints and decomposed the APAR into the APAR of the shaded leaves (APAR<sub>sh</sub>) and sunlit leaves (APAR<sub>su</sub>) with reference to the BEPS model (Chen et al., 1999). The calculation of Form 3 is presented as follows:

Form 3:

$$GPP = \varepsilon_{\max} \times f(W) \times f(T) \times f(CO_2) \times (f(PAR_{su}) \times APAR_{su} + f(PAR_{sh}) \times APAR_{sh})$$
(13)

where  $f(PAR_{su})$  and  $f(PAR_{sh})$  are the radiation scalars for sunlit leaves and shaded leaves, respectively, calculated as Eq. (14) and Eq. (15).

$$f(PAR_{su}) = \frac{1}{a \times PAR_{su} + 1} \tag{14}$$

$$f(PAR_{sh}) = \frac{1}{a \times PAR_{sh} + 1}$$
(15)

where *a* is optimized for different vegetation types.

 $PAR_{su}$ ,  $PAR_{sh}$ ,  $APAR_{su}$  and  $APAR_{sh}$  are calculated based on the BEPS model (Chen et al., 1999) as follows:

$$PAR_{sh} = \frac{PAR_{dif} - PAR_{dif,u}}{LAI} + C$$
(16)

$$PAR_{su} = \frac{PAR_{dir} \times \cos(\beta)}{\cos(\theta)} + PAR_{sh}$$
(17)

$$PAR_{dif,u} = PAR_{dif} \times e^{\left(-0.5 \times \Omega \times \frac{LA}{\cos(\tilde{\theta})}\right)}$$
(18)

$$\cos(\overline{\theta}) = 0.537 - 0.025 \times LAI \tag{19}$$

$$C = 0.07 \times \Omega \times PAR_{dir} \times (1.1 - 0.1 \times LAI) \times e^{(-\cos(\theta))}$$
<sup>(20)</sup>

$$APAR_{su} = (1 - \alpha) \times PAR_{su} \times LAI_{su}$$
<sup>(21)</sup>

$$APAR_{sh} = (1 - \alpha) \times PAR_{sh} \times LAI_{sh}$$
(22)

where  $PAR_{dif}$  and  $PAR_{dir}$  are the diffuse PAR and direct PAR, respectively;  $PAR_{dif}$  was calculated by parameter calibration using the clear sky index according to Chen et al. (1999) and He et al. (2013);  $PAR_{dir}$  is the residual of *PAR* minus  $PAR_{dif}$ ; *C* is the multiple scattering effects of direct radiation;  $\beta$  is set to  $60^{\circ}$  indicating the mean leaf-sun angle;  $\theta$  is the solar zenith angle;  $\overline{\theta}$  is a representative zenith angle for diffuse radiation transmission;  $\Omega$  and  $\alpha$  are the clumping index and canopy albedo, respectively, which were set for different vegetation types according to Tang et al. (2007) and Zhang et al. (2010) (Table S4), respectively; and  $LAI_{su}$  and  $LAI_{sh}$  denote the LAI of sunlit leaves and shaded leaves, respectively, and are calculated as follows:

$$LAI_{su} = 2 \times \cos(\theta) \times \left(1 - e^{\left(-0.5 \times \Omega \times \frac{LM}{\cos(\theta)}\right)}\right)$$
(23)

$$LAI_{sh} = LAI - LAI_{su} \tag{24}$$

Last, we tried to incorporate canopy N concentrations into the LUE model. The  $\varepsilon_{max}$  is determined by the physiological traits of leaves, while canopy N was linearly correlated with the  $\varepsilon_{max}$  regardless of irradiance level (Guan et al., 2021; Peltoniemi et al., 2012). Therefore, we introduced the vegetation index, which characterizes the canopy N concentrations, into the model to calculate  $\varepsilon_{max}$ . The model of Form 4 is expressed as follows:

Form 4:

$$GPP = (p \times [NI] + q) \times f(W) \times f(T) \times f(CO_2)$$
$$\times [f(PAR_{su}) \times APAR_{su} + f(PAR_{sh}) \times APAR_{sh}]$$
(25)

where p and q are parameters optimized for different vegetation types and [NI] represents the vegetation index selected to characterize the canopy N concentrations. We collected the vegetation indices applied in previous studies to estimate canopy N (Table S2), which were used to drive the model. The optimal vegetation index resulting in the best model performance was determined as the nitrogen index in this model.

#### 2.3. Model parameterization and evaluation

The Shuffled Complex Evolution Procedure developed at the University of Arizona (SCE-UA) is a global optimization algorithm whose objective is to determine the optimal values of parameters within a particular range that ensure that the cost function is minimized (Duan et al., 1992). The cost function is expressed as follow:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - x_i)^2}{n}}$$
(28)

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$
(29)

$$IOA = 1 - \frac{\sum_{i=1}^{n} (y_i - x_i)^2}{\sum_{i=1}^{n} (|y_i - \overline{x}| + |x_i - \overline{x}|)^2}$$
(30)

The MOD17 GPP was the first regular, near real-time, and free-access dataset that has been widely employed to monitor terrestrial ecosystems. The revised EC-LUE model considered the effects of temperature and VPD on LUE and further integrated the effects of radiation composition and atmospheric CO<sub>2</sub> concentration. In this study, the MOD17 algorithm (Zhao et al., 2005) and revised EC-LUE (rEC-LUE) model (Zheng et al., 2020) were selected for comparison with our model. We recalibrated the MOD17 model and rEC-LUE model based on the same data and methods utilized in our model. The means of the optimized parameter values of each vegetation type were employed in the models. GPP simulations based on the MOD17 model (GPP<sub>MOD17</sub>) and rEC-LUE model (GPP<sub>rECLUE</sub>) were compared with our GPP estimates on a daily and annual scale.

In addition, to assess the sensitivity of the model to different variables including temperature, radiation, water, LAI,  $CO_2$  concentration and N index, we established two scenarios. The first scenario was the GPP per 8 days ( $GPP_{og}$ ) calculated by the original model. In the second scenario, one of the variables was changed while keeping the other variables unchanged. Specifically, we increased all time series of the variable *i* by a standard deviation, which was separately calculated according to the all time series data of each site, and kept the remaining variables unchanged. In this scenario, the GPP simulations were set to

$$d = 1 - \frac{RMSE^2}{\sum_{i=1}^n (x_i - \bar{x})^2 + \sum_{i=1}^n (y_i - \bar{x})^2 + 2 \times \left| \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \times \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{x})^2}{\sqrt{\sum_{i=1}^n (y_i - \bar{x})^2}}} \right|}$$
(26)

where RMSE is the root-mean-square error, calculated as Equation (28); n is the total number of data used for model parameter optimization;  $x_i$  and  $y_i$  represent in situ GPP and estimated GPP, respectively, from each model; and  $\overline{x}$  and  $\overline{y}$  are the mean values of in situ GPP and estimated GPP.

In this study,  $\varepsilon_{max}$  in Form1 and Form2;  $\varepsilon_{max}$  and *a* in Form3; and *p*, *q* and a in Form4 were optimized for each vegetation type using the SCE-UA algorithm. To ensure the robustness of the model, we applied a 10fold cross-validation method to optimize the parameters and validate the model. We randomly divided the in situ GPP of each vegetation type into 10 groups, using one group of data in turn for validation and the remaining 9 groups of data to optimize the parameters, ensuring that all data were involved in model calibration and validation. The coefficient of determination (R<sup>2</sup>), RMSE, and mean absolute error (MAE) were utilized to evaluate the accuracy of the GPP estimation. Furthermore, the index of agreement (IOA), which measures the degree of agreement between the estimates and the observations, was selected (Peng et al., 2017; Zhang 2015). The logical range of IOA is 0 to 1, where 1 indicates that the model estimates are consistent with the observations, and 0 represents complete disagreement. These metrics are calculated as follows:

$$R^{2} = \frac{\sum_{i=1}^{n} (y_{i} - \bar{x})^{2}}{\sum_{i=1}^{n} (x_{i} - \bar{x})^{2}}$$
(27)

 $GPP_i$ . The sensitivity of that variable was expressed as the difference between the total annual  $GPP_i$  and  $GPP_{og}$ , calculated as follows:

$$Sensitivity_i = \sum GPP_i - \sum GPP_{og}$$
(31)

#### 3. Results

# 3.1. Comparison of the GPP simulation from different indices characterizing N

The vegetation indices characterizing the canopy N concentrations (N-indices) described in Table S2 were used to calculate  $\varepsilon_{\rm max}$ . The performance of these indices was evaluated by comparing the accuracy of the GPP estimated by the models (Form 4 in Section 2.2) developed using these indices; and the results are shown in Fig. 2. The green chlorophyll index (CI<sub>green</sub>) gave the best performance with an R<sup>2</sup> of 0.68, an RMSE of 2.13 gC/m<sup>2</sup>/d, a MAE of 1.36 gC/m<sup>2</sup>/d, and an IOA of 0.915, followed by the chlorophyll/carotenoid index (CCI), green normalized difference vegetation index (GNDVI) and near-infrared reflectance (NIR). Among these indices, the R<sup>2</sup> between the GPP estimated using the modified simple ratio (MSR) and the FLUXNET GPP was the lowest at 0.65, the RMSE was the highest at 2.31 gC/m<sup>2</sup>/d, and its MAE and IOA were 1.45 gC/m<sup>2</sup>/d and 0.889, respectively. Therefore, we selected CI<sub>green</sub> as the spectral index characterizing the canopy N concentrations and introduced it to the LUE model to calculate  $\varepsilon_{\rm max}$  in this study. In the



Fig. 2. (a) Coefficient of determination (R<sup>2</sup>), (b) root-mean-square error (RMSE), (c) mean absolute error (MAE), and (d) index of agreement (IOA) of the GPP estimated model (Form 4) driven by different vegetation indices.



Fig. 3. Relationship between  $CI_{green}$  and canopy N concentrations (%). The pink area indicates the 95% confidence interval.

following results and discussion section, the [NI] in the model of Form 4 was  $CI_{green}$ . The parameters optimized according to  $CI_{green}$  are shown in Table S3.

In this study, we employed measured canopy N concentrations from previous studies and the average  $CI_{green}$  for the period corresponding to the canopy N concentrations to further corroborate that  $CI_{green}$  can be used to characterize canopy N concentrations. As shown in Fig. 3, a strong correlation between the  $CI_{green}$  and the measured canopy N concentration was obtained, with  $CI_{green}$  explaining 68% of the variation

in the canopy N concentrations.

# 3.2. Model performance of different forms

Fig. 4 shows the validations of the estimated GPP from different forms of the models. Fig. 4(a) indicates that the GPP estimates based on the model of Form 1 (GPPForm1), considering only temperature and water, performed the worst among all forms of models. As shown in Fig. 4(b), the estimated GPP (GPP<sub>Form2</sub>) based on the model of Form 2 had less improvement compared to  $\ensuremath{\mathsf{GPP}_{\mathsf{Form1}}}\xspace$  . Although the model of Form 2 further integrated the effect of CO<sub>2</sub> concentration based on Form 1, the estimation accuracy improved less compared to  $\text{GPP}_{\text{Form1}}$  in the site-scale validation. However, the effect of CO<sub>2</sub> concentration may be amplified in global long-term GPP estimates. Sunlit and shaded leaves had a significant impact on the accuracy of GPP estimates, as shown in Fig. 4(c). In terms of each metric, the estimated GPP based on the model of Form 3 (GPPForm3) had an accuracy that was substantially better than that of GPPForm2. The estimated GPP based on the model of Form 4, which simultaneously considered temperature, water, CO<sub>2</sub>, radiation components, and CIgreen, which characterized canopy N concentrations (GPP<sub>Form4</sub>), had the best performance (Fig. 4(d)).

For different vegetation types, the performance of each model is shown in Fig. 5. All forms of model-estimated GPP reflected high  $R^2$ values for CRO, DBF, ENF, GRA, SAV, WET, and WSA greater than 0.6 and IOA values greater than 0.850. For different forms of models, the performances of Form 1 and Form 2 were similar in terms of  $R^2$  and RMSE for each vegetation type, and the performance of Form 2 improved in GPP estimates of DNF, EBF, and WSA according to the MAE and IOA. The model of Form 3 had a significant improvement in CSH, DBF, DNF, and OSH. The model of form 4 had better performance in GPP estimation for different vegetation types, especially CSH, DNF, EBF, and OSH, and the estimation accuracy further improved compared to Form 3.



Fig. 4. Comparisons of the GPP estimates derived from different forms of models. The red dashed line is the 1:1 line, and the solid black line is the regression line.

#### 3.3. Seasonal variation in GPP estimated by different form models

Seasonal variation is also an important indicator for assessing the reliability of GPP simulations. A qualitative assessment of the seasonal variation was also conducted for some representative sites for the 12 vegetation types, as shown in Fig. 6. The seasonal variation in  $\varepsilon_{max}$  in Form 4 for each vegetation type is displayed in Fig. 7, where  $\varepsilon_{max}$  in Form 1, Form 2, and Form 3 was constant. The GPP estimates of the different vegetation types based on different models exhibited seasonal variation. Among these GPP estimates, GPPForm1 and GPPForm2 were similar, and the seasonal variations in  $\ensuremath{\mathsf{GPP}_{\mathsf{Form3}}}$  and  $\ensuremath{\mathsf{GPP}_{\mathsf{Form4}}}$  were closer to that in FLUXNET GPP than that in GPPForm1 and GPPForm2. Without considering the radiation components, GPP<sub>Form1</sub> and GPP<sub>Form2</sub> were significantly underestimated at peak GPP values at the DBF and SAV sites. Compared to FLUXNET GPP, GPPForm3 was overestimated at the CSH, EBF, GRA, and WET sites and underestimated at the DNF and OSH sites. However, at these sites, GPPForm4 performed better due to the dynamic adjustment of  $\varepsilon_{max}$  (Fig. 7).

#### 3.4. Comparison with MOD17 and the revised EC-LUE models

Based on the recalibrated MOD17 and rEC-LUE models, we calculated GPP<sub>MOD17</sub> and GPP<sub>rECLUE</sub> for the complete time series of 64 sites with observations over three years for comparison with GPP<sub>Form4</sub> on a daily and annual scale. Fig. 8 shows the validation of 8-day GPP<sub>Form4</sub>, GPP<sub>rECLUE</sub>, and GPP<sub>MOD17</sub> for daily values. Compared to GPP<sub>MOD17</sub>, GPP<sub>rECLUE</sub> and GPP<sub>Form4</sub> had better performance, especially GPP<sub>Form4</sub>. Fig. 9 shows a comparison of the three annual GPP estimates. GPP<sub>Form4</sub> still performed best among the three GPP estimations, especially the scatter plot, which was closer to the 1:1 line, indicating less underestimation compared to GPP<sub>rECLUE</sub>, and GPP<sub>MOD17</sub>.

#### 3.5. Sensitivities of GPP simulations to multiple variables

The correlation analysis indicated that the environmental and vegetation physiological factors have different roles in estimating daily GPP<sub>Form4</sub> (Fig. 10), and Fig. 11 shows the sensitivities of GPP simulations to multiple variables. Among all variables, GPP simulations were highly sensitive to radiation and moisture conditions, and GPP simulations were significantly correlated with PAR and EF. Since the LAI reflects structural information about the plants, its seasonal variation is similar to that of GPP, and the LAI has the strongest correlation with GPP simulation. However, the LAI was mainly used to calculate the proportion of shaded and sunlit leaves, and the sensitivity of GPP simulations to the LAI was not the highest in our model. The GPP simulations were positively correlated with temperature (r of 0.53), while the sensitivity of GPP to temperature had a high degree of dispersion. For some vegetation types or sites, higher temperatures could reduce annual GPP. On a daily scale, there was a negative correlation between the atmospheric CO<sub>2</sub> concentration and the GPP simulations. The site-scale GPP simulations were relatively insensitive to the atmospheric CO<sub>2</sub> concentration. However, the CO<sub>2</sub> concentration may have a greater contribution to the interannual variation in GPP on a global scale. CIgreen, as the N index selected in this study, is closely related to the chlorophyll and N contents of vegetation, indirectly reflects the vegetation growth status, and has a strong correlation with the GPP simulations. In the site GPP simulation based on our model,  $\ensuremath{\text{CI}_{\text{green}}}$  had a role in regulating LUE and achieving dynamic  $\varepsilon_{max}$ , which made the GPP simulations less sensitive to  $CI_{green}$ .

		(a)	R	2		0.78	(b)	RMSE(g	$(C/m^2/d)$	)	3.62
IGBP	WSA-	0.60	0.61	0.62	0.62		- 1.39	1.38	1.32	1.32	
	WET-	0.74	0.74	0.73	0.73		1.57	1.56	1.52	1.52	
	SAV-	0.71	0.71	0.72	0.72		- 1.27	1.27	1.25	1.25	
	OSH-	0.53	0.52	0.60	0.62		0.82	0.82	0.75	0.73	
	MF-	0.32	0.32	0.34	0.35		3.14	3.14	2.99	2.99	
	GRA -	0.60	0.60	0.60	0.61		- 1.38	1.37	1.42	1.38	
	ENF-	0.63	0.63	0.63	0.64		- 1.74	1.74	1.70	1.70	
	EBF -	0.54	0.54	0.46	0.56		3.27	3.10	3.24	3.01	
	DNF -	0.25	0.25	0.29	0.41		- 1.52	1.51	1.44	1.35	
	DBF -	0.64	0.64	0.78	0.78		- 2.81	2.81	2.06	2.06	
	CSH-	0.46	0.46	0.58	0.61		- 1.67	1.66	1.22	1.17	
	CRO-	0.65	0.64	0.66	0.66		3.62	3.62	3.60	3.60	
		- 1	1	1	1	0.25		1	1	1	0.73
	(c)		$MAE(g C/m^2/d)$		2.68 (d)		IOA		0.939		
		(c)	MAE(g	$C/m^2/d$ )		2.68	(d)	IC	0A		0.939
	WSA -	(c)	MAE(g	$\frac{C/m^2/d}{0.97}$	0.97	2.68	( <b>d</b> ) - 0.869	IC	0A	0.881	0.939
	WSA - WET -	(c) 1.02 1.16	MAE(g 1.01 1.16	$\frac{C/m^2/d}{0.97}$	0.97	2.68	(d) - 0.869 - 0.922	IC 0.871 0.923	0A 0.880 0.919	0.881	0.939
	WSA - WET - SAV -	(c) 1.02 1.16 0.87	MAE(g 1.01 1.16 0.87	C/m <sup>2</sup> /d) 0.97 1.11 0.88	0.97 1.12 0.87	2.68	(d) - 0.869 - 0.922 - 0.911	IC 0.871 0.923 0.910	OA 0.880 0.919 0.913	0.881 0.923 0.913	0.939
	WSA - WET - SAV - OSH -	(c) 1.02 1.16 0.87 0.52	MAE(g 1.01 1.16 0.87 0.52	C/m <sup>2</sup> /d) 0.97 1.11 0.88 0.49	0.97 1.12 0.87 0.47	2.68	(d) 0.869 0.922 0.911 0.833	IC 0.871 0.923 0.910 0.832	OA 0.880 0.919 0.913 0.867	0.881 0.923 0.913 0.875	0.939
	WSA - WET - SAV - OSH - MF -	(c) 1.02 1.16 0.87 0.52 2.19	MAE(g 1.01 1.16 0.87 0.52 2.19	C/m <sup>2</sup> /d) 0.97 1.11 0.88 0.49 2.10	0.97 1.12 0.87 0.47 2.10	2.68	(d) - 0.869 - 0.922 - 0.911 - 0.833 - 0.751	IC 0.871 0.923 0.910 0.832 0.752	0.880 0.919 0.913 0.867 0.766	0.881 0.923 0.913 0.875 0.768	0.939
BP	WSA - WET - SAV - OSH - MF - GRA -	(c) 1.02 1.16 0.87 0.52 2.19 0.73	MAE(g 1.01 1.16 0.87 0.52 2.19 0.73	C/m <sup>2</sup> /d) 0.97 1.11 0.88 0.49 2.10 0.80	0.97 1.12 0.87 0.47 2.10 0.75	2.68	(d) 0.869 0.922 0.911 0.833 0.751 0.873	IC 0.871 0.923 0.910 0.832 0.752 0.874	0.880 0.919 0.913 0.867 0.766 0.869	0.881 0.923 0.913 0.875 0.768 0.874	0.939
IGBP	WSA - WET - SAV - OSH - MF - GRA - ENF -	(c) 1.02 1.16 0.87 0.52 2.19 0.73 1.23	MAE(g 1.01 1.16 0.87 0.52 2.19 0.73 1.23	C/m <sup>2</sup> /d) 0.97 1.11 0.88 0.49 2.10 0.80 1.21	0.97 1.12 0.87 0.47 2.10 0.75 1.20	2.68	(d) - 0.869 - 0.922 - 0.911 - 0.833 - 0.751 - 0.873 - 0.886	IC 0.871 0.923 0.910 0.832 0.752 0.874 0.887	A 0.880 0.919 0.913 0.867 0.766 0.869 0.887	0.881 0.923 0.913 0.875 0.768 0.874 0.889	0.939
IGBP	WSA - WET - SAV - OSH - MF - GRA - ENF - EBF -	(c) 1.02 1.16 0.87 0.52 2.19 0.73 1.23 2.21	MAE(g 1.01 1.16 0.87 0.52 2.19 0.73 1.23 2.14	C/m <sup>2</sup> /d) 0.97 1.11 0.88 0.49 2.10 0.80 1.21 2.36	0.97 1.12 0.87 0.47 2.10 0.75 1.20 1.99	2.68	(d) - 0.869 - 0.922 - 0.911 - 0.833 - 0.751 - 0.873 - 0.886 - 0.820	IC 0.871 0.923 0.910 0.832 0.752 0.874 0.887 0.849	A 0.880 0.919 0.913 0.867 0.766 0.869 0.887 0.813	0.881 0.923 0.913 0.875 0.768 0.874 0.889 0.852	0.939
IGBP	WSA - WET - SAV - OSH - MF - GRA - ENF - EBF - DNF -	(c) 1.02 1.16 0.87 0.52 2.19 0.73 1.23 2.21 1.22	MAE(g 1.01 1.16 0.87 0.52 2.19 0.73 1.23 2.14 1.21	C/m <sup>2</sup> /d) 0.97 1.11 0.88 0.49 2.10 0.80 1.21 2.36 1.17	0.97 1.12 0.87 0.47 2.10 0.75 1.20 1.99 1.03	2.68	(d) - 0.869 - 0.922 - 0.911 - 0.833 - 0.751 - 0.873 - 0.886 - 0.820 - 0.695	IC 0.871 0.923 0.910 0.832 0.752 0.874 0.887 0.887 0.849 0.696	A 0.880 0.919 0.913 0.867 0.766 0.869 0.887 0.813 0.720	0.881 0.923 0.913 0.875 0.768 0.874 0.889 0.852 0.786	0.939
IGBP	WSA - WET - SAV - OSH - MF - GRA - ENF - EBF - DNF - DBF -	(c) 1.02 1.16 0.87 0.52 2.19 0.73 1.23 2.21 1.22 2.04	MAE(g 1.01 1.16 0.87 0.52 2.19 0.73 1.23 2.14 1.21 2.04	C/m <sup>2</sup> /d) 0.97 1.11 0.88 0.49 2.10 0.80 1.21 2.36 1.17 1.42	0.97 1.12 0.87 0.47 2.10 0.75 1.20 1.99 1.03 1.41	2.68	<ul> <li>(d)</li> <li>0.869</li> <li>0.922</li> <li>0.911</li> <li>0.833</li> <li>0.751</li> <li>0.873</li> <li>0.886</li> <li>0.820</li> <li>0.695</li> <li>0.887</li> </ul>	IC 0.871 0.923 0.910 0.832 0.752 0.874 0.887 0.849 0.696 0.886	A 0.880 0.919 0.913 0.867 0.766 0.869 0.887 0.813 0.720 0.938	0.881 0.923 0.913 0.875 0.768 0.874 0.889 0.852 0.786 0.939	0.939
IGBP	WSA - WET - SAV - OSH - GRA - ENF - EBF - DNF - DBF - CSH -	(c) 1.02 1.16 0.87 0.52 2.19 0.73 1.23 2.21 1.22 2.04 1.35	MAE(g 1.01 1.16 0.87 0.52 2.19 0.73 1.23 2.14 1.21 2.04 1.35	C/m <sup>2</sup> /d) 0.97 1.11 0.88 0.49 2.10 0.80 1.21 2.36 1.17 1.42 0.94	0.97 1.12 0.87 0.47 2.10 0.75 1.20 1.99 1.03 1.41 0.91	2.68	<ul> <li>(d)</li> <li>0.869</li> <li>0.922</li> <li>0.911</li> <li>0.833</li> <li>0.751</li> <li>0.873</li> <li>0.886</li> <li>0.820</li> <li>0.695</li> <li>0.887</li> <li>0.794</li> </ul>	IC 0.871 0.923 0.910 0.832 0.752 0.874 0.887 0.849 0.696 0.886 0.794	A 0.880 0.919 0.913 0.867 0.766 0.869 0.887 0.813 0.720 0.938 0.864	0.881 0.923 0.913 0.875 0.768 0.874 0.889 0.852 0.786 0.939 0.876	0.939
IGBP	WSA - WET - SAV - OSH - GRA - ENF - EBF - DNF - DBF - CSH - CRO -	(c) 1.02 1.16 0.87 0.52 2.19 0.73 1.23 2.21 1.22 2.04 1.35 2.61	MAE(g 1.01 1.16 0.87 0.52 2.19 0.73 1.23 2.14 1.21 2.04 1.35 2.62	C/m <sup>2</sup> /d) 0.97 1.11 0.88 0.49 2.10 0.80 1.21 2.36 1.17 1.42 0.94 2.68	0.97 1.12 0.87 0.47 2.10 0.75 1.20 1.99 1.03 1.41 0.91 2.63		(d) 0.869 0.922 0.911 0.833 0.751 0.873 0.886 0.820 0.695 0.887 0.887 0.794 0.857	IC 0.871 0.923 0.910 0.832 0.752 0.874 0.887 0.849 0.696 0.886 0.794 0.858	A 0.880 0.919 0.913 0.867 0.766 0.869 0.887 0.813 0.720 0.938 0.864 0.857	0.881 0.923 0.913 0.875 0.768 0.874 0.889 0.852 0.786 0.939 0.876 0.856	0.939

Fig. 5. Comparison of GPP estimates of different vegetation types by different forms of models.

### 4. Discussion

# 4.1. Impact of multiple environmental and vegetation biophysical factors on the GPP simulations

In this study, based on the LUE model that considered only temperature and water (Model of Form 1), the effect of atmospheric CO<sub>2</sub> concentration (Model of Form 2), the radiation components (Model of Form 3), and a vegetation index characterizing canopy N concentrations (Model of Form 4) were gradually integrated. After comparison and validation, the model of Form 4 performed best. Radiation, water, and temperature were required variables in all forms of models in this study. Since radiation was the most crucial energy driver in the model, the estimated GPPForm4 correlated well with PAR. Temperature influences the stomatal activity of vegetation during photosynthesis (Bao et al., 2022). The optimum temperature could promote vegetation photosynthetic rates, resulting in high LUE, while high temperature could lead to stomatal closure and low temperature could reduce enzyme activity, which also leads to a large difference in the sensitivity of GPP simulations to temperature. The stomatal conductance influenced by plant water stress regulates water and carbon exchange between leaves and the atmosphere (Pei et al., 2022). In this study, the EF was chosen as the probe of water stress in the LUE model. Compared to soil water indicators (e.g., soil water saturation and soil water content) and atmospheric humidity indicators (e.g., VPD), EF is more sensitive to plant water content (Zhang et al., 2015). A lower EF indicates that photosynthetic activity is limited by water, which means that less energy is available for water evaporation and carbon uptake, while a higher EF indicates that more energy is allocated to latent heat, resulting in a stronger water exchange between the ecosystem and the atmosphere (Bao et al., 2022).

The effect of atmospheric  $CO_2$  concentration was integrated into the LUE model. Although as the fuel of photosynthesis, a continued rise in atmospheric  $CO_2$  positively contributed to the increase in GPP (Bao et al., 2022; Kolby Smith et al. 2016), these two have different seasonal patterns, which also resulted in a negative correlation between GPP simulations and  $CO_2$  on a daily scale. The photosynthetic efficiency of vegetation reached its highest value during the growing season, which induced the lowest atmospheric  $CO_2$  concentration in the year, while during the nongrowing season, the photosynthetic efficiency reached its lowest value, resulting in the highest  $CO_2$  concentration of the year. Moreover, the  $CO_2$  fertilization effect is spatially and temporally heterogeneous, which may have a more obvious influence on the quantification of global GPP (Bao et al., 2022; Liu et al., 2016).

The maximum carboxylation rate and electron transport rate are limiting factors for the photosynthetic rates of sunlit leaves and shaded leaves, respectively, in the vegetation canopy (Chen 1999; De Pury and Farquhar 1997; Guan et al., 2021; Propastin et al., 2012). Because shaded leaves can only receive diffuse radiation and are not easily saturated, there are specific differences in APAR and LUE between shaded leaves and sunlit leaves (Guan et al., 2021). After the radiative component was considered in the model, the performance of Form3 was significantly better than that of Form1 and Form2. In addition, compared with other models, the estimation accuracy of rEC-LUE and Form4 considering the radiation component was better than that of MOD17. Insufficient consideration of the contribution of shade leaves leads to underestimation of GPP in areas with dense vegetation cover (Chen et al., 2020; He et al., 2018).

In previous studies,  $CI_{green}$  was considered to have a good relationship with N concentrations. He et al. (2016) obtained an  $R^2$  of 0.60~0.75 for the relationship between canopy N concentrations of winter wheat and  $CI_{green}$  at different viewing zenith angles. Mutowo



Fig. 6. Seasonal variations in the Fluxnet GPP and simulated GPP for different vegetation types. DOY means the day of the year.

et al. (2018) found CIgreen to have the highest importance among multiple vegetation indices when using the random forest algorithm to predict foliar N concentrations of woodland. Clevers and Gitelson (2013) discovered that the  $CI_{green}$  and canopy N of grass and potato were linearly correlated (R<sup>2</sup> values of 0.77 and 0.89, respectively). In this study, we used satellite reflectance to calculate CIgreen, and although N itself may not be the only variable driving the observed pattern, leaf N concentrations influence leaf traits related to photosynthetic capacity that affect reflectance (Ollinger et al., 2008). From the structure of the model, CIgreen was integrated into the LUE model to obtain dynamic  $\varepsilon_{\rm max}$ . Previous studies have also shown that  ${\rm CI}_{\rm green}$  can be used to estimate the maximum photosynthesis velocity at a high PAR level of 2000  $(\mu mol/m^2/s)$ , which can be converted to maximum incident LUE or even  $\varepsilon_{max}$ , and suggested that  $CI_{green}$  has good potential for estimating seasonal changes in  $\varepsilon_{\rm max}$  (Kanako et al., 2014; Thanyapraneedkul et al., 2012). Notably,  $\varepsilon_{max}$  is a large source of uncertainty. The  $\varepsilon_{max}$  of vegetation spatially varies, which previously was partially attributed to differences in plant species (Peltoniemi et al., 2012). Even for the same vegetation,  $\varepsilon_{max}$  may be different (Madani et al., 2014). Lin et al. (2017)

suggested that  $\varepsilon_{\rm max}$  for most vegetation types seasonally varies. Houborg et al. (2009) proposed the importance of considering the seasonal variability in  $\varepsilon_{\rm max}$  in the LUE model and believed that dynamic  $\varepsilon_{\rm max}$ could make the model more suitable because foliage adaptively responds to seasonal fluctuations in environmental conditions.

In this study, we developed a newly introduced LUE framework based on the linear relationship of  $\varepsilon_{max}$  with canopy nitrogen content. Moreno-Martínez et al. (2018) have developed a modular processing chain to derive global high-resolution maps of nitrogen content per dry mass. Although this dataset only have spatial distribution, there was still a significant positive correlation (r = 0.57) between the average value of  $\varepsilon_{max}$  in the growing season from 2001 to 2015 based on CI<sub>green</sub> and N content at the spatial scale (Fig. S1).

Our model linked the N index with dynamic  $\varepsilon_{max}$ , which provided an idea for the LUE model and improved the accuracy of the GPP estimates. In addition, dynamic  $\varepsilon_{max}$  allows the estimated GPP to be as close as possible to the observations in terms of seasonal variability and may have a greater impact on the global GPP estimation.



Fig. 7. Seasonal variations in  $\varepsilon_{max}$  of different forms of models for different vegetation types. DOY means the day of the year.



Fig. 8. Comparison of daily (a) GPP<sub>Form4</sub>, (b) GPP<sub>rECLUE</sub> and (c) GPP<sub>MOD17</sub> at 8-day intervals.



Fig. 9. Comparison of annual (a)  $GPP_{Form4}$ , (b)  $GPP_{recLUE}$  and (c)  $GPP_{MOD17}$ .



Fig. 10. Correlation (r) matrices between daily GPP<sub>Form4</sub> and the environmental and vegetation biophysical factors.

# 4.2. Uncertainties analysis

The above results demonstrated the good performance of the improved LUE model that integrated multiple environmental variables and the N-index by validation against FLUXNET GPP and other satellite GPP models. However, GPP estimated by our model may still be underestimated, which was mostly caused by uncertainty in the models.

(1) Although we used MODIS reflectance data to calculate the vegetation index characterizing canopy N concentrations in this study, the use of hyperspectral sensors likely have greater potential. Many studies used hyperspectral data to calculate the indices and revealed that these indices can be well correlated with foliar or canopy N. For example, Trotter et al. (2002) indicated a strong correlation between PRI and LUE for plant species with significant differences in leaf N content. Loozen et al. (2018) discovered significant log-linear relationships between the Medium Resolution Imaging Spectrometer (MERIS) Terrestrial Chlorophyll Index (MTCI) and canopy N concentration. Chen

et al. (2010) established the Double-peak Canopy Index (DCNI) and suggested that it correlated with N concentrations in cropland. Hyperspectral sensors can obtain more information about plant traits or nitrogen content from narrow reflectance bands. However, the spectral resolution of the sensors and the time span of the available data prevent us from applying these hyperspectral indices to global long time-series studies. In addition, we lacked consideration of the nonuniform leaf N distribution varying with canopy depth. This shortcoming may ultimately lead to an underestimation of GPP. However, leaf N concentrations based on dry weight (%) varied less with canopy depth than leaf N content based on leaf area (g N/m<sup>2</sup>) (Li et al., 2013). He et al. (2020) also found that vegetation indices could not only calculate the N concentrations at the top of the canopy but also have a good relationship with N concentrations throughout the canopy.

(2) Although the model of Form 4 performed well on a site scale, missing or poor-quality reflectance data may prevent the calculation of the nitrogen index of some pixels, which could be



**Fig. 11.** Sensitivity of GPP estimates to different variables, which represents the change in annual GPP due to a standard deviation increase in the all time series of the variable of the daily GPP estimate.

supplemented using the model of Form3. Moreover, other vegetation parameters could be applied to improve the model. For example, the LAI was utilized to scale photosynthetic capacity in the Breathing Earth System Simulator (BESS) model (Ryu et al., 2011). Croft et al. (2017) indicated that leaf chlorophyll content is a proxy for leaf photosynthetic capacity, and Luo et al. (2019) used leaf chlorophyll content to improve the maximum carboxylation rate in the BEPS model. These parameters related to the plant photosynthetic capacity deserve further exploration of their role in LUE models.

- (3) CIgreen was used to characterize canopy N concentrations, while N is the main component of chlorophyll, and both chlorophyll content and canopy N concentrations are related to vegetation growth (LAI, FPAR). Here, we have tested the correlation coefficient (r) between CIgreen and each variable, as shown in Fig. S4. Both FPAR and f (W) are positively correlated with CIgreen, and the r value of CIgreen and FPAR was 0.69. CIgreen may carry the signals of leaf chlorophyll content and canopy water content. Therefore, using vegetation index to calculate  $\varepsilon_{max}$  is difficult to avoid double consideration of effects from FPAR (or APAR) and f (W). In the optimization process, we used SCE-UA algorithm to optimize the parameters, there is a good correlation between  $CI_{green}$  and GPP (r = 0.47). However, f (W) and APAR were calculated directly from the input data, and there were no parameters in f(W) and APAR need to be optimized. The r value between  $CI_{green}$  and  $GPP/(APAR \times f(W))$  was only 0.02, so that the signals of f(W) and APAR might be weakened in the process of parameter optimization for CIgreen. This part needs to be further explored in future research. It may be a good choice to directly use long time-series canopy N concentration product instead of vegetation index, which can eliminate the influence of double consideration caused by vegetation index related to other variables.
- (4) The spatial resolution of the main input data to our model (FPAR and LAI) determined the spatial scale for estimated GPP, while the longitudinal length scale of the flux tower footprints is typically 500~2000 m (Baldocchi 2014). To explore this uncertainty, we first employed 500 m GLASS LAI products and compared it with 5 km LAI data. As shown in Fig. S2, 500 m LAI and 5 km LAI have high consistency in the sites selected in this study. Then, we used the parameters optimized based on 5 km LAI in our manuscript and replaced the input data with 500 m LAI, the GPP simulation accuracy were shown in Fig. S3 (a). In addition, we

re-optimized the parameters based on 500 m LAI data and evaluated the performance of GPP simulations using 500 m LAI data as input data, as shown in Fig. S3 (b). Compared with the GPP estimation accuracy of the model of Form4 (Fig. 4(d)), replacing 5 km LAI with 500 m LAI and using the original parameters reduced the accuracy of GPP estimation. However, the accuracy of GPP estimation using 500 m LAI and parameters re-optimized by 500 m LAI was similar to that of the model in Fig.4(d). Although the scale problem still existed, we tried our best to eliminate this part of the error by ensuring the homogeneity of the underlying surface of the selected sites. And since the purpose of this study is to develop a model that can be used to estimate the global GPP, the input data with a spatial resolution of 0.05 ° may be suitable for this purpose. Some previous studies also used the input data with a spatial resolution of 0.05 °to calibrate models and to estimate global GPP (Bai et al., 2022; Bi et al., 2022; Li and Xiao 2019).

- (5) Generally, the photosynthesis of C4 plants is stronger than that of C3 plants under the same climatic conditions (Wang et al., 2021c). The traditional LUE model with fixed  $\varepsilon_{max}$  without distinguishing these two important plant functional types may lead to the overestimation and underestimation of GPP for C3 plants and C4 plants, respectively. Likewise, in our model, the same coefficients before the N index of C3/C4 plants may also cause uncertainty. In future research, we may need fine biotype products to separately optimize the parameters of C3/C4 vegetation, which would be beneficial to avoid uncertainty for cropland and grassland (Wang et al., 2021c; Yuan et al., 2015).
- (6) In this study, we employed in situ GPP from all the sites of each vegetation type to optimize the model, which is also referred to as joint optimization, and the parameters of joint optimization are more representative and less variable (Huang et al., 2021). However, there are great differences in the optimal parameters among different climatic regions. Huang et al. (2021) discovered significant differences in the parameters of GRA at high latitudes with temperature and in dry areas with high temperature. Lin et al. (2021) have also suggested that there are still differences in photosynthetic characteristics even among the same vegetation types in different climatic regions into various plant functional types could reduce the uncertainty in terrestrial carbon assessment.

Moreover, unbalanced training data may explain any systematic errors. However, we obtained 10 sets of train data based on the 10-fold cross-validation method and each set of training data contained a different number of low, medium and high GPP values, but the final parameters were still within an acceptable range. We believed that the form of the model is the main reason for this phenomenon. Different stress factors in the model were calculated in a fixed form, used the same parameters that were directly set according to previous studies, such as the optimal temperature, the minimum temperature and the carbon dioxide compensation point. Different forms of stress factors greatly affect the GPP simulation (Bao et al., 2022). Since there is no fixed model form for some machine learning models, as long as the appropriate variables are selected, the estimated GPP will be as close to 1:1 as possible. However, the appropriate forms of different stress factors in LUE model should be further explored.

# 5. Conclusions

In this study, we developed an improved LUE model that simultaneously incorporated the effects of temperature, water, atmospheric CO<sub>2</sub>, radiation components, and canopy N concentrations to achieve a daily GPP estimation with an 8-day temporal resolution. According to the linear correlation between canopy N concentrations and  $\varepsilon_{max}$ , we selected a vegetation index that could characterize the canopy N

concentrations and incorporate it into the LUE model. By comparing the models' performance with different indices, we determined that  $CI_{green}$  was the best index introduced to the LUE model by calculating  $\varepsilon_{max}$ . Moreover, using the measurements from previous studies,  $CI_{green}$  was further found to have a good correlation with the canopy N concentrations, and the value of  $R^2$  reached 0.68.

To demonstrate the superiority of the GPP estimates using the LUE model integrating multiple environmental variables and canopy N concentrations ( $GPP_{Form4}$ ), we compared it with the GPP estimation using the LUE model considering only temperature and water (GPP<sub>Form1</sub>) and with the GPP estimation by further stepwise addition of the effects of CO2 and radiation components (GPPForm2 and GPForm3, respectively). The GPP estimations by each model were validated using FLUXNET GPP. The results showed that the accuracy of GPP<sub>Form4</sub> was the highest (R<sup>2</sup> =0.69, RMSE = 2.13 gC/m<sup>2</sup>/d, MAE =1.36 gC/m<sup>2</sup>/d, and IOA =0.915). In addition, GPP<sub>Form4</sub> also had good performance for different vegetation types. It was determined that considering the radiation component greatly improved the accuracy of the GPP estimates, and integrating the vegetation index representing the canopy N concentrations achieved a dynamic  $\varepsilon_{max}$ , which was beneficial to reducing the degrees of underestimation at peak values and those of overestimation in the dormant season. Further comparisons with the MOD17 GPP algorithm and the revised ECLUE model showed that GPPForm4 exhibited higher overall accuracy on a daily and annual scale, showing fewer underestimates compared to other products.

In conclusion, this study is the first attempt to integrate temperature, water, the effects of CO<sub>2</sub> fertilization, radiation components, and N into the LUE model. The vegetation index characterizing canopy N concentrations was introduced to the LUE model to obtain spatiotemporally dynamic  $\varepsilon_{max}$ , which is more satisfying in GPP estimation than with fixed values. The proposed model has the potential to map global long-term GPP. To improve the accuracy of the model, future improvements of the model will include distinguishing between C3 and C4 plants and attempting to integrate other vegetation indices or vegetation parameters that are considered probes of photosynthetic capacity.

#### **Declaration of Competing Interest**

The authors declare no conflict of interest.

## Data availability

Data will be made available on request.

# Acknowledgements

This work was supported by the National Natural Science Foundation of China (42271330). We thank NASA for providing the MODIS reflectance products (MCD43A4 & MYDOCGA) and all fluxnet communities of the eddy covariance data (FLUXNET2015, https://fluxnet.fluxdata. org/data/). We are grateful to Mikko Peltoniemi, Scott V. Ollinger, Takagi Kentaro, and Jeffrey M. Klopatek for the canopy nitrogen concentration data. We also thank the anonymous reviewers for their helpful suggestions on the manuscript.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2023.109359.

#### References

Ainsworth, E.A., Long, S.P., 2005. What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO2. New Phytol. 165, 351–372.

- Alemohammad, S.H., Fang, B., Konings, A.G., et al., 2017. Water, energy, and carbon with artificial neural networks (WECANN): a statistically based estimate of global surface turbulent fluxes and gross primary productivity using solar-induced fluorescence. Biogeosciences 14, 4101–4124.
- Alton, P.B., North, P.R., Los, S.O., 2007. The impact of diffuse sunlight on canopy lightuse efficiency, gross photosynthetic product and net ecosystem exchange in three forest biomes. Glob. Chang. Biol. 13, 776–787.
- Bai, J., Zhang, H., Sun, R., et al., 2022. Estimation of global GPP from GOME-2 and OCO-2 SIF by considering the dynamic variations of GPP-SIF relationship. Agric. For. Meteorol. 326, 109180.
- Baldocchi, D., 2014. Measuring fluxes of trace gases and energy between ecosystems and the atmosphere – the state and future of the eddy covariance method. Glob. Chang. Biol. 20, 3600–3609.
- Balzarolo, M., Valdameri, N., Fu, Y.H., et al., 2019. Different determinants of radiation use efficiency in cold and temperate forests. Global Ecol. Biogeogr. 28, 1649–1667. Bao, S., Wutzler, T., Koirala, S., et al., 2022. Environment-sensitivity functions for gross
- primary productivity in light use efficiency models. Agric. For. Meteorol. 312, 108708.
- Barton, C.V.M., North, P.R.J., 2001. Remote sensing of canopy light use efficiency using the photochemical reflectance index: model and sensitivity analysis. Remote Sens. Environ. 78, 264–273.
- Beer, C., Reichstein, M., Tomelleri, E., et al., 2010. Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. Science 329, 834–838.
- Bi, W., He, W., Zhou, Y., et al., 2022. A global 0.05° dataset for gross primary production of sunlit and shaded vegetation canopies from 1992 to 2020. Sci. Data 9, 213.
- Bonan, G.B., Lombardozzi, D.L., Wieder, W.R., et al., 2019. Model structure and climate data uncertainty in historical simulations of the terrestrial carbon cycle (1850–2014). Glob. Biogeochem. Cycles 33, 1310–1326.
- Butler, E.E., Datta, A., Flores-Moreno, H., et al., 2017. Mapping local and global variability in plant trait distributions. Proc. Natl. Acad. Sci. 114, E10937.
- Chen, B., Arain, M.A., Chen, J.M., et al., 2020. Importance of Shaded Leaf Contribution to the Total GPP of Canadian Terrestrial Ecosystems: evaluation of MODIS GPP. J. Geophys. Res. 125, e2020JG005917.
- Chen, J.M., 1999. Spatial Scaling of a Remotely Sensed Surface Parameter by Contexture. Remote Sens. Environ. 69, 30–42.
- Chen, J.M., Liu, J., Cihlar, J., et al., 1999. Daily canopy photosynthesis model through temporal and spatial scaling for remote sensing applications. Ecol. Modell. 124, 99–119.
- Chen, J.M., Mo, G., Pisek, J., et al., 2012. Effects of foliage clumping on the estimation of global terrestrial gross primary productivity. Glob. Biogeochem. Cycles 26.
- Chen, P., Haboudane, D., Tremblay, N., et al., 2010. New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. Remote Sens. Environ. 114, 1987–1997.
- Chen, Y., Feng, X., Fu, B., et al., 2021. Improved global maps of the optimum growth temperature, maximum light use efficiency, and gross primary production for vegetation. J. Geophys. Res. 126 e2020JG005651.
- Clevers, J.G.P.W., Gitelson, A.A., 2013. Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3. Int. J. Appl. Earth Obs. Geoinf. 23, 344–351.
- Clevers, J.G.P.W., Kooistra, L., 2012. Using hyperspectral remote sensing data for retrieving canopy chlorophyll and nitrogen content. IEEE J. Sel. Top. Appl. Earth Observat. Remote Sens. 5, 574–583.
- Croft, H., Chen, J.M., Luo, X., et al., 2017. Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. Glob. Chang. Biol. 23, 3513–3524.
- De Pury, D.G.G., Farquhar, G.D., 1997. Simple scaling of photosynthesis from leaves to canopies without the errors of big-leaf models. Plant Cell Environ. 20, 537–557.
- Dong, T., Liu, J., Qian, B., et al., 2017. Deriving maximum light use efficiency from crop growth model and satellite data to improve crop biomass estimation. IEEE J. Sel. Top. Appl. Earth Observat. Remote Sens. 10, 104–117.
- Dou, X., Yang, Y., 2018. Estimating forest carbon fluxes using four different data-driven techniques based on long-term eddy covariance measurements: model comparison and evaluation. Sci. Total Environ. 627, 78–94.
- Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour Res 28, 1015–1031.
- Fang, J., Yu, G., Liu, L., et al., 2018. Climate change, human impacts, and carbon sequestration in China. Proc. Natl. Acad. Sci. 115, 4015–4020.
- Farquhar, G.D., von Caemmerer, S., Berry, J.A., 1980. A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. Planta 149, 78–90.
- Guan, X., Chen, J.M., Shen, H., et al., 2021. A modified two-leaf light use efficiency model for improving the simulation of GPP using a radiation scalar. Agric. For. Meteorol. 307, 108546.
- Guo, J., Trotter, C.M., 2004. Estimating photosynthetic light-use efficiency using the photochemical reflectance index: variations among species. Funct. Plant Biol. 31, 255–265.
- He, J., Zhang, X., Guo, W., et al., 2020. Estimation of vertical leaf nitrogen distribution within a rice canopy based on hyperspectral data. Front Plant Sci 10.
- He, L., Chen, J.M., Gonsamo, A., et al., 2018. Changes in the shadow: the shifting role of shaded leaves in global carbon and water cycles under climate change. Geophys. Res. Lett. 45, 5052–5061.
- He, L., Zhang, H., Zhang, Y., et al., 2016. Estimating canopy leaf nitrogen concentration in winter wheat based on multi-angular hyperspectral remote sensing. Eur. J. Agron. 73, 170–185.
- He, M., Ju, W., Zhou, Y., 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.

Heinsch, F.A., Reeves, M., Votava, P., et al., 2003. User's guide GPP and NPP (MOD17A2/A3) products NASA MODIS land algorithm. Version 2, 666–684.

Hikosaka, K., 2004. Interspecific difference in the photosynthesis-nitrogen relationship: patterns, physiological causes, and ecological importance. J. Plant Res. 117,

- 481–494. Houborg, R., Anderson, M.C., Norman, J.M., et al., 2009. Intercomparison of a 'bottomup' and 'top-down' modeling paradigm for estimating carbon and energy fluxes over a variety of vegetative regimes across the U.S. Agric. For. Meteorol. 149, 1875–1895.
- Huang, M., Piao, S., Ciais, P., et al., 2019. Air temperature optima of vegetation productivity across global biomes. Nat. Ecol. Evol. 3, 772–779.
- Huang, X., Xiao, J., Wang, X., et al., 2021. Improving the global MODIS GPP model by optimizing parameters with FLUXNET data. Agric. For. Meteorol. 300, 108314.
- Hunt Jr, E.R., Piper, S.C., Nemani, R., et al., 1996. Global net carbon exchange and intraannual atmospheric CO2 concentrations predicted by an ecosystem process model and three-dimensional atmospheric transport model. Glob. Biogeochem. Cycles 10, 431–456.
- Ichii, K., Ueyama, M., Kondo, M., et al., 2017. New data-driven estimation of terrestrial CO2 fluxes in Asia using a standardized database of eddy covariance measurements, remote sensing data, and support vector regression. J. Geophys. Res.: Biogeosci. 122, 767–795.
- Jacobson, A.R., Schuldt, K.N., Miller, J.B., et al., 2020. CarbonTracker CT2019B. NOAA Global Monitoring Laboratory.
- Kalliokoski, T., Mäkelä, A., Fronzek, S., et al., 2018. Decomposing sources of uncertainty in climate change projections of boreal forest primary production. Agric. For. Meteorol. 262, 192–205.
- Kanako, M., Shinobu, F., Noriko, S., et al., 2014. Estimating the seasonal maximum light use efficiency. In: Proc.SPIE, p. 92603R.
- Keenan, T.F., Luo, X., De Kauwe, M.G., et al., 2021. A constraint on historic growth in global photosynthesis due to increasing CO2. Nature 600, 253–258.
   Kolby Smith, W., Reed, S.C., Cleveland, C.C., et al., 2016. Large divergence of satellite
- Kolby Smith, W., Reed, S.C., Cleveland, C.C., et al., 2016. Large divergence of satellite and Earth system model estimates of global terrestrial CO2 fertilization. Nat. Clim. Chang 6, 306–310.
- Koyama, K., Kikuzawa, K., 2010. Geometrical similarity analysis of photosynthetic light response curves, light saturation and light use efficiency. Oecologia 164, 53–63.
- Kurc, S.A., Small, E.E., 2004. Dynamics of evapotranspiration in semiarid grassland and shrubland ecosystems during the summer monsoon season, central New Mexico. Water Resour. Res. 40.
- Lepine, L.C., Ollinger, S.V., Ouimette, A.P., et al., 2016. Examining spectral reflectance features related to foliar nitrogen in forests: implications for broad-scale nitrogen mapping. Remote Sens. Environ. 173, 174–186.
- Leuning, R., 1995. A critical appraisal of a combined stomatal-photosynthesis model for C3 plants. Plant Cell Environ. 18, 339–355.
- Leverenz, J.W., 1987. Chlorophyll content and the light response curve of shade-adapted conifer needles. Physiol. Plant 71, 20–29.
- Li, H., Zhao, C., Huang, W., et al., 2013. Non-uniform vertical nitrogen distribution within plant canopy and its estimation by remote sensing: a review. Field Crops Res. 142, 75–84.
- Li, X., Xiao, J., 2019. Mapping photosynthesis solely from solar-induced chlorophyll fluorescence: a global, fine-resolution dataset of gross primary production derived from OCO-2. Remote Sens. (Basel) 11.
- Liang, S., Zhao, X., Liu, S., et al., 2013. A long-term Global LAnd Surface Satellite (GLASS) data-set for environmental studies. Int. J. Digital Earth 6, 5–33.
- Lin, S., Li, J., Liu, Q., et al., 2021. Improved global estimations of gross primary productivity of natural vegetation types by incorporating plant functional type. Int. J. Appl. Earth Obs. Geoinf. 100, 102328.
- Lin, X., Chen, B., Chen, J., 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.M., Cihlar, J., et al., 1997. A process-based boreal ecosystem productivity simulator using remote sensing inputs. Remote Sens. Environ. 62, 158–175.
- Liu, S., Zhuang, Q., He, Y., et al., 2016. Evaluating atmospheric CO2 effects on gross primary productivity and net ecosystem exchanges of terrestrial ecosystems in the conterminous United States using the AmeriFlux data and an artificial neural network approach. Agric. For. Meteorol. 220, 38–49.
- Liu, X., Liu, Z., Liu, L., et al., 2021. Modelling the influence of incident radiation on the SIF-based GPP estimation for maize. Agric. For. Meteorol. 307, 108522.
- Loozen, Y., Rebel, K.T., de Jong, S.M., et al., 2020. Mapping canopy nitrogen in European forests using remote sensing and environmental variables with the random forests method. Remote Sens. Environ. 247, 111933.
- Loozen, Y., Rebel, K.T., Karssenberg, D., et al., 2018. Remote sensing of canopy nitrogen at regional scale in Mediterranean forests using the spaceborne MERIS Terrestrial Chlorophyll Index. Biogeosciences 15, 2723–2742.
- Luo, X., Croft, H., Chen, J.M., et al., 2019. Improved estimates of global terrestrial photosynthesis using information on leaf chlorophyll content. Glob. Chang. Biol. 25, 2499–2514.
- Madani, N., Kimball, J.S., Affleck, D.L.R., et al., 2014. Improving ecosystem productivity modeling through spatially explicit estimation of optimal light use efficiency. J. Geophys. Res.: Biogeosci. 119, 1755–1769.
- MÄKelä, A., Pulkkinen, M., Kolari, P., 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, 92–108.
- Marie, W., Frédéric, B., Ranga, B.M., et al., 2000. Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data. Agronomie 20, 3–22.

- Medlyn, B.E., Dreyer, E., Ellsworth, D., et al., 2002. Temperature response of parameters of a biochemically based model of photosynthesis. II. A review of experimental data. Plant Cell Environ. 25, 1167–1179.
- Melillo, J.M., McGuire, A.D., Kicklighter, D.W., et al., 1993. Global climate change and terrestrial net primary production. Nature 363, 234–240.
- Monteith, J.L., 1972. Solar Radiation and Productivity in Tropical Ecosystems. J. Appl. Ecol. 9, 747–766.
- Monteith, J.L., Moss, C.J., Cooke, G.W., et al., 1977. Climate and the efficiency of crop production in Britain. Philos. Trans. R. Soc. Lond. B, Biol. Sci. 281, 277–294.
- Moreno-Martínez, Á., Camps-Valls, G., Kattge, J., et al., 2018. A methodology to derive global maps of leaf traits using remote sensing and climate data. Remote Sens. Environ. 218, 69–88.
- Mutowo, G., Mutanga, O., Masocha, M., 2018. Evaluating the applications of the nearinfrared region in mapping foliar N in the miombo woodlands. Remote Sens. (Basel) 10.
- Norby, R.J., DeLucia, E.H., Gielen, B., et al., 2005. Forest response to elevated CO< sub>2</sub&gt; is conserved across a broad range of productivity. Proc. Natl. Acad. Sci. U.S.A. 102, 18052.
- Norby, R.J., Wullschleger, S.D., Gunderson, C.A., et al., 1999. Tree responses to rising CO2 in field experiments: implications for the future forest. Plant Cell Environ. 22, 683–714.
- Ollinger, S.V., Richardson, A.D., Martin, M.E., et al., 2008. Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: functional relations and potential climate feedbacks. Proc. Natl. Acad. Sci. 105, 19336.
- Parton, W.J., Scurlock, J.M.O., Ojima, D.S., et al., 1993. Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide. Glob. Biogeochem. Cycles 7, 785–809.
- Pei, Y., Dong, J., Zhang, Y., et al., 2022. Evolution of light use efficiency models: improvement, uncertainties, and implications. Agric. For. Meteorol. 317, 108905.
- Peltoniemi, M., Pulkkinen, M., Kolari, P., et al., 2012. Does canopy mean nitrogen concentration explain variation in canopy light use efficiency across 14 contrasting forest sites? Tree Physiol. 32, 200–218.
- Peng, D., Zhang, X., Wu, C., et al., 2017. Intercomparison and evaluation of spring phenology products using National Phenology Network and AmeriFlux observations in the contiguous United States. Agric. For. Meteorol. 242, 33–46.
- Potter, C.S., Randerson, J.T., Field, C.B., et al., 1993. Terrestrial ecosystem production: a process model based on global satellite and surface data. Glob. Biogeochem. Cycles 7, 811–841.
- Propastin, P., Ibrom, A., Knohl, A., et al., 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.
- Raich, J.W., Rastetter, E.B., Melillo, J.M., et al., 1991. Potential Net Primary Productivity in South America: application of a Global Model. Ecol. Appl. 1, 399–429.
- Reich, P.B., 2012. Key canopy traits drive forest productivity. Proc. R. Soc. B 279, 2128–2134.
- 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, 125–154.
- Ryu, Y., Baldocchi, D.D., Kobayashi, H., et al., 2011. Integration of MODIS land and atmosphere products with a coupled-process model to estimate gross primary productivity and evapotranspiration from 1km to global scales. Global. Biogeochem. Cycles 25.
- Ryu, Y., Berry, J.A., Baldocchi, D.D., 2019. What is global photosynthesis? History, uncertainties and opportunities. Remote Sens. Environ. 223, 95–114.
- Savitzky, A., Golay, M.J.E., 1964. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal. Chem. 36, 1627–1639.
- Sharpe, D.M., 1975. Methods of assessing the primary production of regions. In: Lieth, H., Whittaker, & R.H. (Eds.), Primary Productivity of the Biosphere. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 147–166.
- Stocker, B.D., Wang, H., Smith, N.G., et al., 2020. P-model v1.0: an optimality-based light use efficiency model for simulating ecosystem gross primary production. Geosci. Model Dev. 13, 1545–1581.
- Stocker, B.D., Zscheischler, J., Keenan, T.F., et al., 2019. Drought impacts on terrestrial primary production underestimated by satellite monitoring. Nat. Geosci. 12, 264–270.
- Suleiman, A., Crago, R., 2004. Hourly and daytime evapotranspiration from grassland using radiometric surface temperatures. Agron. J. 96, 384–390.
- Swinnen, E., Toté, C., & Van Hoolst, R. (2021). Global Land component of the Copernicus Land Service: dry Matter Productivity product (1km resolution, Version 1). In: https://land.copernicus.eu/global/products/dmp.
- Tang, S., Chen, J.M., Zhu, Q., et al., 2007. LAI inversion algorithm based on directional reflectance kernels. J. Environ. Manage. 85, 638–648.
- Tang, Z., Xu, W., Zhou, G., et al., 2018. Patterns of plant carbon, nitrogen, and phosphorus concentration in relation to productivity in China's terrestrial ecosystems. Proc. Natl. Acad. Sci. 115, 4033.
- Thanyapraneedkul, J., Muramatsu, K., Daigo, M., et al., 2012. A Vegetation index to estimate terrestrial gross primary production capacity for the global change observation mission-climate (GCOM-C)/Second-Generation global imager (SGLI) satellite sensor. Remote Sens (Basel) 3689–3720.
- Trotter, G.M., Whitehead, D., Pinkney, E.J., 2002. The photochemical reflectance index as a measure of photosynthetic light use efficiency for plants with varying foliar nitrogen contents. Int. J. Remote Sens. 23, 1207–1212.
- Uchijima, Z., Seino, H., 1985. Agro climatic evaluation of net primary productivity of natural vegetation I. Chikugo model for evaluating productivity. J. Agric. Meteorol. (Japan) 40, 343–352.

#### H. Zhang et al.

- Veroustraete, F., Sabbe, H., Eerens, H., 2002. Estimation of carbon mass fluxes over Europe using the C-Fix model and Euroflux data. Remote Sens. Environ. 83, 376–399.
- Verrelst, J., Rivera-Caicedo, J.P., Reyes-Muñoz, P., et al., 2021. Mapping landscape canopy nitrogen content from space using PRISMA data. ISPRS J. Photogramm. Remote Sens. 178, 382–395.
- Wang, F., Harindintwali, J.D., Yuan, Z., et al., 2021a. Technologies and perspectives for achieving carbon neutrality. The Innovation 2.
- Wang, J., Sun, R., Zhang, H., et al., 2021b. New Global MuSyQ GPP/NPP Remote Sensing Products From 1981 to 2018. IEEE J. Sel. Top. Appl. Earth Observat. Remote Sens. 14, 5596–5612.
- Wang, S., Huang, K., Yan, H., 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., Zhang, Y., Ju, W., et al., 2020. Recent global decline of CO2 fertilization effects on vegetation photosynthesis. Science 370, 1295–1300.
- Wang, Y., Li, R., Hu, J., et al., 2021c. Daily estimation of gross primary production under all sky using a light use efficiency model coupled with satellite passive microwave measurements. Remote Sens. Environ. 267, 112721.
- Wen, P., He, J., Ning, F., et al., 2019. Estimating leaf nitrogen concentration considering unsynchronized maize growth stages with canopy hyperspectral technique. Ecol. Indic. 107, 105590.
- Xiao, J., Ollinger, S.V., Frolking, S., et al., 2014. Data-driven diagnostics of terrestrial carbon dynamics over North America. Agric. For. Meteorol. 197, 142–157.
- Xiao, J., Zhuang, Q., Baldocchi, D.D., et al., 2008. Estimation of net ecosystem carbon exchange for the conterminous United States by combining MODIS and AmeriFlux data. Agric. For. Meteorol. 148, 1827–1847.
- Xiao, X., Zhang, Q., Braswell, B., et al., 2004. Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data. Remote Sens. Environ. 91, 256–270.
- Xiao, Z., Liang, S., Jiang, B., 2017. Evaluation of four long time-series global leaf area index products. Agric. For. Meteorol. 246, 218–230.
- Xiao, Z., Liang, S., Sun, R., 2018. Evaluation of three long time series for global fraction of absorbed photosynthetically active radiation (FAPAR) products. IEEE Trans. Geosci. Remote Sens. 56, 5509–5524.

- Yang, D., Xu, X., Xiao, F., et al., 2021. Improving modeling of ecosystem gross primary productivity through re-optimizing temperature restrictions on photosynthesis. Sci. Total Environ. 788, 147805.
- Yuan, W., Cai, W., Nguy-Robertson, A.L., et al., 2015. Uncertainty in simulating gross primary production of cropland ecosystem from satellite-based models. Agric. For. Meteorol. 207, 48–57.
- Yuan, W., Cai, W., Xia, J., 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., Liu, S., Zhou, G., 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, 189–207.
- Zhang, M., Yu, G.-R., Zhuang, J., et al., 2011. Effects of cloudiness change on net ecosystem exchange, light use efficiency, and water use efficiency in typical ecosystems of China. Agric. For. Meteorol. 151, 803–816.
- Zhang, X., 2015. Reconstruction of a complete global time series of daily vegetation index trajectory from long-term AVHRR data. Remote Sens. Environ. 156, 457–472.
- Zhang, X., Liang, S., Wang, K., et al., 2010. Analysis of Global Land Surface Shortwave Broadband Albedo From Multiple Data Sources. IEEE J. Sel. Top. Appl. Earth Observat. Remote Sens. 3, 296–305.
- Zhang, Y., Song, C., Sun, G., et al., 2015. Understanding moisture stress on light use efficiency across terrestrial ecosystems based on global flux and remote-sensing data. J. Geophys. Res.: Biogeosci. 120, 2053–2066.
- Zhao, M., Heinsch, F.A., Nemani, R.R., et al., 2005. Improvements of the MODIS terrestrial gross and net primary production global data set. Remote Sens. Environ. 95, 164–176.
- Zheng, Y., Shen, R., Wang, Y., et al., 2020. Improved estimate of global gross primary production for reproducing its long-term variation, 1982–2017. Earth Syst. Sci. Data 12, 2725–2746.
- Zhou, X., Xin, Q., 2019. Improving satellite-based modelling of gross primary production in deciduous broadleaf forests by accounting for seasonality in light use efficiency. Int. J. Remote Sens. 40, 931–955.