



OPEN

DATA DESCRIPTOR

A global 0.05° dataset for gross primary production of sunlit and shaded vegetation canopies from 1992 to 2020

Wenjun Bi¹, Wei He^{2,3}, Yanlian Zhou^{1,4}✉, Weimin Ju^{1,4}, Yibo Liu⁵, Yang Liu⁶, Xiaoyu Zhang¹, Xiaonan Wei¹ & Nuo Cheng¹

Distinguishing gross primary production of sunlit and shaded leaves (GPP_{sun} and GPP_{shade}) is crucial for improving our understanding of the underlying mechanisms regulating long-term GPP variations. Here we produce a global 0.05°, 8-day dataset for GPP , GPP_{shade} and GPP_{sun} over 1992–2020 using an updated two-leaf light use efficiency model (TL-LUE), which is driven by the GLOBMAP leaf area index, CRUJRA meteorology, and ESA-CCI land cover. Our products estimate the mean annual totals of global GPP , GPP_{sun} and GPP_{shade} over 1992–2020 at 125.0 ± 3.8 (mean \pm std) $Pg\ C\ a^{-1}$, $50.5 \pm 1.2\ Pg\ C\ a^{-1}$, and $74.5 \pm 2.6\ Pg\ C\ a^{-1}$, respectively, in which EBF (evergreen broadleaf forest) and CRO (crops) contribute more than half of the totals. They show clear increasing trends over time, in which the trend of GPP (also GPP_{sun} and GPP_{shade}) for CRO is distinctively greatest, and that for DBF (deciduous broadleaf forest) is relatively large and GPP_{shade} overwhelmingly outweighs GPP_{sun} . This new dataset advances our in-depth understanding of large-scale carbon cycle processes and dynamics.

Background & Summary

Gross primary production (GPP) is a vital component of the terrestrial carbon budget and plays a prominent role in the global carbon cycle^{1–4}. Accurate estimation of terrestrial GPP is critical for understanding the interaction between the terrestrial biosphere and the atmosphere in the context of global climate change^{5,6}, projecting future change⁷, and informing climate policy decisions⁸. Therefore, characterizing the spatiotemporal variation of GPP⁹ is a key issue in the climate change study.

GPP is closely related to vegetation types^{10–12}, meteorological factors^{13–17}, soil moisture^{18,19}, and other factors. In particular, GPP is affected by vegetation canopy structures^{12,20,21}, e.g., sunlit and shaded leaves. Sunlit leaves can absorb direct and diffuse radiation simultaneously, and light saturation is easy to occur when the radiation is high, while shaded leaves can only absorb diffuse radiation and the absorbed radiation intensity is generally between the light compensation point and the light saturation point^{22–24}. The two components of GPP derived from sunlit (GPP_{sun}) and shaded leaves (GPP_{shade}) have drawn increasing attentions recently due to two reasons. First, commonly used sun-induced chlorophyll fluorescence (SIF), which is strongly correlated with GPP at various scales^{25–30}, is mainly emitted from sunlit leaves^{31–34} and GPP_{sun} can also be used to retrieve key photosynthetic parameters such as maximum carboxylation velocity (V_{cmax})³⁵. Besides, the contribution of GPP_{shade} to the total GPP increased with the increases of leaf area index (LAI) and diffuse radiation ratio^{36,37}. Therefore, it is of great

¹Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Key Laboratory for Land Satellite Remote Sensing Applications of Ministry of Natural Resources, School of Geography and Ocean Science, Nanjing University, Nanjing, Jiangsu, 210023, China. ²International Institute for Earth System Science, Nanjing University, Nanjing, 210023, China. ³State Key Laboratory of Remote Sensing Science Jointly Sponsored by Beijing Normal University and Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100854, China. ⁴Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, Jiangsu, 210023, China. ⁵Jiangsu Key Laboratory of Agricultural Meteorology, School of Applied Meteorology, Nanjing University of Information Science and Technology, Nanjing, 210044, China. ⁶State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China. ✉e-mail: zhouyl@nju.edu.cn

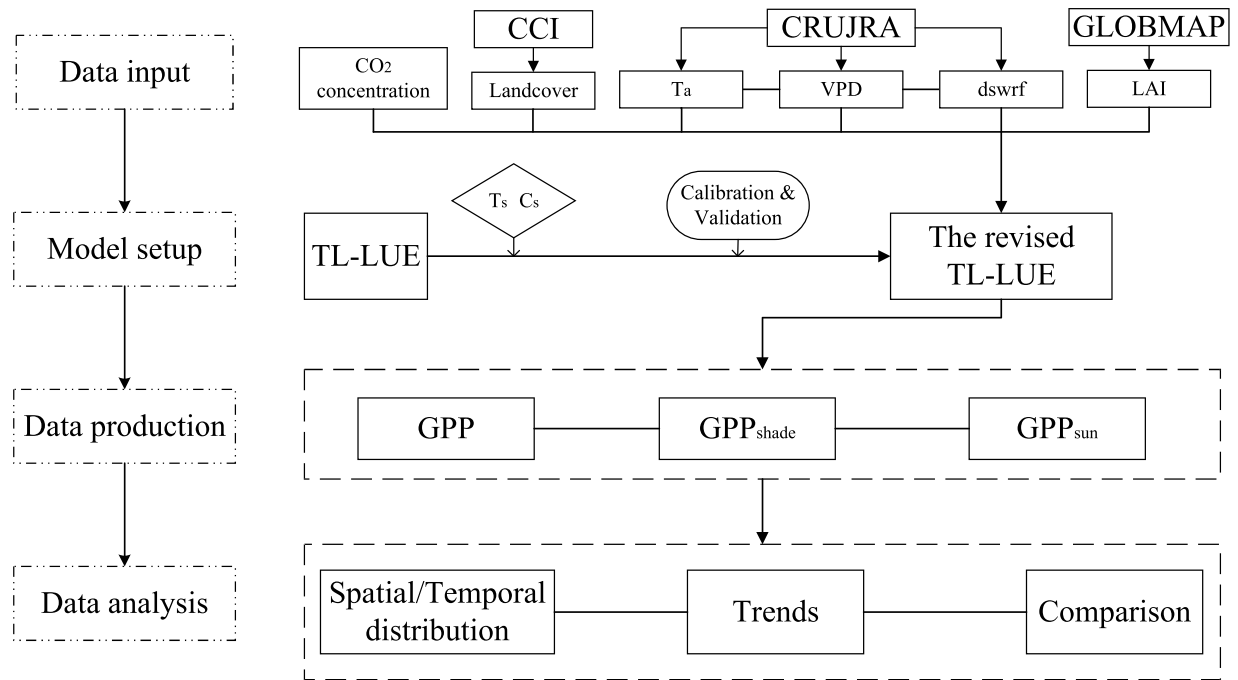


Fig. 1 Workflow of the study. CCI represents European Space Agency Climate Change Initiative Land Cover. CRUJRA represents the Climatic Research Unit and Japanese reanalysis. VPD is vapor pressure deficit. T_a is air temperature. T_s and C_s are regulation scalars for temperature and CO_2 concentration. The dswrf is downward solar radiation flux. TL-LUE is two-leaf light use efficiency model. GPP_{sun} and GPP_{shade} are GPP derived by sunlit and shaded leaves.

importance to distinguish GPP_{shade} and GPP_{sun} respectively for building an improved SIF-GPP relationship and for obtaining high-precision photosynthetic parameters to feed carbon cycle models. Second, some process-based terrestrial biosphere models simulate GPP_{sun} and GPP_{shade} individually, but there is a lack of credible, large-scale, and long-time series GPP_{sun} and GPP_{shade} products for validating model outputs. Thus, such products can support exploring the similarities and differences of sunlit and shaded leaves contributing to GPP or SIF, to further excavate the interior ecological mechanism of different carbon cycle processes and advance carbon cycle modelling.

Light use efficiency (LUE) models have the advantages of few required parameters, and easy to digest remote sensing data^{38,39}. They have been established as a popular method to estimate regional and global carbon fluxes^{40–44}. By incorporating satellite-derived land surface variables into LUE models, several global data products (e.g. MOD17A2) have been produced^{45–47}. Considering the difference in solar radiation absorption and LUE of leaves within a canopy, the two-leaf LUE (TL-LUE) model divides the vegetation canopy into shaded and sunlit leaves and calculates GPP separately for these two portions⁴⁸. Therefore, the TL-LUE model significantly reduces the sensitivity to sky conditions, effectively alleviates the systematic underestimation of GPP under low solar radiation by the MOD17 model, and improves the simulation accuracy of GPP^{48,49}.

In the previous version of the TL-LUE model, the CO_2 fertilization effect (CFE), the enhancement of vegetation productivity by the increase of CO_2 concentration, is not included. It's well-known that global atmospheric CO_2 concentration has continued to rise, increasing about 17% during 1992–2020, and the increase of atmospheric CO_2 substantially enhance global GPP¹⁶. CFE on the global terrestrial carbon exchange has attracted widespread attentions^{4,16,50}, but has rarely been considered in LUE models. Thus, it is imperative to include the change of atmospheric CO_2 concentration in GPP estimation with LUE models. Based on the previous version of TL-LUE model, this study additionally includes atmospheric CO_2 concentration regulation scalar (C_s) and replaces temperature regulation scalar (T_s) with that used in the Terrestrial Ecosystem Model (TEM) model⁵¹ to capture negative effects of low and high temperatures on GPP. Eddy covariance carbon flux measurements of 68 sites (480 site years) and 25 sites (170 site years) from the FLUXNET2015 dataset were used to calibrate and validate the revised TL-LUE model, respectively.

This study employs various remote sensing data as model inputs, including European Space Agency Climate Change Initiative Land Cover (ESA-CCI land cover) from 1992 to 2020 and GLOBMAP leaf area index (GLOBMAP-LAI), in conjunction with meteorological data provided by the Climatic Research Unit and Japanese reanalysis (CRUJRA v2.2) dataset. The GPPs are calculated by the revised TL-LUE model (Fig. 1). The temporal resolutions of the dataset are 8-day, monthly and annual, and the spatial resolution is $0.05^\circ \times 0.05^\circ$. During 1992–2020, the estimated mean annual totals of global GPP, GPP_{sun} , and GPP_{shade} are $125.0 \text{ Pg C a}^{-1}$, 50.5 Pg C a^{-1} , and 74.5 Pg C a^{-1} , respectively.

Methods

Model description. The model used in this study is a revised version of the two-leaf light use efficiency (TL-LUE) model⁴⁸. The revised TL-LUE model adds the atmospheric CO_2 concentration regulation scalar and modifies air temperature regulation scalar. GPP is divided into GPP_{shade} and GPP_{sun} ⁵². It is described as Eqs. (1–3):

Vegetation Type	DBF	EBF	ENF	MF	CRO	GRA	OSH	SAV	WET	WSA
$\epsilon_{msh}(\text{gC MJ}^{-1})$	3.75 ± 0.52	3.26 ± 0.93	3.40 ± 1.19	3.00 ± 0.66	4.80 ± 1.94	4.57 ± 1.67	3.10 ± 0.42	4.65 ± 0.64	2.53 ± 1.02	2.70
$\epsilon_{msu}(\text{gC MJ}^{-1})$	0.92 ± 0.29	1.44 ± 0.64	0.89 ± 0.49	0.80 ± 0.41	1.43 ± 0.75	1.16 ± 0.45	0.65 ± 0.07	3.45 ± 0.64	1.23 ± 0.92	2.60
VPD _{max} (kPa)	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1
VPD _{min} (kPa)	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
T _{opt} (°C)	23.1	25.8	19.7	24.5	23.5	20.9	22.3	25.8	24.2	26.2
albedo(α) ^{49,95,96}	0.18	0.18	0.15	0.17	0.23	0.23	0.16	0.18	0.23	0.23
Clumping index(Ω) ⁹⁷	0.8	0.8	0.6	0.7	0.9	0.9	0.8	0.8	0.9	0.8

Table 1. Parameters used in the revised TL-LUE model. DBF: deciduous broadleaf forest; EBF: evergreen broadleaf forest; ENF: evergreen needleleaf forest; MF: mixed forest; CRO: cropland; GRA: grasslands; OSH: open shrublands; SAV: savannas; WET: wetlands; WSA: woody savannas.

$$\text{GPP} = \text{GPP}_{\text{shade}} + \text{GPP}_{\text{sun}} \quad (1)$$

$$\text{GPP}_{\text{shade}} = (\epsilon_{\text{msh}} \times \text{APAR}_{\text{sh}}) \times T_s \times W_s \times C_s \quad (2)$$

$$\text{GPP}_{\text{sun}} = (\epsilon_{\text{msu}} \times \text{APAR}_{\text{su}}) \times T_s \times W_s \times C_s \quad (3)$$

where ϵ_{msu} and ϵ_{msh} are the maximum light use efficiency of sunlit and shaded leaves (detailed in Table 1), respectively. APAR_{su} and APAR_{sh} are the absorbed PAR of sunlit and shaded leaves, respectively. They were expressed as:

$$\text{APAR}_{\text{sh}} = (1 - \alpha) \times \left(\frac{\text{PAR}_{\text{dif}} - \text{PAR}_{\text{dif,u}}}{\text{LAI}} + C \right) \times \text{LAI}_{\text{sh}} \quad (4)$$

$$\text{APAR}_{\text{su}} = (1 - \alpha) \times \left(\frac{\text{PAR}_{\text{dir}} \times \cos \beta}{\cos \theta} + \frac{\text{PAR}_{\text{dif}} - \text{PAR}_{\text{dif,u}}}{\text{LAI}} + C \right) \times \text{LAI}_{\text{su}} \quad (5)$$

$$\text{LAI}_{\text{su}} = 2 \cos \theta \times \left(1 - e^{-\frac{\text{LAI} \times \Omega}{2 \cos \theta}} \right) \quad (6)$$

$$\text{LAI}_{\text{sh}} = \text{LAI} - \text{LAI}_{\text{su}} \quad (7)$$

$$C = 0.07 \times \Omega \times \text{PAR}_{\text{dir}} \times (1.1 - 0.1\text{LAI})e^{-\cos \theta} \quad (8)$$

$$\text{PAR}_{\text{dif}} = \text{PAR} \times (0.7527 + 3.8453R - 16.316R^2 + 18.962R^3 - 7.0802R^4) \quad (9)$$

$$\text{PAR}_{\text{dir}} = \text{PAR} - \text{PAR}_{\text{dif}} \quad (10)$$

$$\text{PAR}_{\text{dif,u}} = \text{PAR}_{\text{dif}} \times e^{-\frac{0.5\Omega\text{LAI}}{\cos \bar{\theta}}} \quad (11)$$

$$\cos \bar{\theta} = 0.537 + 0.025 \times \text{LAI} \quad (12)$$

where α is albedo; θ is the solar zenith angle; β is the leaf angle, which is set as 60°; Ω is clumping index (detailed in Table 1); C is multiple scattered radiation (unit: W m^{-2}); PAR_{dir} and PAR_{dif} (unit: W m^{-2}) are the incoming direct and diffuse photosynthetically active radiation above the canopy; $\text{PAR}_{\text{dif,u}}$ (unit: W m^{-2}) denotes diffuse PAR under the canopy^{48,53}; LAI_{su} and LAI_{sh} are the LAI of sunlit and shaded leaves; R represents the sky clearness index and equals to $S/(S_0 \cos \theta)$, where S is solar radiation in W m^{-2} , and S_0 is the solar constant (1367 W m^{-2}); $\bar{\theta}$ is the representative zenith angle for diffuse radiation transmission⁵³.

The regulation scalars of temperature (T_s)⁵², water stress (W_s)⁴⁸, and atmospheric CO_2 concentration (C_s)⁵⁴ are calculated as follows:

$$T_s = \frac{(T - T_{\text{max}}) \times (T - T_{\text{min}})}{(T - T_{\text{max}}) \times (T - T_{\text{min}}) - (T - T_{\text{opt}})^2} \quad (13)$$

$$W_s = \frac{VPD_{\max} - VPD}{VPD_{\max} - VPD_{\min}} \quad (14)$$

$$C_s = \frac{C_i - \Gamma^*}{C_i + 2\Gamma^*} \quad (15)$$

where the maximum (T_{\max}) and minimum temperatures for vegetation photosynthesis (T_{\min}) were set to 313.15 K and 273.15 K, respectively⁴⁹. The optimum temperature for vegetation photosynthesis (T_{opt}) is set as the average of different types in Huang *et al.*⁵⁵ (details in Table 1). VPD_{\max} and VPD_{\min} are the VPD when GPP reaches the maximum and minimum, respectively³⁸. If the value of VPD is greater than or equal to VPD_{\max} , W_s is equal to 0 and if the value of VPD is less than or equal to VPD_{\min} , W_s is set to 1⁴⁸. Γ^* is the CO_2 compensation point in the absence of dark respiration (calculated by Eq. 16)⁵⁴; C_i is the intercellular concentration of CO_2 (ppm).

$$\Gamma^* = 4.22 \times e^{\frac{37830(T-298.15)}{298.15RT}} \quad (16)$$

$$C_i = C_a \times \chi \quad (17)$$

$$\chi = \frac{\xi}{\xi + \sqrt{VPD}} \quad (18)$$

$$\xi = \sqrt{\frac{356.51K}{1.6\eta^*}} \quad (19)$$

$$K = K_c \times \left(1 + \frac{P_o}{K_o}\right) \quad (20)$$

$$K_c = 39.97 \times e^{\frac{79.43 \times (T-298.15)}{298.15RT}} \quad (21)$$

$$K_o = 27480 \times e^{\frac{36.38 \times (T-298.15)}{298.15RT}} \quad (22)$$

where C_a is the atmospheric CO_2 concentration (using NOAA global monthly mean CO_2 concentration at the unit of ppm), and χ is the ratio of leaf-internal to ambient CO_2 ⁵⁶. K is the Michaelis-Menten coefficient of Rubisco and η^* is the viscosity of water relative to its value at 25 °C (0.8903)⁵⁷. K_c and K_o are the Michaelis-Menten coefficients of Rubisco for CO_2 and O_2 , respectively, and P_o is the partial pressure of O_2 (21 kPa)⁵⁶. R is the molar gas constant (8.314 J mol⁻¹ K⁻¹).

Input data and processing. Eddy covariance flux measurements from the FULLSET daily product in the FLUXNET2015 dataset (<https://fluxnet.org>) were used for model calibration and validation. We selected site years data according to the following requirements: the missing observations of air temperature (TA_F), vapor pressure deficit (VPD_F), CO_2 mole fraction (CO2_F_MDS), incoming photosynthetic photon flux density (PPFD_IN), or shortwave radiation (SW_IN_F), gross primary production (GPP_DT_CUT_MEAN) in one year are less than 2 months. A linear interpolation was applied to fill the individual missing values, which accounted for about 2% of the total measurements. About 75% of the sites were randomly selected to calibrate the revised TL-LUE model parameters for each vegetation type, and the remaining sites were applied for parameters validation. The sites and years selected for calibration and validation are detailed in Table S1. The shortwave-to-PAR conversion parameter in global was estimated to vary between 0.39 to 0.53^{58–60}. With the comparison between measurements of PAR and shortwave radiation in sites, 0.43 is most suitable for this study. The spatial distribution of calibration and validation sites is shown in Figure S1.

The land cover dataset we used is European Space Agency Climate Change Initiative Land Cover (ESA-CCI land cover) at a 300 m spatial resolution for every year from 1992 to 2020 (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-land-cover>). We resampled the raw data to $0.05^\circ \times 0.05^\circ$ using nearest neighbour resampling. ESA-CCI land cover dataset uses the United Nations Land Cover Classification System (LCCS), thus, we converted them to match the International Geosphere-Biosphere Program land cover scheme (IGBP)⁶¹. In particular, ESA-CCI land cover provides land cover data for the years before 2000, which makes it possible to study changes in GPP caused by changes in land cover types over long-term series.

GLOBMAP leaf area index (GLOBMAP-LAI)⁶² as a model input is available at 0.0727° spatial resolution for every 8 days (2001–2020) and half-month (1992–2000) from 1992 to 2020. GLOBMAP LAI (Version 3) provides a consistent long-time LAI product (1981–2020, continuously updated) by quantitative fusion of Moderate Resolution Imaging Spectroradiometer (MODIS) and historical Advanced Very High Resolution Radiometer

(AVHRR) data. The long-term LAI series was made up by combination of AVHRR LAI (1981–2000) and MODIS LAI (2001–2020). MODIS LAI series was generated from MODIS land surface reflectance data (MOD09A1)⁶³ based on the GLOBCARBON LAI algorithm⁶⁴. The relationships between GIMMS NDVI and MODIS LAI were established pixel by pixel using the two data series during overlapped period (2000–2006). And the AVHRR LAI⁶⁵ back to 1981 was estimated from historical AVHRR observations based on these pixel-level relationships. The clumping effects was considered at the pixel level by employing global clumping index map at 500 m resolution⁶⁶. The cloud mask for the MOD09A1 data were created by a new cloud detection algorithm based on time series surface reflectance observations⁶⁷. GLOBMAP-LAI has been smoothed by locally adjusted cubic spline capping approach⁶⁸. We resampled them to $0.05^\circ \times 0.05^\circ$ for the model. Additionally, we extracted the LAI of each site from GLOBMAP-LAI (500 m, 8-day) for model calibration and validation.

The Climatic Research Unit and Japanese reanalysis (CRUJRA) version 2.2 data⁶⁹ provided 6-hourly at 0.5° resolution meteorological variables, such as downward solar radiation flux (dswrf, unit: $\text{J m}^{-2} \text{h}^{-1}$), specific humidity (spfh, unit: kg kg^{-1}), the temperature at 2 m (tmp, unit: K), pressure (pres, unit: Pa), from 1992 to 2020. Daily dswrf used in this study (unit: $\text{J m}^{-2} \text{d}^{-1}$) was converted from the CRUJRA dswrf dataset by summing four 6-hourly data per day. Temperature (unit: K) and vapor pressure deficit (VPD, unit: kPa) by taking the mean of the 6-hourly data of each day in CRUJRA. The daily dswrf was corrected with site shortwave radiation data. The global monthly CO_2 concentration (unit: ppm) is available on www.esrl.noaa.gov/gmd/ccgg/trends/. VPD was calculated using specific humidity, pressure, and temperature as Eqs. (23, 24):

$$\text{VPD} = \text{VP}_{\text{sat}} - \text{spfh} \times \frac{\text{pres}/1000}{0.622 + \text{spfh} \times 0.378} \quad (23)$$

$$\text{VP}_{\text{sat}} = 0.61121 \times e^{\left(18.678 - \frac{t_{\text{air}}}{234.5}\right) \times \frac{t_{\text{air}}}{257.14 + t_{\text{air}}}} \quad (24)$$

where t_{air} is the air temperature at the unit of $^\circ\text{C}$. VP_{sat} means the saturated vapor pressure (kPa); spfh represents specific humidity (kg kg^{-1}); pres is pressure (Pa).

Calibration of model parameters. The maximum LUEs of the sunlit (ϵ_{msu}) and shaded (ϵ_{msh}) leaves spatially differ due to the changes in vegetation canopy structures and vegetation species types⁷⁰, which leads to the distinct ϵ_{msh} and ϵ_{msu} of different vegetation types. To reduce the GPP simulation bias caused by ϵ_{msh} and ϵ_{msu} of sunlit and shaded leaves, the daily data of 68 sites (480 site years) in the FLUXNET2015 dataset (details in Table S1) were used for parameter optimization with the shuffled complex Evolution-University of Arizona method^{49,71}. The agreement index (d) was used as the optimization criterion. This index was widely used in parameter optimization^{49,72}. It ranges from 0 (complete disagreement) to 1 (complete agreement). Parameter values identified when d maximizes are optimization results. The calculation of d is as Eq. (25):

$$d = 1 - \frac{\sum_{n=1}^m (E_n - M_n)^2}{\sum_{n=1}^m (|E_n - \bar{M}| + |M_n - \bar{M}|)^2} \quad (25)$$

where m is the number of all measurements; E_n and M_n are the n^{th} estimation and measured GPP, respectively. \bar{M} represents the mean of all measured GPP values.

The average ϵ_{msh} and ϵ_{msu} values with standard deviation of all vegetation types are shown in Table 1. The parameters of deciduous needleleaf forests (DNF) were consistent with DBF settings in modelling. The R^2 of observed GPP (GPP_{EC}) and estimated GPP (GPP) in WSA and SAV were 0.39 and 0.41, respectively, and the R^2 values of other vegetation types were all above 0.5, in which the R^2 of DBF was the highest (0.89). The calibration results of different vegetation types are shown in Figure S2.

Data Records

The dataset provides global gridded 0.05° GPP, $\text{GPP}_{\text{shade}}$ and GPP_{sun} at three temporal resolutions (8-day, monthly, annual) from 1992 to 2020. The units are $\text{g C m}^{-2} (\text{8day})^{-1}$, $\text{g C m}^{-2} \text{month}^{-1}$ and $\text{g C m}^{-2} \text{a}^{-1}$, respectively. We divided the dataset into 29 folders by year, where each folder contains the data for the year at three temporal scales. The files were named as “GGGG_v21_TTTT.tif” and stored in the GeoTiff format, where GGGG represents GPP, $\text{GPP}_{\text{shade}}$ and GPP_{sun} . For the 8-day scale, TTTT represents year and DOY (e.g. Shade_GPP_v21_1999_249.tif). For the one-month scale, TTTT represents year and month (e.g. Sun_GPP_v21_1999_01.tif). For the annual scale, TTTT represents only year (e.g. GPP_v21_1999.tif). The scale factor of the monthly data is 0.1, that of the 8-day data is 0.01. The data type of the monthly and 8-day data is 16-bit integer, and that of the annual data is double. All datasets are publicly available from the DRYAD repository <https://doi.org/10.5061/dryad.dfn2z352k>⁷³.

The high annual average GPP values were mainly distributed in the Amazon, Indonesia and Congo Basin in the low latitude regions, which was about $3500 \text{ g C m}^{-2} \text{a}^{-1}$. The relatively high GPP ($\sim 2000 \text{ g C m}^{-2} \text{a}^{-1}$) occurred in Southeast Asia, the southeast of North and South America, and south central Africa, and the GPP that ranged from 0 to $1500 \text{ g C m}^{-2} \text{a}^{-1}$ accounted for 74.7% of global vegetation cover. The shaded and sunlit GPP ($\text{GPP}_{\text{shade}}$ and GPP_{sun}) had similar spatial distribution with GPP, but the value of GPP_{sun} was lower than $\text{GPP}_{\text{shade}}$ near the equator (Fig. 2b,c). $\text{GPP}_{\text{shade}}$ that ranged from 0 to $1000 \text{ g C m}^{-2} \text{a}^{-1}$ accounted for 78.7% of global vegetation cover, and GPP_{sun} that ranged from 0 to $750 \text{ g C m}^{-2} \text{a}^{-1}$ accounted for 80.8% of global vegetation cover.

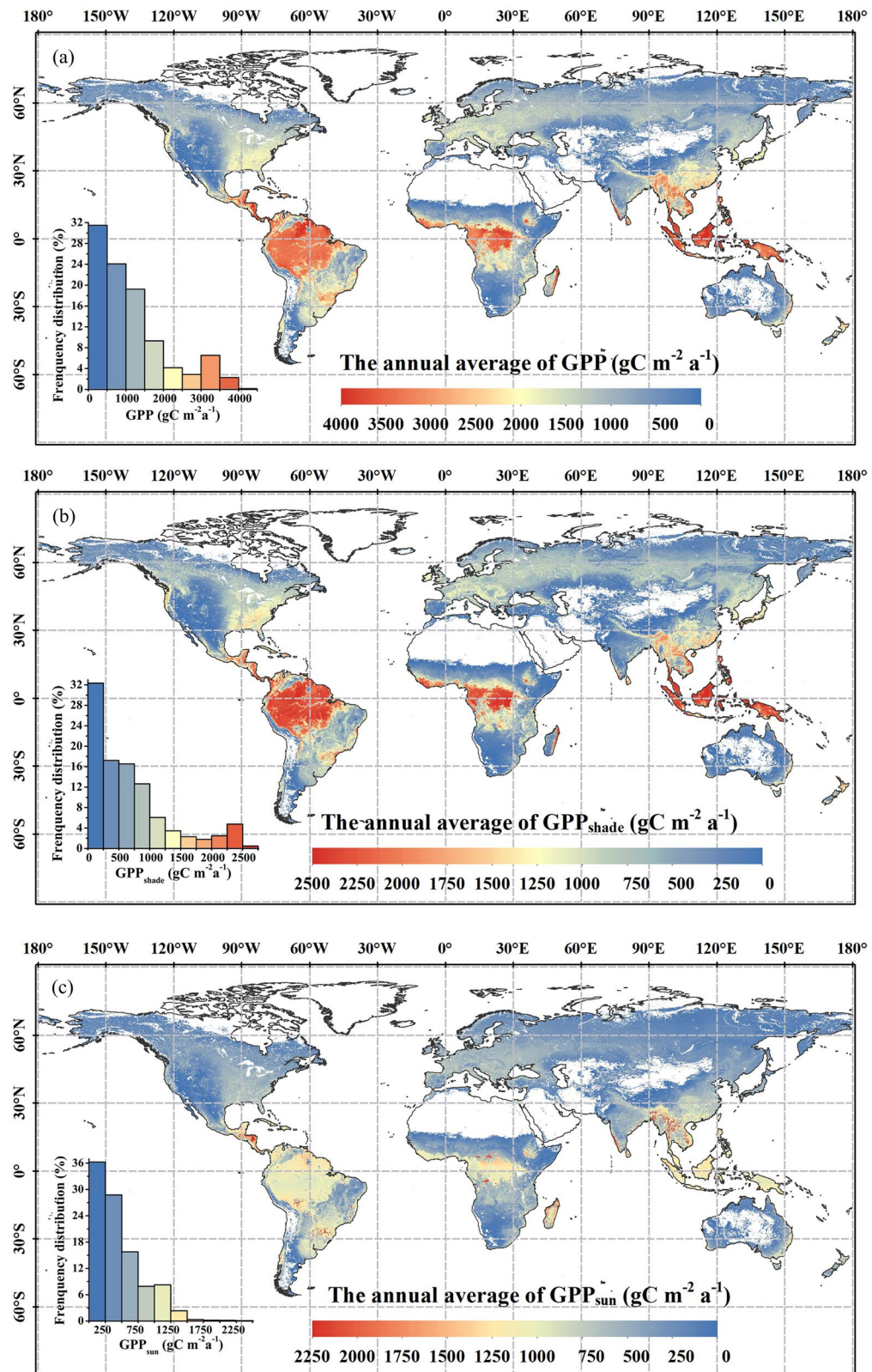


Fig. 2 Spatial distribution of the annual average (a) GPP, (b) GPP_{shade} and (c) GPP_{sun} from 1992 to 2020.

From 1992 to 2020, the GPP of Malaysia, Southeast Asia, Indian Peninsula, central Africa, north and southeast South America, and central Europe all showed an obvious increasing trend, and the rate was close to $20 \text{ gC m}^{-2} \text{a}^{-2}$. The GPP scattered in central South America, east Africa, Central Asia, and Southeast Asia showed a significant decreasing trend, and the change rate was near to $-20 \text{ gC m}^{-2} \text{a}^{-2}$ (Fig. 3a). For

GPP_{shade}, eastern and southern Asia, central Europe, central and western Africa, northwest and southeast South America showed a significant increasing trend, and the change rate was above $10 \text{ g C m}^{-2} \text{ a}^{-2}$, which was consistent with the trend of GPP. Besides, areas with reduced GPP_{shade} (trend $< -10 \text{ g C m}^{-2} \text{ a}^{-2}$) were sporadically distributed in Central Asia, Central South America, Eastern Africa, and Southeast Asia (Fig. 3b). For GPP_{sun}, central and southeastern South America, southeastern Asia, and Europe showed the most obvious increase, with a rate of change of around $10 \text{ g C m}^{-2} \text{ a}^{-2}$, which was similar to the spatial distribution of GPP and GPP_{shade} trends. However, the value of GPP_{sun} was lower than GPP_{shade} (Fig. 3b,c). The vast majority of global vegetation cover with significant change exhibits an increasing trend (GPP trend > 0 : 92.0%, GPP_{shade} trend > 0 : 91.2%, and GPP_{sun} trend > 0 : 88.7%). In general, most global GPP, GPP_{shade}, and GPP_{sun} revealed increasing trends (Fig. 3).

The mean annual totals of global GPP, GPP_{sun}, and GPP_{shade} from 1992 to 2020 are $125.0 \text{ Pg C a}^{-1}$, 50.5 Pg C a^{-1} and 74.5 Pg C a^{-1} , respectively. Overall, the GPP proportions of individual vegetation types to total were similar for global GPP, GPP_{sun}, and GPP_{shade}. Among the 11 vegetation types, EBF contributed most GPP, followed by CRO. These two types accounted for more than half of the total GPP, GPP_{sun}, and GPP_{shade} (Fig. 4). The GPP of SAV, MF, WET, and WSA were 1.0 Pg C a^{-1} , 2.0 Pg C a^{-1} , 2.8 Pg C a^{-1} , and 2.9 Pg C a^{-1} , respectively, which were relatively low (Fig. 4a). The GPP_{sun} were close to GPP_{shade} for CRO, GRA, and WET, while the GPP_{sun} were higher than GPP_{shade} for SAV and WSA. The GPP_{shade} of ENF, EBF, DNF, DBF, MF, and OSH with relatively complicated canopy structures were much higher than their GPP_{sun} (Fig. 4b,c). Overall, GPP_{shade} played a key role in GPP for forest types, while GPP_{sun} was greater than GPP_{shade} for non-forest types. In total, GPP_{shade} contributes more to GPP than GPP_{sun}. In addition, the GPP of all vegetation types showed an increasing trend. With one exception that the GPP_{sun} of SAV showed a decreasing trend, the GPP_{sun} and GPP_{shade} of other vegetation types showed an increasing trend, among which GPP (also GPP_{sun}, and GPP_{shade}) of CRO showed the distinctively greatest increasing trend. Except WSA, the increasing trend of GPP_{shade} of other vegetation types is greater than that of GPP_{sun} (Fig. 4d). It's worth noting that for DBF, the increasing trend of GPP is relatively large, and GPP_{shade} overwhelmingly outweighs GPP_{sun}. MF shows a similar phenomenon, but the overall trend is smaller.

Technical Validation

Validation of model parameters. Carbon flux data of 25 sites (170 site years) from the FLUXNET2015 dataset were selected (detailed in Table S1) for model validation. The comparison between GPP that estimated by the revised TL-LUE model with optimized ϵ_{msu} and ϵ_{msh} , and GPP measurements (GPP_{EC}) at each flux site is shown in Fig. 5. The revised TL-LUE model performed well in estimation of GPP for all vegetation types. All sites have R^2 above 0.5, except for AU-Gin (WSA) ($R^2 = 0.49$) and BR-Sa3 (EBF) ($R^2 = 0.36$).

Comparisons with other global GPP products. Previous studies have shown that the differences in GPP estimation are usually caused by different model structures, parameter settings, and input data^{74–76}. Here, we compare our global annual GPP with several global GPP products derived from remote sensing models, including data-driven models and LUE models (Fig. 6).

For data-driven GPP products, Li *et al.*⁷⁷ used SIF-GPP relationship and SIF observed by the Orbiting Carbon Observatory-2 (OCO-2) to obtain global GPP, which named GOSIF GPP ($135.5 \pm 8.8 \text{ Pg C a}^{-1}$, 2001–2017). And the WECANN product was produced using an artificial neural network (ANN) with SIF and other data sources as inputs⁷⁸. The WECANN-GPP showed an obvious decreasing trend between 2007 and 2015, with a range of 110.1 to $118.2 \text{ Pg C a}^{-1}$. FluxSat GPP⁷⁹, which was generated using satellite data-driven models based on the LUE framework, ranged from 148.1 to $156.7 \text{ Pg C a}^{-1}$ during 2001 to 2019. In addition, a few recent studies have identified that there is a strong spatio-temporal correlation between near-infrared reflectance (NIRv) and GPP, suggesting that NIRv can be employed to estimate the GPP of vegetation^{13,79}. The global total of GPP estimated from NIRv by Wang *et al.*⁸⁰ was $130.5 \pm 3.2 \text{ Pg C a}^{-1}$ during 1992–2018, close to our estimate. For the LUE models, the range of GPP obtained by the improved EC-LUE model⁸¹ was 106.4 – $118.3 \text{ Pg C a}^{-1}$ (1992–2018), and that by VPM model was 131.2 – $140.0 \text{ Pg C a}^{-1}$ (2000–2016). The range of MOD17A2H.006⁸² was 94.8 – $106.2 \text{ Pg C a}^{-1}$ (2001–2020). The global annual GPP of FluxSat was the highest, and that of MOD17A2 was the lowest. Our estimated global GPP ranged from 120.02 to $132.65 \text{ Pg C a}^{-1}$, placing at the middle of the various GPP products. Anav *et al.*⁷⁶ suggested that previous observation-constrained estimates of global GPP^{45,83} (e.g. based on either $\delta^{18}\text{O}$ measurements of atmospheric CO_2 or eddy covariance flux upscaling) was at 120 Pg C a^{-1} for the period before 2010. Thus, our estimate agreed reasonably well with such estimates.

In addition, our estimated global GPP showed an overall increasing trend, which is consistent with most other GPP products. Only WECANN-GPP and EC-LUE GPP (after 2000) showed a significant declining trend. The declining trend of WECANN obviously associated with the degradation of GOME-2 SIF sensor, and the GOME-2 SIF data used for training was not corrected⁸⁴.

Uncertainties. Previous studies have indicated that different LAI products lead to clear differences in estimated GPP⁸⁵, and the uncertainties of various LAI products are higher in low-latitude areas^{86,87}. Since the land cover used by GLOBMAP-LAI is different from that used in this study, the corresponding LAI values of a small amount pixels ($< 0.01\%$) for SAV and WSA in low latitude areas are relatively high, which lead to abnormally high estimated GPP. These anomalies are not processed because of remaining the authenticity of data.

The model parameters varied in different areas, due to the plant species included in the same vegetation type. In particular, the LUE of C3 and C4 crops was greatly discrepant as many previous studies proved^{88–91}. According to the restriction of ESA-CCI land cover data, C3 and C4 plants could not be distinguished, so the GPP of C3 and C4 crops cannot be gained separately. The LAI of each site extracted from GLOBMAP-LAI with a

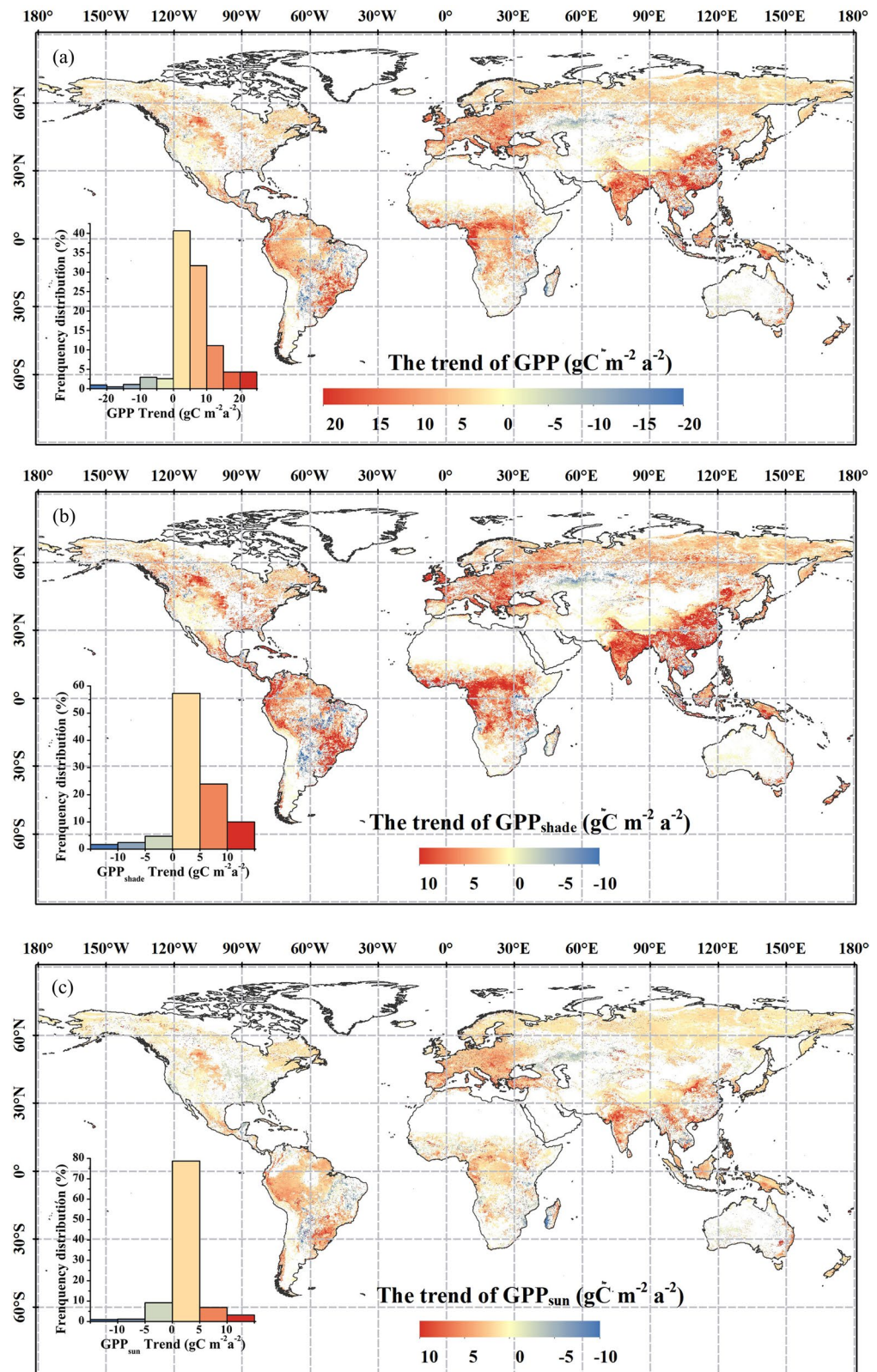


Fig. 3 Spatial distribution of the trend of (a) GPP, (b) $\text{GPP}_{\text{shade}}$, and (c) GPP_{sun} during 1992 to 2020. The results have removed the value which is not significant ($p > 0.05$).

resolution of 500 m cannot completely match the flux tower data scale, which cause uncertainty in the parameter optimization. Simultaneously, the eddy covariance measurements also have some uncertainties, which inevitably affected the parameter optimization.

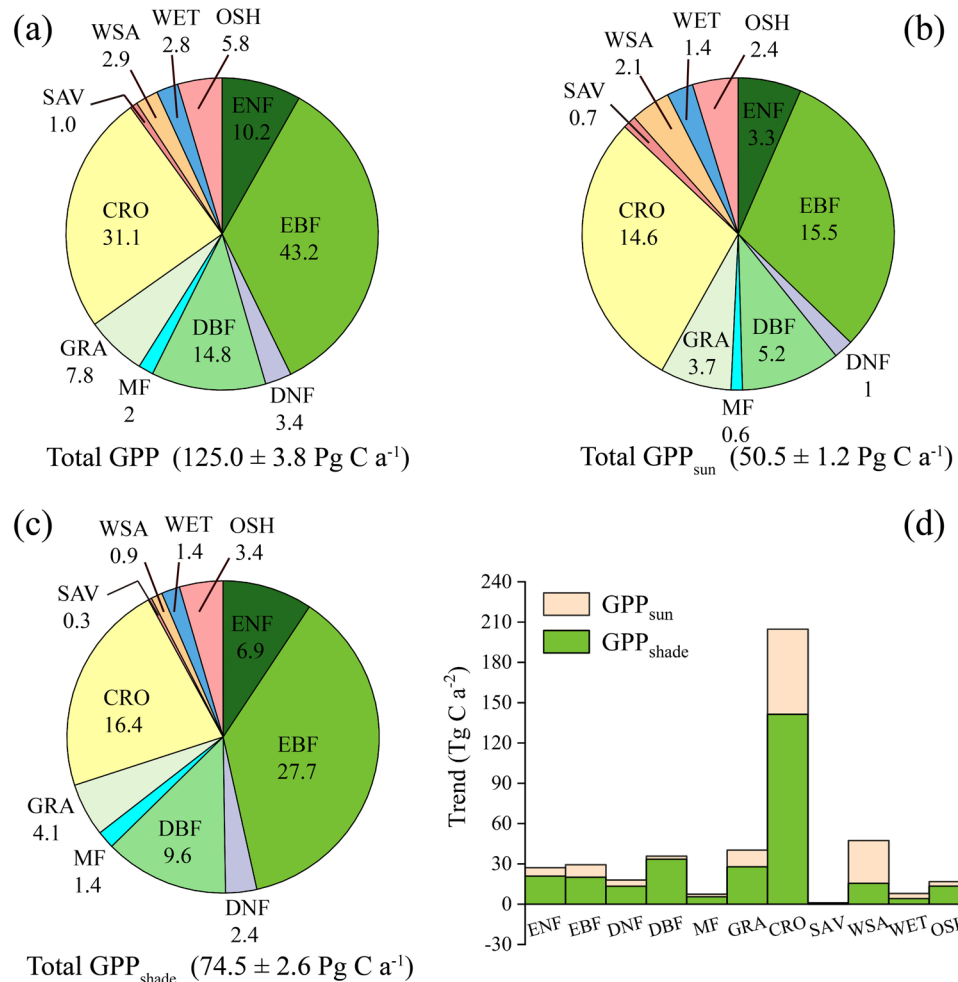


Fig. 4 The mean annual totals of (a) GPP, (b) GPP_{sun}, (c) GPP_{shade} and (d) the trend of GPP_{shade} and GPP_{sun} from 1992 to 2020 for different vegetation types over the globe.

Potential benefits and usages of this dataset. *Facilitate researches on the relationship between SIF and GPP.* As is known, SIF signals come mainly from sunlit leaves^{31,92}. Most of the current studies on the relationship between SIF and the photosynthesis of sunlit leaves are at canopy and leaf scales, while similar studies at large scales are currently rare due to the lack of publicly available global or regional GPP datasets for sunlit leaves. In addition, it is known that there is a link between SIF and GPP across biome types³⁴, but the relationship is not well quantified. The GPP_{sun} and GPP provided in this study may help to explore the relationship between SIF and GPP_{sun} or GPP in different ecosystem types at large scales.

Facilitate researches on the interactions between solar radiation and terrestrial carbon cycling. Compared with direct radiation, the increase of diffuse radiation can effectively promote carbon fixation⁹³. Shaded leaves make more effective use of diffuse radiation⁴⁸, and GPP_{shade} plays a major role in the vegetation areas with more sheltered leaves or cloudy conditions⁹⁴. The GPP_{shade} and GPP_{sun} cannot be directly measured over large regions. The GPP_{shade} and GPP_{sun} estimated by the revised TL-LUE model can capture the contribution of sunlit and shaded leaves to GPP in long-term and at large scales, and make it possible to quantify the carbon fixation increase (or decrease) influenced by the change in radiation fraction over a long period and at large scales, which facilitates further investigations on the interactions between radiation and terrestrial carbon cycling.

Facilitate researches on the dynamics, processes and drivers of GPP at large scales. This study provides 8-day and monthly GPP from 1992 to 2020, which allows for studying the changes in seasonal cycles (e.g. amplitude and phase changes of growing season) of GPP and its processes (GPP_{sun} and GPP_{shade}) over many years. In addition, we employed a long-term ESA-CCI land cover data since 1992 (while most other dataset uses MODIS land cover since 2000) with the consideration of year to year land cover changes, enabling it to characterize the impact of land cover change on GPP. The dataset would help to dig the underlying mechanisms of climate and human impacts on global terrestrial GPP.

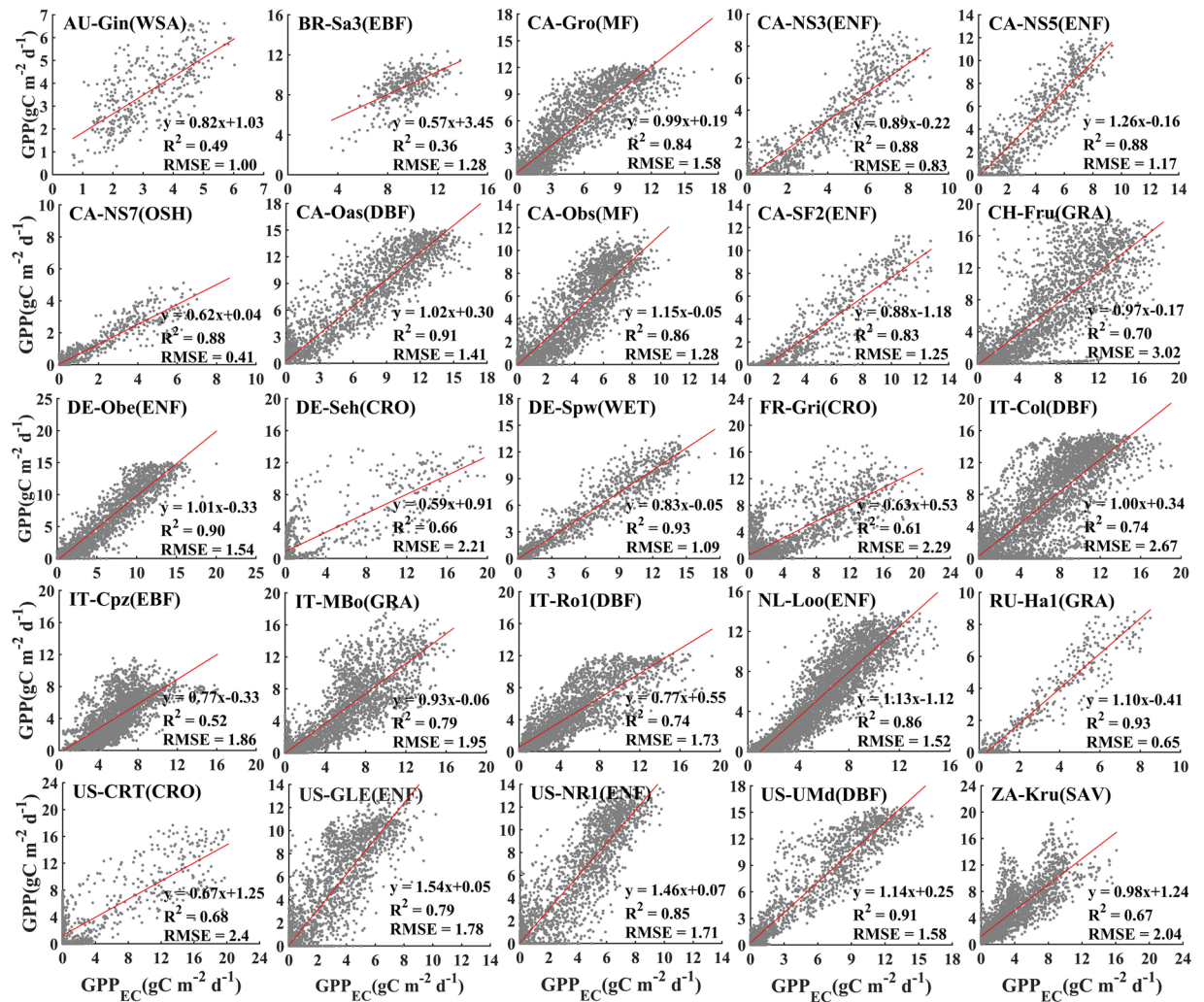


Fig. 5 Validation of daily GPP estimated by the revised TL-LUE (GPP) model with tower measurements (GPP_{EC}) at 25 FLUXNET sites. The revised TL-LUE model was driven by average optimized parameters for different vegetation types, tower-based meteorological data, and smoothed LAI.

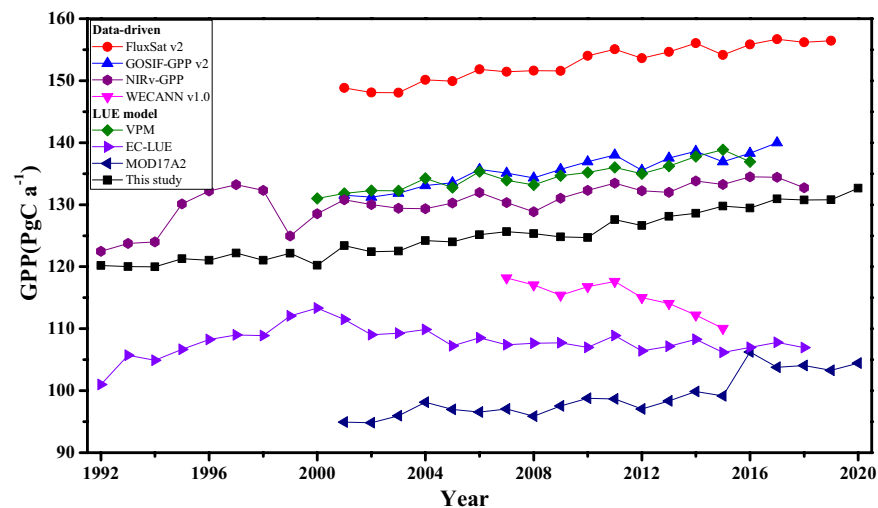


Fig. 6 Comparison of global annual GPP totals estimated from a set of global remote sensing GPP products.

Usage Notes

In the dataset, in order to ensure the authenticity, we did not delete or modify a small number of abnormally high values. Therefore, when using this dataset, you can set thresholds to remove the anomalies.

Code availability

We used the MATLAB 2020b for data processing. The core codes for the study are available at https://github.com/BiWenjunnju/code_TL.git.

Received: 15 December 2021; Accepted: 4 April 2022;

Published online: 16 May 2022

References

- Cox, P. & Jones, C. Climate change - Illuminating the modern dance of climate and CO₂. *Science* **321**, 1642–1644 (2008).
- Gilmanov, T. G. *et al.* Gross primary production and light response parameters of four Southern Plains ecosystems estimated using long-term CO₂-flux tower measurements. *Glob. Biogeochem. Cycle* **17**, 1071 (2003).
- Running, S. W. Climate change - Ecosystem disturbance, carbon, and climate. *Science* **321**, 652–653 (2008).
- Sun, Z. *et al.* Spatial pattern of GPP variations in terrestrial ecosystems and its drivers: Climatic factors, CO₂ concentration and land-cover change, 1982–2015. *Ecol. Inform.* **46**, 156–165 (2018).
- Running, S. W. *et al.* A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. *Remote Sens. Environ.* **70**, 108–127 (1999).
- Madani, N. *et al.* The Impacts of Climate and Wildfire on Ecosystem Gross Primary Productivity in Alaska. *J. Geophys. Res.-Biogeosci.* **126**, e2020JG006078 (2021).
- Morales, P. *et al.* Comparing and evaluating process-based ecosystem model predictions of carbon and water fluxes in major European forest biomes. *Glob. Change Biol.* **11**, 2211–2233 (2005).
- Tramontana, G., Ichii, K., Camps-Valls, G., Tomelleri, E. & Papale, D. Uncertainty analysis of gross primary production upscaling using Random Forests, remote sensing and eddy covariance data. *Remote Sens. Environ.* **168**, 360–373 (2015).
- Canadell, J. G. *et al.* Carbon metabolism of the terrestrial biosphere: A multitechnique approach for improved understanding. *Ecosystems* **3**, 115–130 (2000).
- Fletcher, B. J. *et al.* Photosynthesis and productivity in heterogeneous arctic tundra: consequences for ecosystem function of mixing vegetation types at stand edges. *J. Ecol.* **100**, 441–451 (2012).
- Liu, L., Guan, L. & Liu, X. Directly estimating diurnal changes in GPP for C3 and C4 crops using far-red sun-induced chlorophyll fluorescence. *Agr. Forest Meteorol.* **232**, 1–9 (2017).
- Xu, X. *et al.* Long-term trend in vegetation gross primary production, phenology and their relationships inferred from the FLUXNET data. *J. Environ. Manage.* **246**, 605–616 (2019).
- Baldocchi, D. D. How eddy covariance flux measurements have contributed to our understanding of Global Change Biology. *Glob. Change Biol.* **26**, 242–260 (2020).
- He, L., Chen, J. M., Liu, J., Belair, S. & Luo, X. Assessment of SMAP soil moisture for global simulation of gross primary production. *J. Geophys. Res.-Biogeosci.* **122**, 1549–1563 (2017).
- Wang, S., Ibrom, A., Bauer-Gottwein, P. & Garcia, M. Incorporating diffuse radiation into a light use efficiency and evapotranspiration model: An 11-year study in a high latitude deciduous forest. *Agr. Forest Meteorol.* **248**, 479–493 (2018).
- Wang, S. *et al.* Recent global decline of CO₂ fertilization effects on vegetation photosynthesis. *Science* **370**, 1295–1300 (2020).
- Yu, G., Fu, Y., Sun, X., Wen, X. & Zhang, L. Recent progress and future directions of ChinaFLUX. *Sci. China Ser. D-Earth Sci.* **49**, 1–23 (2006).
- McCallum, I. *et al.* Improved light and temperature responses for light-use-efficiency-based GPP models. *Biogeosciences* **10**, 6577–6590 (2013).
- Stocker, B. D. *et al.* Drought impacts on terrestrial primary production underestimated by satellite monitoring. *Nature Geoscience* **12**, 264+ (2019).
- Cheng, S. J. *et al.* Variations in the influence of diffuse light on gross primary productivity in temperate ecosystems. *Agr. Forest Meteorol.* **201**, 98–110 (2015).
- Zhang, M. *et al.* Effects of cloudiness change on net ecosystem exchange, light use efficiency, and water use efficiency in typical ecosystems of China. *Agr. Forest Meteorol.* **151**, 803–816 (2011).
- Oliphant, A. J. *et al.* The role of sky conditions on gross primary production in a mixed deciduous forest. *Agr. Forest Meteorol.* **151**, 781–791 (2011).
- Urban, O. *et al.* Ecophysiological controls over the net ecosystem exchange of mountain spruce stand. Comparison of the response in direct vs. diffuse solar radiation. *Glob. Change Biol.* **13**, 157–168 (2007).
- Zhou, H. *et al.* Large contributions of diffuse radiation to global gross primary productivity during 1981–2015. *Glob. Biogeochem. Cycle* **35**, e2021GB006957 (2021).
- Guanter, L. *et al.* Retrieval and global assessment of terrestrial chlorophyll fluorescence from GOSAT space measurements. *Remote Sens. Environ.* **121**, 236–251 (2012).
- Guanter, L. *et al.* Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. *Proc. Natl. Acad. Sci. USA* **111**, E1327–E1333 (2014).
- Liu, L. & Cheng, Z. Detection of vegetation light-use efficiency based on solar-induced chlorophyll fluorescence separated from canopy radiance spectrum. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **3**, 306–312 (2010).
- MacBean, N. *et al.* Strong constraint on modelled global carbon uptake using solar-induced chlorophyll fluorescence data (vol 8, 1973, 2018). *Sci. Rep.* **8**, 10420 (2018).
- Meroni, M. *et al.* Remote sensing of solar-induced chlorophyll fluorescence: Review of methods and applications. *Remote Sens. Environ.* **113**, 2037–2051 (2009).
- Zheng, T. & Chen, J. M. Photochemical reflectance ratio for tracking light use efficiency for sunlit leaves in two forest types. *ISPRS-J. Photogramm. Remote Sens.* **123**, 47–61 (2017).
- Damm, A. *et al.* Remote sensing of sun-induced fluorescence to improve modeling of diurnal courses of gross primary production (GPP). *Glob. Change Biol.* **16**, 171–186 (2010).
- Lee, J. E. *et al.* Simulations of chlorophyll fluorescence incorporated into the Community Land Model version 4. *Glob. Change Biol.* **21**, 3469–3477 (2015).
- Pinto, F. *et al.* Sun-induced chlorophyll fluorescence from high-resolution imaging spectroscopy data to quantify spatio-temporal patterns of photosynthetic function in crop canopies. *Plant Cell Environ.* **39**, 1500–1512 (2016).
- Porcar-Castell, A. *et al.* Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: mechanisms and challenges. *J. Exp. Bot.* **65**, 4065–4095 (2014).

35. Xie, X., Li, A., Jin, H., Yin, G. & Nan, X. Derivation of temporally continuous leaf maximum carboxylation rate (V-cmax) from the sunlit leaf gross photosynthesis productivity through combining BEPS model with light response curve at tower flux sites. *Agr. Forest Meteorol.* **259**, 82–94 (2018).
36. Chen, J. M., Liu, J., Leblanc, S. G., Lacaze, R. & Roujean, J. L. Multi-angular optical remote sensing for assessing vegetation structure and carbon absorption. *Remote Sens. Environ.* **84**, 516–525 (2003).
37. Chen, J. M. *et al.* Effects of foliage clumping on the estimation of global terrestrial gross primary productivity. *Glob. Biogeochem. Cycle* **26**, GB1019 (2012).
38. Running, S. W., Thornton, P. E., Nemani, R. & Glassy, J. M. in *Methods in Ecosystem Science*. Ch.3 (Springer, New York, NY, Press, 2000).
39. Wu, C., Munger, J. W., Niu, Z. & Kuang, D. Comparison of multiple models for estimating gross primary production using MODIS and eddy covariance data in Harvard Forest. *Remote Sens. Environ.* **114**, 2925–2939 (2010).
40. Makela, A. *et al.* Developing an empirical model of stand GPP with the LUE approach: analysis of eddy covariance data at five contrasting conifer sites in Europe. *Glob. Change Biol.* **14**, 92–108 (2008).
41. McCallum, I. *et al.* Satellite-based terrestrial production efficiency modeling. *Carbon Balanc. Manag.* **4**, 8–8 (2009).
42. Wang, H. *et al.* Deriving maximal light use efficiency from coordinated flux measurements and satellite data for regional gross primary production modeling. *Remote Sens. Environ.* **114**, 2248–2258 (2010).
43. Yu, R. An improved estimation of net primary productivity of grassland in the Qinghai-Tibet region using light use efficiency with vegetation photosynthesis model. *Ecol. Model.* **431**, 109121 (2020).
44. Yuan, W. *et al.* Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. *Agr. Forest Meteorol.* **143**, 189–207 (2007).
45. Beer, C. *et al.* Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science* **329**, 834–838 (2010).
46. Running, S. W. *et al.* A continuous satellite-derived measure of global terrestrial primary production. *Bioscience* **54**, 547–560 (2004).
47. Zhang, Y. *et al.* Development of a coupled carbon and water model for estimating global gross primary productivity and evapotranspiration based on eddy flux and remote sensing data. *Agr. Forest Meteorol.* **223**, 116–131 (2016).
48. He, M. *et al.* Development of a two-leaf light use efficiency model for improving the calculation of terrestrial gross primary productivity. *Agr. Forest Meteorol.* **173**, 28–39 (2013).
49. Zhou, Y. *et al.* Global parameterization and validation of a two-leaf light use efficiency model for predicting gross primary production across FLUXNET sites. *J. Geophys. Res.-Biogeosci.* **121**, 1045–1072 (2016).
50. Friedlingstein, P. *et al.* Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks. *J. Clim.* **27**, 511–526 (2014).
51. Raich, J. W. *et al.* Potential net primary productivity in South-America - application of a global-model. *Ecol. Appl.* **1**, 399–429 (1991).
52. Li, J. *et al.* An algorithm differentiating sunlit and shaded leaves for improving canopy conductance and vapotranspiration estimates. *J. Geophys. Res.-Biogeosci.* **124**, 807–824 (2019).
53. Chen, J. M., Liu, J., Cihlar, J. & Goulden, M. L. Daily canopy photosynthesis model through temporal and spatial scaling for remote sensing applications. *Ecol. Model.* **124**, 99–119 (1999).
54. Keenan, T. F. *et al.* Recent pause in the growth rate of atmospheric CO₂ due to enhanced terrestrial carbon uptake. *Nat. Commun.* **7**, 13428 (2016).
55. Huang, M. *et al.* Air temperature optima of vegetation productivity across global biomes. *Nat. Ecol. Evol.* **3**, 772–779 (2019).
56. Prentice, I. C., Dong, N., Gleason, S. M., Maire, V. & Wright, I. J. Balancing the costs of carbon gain and water transport: testing a new theoretical framework for plant functional ecology. *Ecol. Lett.* **17**, 82–91 (2014).
57. Korson, L., Drosthan, W. & Millero, F. J. Viscosity of water at various temperatures. *J. Phys. Chem.* **73**, 34–39 (1969).
58. Olofsson, P., Van Laake, P. E. & Eklundh, L. Estimation of absorbed PAR across Scandinavia from satellite measurements Part I: Incident PAR. *Remote Sens. Environ.* **110**, 252–261 (2007).
59. González, J. A. & Calbó, J. Modelled and measured ratio of PAR to global radiation under cloudless skies. *Agr. Forest Meteorol.* **110**, 319–325 (2002).
60. Zhang, X., Zhang, Y. & Zhou, Y. Measuring and modelling photosynthetically active radiation in Tibet Plateau during April–October. *Agr. Forest Meteorol.* **102**, 207–212 (2000).
61. Yang, Y., Xiao, P., Feng, X. & Li, H. Accuracy assessment of seven global land cover datasets over China. *ISPRS-J. Photogramm. Remote Sens.* **125**, 156–173 (2017).
62. Liu, Y., Liu, R. & Chen, J. M. GLOBMAP global Leaf Area Index since 1981. *Zenodo* <https://doi.org/10.5281/zenodo.4700264> (2019).
63. Vermote, E. MOD09A1 MODIS/Terra Surface Reflectance 8-Day L3 Global 500m SIN Grid V006. NASA EOSDIS Land Processes DAAC <https://doi.org/10.5067/MODIS/MOD09A1.006> (2015).
64. Deng, F., Chen, J. M., Plummer, S., Chen, M. & Pisek, J. Algorithm for global leaf area index retrieval using satellite imagery. *IEEE Trans. Geosci. Remote Sens.* **44**, 2219–2229 (2006).
65. Vermote, E. NOAA CDR Program. NOAA Climate Data Record (CDR) of AVHRR Leaf Area Index (LAI) and Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), Version 5. LAI. NOAA National Centers for Environmental Information <https://doi.org/10.7289/V5TT4P69> (2019).
66. He, L., Chen, J. M., Pisek, J., Schaaf, C. & Strahler, A. Global clumping index map derived from the MODIS BRDF product. *Remote Sens. Environ.* **119**, 118–130 (2012).
67. Liu, R. G. & Liu, Y. Generation of new cloud masks from MODIS land surface reflectance products. *Remote Sens. Environ.* **133**, 21–37 (2013).
68. Chen, J. M., Deng, F. & Chen, M. Locally adjusted cubic-spline capping for reconstructing seasonal trajectories of a satellite-derived surface parameter. *IEEE Trans. Geosci. Remote Sens.* **44**, 2230–2238 (2006).
69. Harris, I. C. CRU JRA: Collection of CRU JRA forcing datasets of gridded land surface blend of Climatic Research Unit (CRU) and Japanese reanalysis (JRA) data. *Centre for Environmental Data Analysis* <http://catalogue.ceda.ac.uk/uuid/863a47a6d8414b6982e1396c69a9efe8> (2019).
70. Li, X., Liang, H. & Cheng, W. Evaluation and comparison of light use efficiency models for their sensitivity to the diffuse PAR fraction and aerosol loading in China. *Int. J. Appl. Earth Obs. Geoinf.* **95**, 102269 (2021).
71. Duan, Q. Y., Sorooshian, S. & Gupta, V. Effective and efficient global optimization for conceptual rain full-runoff models. *Water Resour. Res.* **28**, 1015–1031 (1992).
72. Gu, L. H. *et al.* Advantages of diffuse radiation for terrestrial ecosystem productivity. *J. Geophys. Res.-Atmos.* **107**, 4050 (2002).
73. Bi, W. & Zhou, Y. A global 0.05° dataset for gross primary production of sunlit and shaded vegetation canopies (1992–2020). *Dryad* <https://doi.org/10.5061/dryad.dfn2z352k> (2022).
74. Ogutu, B. O. & Dash, J. Assessing the capacity of three production efficiency models in simulating gross carbon uptake across multiple biomes in conterminous USA. *Agr. Forest Meteorol.* **174**, 158–169 (2013).
75. Cai, W. *et al.* Large differences in terrestrial vegetation production derived from satellite-based light use efficiency models. *Remote Sens.* **6**, 8945–8965 (2014).
76. Anav, A. *et al.* Spatiotemporal patterns of terrestrial gross primary production: a review. *Rev. Geophys.* **53**, 785–818 (2015).
77. Li, X. & Xiao, J. Mapping photosynthesis solely from solar-induced chlorophyll fluorescence: A global, fine-resolution dataset of gross primary production derived from OCO-2. *Remote Sens.* **11**, 2563 (2019).
78. Alemohammad, S. H. *et al.* 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 (2017).

79. Joiner, J. *et al.* Estimation of terrestrial global gross primary production (GPP) with satellite data-driven models and eddy covariance flux data. *Remote Sens.* **10**, 1346 (2018).
80. Wang, S., Zhang, Y., Ju, W., Qiu, B. & Zhang, Z. Tracking the seasonal and inter-annual variations of global gross primary production during last four decades using satellite near-infrared reflectance data. *Sci. Total Environ.* **755**, 142569 (2021).
81. Zheng, Y. *et al.* Improved estimate of global gross primary production for reproducing its long-term variation, 1982–2017. *Earth Syst. Sci. Data* **12**, 2725–2746 (2020).
82. Running, S., Mu, Q. & Zhao, M. MOD17A2H MODIS/Terra Gross Primary Productivity 8-Day L4 Global 500m SIN Grid V006. NASA EOSDIS Land Processes DAAC <https://doi.org/10.5067/MODIS/MOD17A2H.006> (2015).
83. Ciais, P. *et al.* A three-dimensional synthesis study of delta O-18 in atmospheric CO₂. 1. Surface fluxes. *J. Geophys. Res.-Atmos.* **102**, 5857–5872 (1997).
84. Zhang, Y., Joiner, J., Gentine, P. & Zhou, S. Reduced solar-induced chlorophyll fluorescence from GOME-2 during Amazon drought caused by dataset artifacts. *Glob. Change Biol.* **24**, 2229–2230 (2018).
85. Xie, X. *et al.* Assessment of five satellite-derived LAI datasets for GPP estimations through ecosystem models. *Sci. Total Environ.* **690**, 1120–1130 (2019).
86. Fang, H., Wei, S., Jiang, C. & Scipal, K. Theoretical uncertainty analysis of global MODIS, CYCLOPES, and GLOBCARBON LAI products using a triple collocation method. *Remote Sens. Environ.* **124**, 610–621 (2012).
87. Camacho, F., Cemicarho, J., Lacaze, R., Baret, F. & Weiss, M. GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products. *Remote Sens. Environ.* **137**, 310–329 (2013).
88. Prince, S. D. & Goward, S. N. Global primary production: A remote sensing approach. *J. Biogeogr.* **22**, 815–835 (1995).
89. Verma, S. B. *et al.* Annual carbon dioxide exchange in irrigated and rainfed maize-based agroecosystems. *Agr. Forest Meteorol.* **131**, 77–96 (2005).
90. Yan, H. *et al.* Improved global simulations of gross primary product based on a new definition of water stress factor and a separate treatment of C3 and C4 plants. *Ecol. Model.* **297**, 42–59 (2015).
91. Jiang, S. *et al.* Comparison of satellite-based models for estimating gross primary productivity in agroecosystems. *Agr. Forest Meteorol.* **297**, 108253 (2021).
92. Yang, X. *et al.* Solar-induced chlorophyll fluorescence that correlates with canopy photosynthesis on diurnal and seasonal scales in a temperate deciduous forest. *Geophys. Res. Lett.* **42**, 2977–2987 (2015).
93. Zhou, H. *et al.* Responses of gross primary productivity to diffuse radiation at global FLUXNET sites. *Atmos. Environ.* **244**, 117905 (2021).
94. Han, J. *et al.* Effects of diffuse photosynthetically active radiation on gross primary productivity in a subtropical coniferous plantation vary in different timescales. *Ecol. Indic.* **115**, 106403 (2020).
95. Grant, I. F., Prata, A. J. & Cechet, R. P. The impact of the diurnal variation of albedo on the remote sensing of the daily mean albedo of grassland. *J. Appl. Meteorol.* **39**, 231–244 (2000).
96. Singarayer, J. S., Ridgwell, A. & Irvine, P. Assessing the benefits of crop albedo bio-geoengineering. *Environ. Res. Lett.* **4**, 045110 (2009).
97. Tang, S. *et al.* LAI inversion algorithm based on directional reflectance kernels. *J. Environ. Manage.* **85**, 638–648 (2007).

Acknowledgements

This work was financially supported by the National key Research and Development Program of China (2019YFA0606604), the National Natural Science Foundation of China (42077419), and the Open Funding Project of the State Key Laboratory of Remote Sensing Science (OFSLRSS202012). We also thank the FLUXNET community (<https://fluxnet.fluxdata.org/>) that provided us eddy covariance data, the European Space Agency (ESA) CCI projects (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/satellite-land-cover>) that provided us land cover data, the Climatic Research Unit (CRU) and the Japanese Meteorological Agency (JMA) that provided meteorological data, the National Oceanic and Atmospheric Administration that provided global CO₂ concentration data (www.esrl.noaa.gov/gmd/ccgg/trends/), and Liu Yang from institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, that provided GLOBMAP LAI.

Author contributions

Wenjun Bi: Data processing, modelling, paper writing; Yanlian Zhou, Wei He and Weimin Ju: the study design; Yibo Liu: providing core code; Yang Liu: providing and processing LAI data; Xiaonan Wei, Nuo Cheng, and Xiaoyu Zhang: the original data collecting and processing. All authors reviewed and edited this manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-022-01309-2>.

Correspondence and requests for materials should be addressed to Y.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022