



Evaluation and Fusion of Multi-Source Gross Primary Productivity Products over China under Widespread Greening

Bingxiao Liu ^a, Wen Wang ^{a,*}, Wei Wu ^a, Yujie Liu ^{a,b}

^a State Key Laboratory of Water Disaster Prevention, Hohai University, Nanjing 210098, China

^b Hydrology and Water Resources Monitoring Center in the Upper Ganjiang River, Ganzhou 341000, China

* Corresponding author, Email address: wangwen@hhu.edu.cn

Abstract: Vegetation greening has been widely observed across China, yet whether this structural change has translated into stable terrestrial carbon uptake remains uncertain because gross primary productivity (GPP) estimates differ substantially among products. Here we systematically evaluated 12 widely used multi-source GPP products at a common 0.05° monthly resolution over 2001–2020, using paired observations from 33 ChinaFLUX eddy-covariance towers spanning the major vegetation types in China. We also characterized the greening signal with six vegetation indicators derived from satellite leaf area index (LAI), dynamic global vegetation model (DGVM) LAI, and near-infrared reflectance of vegetation (NIRv). Inter-product discrepancies were largest in high-productivity regions and wetlands, where about half of the products systematically underestimated monthly GPP exceeding 300 g C m⁻². Differences in greenness inputs and model structures further amplified the divergence among product categories. Among individual products, VPM, PML-V2, and GOSIF showed the strongest overall agreement with flux-tower observations. Using inter-product consistency as a selection constraint, we tested three fusion approaches on multiple subsets of source products. The arithmetic mean of five products (GLASS, VPM, TL-LUE, PML-V2, and GOSIF) achieved the best overall agreement with tower observations ($r = 0.87$; RMSE = 53.9 g C m⁻² month⁻¹) and produced a national multi-year mean GPP of 7.72 ± 2.03 Pg C yr⁻¹. Joint analysis



23 of the GPP and LAI ensembles indicates that approximately 27.21% of greening areas showed an
24 effective carbon-uptake response across China, but the relationship is spatially heterogeneous. These
25 findings demonstrate that China's greening is functionally meaningful at the national scale, but its
26 carbon-sink effectiveness depends strongly on regional environmental constraints and structure-
27 function coupling.

28 **Keywords:** Gross primary productivity; Carbon sink; Vegetation greening; Light use efficiency model;
29 Dynamic global vegetation model; Eddy covariance flux tower

30 **1. Introduction**

31 Widespread global greening has been detected since the 1980s through satellite observations,
32 with ecosystem model attribution suggesting that rising atmospheric CO₂ is its dominant driver (Zhu
33 et al., 2016). China's greening has been particularly pronounced since the beginning of the 21st
34 century, contributing 25% of the global net increase in leaf area despite containing only 6.6% of the
35 world's vegetated area, mainly through forest expansion and cropland intensification (Chen et al.,
36 2019a). Consequently, the underlying carbon-sink processes are likely to be more complex than
37 those driven primarily by climatic variability. The contribution of afforestation to carbon
38 sequestration therefore warrants rigorous evaluation (Cheng et al., 2024).

39 Terrestrial gross primary productivity (GPP), defined as the carbon assimilated by vegetation
40 through photosynthesis per unit area and time, is the most direct functional indicator linking
41 vegetation structure to the carbon cycle. However, direct observation of GPP at regional to
42 continental scales remains a major challenge (Hu et al., 2022). Driven by the growing availability of
43 remote-sensing observations and flux-tower networks, statistical methods, light use efficiency (LUE)



44 framework, land surface models (LSMs), and machine-learning (ML) techniques now underpins
45 global GPP estimation (Liao et al., 2023). Differences in parameterization, structural assumptions,
46 and radiative-transfer treatments still produce pronounced inter-model divergence, particularly over
47 heterogeneous and human-managed landscapes. Although energy-balance closure issues remain, the
48 eddy-covariance (EC) approach is still the most widely accepted reference for quantifying carbon
49 fluxes and for evaluating gridded GPP estimates (Xiao et al., 2025; Yu et al., 2006; Zhan et al.,
50 2024).

51 Despite growing attention to greening-carbon relationships, several gaps remain unresolved.
52 Attributing carbon-sink effects of greening is complicated by the dependence of many GPP products
53 on greenness-related inputs (Verma et al., 2014; Zhao et al., 2005), which limits the availability of
54 greenness-independent GPP datasets for attribution analysis (Hu et al., 2022). A robust, multi-source
55 GPP ensemble that integrates products with diverse model structures and input data is therefore
56 needed to provide a more independent basis for greening–productivity attribution (Zhang and Ye,
57 2022). While several global GPP intercomparison studies exist, systematic evaluation of recent
58 products against a regionally dense flux-tower network in China, including products derived from
59 solar-induced chlorophyll fluorescence (SIF) and vegetation optical depth (VOD), has been limited.
60 Even where greening signals are robust, their conversion into stable carbon uptake is shaped by
61 water, temperature, and radiation constraints (Chen et al., 2021; Kannenberg et al., 2022). The
62 underlying mechanism is a decoupling between vegetation structure and function, yet the drivers of
63 this decoupling remain poorly constrained (Zhang et al., 2021). Determining whether, where and
64 why greening signals are converted into effective carbon uptake therefore remains a key challenge



65 across China.

66 Building on these gaps, this study addresses four research questions:

67 (a) How well do widely used multi-source GPP products represent the spatial pattern, long-term
68 trend, and inter-annual variability of GPP in China, and where do they diverge most?

69 (b) How do GPP estimates differ within and between product categories, and are these
70 discrepancies driven mainly by differences in input data or in model structure?

71 (c) Can a robust GPP ensemble be constructed by combining products of complementary types,
72 and which fusion method performs best when evaluated against independent flux-tower
73 observations?

74 (d) Where, and under what environmental controls, has China's greening been accompanied by
75 a proportional increase in productivity, and where has the structural-functional coupling
76 broken down?

77 To address these questions, we used six vegetation indicators to characterize the greening signal
78 across China and systematically evaluated 12 multi-source GPP products against monthly
79 observations from 33 ChinaFLUX towers at a common 0.05° resolution over 2001–2020. We then
80 quantified product accuracy with Taylor-based metrics, estimated relative product uncertainty using
81 the three-cornered-hat (TCH) method, and compared three fusion strategies. By jointly analyzing the
82 best-performing GPP ensemble with the LAI ensemble and applying SHAP-based attribution, we
83 assessed where, and why, vegetation greening has produced stable carbon uptake across China.

84 **2. Materials**



85 **2.1 Data**

86 **2.1.1 EC flux observations**

87 Flux observation towers (**Fig. 1**) are widely distributed across diverse biomes and climatic zones,
88 providing a foundation for large-scale carbon flux estimation and model validation. The datasets of
89 33 carbon and water flux tower released by ChinaFLUX (<http://www.chinaflux.org/>) were used to
90 evaluate the GPP data products. The tower-based GPP was derived from the measured net ecosystem
91 exchange (NEE) and ecosystem respiration (Re) using the nighttime flux partitioning algorithm
92 (Reichstein et al., 2005), and GPP was calculated as $GPP_{month} = -NEE_{month} + R_{e,month}$.

93 The stations are distributed in 5 land cover types, including cropland (6 stations), forest (14
94 stations), grassland (7 stations), shrubland (3 stations), and wetland (3 stations). Detailed information
95 of the 33 flux observation stations, including geographic location, temporal resolution, and reference
96 is provided in Table S1 (see supplementary material).

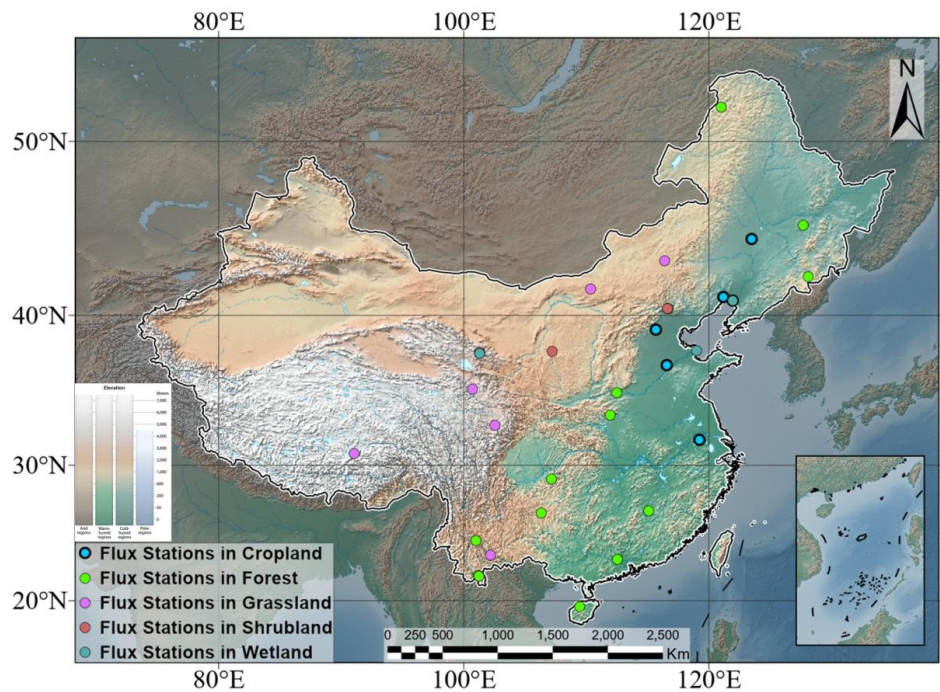


Fig. 1. Locations of EC flux observation towers across China, grouped by ecosystem type (wetland, shrubland, grassland, forest, and cropland). The background hill shading denotes topography. Basemap data from Natural Earth, 1:10m Cross-blended Hypsometric Tints.

2.1.2 GPP products

Twelve GPP products were evaluated in this study (**Table 1**). These products were grouped into three categories according to their underlying modeling frameworks: LUE Models, LSMs, and Observation-constrained Products. There are four products based on the LUE framework, namely MODIS MOD17A2H (Running et al., 2021), GLASS (Zheng et al., 2020), VPM (Zhang et al., 2017), and TL-LUE (Bi et al., 2022). Among the 21 TRENDY v12 (Sitch et al., 2024) dynamic global vegetation model (DGVM) products, CLASSIC, O-CN, and YIBs showed the highest consistency with regional flux-tower observations in China and were therefore selected as



representative DGVM products for evaluation (**Fig. S1**). BEPS (Ju and Zhou, 2021) was included as a widely used process-based ecosystem model, while FLUXCOM (Jung et al., 2020) represents a FLUXNET upscaling product generated using ML methods and meteorological remote-sensing predictors. PML-V2 (Zhang et al., 2016) was selected as a water-carbon coupled biophysical product. GOSIF GPP (Li and Xiao, 2019) and VODCA2GPP (Wild et al., 2022) were included to represent GPP estimates derived from SIF and VOD, respectively. Together, these products cover different modeling frameworks, input data sources, and assumptions for estimating terrestrial GPP. More detailed information on these products is provided in the supplementary materials.

Table 1. Information on GPP products

Categories	Product	Resolution	Source
Light Use Efficiency Models	MODIS	500 m / 8-day	https://lpdaac.usgs.gov/products/mod17a2hv061/
	GLASS	500 m / 8-day	https://www.glass.hku.hk/archive/GPP
	VPM	500 m / 8-day	https://www.tpdac.ac.cn/en/data/582663f5-3be7-4f26-bc45-b56a3c4fc3b7/
	TL-LUE	0.05° / month	https://datadryad.org/dataset/doi:10.5061/dryad.dfn2z352k
Land Surface Models	BEPS	8 km / day	https://www.nesdc.org.cn/sdo/detail?id=612f42ee7e28172cbcd3d809
	CLASSIC	0.5° / month	https://mdosullivan.github.io/GCB/
	O-CN	0.5° / month	https://mdosullivan.github.io/GCB/
	YIBs	0.5° / month	https://mdosullivan.github.io/GCB/
Observation-constrained Products	PML-V2	500 m / 8-day	https://zenodo.org/records/10647618
	FLUXCOM	0.5° / month	https://www.bgc-jena.mpg.de/geodb/projects/Data.php
	GOSIF	0.05° / month	https://data.globalecology.unh.edu/data/GOSIF-GPP_v2/Monthly/Mean/
	VODCA2GPP	0.25° / 8-day	https://researchdata.tuwien.ac.at/records/1k7aj-bdz35

2.1.3 Vegetation greenness and structural indicators

In this study, LAI and near-infrared reflectance of vegetation (NIRv) were used to characterize vegetation structural change and canopy photosynthetic activity. Satellite-based LAI was derived from two long-term products, GIMMS LAI4g and GLASS LAI. GIMMS LAI4g provides global



half-monthly LAI at 1/12° resolution from 1982 to 2020 and was specifically developed to reduce uncertainties caused by NOAA satellite orbital drift, AVHRR sensor degradation, and limited historical LAI reference data (Cao et al., 2023). GLASS LAI provides global 8-day LAI, with 0.05° AVHRR-based data before 2000 and MODIS-based data thereafter, forming a consistent long-term LAI record (Liang et al., 2021; Xiao et al., 2016). Although these products are widely used, satellite LAI still contains uncertainties related to atmospheric contamination, retrieval algorithms, sensor degradation, and transitions between AVHRR and MODIS sensors (Fang et al., 2019). To reduce product-specific biases and improve robustness, their arithmetic mean was then calculated as the satellite-based LAI ensemble.

DGVM-based LAI was obtained from CLASSIC, O-CN, and YIBs simulations from TRENDY v12. These models provide process-based representations of vegetation dynamics and carbon cycling, but are affected by model structure, parameterization, and forcing data uncertainties (Friedlingstein et al., 2023).

NIRv was used as an additional satellite indicator of vegetation photosynthetic activity. NIRv is calculated from near-infrared reflectance and vegetation greenness and has been shown to correlate well with solar-induced chlorophyll fluorescence and GPP (Badgley et al., 2017). In this study, NIRv was derived from MODIS reflectance products (MCD43A4_Version6.1) at 0.05° resolution.

2.1.4 Meteorological factors

The meteorological predictors used in this study were derived from the China Meteorological Forcing Dataset (CMFD), a high-resolution gridded dataset developed for land-surface, hydrological,



143 and ecological studies over China (He et al., 2020). CMFD integrates remote-sensing products,
144 reanalysis data, and in situ meteorological observations, and provides 3-hourly forcing variables at
145 0.1° spatial resolution from 1979 onward. The CMFD variables used here include precipitation
146 (PREC), shortwave radiation (SRAD), air temperature (TEMP), and variables required to derive
147 vapor pressure deficit (VPD). Soil water content (SWVL) was derived from ERA5-Land, which
148 provides hourly land surface variables at approximately 0.1° resolution (Muñoz-Sabater et al., 2021).
149 We used volumetric soil water from the 0–7, 7–28, and 28–100 cm layers to represent root-zone soil
150 moisture within 0–100 cm. Atmospheric CO₂ concentration was obtained from NOAA
151 CarbonTracker CT2022, a CO₂ measurement and modeling system that assimilates atmospheric CO₂
152 observations with atmospheric transport simulations and provides global CO₂ estimates (Jacobson et
153 al., 2023).

154 For this study, we derived annual mean climate conditions, interannual variability, and long-
155 term trends from during the study period. For each variable, annual means were calculated to
156 describe background environmental conditions, coefficients of variation were computed to quantify
157 temporal variability, and linear slopes were estimated to characterize long-term changes. These
158 predictors allow the ML model to evaluate both the relative importance and direction of climatic
159 controls on LAI-GPP coupling changes across China.

160 **2.2 Methods**

161 **2.2.1 Scaling GPP products to common spatial resolution**

162 The footprint of eddy-covariance flux towers usually ranges from several hundred meters to



163 several kilometers, depending on tower height, canopy structure, and atmospheric conditions.
164 Therefore, selecting an appropriate product resolution for tower-based evaluation can improve the
165 spatial consistency between gridded GPP products and flux-tower observations. Previous studies
166 have also shown that matching satellite observations with flux-tower footprints can improve the
167 agreement between remotely sensed photosynthesis estimates and tower-based measurements (Wang
168 et al., 2023).

169 For relatively fine-resolution products, such as PML-V2, GLASS, and MODIS, the original data
170 were aggregated to the target resolution using area-weighted conservative remapping. However, for
171 coarse-resolution products, such as TRENDY products with a spatial resolution of approximately
172 $0.5^\circ \times 0.5^\circ$, directly resampling them to a few-kilometer scale may introduce substantial spatial
173 scaling errors, particularly in heterogeneous landscapes. Previous studies have shown that vegetation
174 heterogeneity and sub-pixel variability can strongly affect the spatial scaling of GPP estimates (Xie
175 et al., 2018).

176 Because the purpose of this downscaling procedure was to place all GPP products on a common
177 spatial scale for evaluation, rather than to generate a new independent fine-resolution GPP product,
178 complex methods such as ML downscaling framework were not adopted. Such methods would
179 require reliable fine-resolution reference GPP data for training, which are generally unavailable at
180 large scales. In addition, the use of complex model structures and multiple climatic or topographic
181 predictors may introduce additional uncertainty propagation, which could compromise the fairness of
182 the inter-product comparison (Guan et al., 2025). Therefore, we used a vegetation-index-based
183 weighted downscaling approach to assign larger weights to sub-grid areas with stronger vegetation



184 photosynthetic activity. To examine the sensitivity of the evaluation results to the spatial treatment
185 and auxiliary variable, we conducted several sensitivity analyses. We compared results obtained at
186 the original product resolution, after downscaling to 0.05° , after downscaling to 0.01° , and after
187 aggregating all products to 0.5° . We also tested both LAI and NIRv as auxiliary variables for
188 weighted downscaling. The results showed that using NIRv to downscale the products to 0.05°
189 produced the highest correlation and the lowest RMSE against the flux-tower observations. This
190 choice is supported by the physical interpretation of NIRv as a proxy for canopy photosynthesis.
191 Badgley et al. (2017) showed that NIRv is strongly correlated with SIF as well as both site-level and
192 gridded GPP estimates, suggesting that NIRv can serve as an effective reflectance-based proxy for
193 terrestrial photosynthesis. Cao et al. (2025) further indicated that NIRv generally exhibits stronger
194 correlations with tower-based GPP than other vegetation indices across multiple ecosystems.

195 Although uncertainty in auxiliary NIRv may affect local tower-grid mismatches in highly
196 heterogeneous regions, the weighted downscaling is conservative and preserves the original regional
197 GPP totals and large-scale mean values while maintaining the independence of the evaluated GPP
198 products. And for each coarse grid cell, the downscaled GPP at fine grid cell i is computed as
199 follows:

$$200 \quad G_{\text{down},i} = G_{\text{coarse}} \frac{P_{\text{fine},i}}{\bar{P}_{\text{coarse}}} \quad (1)$$

201 where $G_{\text{down},i}$ is the downscaled GPP value at fine grid cell i , G_{coarse} is the original coarse-
202 resolution GPP value for the corresponding coarse grid cell, $P_{\text{fine},i}$ is the fine-resolution NIRv value
203 at fine grid cell i , and $\bar{P}_{\text{coarse}}$ is the mean NIRv value over all fine grid cells within the



204 corresponding coarse grid cell.

205 **2.2.2 Taylor diagram**

206 The Taylor diagram is a graphical tool that summarizes how well model outputs match
 207 reference observations by jointly displaying their standard deviation (SD), Pearson's correlation
 208 coefficient (r), and root-mean-square error (RMSE) difference in a single polar plot. It is widely used
 209 for comparison between flux-tower observations and product estimates (Zuo et al., 2025). The
 210 evaluation metrics are calculated as follows:

$$211 \quad SD_T = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (T_i - \bar{T})^2} \quad (2)$$

$$212 \quad SD_G = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (G_i - \bar{G})^2} \quad (3)$$

$$213 \quad r = \frac{\sum_{i=1}^N (T_i - \bar{T})(G_i - \bar{G})}{\sqrt{\sum_{i=1}^N (T_i - \bar{T})^2 \sum_{i=1}^N (G_i - \bar{G})^2}} \quad (4)$$

$$214 \quad RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (G_i - T_i)^2} \quad (5)$$

215 where T_i denotes the flux-tower station GPP at sample i and G_i denotes the gridded GPP at the
 216 same sample. $\bar{T}$ and $\bar{G}$ are the arithmetic means of T_i and G_i over all N paired samples;
 217 SD_T and SD_G are the standard deviations of T_i and G_i ; All calculations use the same set of
 218 paired samples $i = 1, \dots, N$ to ensure consistency across metrics.

219 **2.2.3 Trend analysis**



220 The Mann-Kendall (MK) nonparametric test was used to detect monotonic trends in the annual
 221 GPP time series. The test statistic S is defined as:

$$222 \quad S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(x_j - x_i) \quad (6)$$

223 Where x_i are the annual GPP values. Under the null hypothesis of no trend, and the variance of S
 224 is calculated as follows:

$$225 \quad \text{Var}(S) = \frac{1}{18} \left[n(n-1)(2n+5) - \sum_p t_p(t_p-1)(2t_p+5) \right] \quad (7)$$

226 where n is the number of years and t_p is the number of tied values in the p -th. The statistics are
 227 then normalized to a standard normal Z -value as follows:

$$228 \quad Z = \begin{cases} \frac{S-1}{\sqrt{\text{Var}(S)}} & S > 0 \\ [6pt] 0 & S = 0 \\ [6pt] \frac{S+1}{\sqrt{\text{Var}(S)}} & S < 0 \end{cases} \quad (8)$$

229 where the observed trend is considered statistically significant if $|Z| \geq 1.96$ (corresponding to a 95%
 230 confidence level). Kendall's τ -value is then computed as follows:

$$231 \quad \tau = \frac{S}{\binom{n}{2}} = \frac{2S}{n(n-1)} \quad (9)$$

232 where τ is Kendall's rank correlation (ranging from -1 to +1), S is the Mann-Kendall statistic
 233 defined above, and n is the number of years.

234 Sen's slope (also known as the Theil-Sen estimator) was used to estimate the median rate of



235 change in GPP. It is computed as the median of all pairwise slopes between data points:

$$236 \quad \beta = \text{median} \left\{ \frac{x_j - x_i}{j - i} : 1 \leq i < j \leq n \right\} \quad (10)$$

237 where x_i and x_j are the ordered annual GPP series. This median slope β represents the typical
238 annual change in GPP.

239 **2.2.4 TCH method**

240 The three-cornered hat (TCH) method (McColl et al., 2014; Stoffelen, 1998) is used to estimate
241 the random error variance of multiple datasets when an unknown true reference is unavailable. The
242 basic assumption is that each product consists of the same true signal and an independent random
243 error term. The variance of the difference between any two products can be expressed as the sum of
244 their individual error variances. This allows the relative uncertainty of each product to be estimated
245 from all pairwise differences among products (Sjoberg et al., 2021). It should be noted that several
246 products in this study share common input data, which may introduce error correlations and thus
247 cause TCH to underestimate the true uncertainty of the ensemble estimate. For each i th product at
248 time t , the observation can be written as:

$$249 \quad X_i(t) = T(t) + \varepsilon_i(t), \quad (11)$$

250 where $T(t)$ is the unknown true signal and $\varepsilon_i(t)$ is the random error of product i . If the errors of
251 products i and j are independent, the variance of their difference is:

$$252 \quad \text{Var}[X_i(t) - X_j(t)] = \sigma_i^2 + \sigma_j^2, \quad (12)$$

253 where σ_i^2 and σ_j^2 are the error variances of products i and j . Using the 12 products, the



variances of all unique product pairs were combined to form a linear system:

$$\mathbf{A}\mathbf{v} = \mathbf{b}, \quad (13)$$

where $\mathbf{v} = [\sigma_1^2, \sigma_2^2, \dots, \sigma_{12}^2]^T$, $\mathbf{b}$ is the vector of sample variances of all pairwise product differences, and $\mathbf{A}$ is a coefficient matrix composed of 0 and 1. For each product pair (i, j) , the corresponding row of $\mathbf{A}$ has values of 1 in the i th and j th columns and 0 elsewhere. Thus, each equation in the system represents the relationship $s_{ij}^2 = \sigma_i^2 + \sigma_j^2$. The error variance vector was then estimated using the Moore-Penrose pseudoinverse:

$$\mathbf{v} = \mathbf{A}^+ \mathbf{b}. \quad (14)$$

For each grid cell, only years with valid values for all products were retained. The sample variance of the difference between products i and j was calculated as:

$$s_{ij}^2 = \frac{1}{n-1} \sum_{t=1}^n (d_{ij}(t) - \bar{d}_{ij})^2, \quad (15)$$

where s_{ij}^2 is the sample variance of the difference series between products i and j , n is the number of valid years, $d_{ij}(t)$ is the difference between the two products at time t , and $\bar{d}_{ij}$ is the mean of $d_{ij}(t)$ over the valid years. The difference series is defined as:

$$d_{ij}(t) = X_i(t) - X_j(t), \quad (16)$$

2.2.5 Bayesian Model Averaging (BMA)

Bayesian Model Averaging (BMA) is a probabilistic ensemble method used to combine multiple candidate predictors into one calibrated estimate (Hoeting et al., 1999; Raftery et al., 2005). BMA assigns a posterior weight to each candidate model according to its support from the calibration data,



then produces a weighted prediction that accounts for model uncertainty. For a target variable y , candidate models M_k , and input vector $\mathbf{x}$, the BMA predictive distribution is

$$p(y | \mathbf{x}, D) = \sum_{k=1}^K p(y | \mathbf{x}, M_k, D) p(M_k | D), \quad (17)$$

where D is the calibration dataset, $p(y | \mathbf{x}, M_k, D)$ is the predictive distribution from model M_k , $p(M_k | D)$ is the posterior model probability. In deterministic fusion, the posterior mean is used as the fused estimate:

$$\hat{y}_{\text{BMA}} = \beta_0 + \sum_{k=1}^K w_k x_k, \quad w_k \geq 0, \quad \sum_{k=1}^K w_k = 1, \quad (18)$$

where x_k is the prediction from product k , w_k is its normalized weight, and β_0 is the fitted intercept. Missing predictors are replaced during calibration, and available predictors are reweighted during prediction. BMA is widely used for model combination and ensemble calibration because it converts relative model skill into explicit probabilistic weights (Tian et al., 2023).

2.2.6 Random Forest (RF)

Random Forest (RF) is an ensemble learning algorithm that integrates multiple decision trees to improve predictive accuracy, robustness, and generalization (Breiman, 2001). For regression tasks, each tree is constructed from a bootstrap sample of the training data, and a random subset of predictors is selected at each node to determine the optimal split. The final output is obtained by averaging predictions from all trees, which reduces variance and limits the instability commonly associated with single decision trees. Owing to its nonparametric structure, RF is well suited for modeling nonlinear relationships, high-dimensional predictors, threshold responses, and complex



292 interactions among variables (Van Jaarsveld et al., 2024). In this study, RF regression was used for
293 both data fusion and factor interpretation.

294 The predictive performance of RF is mainly controlled by several hyperparameters, including
295 the number of trees, maximum tree depth, number of predictors considered at each split, minimum
296 samples required to split a node, and minimum samples retained in terminal leaves. A larger number
297 of trees generally improves model stability, whereas constraints on tree depth and leaf size help
298 prevent overfitting. The number of randomly selected predictors regulates the trade-off between tree
299 diversity and individual tree strength. In addition to prediction, RF provides measures of variable
300 importance based on impurity reduction or permutation-based performance loss. To improve
301 interpretability, SHapley Additive exPlanations (SHAP) can be applied to quantify the contribution
302 of each predictor to model outputs. SHAP values decompose individual predictions into additive
303 feature contributions, indicating both the magnitude and direction of each variable's influence. Mean
304 absolute SHAP values further provide a global ranking of predictor importance, enabling more
305 transparent interpretation of RF model behavior (Cappelli et al., 2024; He et al., 2023).

306 **3. Results**

307 **3.1 Widespread greening across China and its potential implications** 308 **for GPP evaluation**

309 China has been identified as one of the fastest and most extensive greening regions worldwide.
310 Satellite evidence shows that China accounts for about one quarter of the global net increase in leaf
311 area despite having only 6.6% of the global vegetated area, with forest expansion and cropland



intensification as the two dominant contributors (Chen et al., 2019a). This greening signal is consistently reflected in the interannual variations of DGVM-derived LAI, satellite-based LAI, and NIRv during 2001 to 2020 (**Fig. 2**). The spatial correlation between the two LAI products derived from different principles reaches 0.877, while their correlations with NIRv are 0.931 for satellite-based LAI and 0.887 for DGVM-based LAI, respectively. These high correlations indicate that the long-term mean spatial patterns of vegetation greenness are broadly consistent among the three datasets.

However, the interpretation of greening in China is not straightforward. Previous studies have shown that China's greening is strongly shaped by human land-use management, including afforestation, ecological restoration, and agricultural intensification, rather than by climate or CO₂ fertilization alone (Piao et al., 2015). Unlike greening driven primarily by climatic change, human-managed greening may introduce additional uncertainty into the evaluation of carbon-sink effectiveness, because vegetation expansion does not necessarily translate into a stable or persistent carbon sink. Therefore, accurately identifying the greening signal is critical, as both the uncertainty and temporal stability of LAI can directly affect conclusions about terrestrial carbon uptake.

China's complex topography, climate gradients, and land-use history lead to strong spatial heterogeneity in LAI trends. In our results, NIRv shows a clear increasing trend across most land areas, especially in southern China, northeastern China, and parts of the Loess Plateau, whereas weak negative trends occur locally in arid northwestern China and some fragmented regions of northern and eastern China. The DGVM ensemble broadly captures the direction of greening, but its spatial pattern is less consistent with satellite-based evidence. Specifically, the DGVM ensemble tends to



333 produce smoother and stronger greening signals in parts of northern and northeastern China, while
334 underrepresenting the observed spatial heterogeneity in southern and southwestern China. Across
335 China, 92.85% of valid satellite-based LAI pixels, 92.03% of DGVM-based LAI pixels, and 88.13%
336 of NIRv pixels exhibit positive trends, while negative trends account for 7.15%, 7.97%, and 11.87%,
337 respectively.

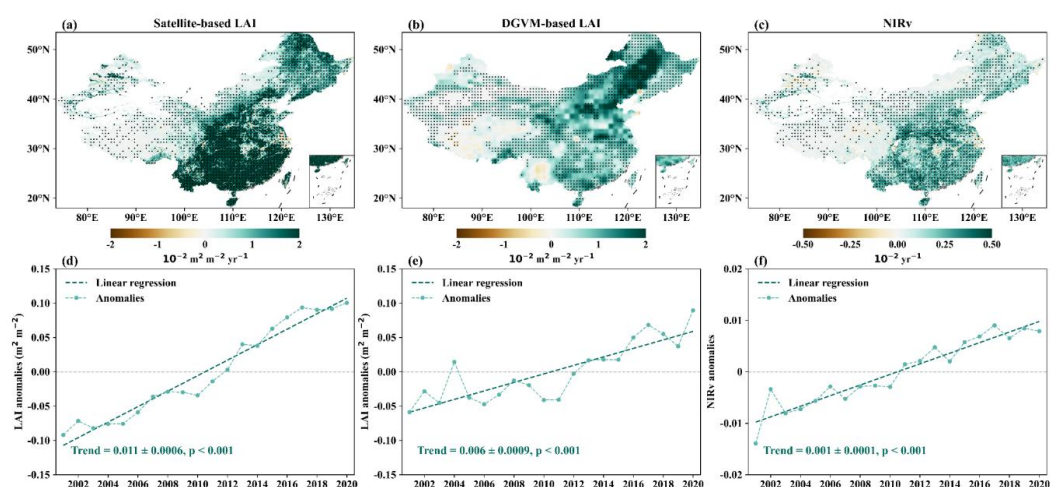
338 The trend comparison further reveals a clear dataset-dependent divergence. The spatial
339 correlation of trends is only 0.196 between satellite-based LAI and DGVM-based LAI, but increases
340 to 0.773 between satellite-based LAI and NIRv. By contrast, the correlation between DGVM-based
341 LAI trends and NIRv trends is only 0.222. This indicates that, although DGVM-based LAI can
342 reproduce the climatological mean greenness pattern reasonably well, it performs less well in
343 capturing the spatial pattern of vegetation change. Satellite-based LAI is much more consistent with
344 NIRv in terms of trend magnitude, spatial distribution, and the distribution of significant greening
345 signals.

346 This discrepancy is important for GPP evaluation. DGVM-based LAI is often smoother and
347 may appear more temporally stable, whereas satellite LAI can be affected by retrieval uncertainties
348 related to cloud contamination, atmospheric correction, sensor degradation, snow cover, topographic
349 effects, and signal saturation in dense canopies (Chen et al., 2023). Nevertheless, in China, the
350 stronger consistency between satellite-based LAI and NIRv suggests that satellite-derived greenness
351 changes better capture the regional greening signal than the DGVM ensemble. The DGVM ensemble
352 appears to miss or misrepresent greening in several heterogeneous regions, probably because land
353 management, crop rotation, ecological restoration, irrigation, and vegetation structural changes are



354 still simplified in many model frameworks. This regional mismatch can be overlooked in global-
 355 scale assessments, but it is particularly important in China, where greening is spatially complex and
 356 strongly influenced by human activities (Yang et al., 2022).

357 Since LAI is a key structural variable controlling canopy light absorption, photosynthesis, water
 358 fluxes, and terrestrial carbon uptake (Badgley et al., 2017), its uncertainty can propagate directly into
 359 GPP products. Previous work has shown that vegetation structural change, as represented by LAI,
 360 contributes significantly to the terrestrial carbon sink, and that LAI errors can strongly influence
 361 modeled GPP (Chen et al., 2019b). In this context, the inconsistency between DGVM-based LAI and
 362 satellite/NIRv greening trends indicates a potential source of uncertainty in GPP evaluation over
 363 China. Therefore, integrating satellite-based LAI, NIRv, and process-based simulations provides a
 364 more robust framework for diagnosing China's greening and for constraining GPP uncertainty in
 365 regions with strong human management and complex environmental conditions.



366
 367 Fig. 2. Spatial patterns and temporal variations of vegetation greening across China during 2001 to 2020 derived
 368 from satellite-based LAI, DGVM-based LAI, and NIRv. The satellite-based LAI is calculated as the ensemble mean



of GLASS and GIMMS LAI products, while the DGVM-based LAI is calculated as the ensemble mean of CLASSIC, O-CN, and YIBs simulations. Panels (a) to (c) show the spatial distributions of linear trends, with stippling indicating grid cells where the trend is statistically significant. Panels (d) to (f) show the corresponding annual anomalies and linear regression trends averaged over China.

3.2 Accuracy, uncertainty, and spatiotemporal patterns across GPP products

All 12 products capture the basic spatial pattern of GPP in China, characterized by higher values in the southeast and lower values in the northwest (Fig. 3). However, they differ substantially in the estimated magnitude of high-productivity regions. MODIS gives the lowest estimate, whereas VODCA2GPP gives the highest. The mean annual accumulated GPP is approximately $833 \text{ g C m}^{-2} \text{ yr}^{-1}$ for MODIS and reaches $1901 \text{ g C m}^{-2} \text{ yr}^{-1}$ for VODCA2GPP. These differences are mainly concentrated in humid southeastern China, South China, parts of Northeast China, and the margins of the Tibetan Plateau. VODCA2GPP and O-CN show clear overestimation, whereas MODIS, FLUXCOM, and VPM tend to produce lower estimates. GLASS, BEPS, PML-V2, and GOSIF fall at intermediate levels. The inter-product discrepancies are not randomly distributed but are concentrated in regions with high GPP, indicating that uncertainty in estimating terrestrial carbon uptake in China mainly originates from highly productive ecosystems.

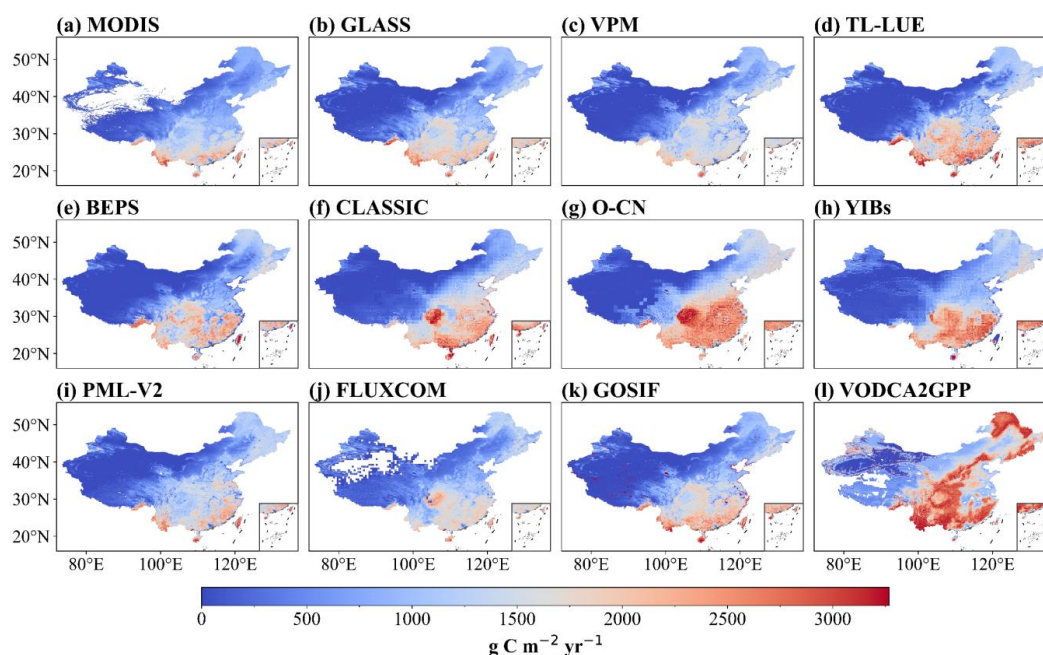
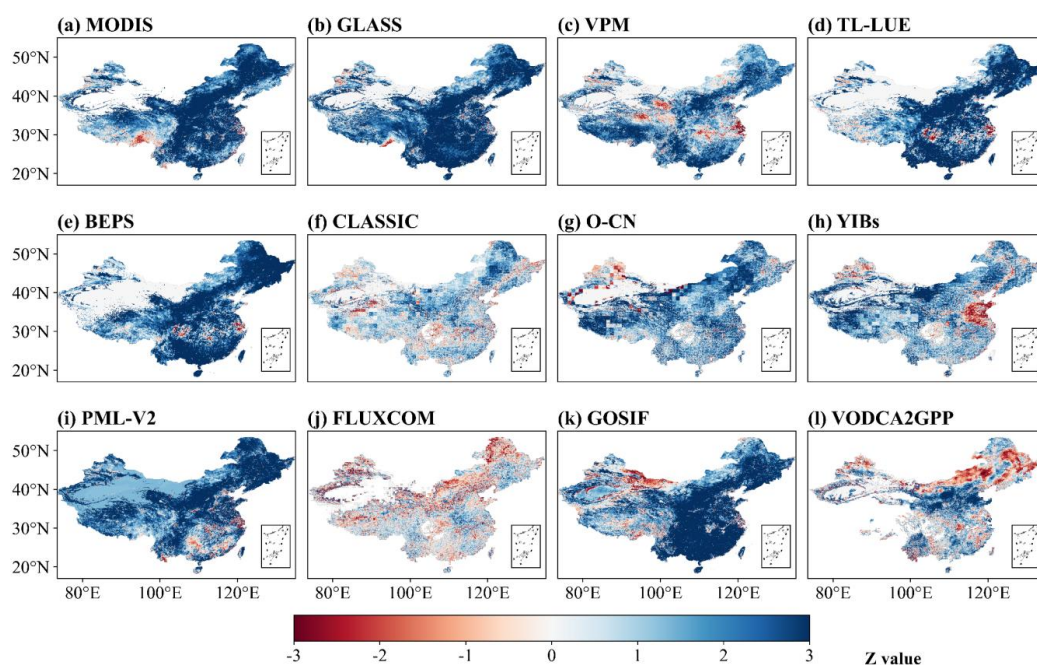


Fig. 3. Spatial distribution of multiyear mean annual GPP across China from 12 products (a–l).

China's GPP generally increased from 2001 to 2020 (Fig. 4, Fig. 5), but the products differ markedly in their estimates of trend magnitude and spatial extent. GOSIF, GLASS, MODIS, BEPS, TL-LUE, and PML-V2 show the most widespread and statistically significant positive trends. GOSIF exhibits the strongest increasing trend, while GLASS and MODIS show relatively large areas of significant increase. VPM and O-CN indicate moderate growth. YIBs, CLASSIC, and VODCA2GPP show weaker increases, with VODCA2GPP also exhibiting more areas of decline. FLUXCOM is the most distinctive product, with an overall Z value close to zero and comparable areas of significant increase and decrease, indicating the weakest long-term trend signal. This is mainly because its machine-learning upscaling framework does not explicitly account for the CO₂ fertilization effect (Yang et al., 2022), and because data-driven products have limited capacity to characterize



398 interannual variability and long-term trends (Nelson et al., 2024). Therefore, these differences reflect
 399 structural differences among methods and their representation of driving factors.



400
 401 **Fig. 4.** Spatial distribution of the Mann-Kendall Z statistic for long-term trends in annual mean GPP across China
 402 derived from 12 products (a–l). Positive Z denotes a monotonic increase in GPP, negative Z a monotonic decrease,
 403 and values near zero indicate no detectable monotonic trend. Larger $|Z|$ implies stronger evidence for a trend ($|Z| \geq$
 404 1.96 and ≥ 2.58 correspond approximately to the 0.05 and 0.01 significance levels, respectively).

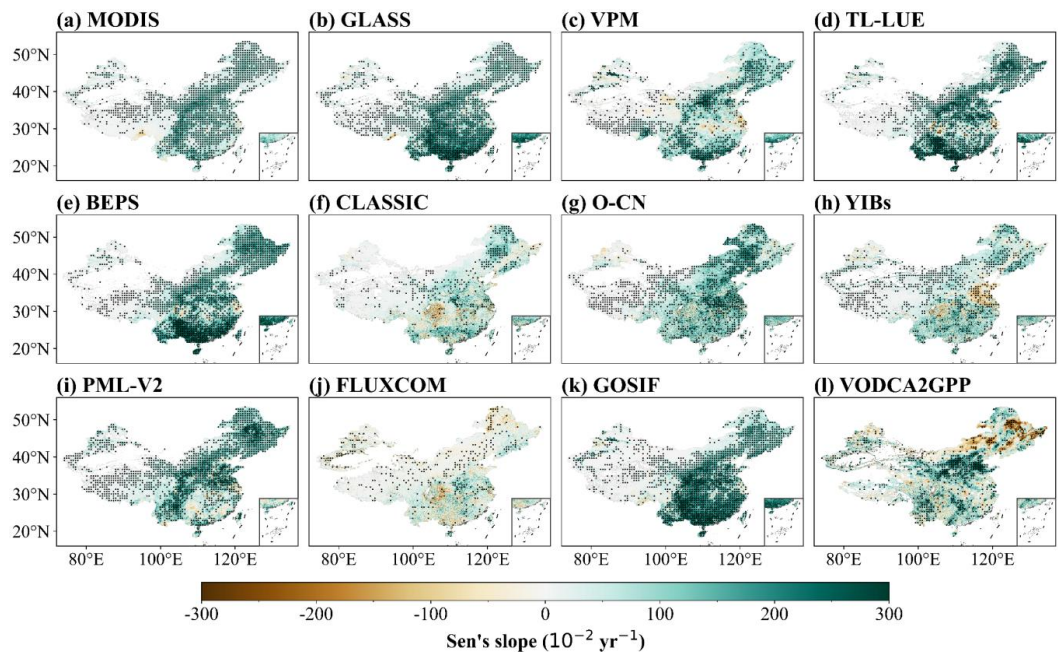


Fig. 5. Spatial distribution of Sen's slope for long-term trends in annual mean GPP across China during 2001–2020, derived from 12 GPP products (a–l). Positive and negative values indicate increasing and decreasing GPP trends, respectively, with larger absolute values representing faster rates of change. Black dots denote grid cells with statistically significant trends at the 0.05 level.

Accuracy assessment against flux-tower observations shows that product performance is strongly dependent on vegetation type, and no product is universally optimal across all ecosystems. VPM shows the best agreement with flux-tower observations, whereas VODCA2GPP performs the weakest. PML-V2 and GOSIF also show stable performance. Across vegetation types, VPM performs particularly well in croplands and grasslands, FLUXCOM performs best in forests, GLASS is the most balanced in shrublands, and GOSIF performs relatively well in wetlands. Wetlands are the most difficult vegetation type for all products to simulate, with generally high RMSE values and clear deviations in standard deviation. The products are generally consistent at low GPP values. However, when observed GPP exceeds approximately $300 \text{ g C m}^{-2} \text{ month}^{-1}$, products such as



MODIS, GLASS, BEPS, and FLUXCOM tend to underestimate high values and show greater dispersion (Fig. 7).

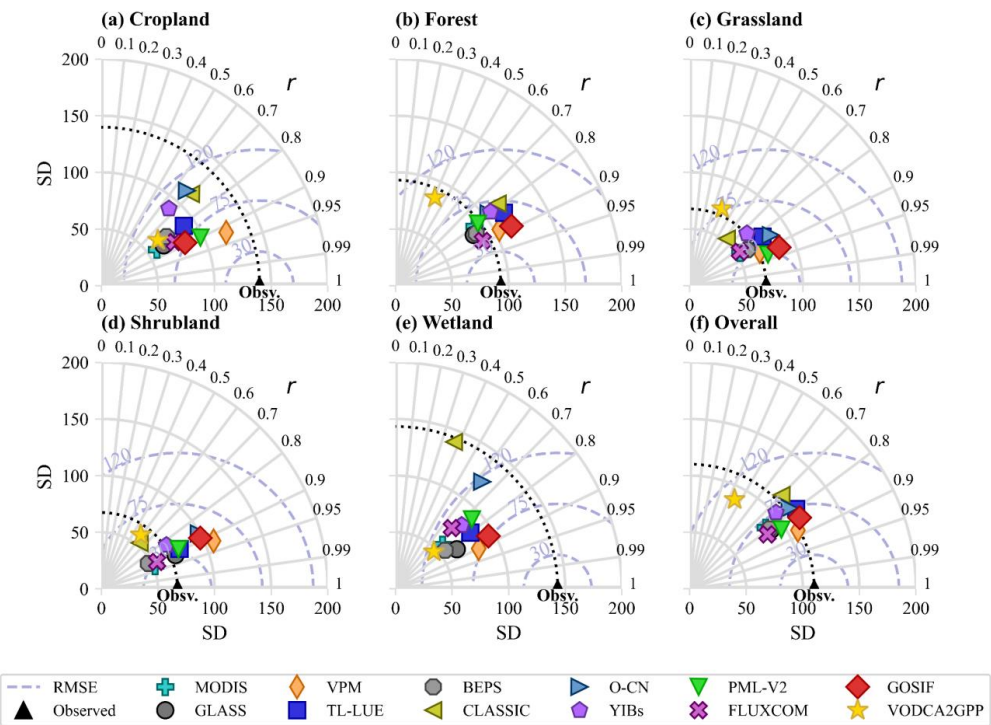
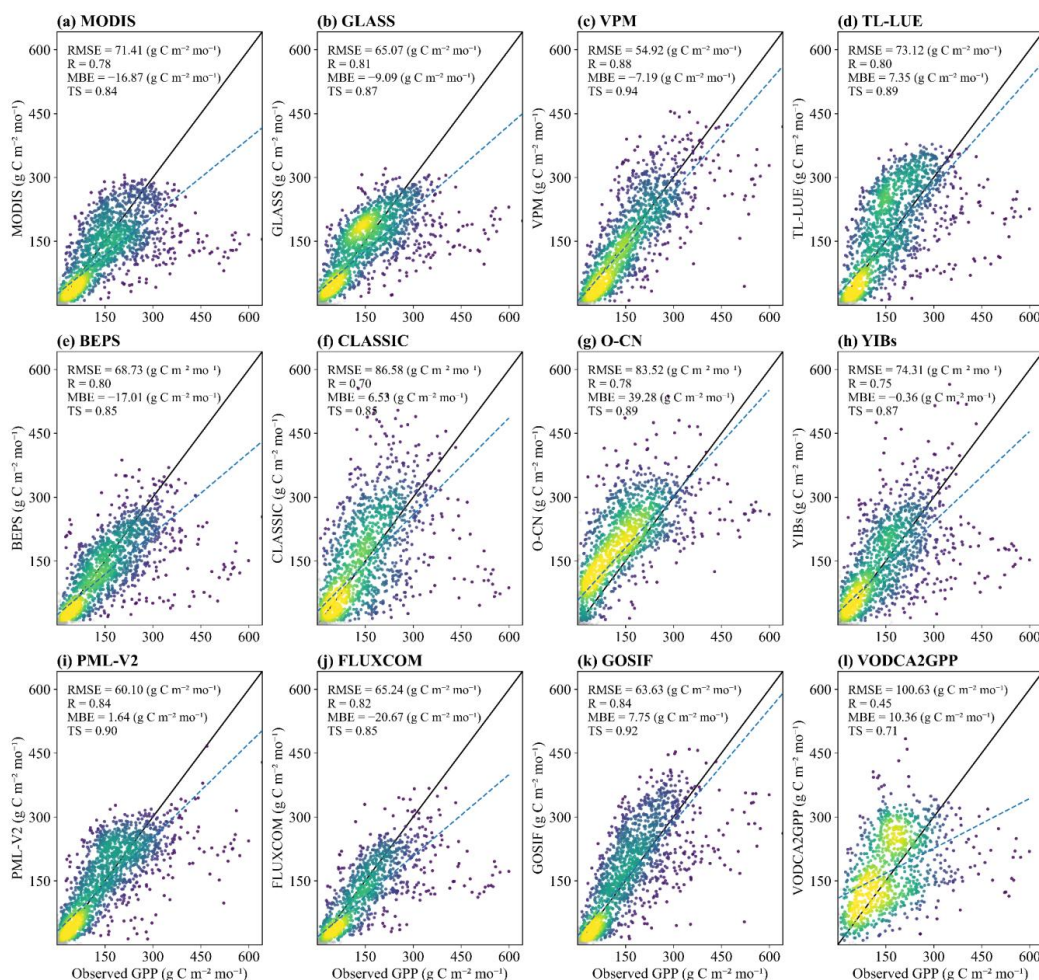


Fig. 6. Taylor diagram comparing 12 products with flux-tower observations across (a) cropland, (b) forest, (c) grassland, (d) shrubland, (e) wetland, and (f) overall performance across all ecosystem types; displayed metrics include r , SD ($\text{g C m}^{-2} \text{ month}^{-1}$), and RMSE ($\text{g C m}^{-2} \text{ month}^{-1}$), where points near the observation indicate higher r , SD closer to the reference, and lower RMSE.



426

427 **Fig. 7.** Tower-product comparisons of monthly GPP across China for 12 products (a-l). Dashed blue indicates the
 428 fitted relationship and black denotes the 1:1 line. Points are colored by kernel-density estimation to reduce
 429 overplotting; densities were computed from samples $\geq 10 \text{ g C m}^{-2} \text{ month}^{-1}$. Summary statistics, including RMSE, r,
 430 mean bias error (MBE), and Taylor skill score (TS), are shown in each panel.

431 The product with the best site-level validation performance is not necessarily the product with
 432 the lowest uncertainty (**Fig. 8**). MODIS and GLASS have the lowest TCH uncertainty, indicating the
 433 strongest consistency with the multi-product ensemble. GOSIF, FLUXCOM, and TL-LUE form a
 434 second tier. VODCA2GPP and O-CN show the highest TCH uncertainty, especially in eastern China,



northeastern China, and some highly productive regions. TCH analysis and site validation reveal two types of reliability. MODIS and GLASS are stable products with high multi-product consistency, whereas VPM and PML-V2 are closer to flux-tower observations but show stronger spatial variability.

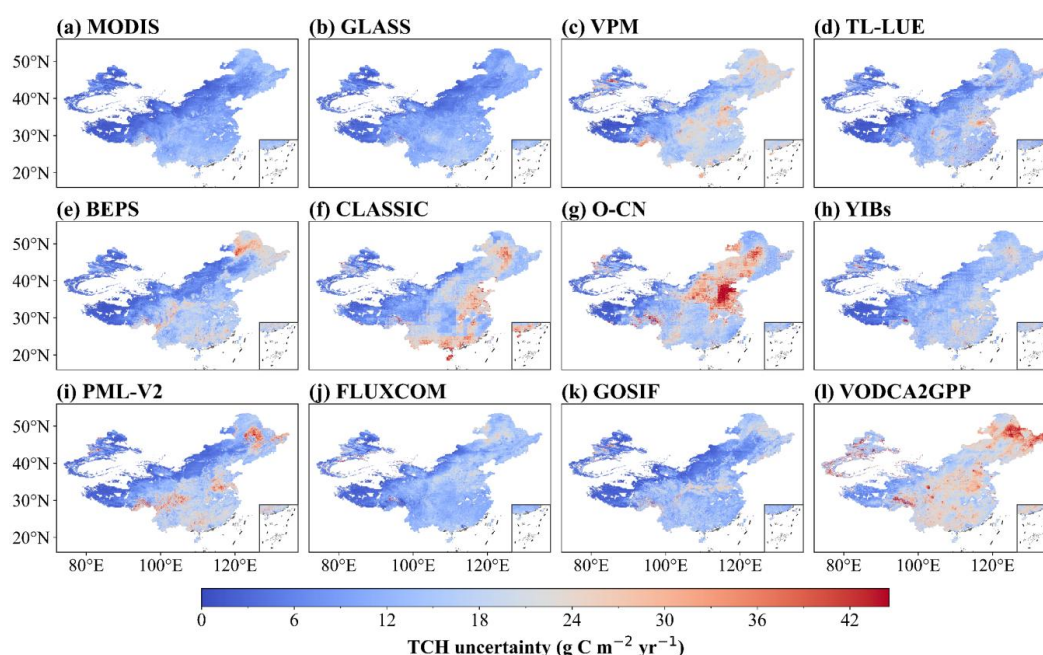


Fig. 8. Spatial distribution of error standard deviation estimated using the TCH method for annual mean GPP across China, derived from 12 GPP products (a-l).

3.3 Discrepancies driven by greenness input data and model structures

The vegetation indices (VIs) input and model structures can significantly affect the accuracy of GPP estimation. For example, filtering VI time series can systematically reduce GPP underestimation caused by cloud contamination and noise, with the most pronounced improvements observed in low latitude regions and during the growing season (Sun et al., 2023). TCH-based uncertainty exhibits



447 pronounced spatial heterogeneity across China (**Fig. 9**), suggesting that GPP estimation remains
448 particularly challenging in regions characterized by complex vegetation dynamics (He et al., 2022).
449 Products driven by satellite-derived LAI inputs exhibit the lowest and most spatially coherent
450 uncertainty, indicating that shared satellite-based canopy structural constraints can substantially
451 reduce inter-product divergence (Xie et al., 2019). Products employing different vegetation proxy
452 variables show greater divergence. Such products may rely on normalized difference vegetation
453 index (NDVI), enhanced vegetation index (EVI), solar-induced chlorophyll fluorescence (SIF), or
454 other remote-sensing proxies. Because these proxies differ in their sensitivity to canopy structure,
455 photosynthetic function, background reflectance, and atmospheric correction, they are more prone to
456 generating divergent GPP estimates (Yang et al., 2022). The absence of vegetation-proxy inputs leads
457 to substantially higher uncertainty in humid and intensively managed regions, highlighting the
458 importance of explicit vegetation observational constraints for reducing GPP uncertainty. The
459 relatively low TCH variance among LUE-based products indicates high internal consistency.
460 However, this should not be interpreted as evidence of higher absolute accuracy, because shared
461 model assumptions and common remote-sensing inputs may introduce correlated errors. LSM-based
462 products exhibit the highest uncertainty, particularly in eastern and southern China, suggesting that
463 differences in process representation and parameterization remain major sources of GPP uncertainty
464 in land surface models (Sitch et al., 2024). Observation-constrained products still show considerable
465 uncertainty in regions with sparse flux-tower coverage or complex environmental gradients,
466 indicating that observational constraints do not fully eliminate extrapolation uncertainty.

467 As the most established framework for GPP estimation, LUE models driven by satellite LAI



468 exhibit high internal consistency and minimal intra-group divergence, representing a relatively stable
469 class within multi-product ensembles (**Fig. 10**). Uncertainty within the DGVM ensemble is spatially
470 heterogeneous, suggesting that inter-model spread in TRENDY is strongly dependent on regional
471 ecoclimatic context and model structure, and that the weak regions of individual models do not fully
472 overlap. For instance, O-CN shows markedly elevated values in parts of northern and central China
473 where intensive cultivation is prevalent, likely reflecting inadequate or idiosyncratic representations
474 of agricultural management processes (Zaehle et al., 2010). VODCA2GPP derives its signal from
475 microwave vegetation optical depth (VOD), which is physically distinct from conventional optical
476 remote-sensing products. VOD is more sensitive to vegetation water content, woody biomass, and
477 canopy moisture status than to greenness or leaf area, and may therefore capture ecological
478 information that is complementary to that of optical products (Wild et al., 2022).

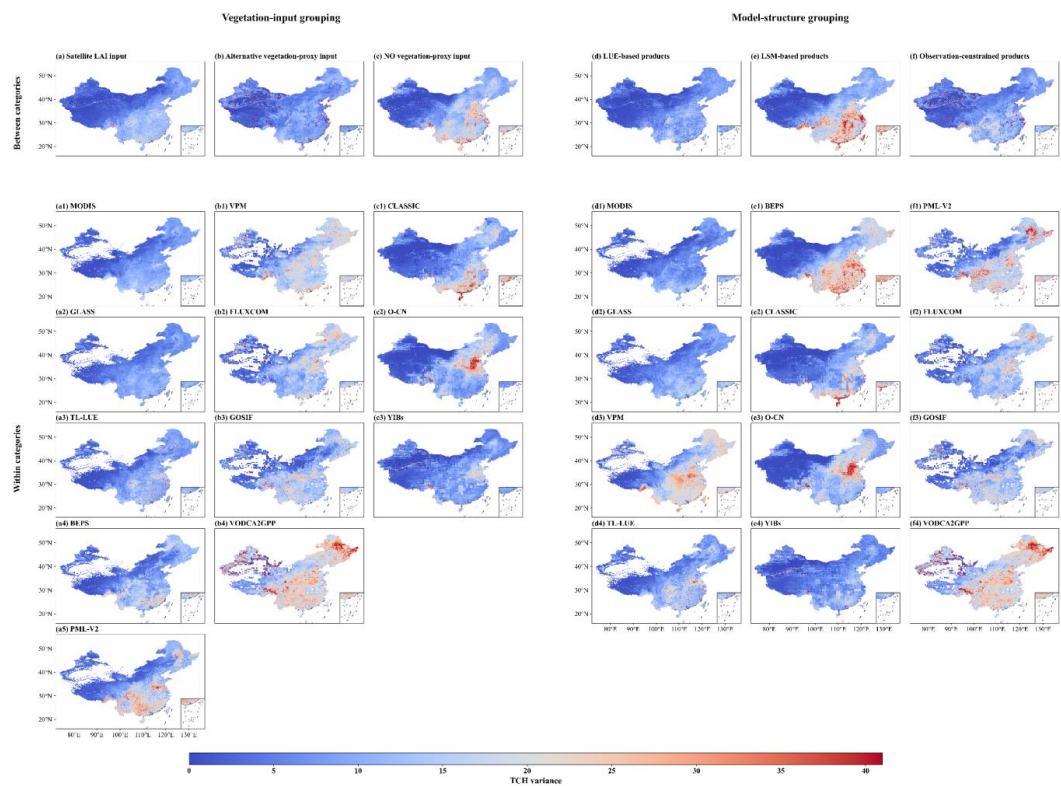


Fig. 9. Uncertainty within and between GPP product categories driven by differences in greenness input data or model structure

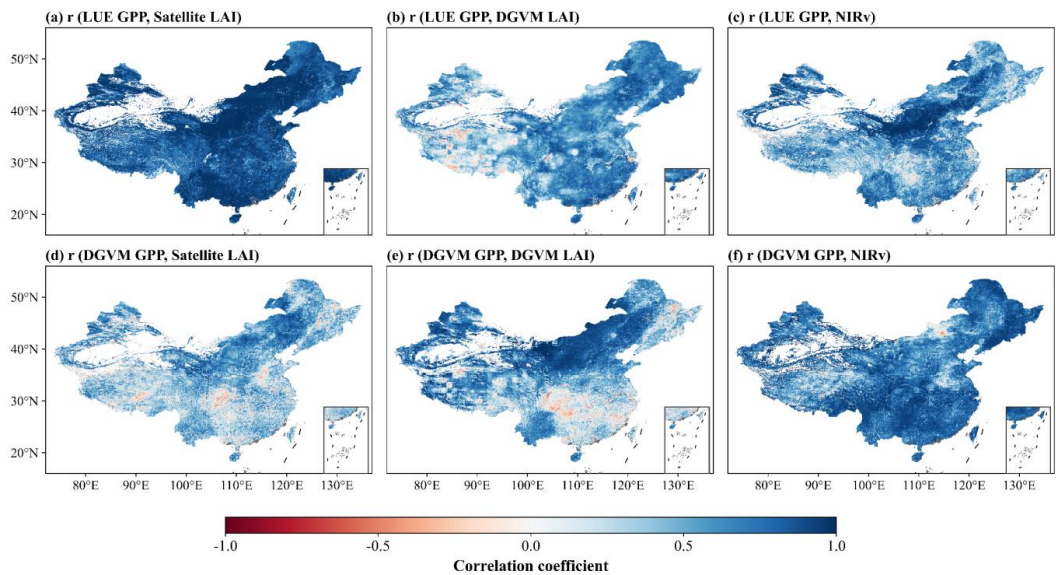


Fig. 10. Spatial coupling between GPP and vegetation proxies and its dependence on greenness input data across product categories

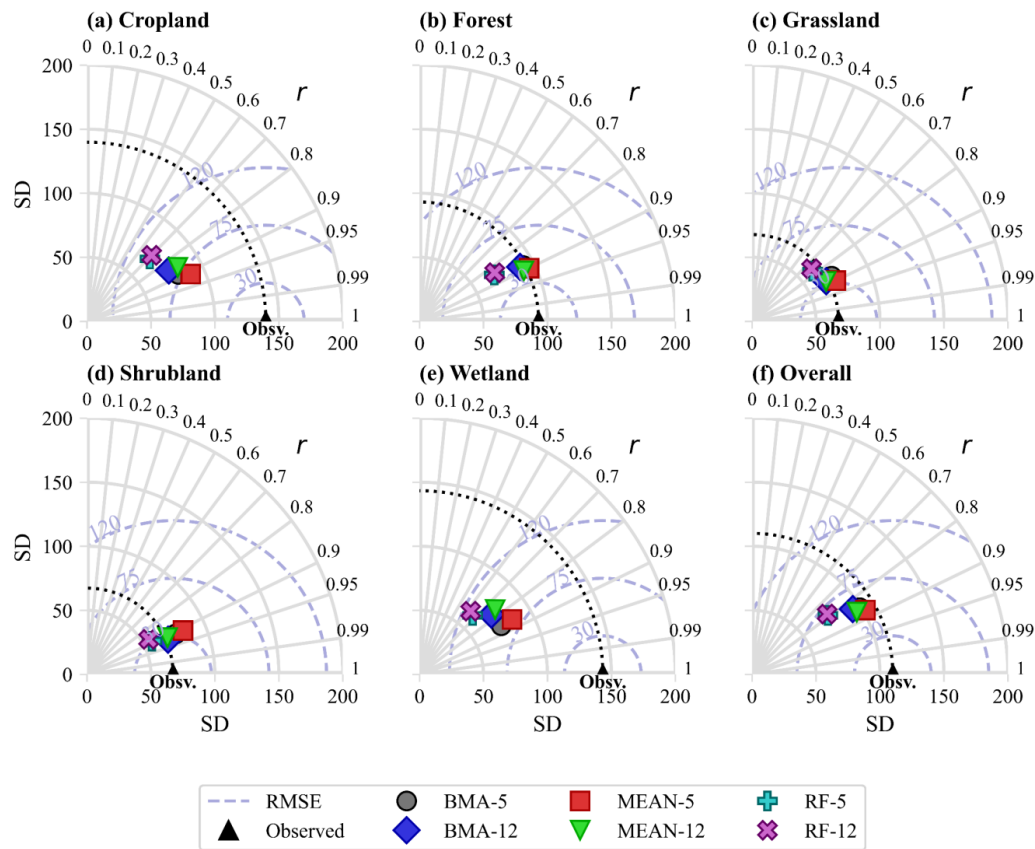
3.4 Product fusion

We employed three product fusion approaches, namely the arithmetic mean, BMA, and RF (Fig. 11). Fusion experiments were conducted using both the full set of available products and subsets of products exhibiting relatively strong consistency. Overall, the fusion based on five products (GLASS, VPM, TL-LUE, PML-V2, and GOSIF) which represent sufficiently diverse model structures and greenness related data sources, achieved the best comprehensive performance in terms of accuracy, trend characterization, and spatiotemporal correlations with greenness indicators. Among the three fusion approaches, the arithmetic mean method demonstrated particularly strong performance, whereas the RF based fusion showed clear limitations in capturing long term trends compared with the arithmetic mean and BMA approaches (Fig. 12).



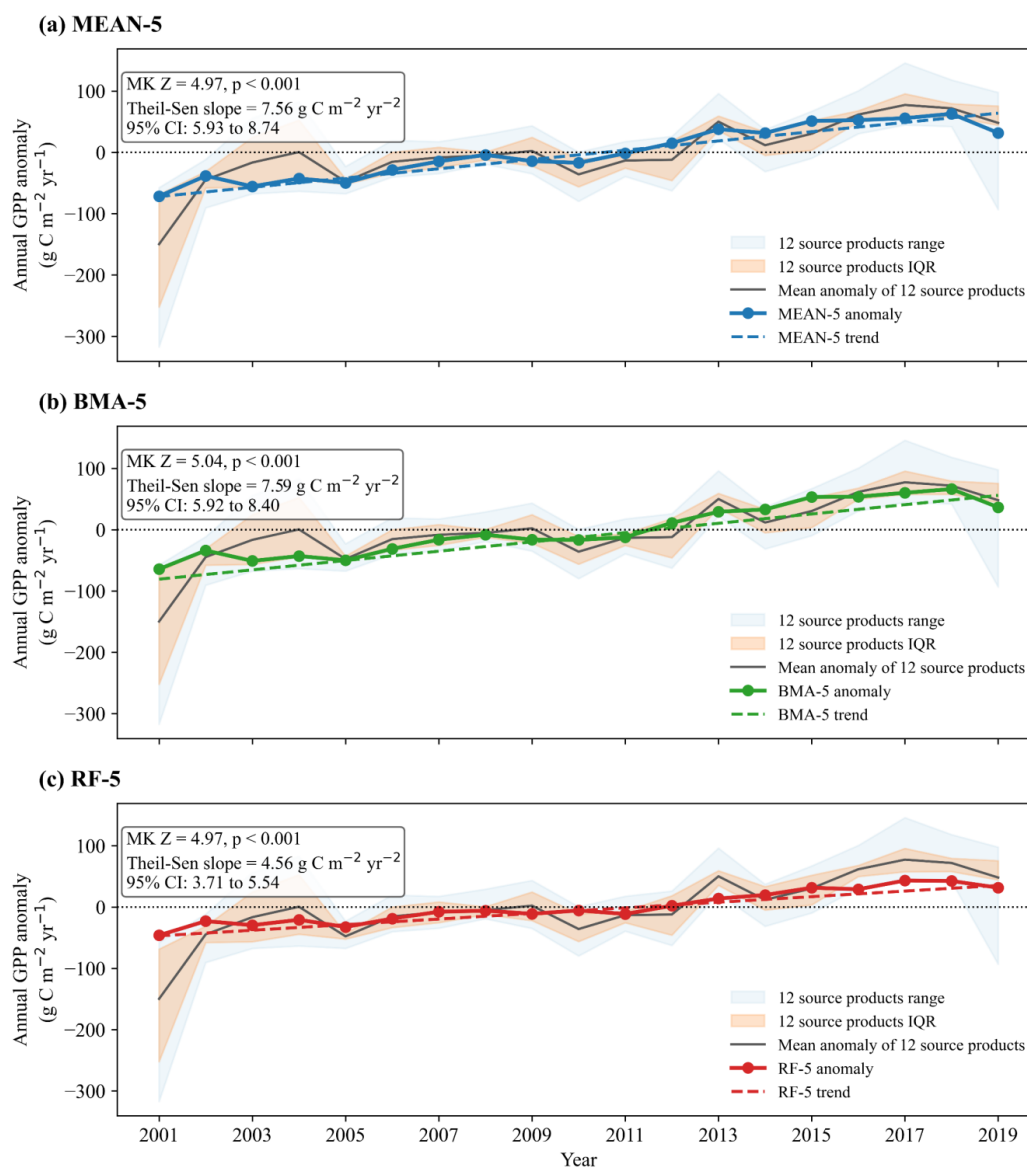
496 The best performing fusion product was MEAN-5, derived by applying the arithmetic mean
497 method to the five selected products (**Fig. 11**). At a monthly temporal resolution and a spatial
498 resolution of $0.05^\circ \times 0.05^\circ$, MEAN-5 exhibited strong agreement with independent flux tower
499 observations, with an overall correlation coefficient of 0.87 and an RMSE of $53.89 \text{ g C m}^{-2} \text{ month}^{-1}$.
500 The predicted mean GPP was $110.99 \text{ g C m}^{-2} \text{ month}^{-1}$, closely matching the flux tower observed
501 mean of $110.63 \text{ g C m}^{-2} \text{ month}^{-1}$. Its predicted standard deviation was 101.87, which was also
502 comparable to the observed standard deviation of 110.07. These results indicate that MEAN-5 not
503 only accurately represents the overall magnitude of carbon uptake, but also effectively preserves its
504 spatiotemporal variability.

505 MEAN-5 also demonstrated robust performance across different vegetation types and exhibited
506 a clear increasing trend over the study period. In addition, it showed strong correlations with both
507 LAI and NIRv, with correlation coefficients exceeding 0.9. Its estimated multi-year mean carbon
508 uptake was $7.72 \pm 2.03 \text{ Pg C yr}^{-1}$, which falls within the range of product-based estimates derived in
509 this study $6.19 \text{ Pg C yr}^{-1}$ (MODIS) to $13.99 \text{ Pg C yr}^{-1}$ (VODCA2GPP) and is close to the 12-product
510 median $7.79 \text{ Pg C yr}^{-1}$ (TL_LUE). Notably, this estimate is closer to values reported in previous
511 studies than the 12-product mean of $8.32 \text{ Pg C yr}^{-1}$ (Lin et al., 2024; Xu et al., 2025). Therefore,
512 MEAN-5 can be regarded as a stable, accurate, and trend consistent GPP fusion product, making it
513 well suited for long term trend analysis, attribution analysis, and regional carbon uptake estimation.



514

515 **Fig. 11.** Taylor diagram comparing ensemble estimates generated using arithmetic mean, BMA, and RF with flux-tower
516 observations across (a) cropland, (b) forest, (c) grassland, (d) shrubland, (e) wetland, and (f) overall performance across
517 all ecosystem types. Each ensemble estimate was derived from either 5-product or 12-product combinations. Displayed
518 metrics include r , SD, and RMSE; points closer to the observation reference indicate higher r , SD closer to the observed
519 value, and lower RMSE.



520

521 **Fig. 12.** Temporal trends in annual GPP anomalies derived from 5-product ensemble estimates using (a) arithmetic mean,
 522 (b) BMA, and (c) RF. Shaded areas indicate the range and interquartile range (IQR) of the 12 source products, the black
 523 line represents the mean anomaly of the 12 source products, and dashed lines show Theil-Sen trends for each ensemble
 524 method.

525 4. Discussion



4.1 Whether, where and why greening turns into effective carbon uptake

LAI shows widespread statistically significant greening covering 67.49% of China's land area, with the GPP ensemble indicating a slightly smaller extent at 57.68%, while significant decreases (browning) are limited to 1.50% and 1.48% for LAI and the GPP ensemble, respectively (**Fig. 13a, b**). Approximately 27.21% of greening areas showed an enhanced carbon-uptake response, defined as areas where the GPP trend exceeded the LAI greening trend. Consistent with the dominant role of CO₂ fertilization in global greening (Zhu et al., 2016), our SHAP results indicate that long-term mean climatic and atmospheric conditions, especially CO₂, precipitation and soil moisture, explain LAI-GPP growth differences better than climate trends or variability (**Fig. 14**).

Concurrent declines of LAI and GPP are rare and limited to scattered pockets of the highly urbanized eastern seaboard, notably the Yangtze River Delta (**Fig. 13d, black**). These patterns are consistent with the ecological compression caused by rapid construction-land expansion and cropland conversion (Li et al., 2020). Beyond these limited zones of co-decline, our results reveal substantial heterogeneity in the relative rates and directions of LAI and GPP change across ecological regions (**Fig. 13c**). The broad regions where LAI increases more rapidly than GPP (**Fig. 13e, blue**) largely reflect China's afforestation and land restoration efforts, such as the Three-North Shelterbelt and the Grain for Green program (**Fig. 13d**), which have markedly expanded vegetation cover across nearly half of the country's drylands since the 1980s (Li et al., 2021). Water limitations are considered the primary constraint on improvements in photosynthetic efficiency with proliferation of foliage in regions such as the Loess Plateau and the Hexi Corridor (Jiao et al., 2021;

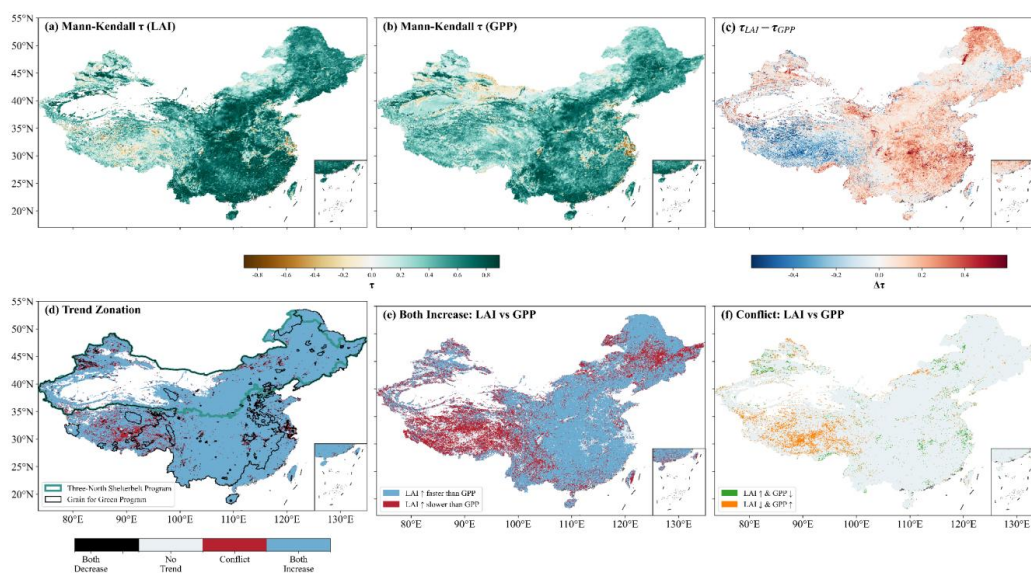


547 Lu et al., 2018), whereas radiation constraints due to canopy saturation are regarded as the dominant
 548 factor in the hilly areas of southern China (Hu et al., 2022) (**Fig. 14**). In contrast, areas where GPP
 549 increases faster than LAI (**Fig. 13e, red**) are mainly concentrated in the Tibetan Plateau and the
 550 Northeast Plain.

551 The conflict between LAI and GPP trends (**Fig. 13f**) represents a structure-function decoupling
 552 (Kannenberget al., 2022), which is an important signal of changes in the terrestrial carbon cycle.
 553 Regions where LAI increases but GPP decreases occur as scattered patches in the mountains of
 554 northeastern China and localized areas of northern and eastern Xinjiang. This pattern of shrub
 555 encroachment and low-efficiency greening highlights the challenges of implementing shelterbelt and
 556 desertification-control afforestation projects in arid regions. Regions where LAI decreases but GPP
 557 increases are mainly located in high-elevation, cold-temperate forest and grassland regions, including
 558 the Tibetan Plateau (especially its high-altitude zones) and the Tianshan Mountains. This pattern
 559 indicates that, in some high-elevation and dryland regions, a decline in vegetation greenness does not
 560 necessarily imply reduced productivity. In these climate-sensitive regions, long-term vegetation
 561 dynamics provide essential evidence for evaluating ecosystem responses to climate change (Gao et
 562 al., 2019). On the Tibetan Plateau, the widespread greening over the past several decades has been
 563 well documented (Huang et al., 2025), attributing to warming-wetting climate change (Huang et al.,
 564 2022; Maina et al., 2022). But the interactions among vegetation, elevation, and climate become
 565 increasingly complex with altitude (Pu et al., 2025). The trade-off between water limitation and
 566 warming effects contributes to the localized emergence of decoupling phenomena between
 567 vegetation greenness and productivity (Liu et al., 2019).



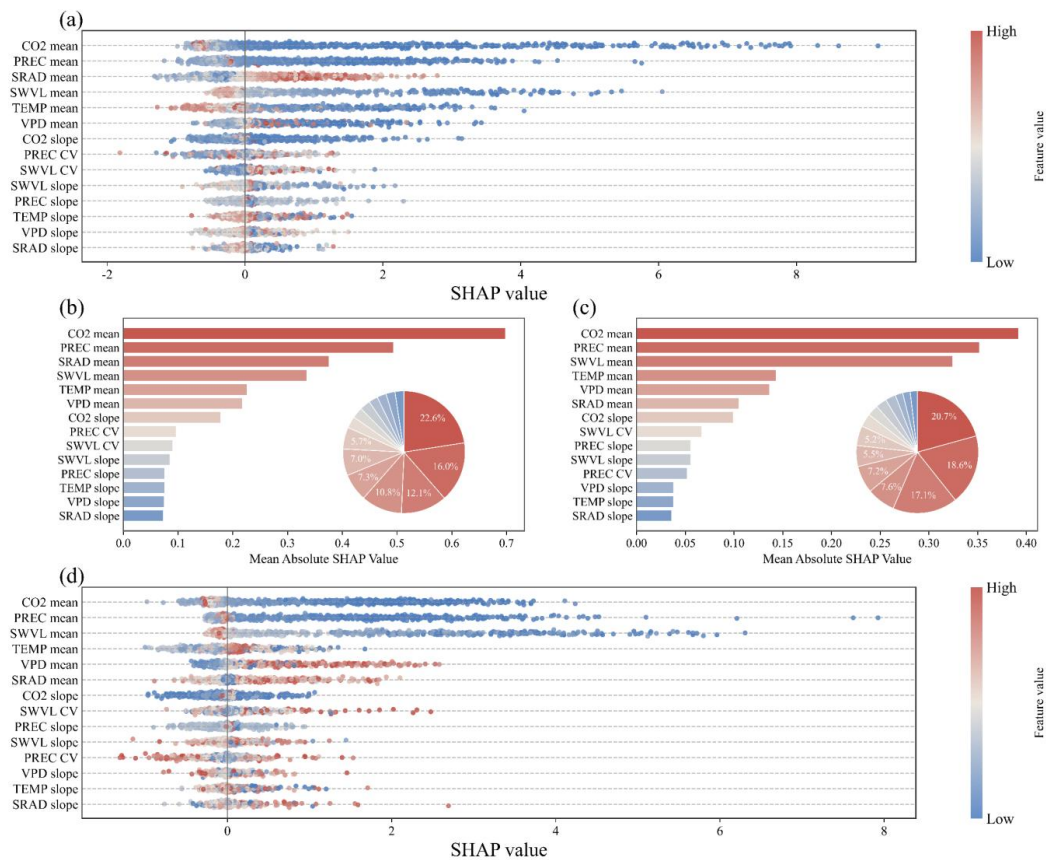
568 The differences between LAI and GPP ensemble suggest that beyond the drivers of greening,
 569 vegetation exhibits long-term transition from structural changes to functional changes. China's
 570 greening is not merely a cosmetic increase in leaf area but is functionally meaningful, leading to
 571 greater gross primary productivity and carbon fixation which can help offset anthropogenic CO₂
 572 emissions and mitigate climate warming (Li et al., 2023). China's land carbon sink has been driven
 573 largely by policy-enabled expansion of planted forests, which has also enhanced carbon sequestration
 574 and aligns with the country's "dual-carbon" targets (peaking before 2030 and neutrality thereafter)
 575 (Cheng et al., 2024).



576
 577 **Fig. 13.** Spatial patterns of long-term trend strength and cross-variable relationship across China based on satellite-
 578 based LAI ensemble and the MEAN-5 GPP ensemble. Mann-Kendall τ for LAI (a) and GPP ensemble (b) trends
 579 (positive/negative indicates increases/decreases; magnitude reflects the strength of trend evidence). (c) Difference
 580 in τ between LAI and GPP ensemble. (d) Trend Zonation of LAI and GPP ensemble trends results in four
 581 classifications: both decrease (black), no trend (light gray), conflict/opposite signs (red; detailed in panel f), and
 582 both increase (blue; detailed in panel e). Overlaid outlines denote the boundaries of the Three-North Shelterbelt
 583 Program and the Grain for Green Program. Boundaries were obtained from the National Earth System Science Data
 584 Center (NESDC), National Science & Technology Infrastructure of China (<http://www.nesdc.org.cn>) (Shao et al.,



585 2022)



586

587 **Fig. 14.** Climatic and atmospheric controls on the rate difference between LAI and GPP trends. SHAP value distributions
588 and feature importance are shown for areas where LAI increases faster than GPP (a, b) and where GPP increases faster
589 than LAI (c, d). Beeswarm plots (a, d) show the direction and magnitude of feature effects, with point color indicating
590 feature magnitude. Bar charts (b, c) rank predictors by mean absolute SHAP values, with inset pie charts showing their
591 proportional contributions.

592 4.2 Causes of model performance differences

593 The tower-product comparison highlights several problems common to nearly all GPP products.

594 The most significant problem lies in the systematic underestimation of high field-measured

595 productivity. Studies of West African tropical forests indicated that global GPP products show a



596 significant systematic underestimation of biometric GPP measurements by 56.3% on average. This
597 tropical feature is considered a general model-data discrepancies and can also present in other high-
598 productivity regions (Zhang-Zheng et al., 2024). The causes are slightly different: the LUE
599 framework is mainly limited by conservative parameter estimation rather than the model structure
600 (Bao et al., 2023; Wang et al., 2017); tower-based ML products tend to underestimate GPP mainly
601 due to the scarcity of tropical flux towers, which leads to insufficient training data (Jian et al., 2022);
602 remote sensing products systematically underestimate fAPAR due to cloud contamination issues
603 (Zhang et al., 2025). TRENDY DGVMs can capture high GPP values more prominently (Yang et al.,
604 2022; Zhang-Zheng et al., 2024), while large differences among DGVMs reflect their greater
605 structural and parameter complexity. Unlike LUE models, which share a relatively fixed framework,
606 DGVMs implement more complex process formulations. In fact, our results show that half of the
607 products systematically underestimate monthly cumulative GPP values exceeding 300 g C m^{-2}
608 month^{-1} . The underlying cause remains an overly conservative response to high GPP. Among the
609 products, GOSIF exhibits the smallest underestimation, suggesting that the SIF-based, observation-
610 constrained estimates provide a more direct link to photosynthetic activity and capture high GPP
611 values better than relying on empirical parameterizations or models trained on flux-tower data. As a
612 result, SIF-based product yields a stronger linear relationship with GPP across time and space
613 (Shekhar et al., 2022). SIF primarily reflects photosynthetic function (physiological activity) (Ruehr
614 et al., 2024), whereas VOD primarily reflects canopy structure and water status (biomass and water
615 content) (Konings et al., 2021). VODCA2GPP has higher analog values in South China, which
616 demonstrates its higher potential for addressing the underestimation of high values. However, its



617 performance compared with flux-tower observations is not outstanding, possibly due to the indirect
618 linkage between VOD and photosynthesis (Yao et al., 2024), retrieval uncertainties (Konkathi and
619 Karthikeyan, 2022), and soil influences (Li et al., 2022). This indicates the importance of multi-
620 source fusion to combine strengths and offset weaknesses, instead of a crude truncation of high
621 values.

622 Another common issue across products is that RMSE over wetlands is two to three times as
623 high as that for the other four land-cover types. Nevertheless, the correlation coefficients remain
624 acceptable ($r > 0.75$), suggesting that the model structures and their temporal dynamics are credible;
625 instead, the discrepancies likely stem from an insufficient representation of wetland-specific physical
626 processes or spatial heterogeneity (Dąbrowska-Zielińska et al., 2022). The models that perform
627 relatively well in wetlands are GOSIF and TRENDY, both of which are good at capturing the details
628 of GPP variations. Those that explicitly include wetland-specific parameterizations or more detailed
629 representations of the unique physiology and hydrology of different plant function types (e.g., water-
630 table dynamics and inundation extent, microtopography, and organic matter accumulation) (Graham
631 et al., 2022) tend to perform better, whereas models that neglect such processes perform worse.

632 Beyond these common issues, the cross-product comparison reveals some model-specific
633 anomalies in some models. FLUXCOM exhibits a distinct salt-and-pepper pattern, indicating strong
634 spatial heterogeneity and instability in the simulated trends. Such behavior highlights that
635 FLUXCOM 'RS+METEO' product should be applied with caution for forecasting and long-term
636 analysis (Nelson et al., 2024; O'Sullivan et al., 2020). O-CN shows markedly elevated GPP in parts
637 of northern and central China where intensive cultivation is prevalent. Such product-specific



anomalies, when systematically identified, also help inform which products should and should not be included in fusion ensembles. The relative performance of the three fusion methods is noteworthy: the arithmetic mean of five selected products outperformed both BMA and RF, with RF failing to reproduce the long-term GPP trend. It suggests that the small calibration sample (33 towers) over a diverse national domain makes BMA noise-sensitive and RF prone to overfitting. While MEAN-5's structurally diverse products yield compensating errors, making equal weighting near-optimal. This suggests a practical guideline that prioritizes selecting a small, structurally diverse, reliable product set and favoring simple averaging until the calibration sample is large enough for complex weighting.

4.3 Limitations and outlooks

Several limitations should be considered when interpreting these results. The ChinaFLUX validation network is relatively dense for croplands ($n = 6$) and forests ($n = 14$) but sparse for shrublands ($n = 3$) and wetlands ($n = 3$), so tower-based statistics for the latter two ecosystems are correspondingly less robust. The consistently elevated wetland RMSE likely reflects both genuine model deficiencies in representing wetland-specific hydrology and limited statistical leverage, rather than ecosystem-specific accuracy alone. The network also underrepresents the western arid regions and high-elevation ecosystems of the Tibetan Plateau, where some of the strongest LAI-GPP decoupling signals were identified, so conclusions in these regions should be regarded as data-constrained inferences.

The validation is best described as semi-independent rather than fully independent. Several products entering MEAN-5 (e.g., VPM, PML-V2, GOSIF) were originally parameterized or



659 empirically tuned using FLUXNET data, and the overlap between training and validation sites
660 cannot be fully removed. Reported skill metrics should therefore be interpreted as an upper bound on
661 truly independent performance. Relatedly, the TCH method assumes that random errors are
662 independent across products, but shared inputs (MODIS reflectance and LAI across multiple LUE
663 products, common TRENDY forcing fields, and MODIS predictors in FLUXCOM) introduce
664 correlated errors that may cause TCH to underestimate the true uncertainty of optically driven
665 products. The relative ordering of product uncertainty is likely robust, but absolute error magnitudes
666 should be treated as conservative estimates.

667 Selecting MEAN-5 from candidate subsets against the validation network raises a residual risk
668 of selection bias, which we mitigated by requiring diverse model structures and greenness inputs and
669 by reporting comparisons across multiple subsets; an out-of-sample assessment over a
670 geographically distinct flux network would further strengthen the result. The LAI ensemble
671 combines weakly correlated satellite-derived and DGVM-based products (spatial trend correlation
672 ≈ 0.196), so coupling conclusions rely primarily on the satellite ensemble, with DGVM-based LAI
673 retained as a complementary perspective rather than an independent reference.

674 Finally, the analysis emphasizes long-term means, trends, and interannual variability and does
675 not explicitly resolve the imprint of extreme events such as droughts, heatwaves, fire disturbance, or
676 large-scale land-use conversion. The meteorological forcing used in the SHAP attribution (CMFD,
677 ERA5-Land, CarbonTracker CT2022) also carries its own uncertainties, particularly over the Tibetan
678 Plateau and other data-sparse regions, which propagate into the relative-importance estimates.
679 Longer flux-tower records, additional independent towers, event-resolved analysis, and explicit



680 propagation of forcing uncertainty into the GPP ensemble would help refine the conclusions drawn
681 here.

682 **5. Conclusion**

683 This study provides a systematic evaluation of 12 widely used multi-source GPP products over
684 China and uses the resulting consistency information to build a robust GPP ensemble. Evaluated at a
685 common 0.05° monthly resolution against 33 ChinaFLUX towers during 2001–2020, all products
686 reproduce the large-scale southeast–northwest gradient and the seasonal dynamics of productivity,
687 but they diverge substantially in magnitude (mean annual GPP ranging from approximately 833 to
688 $1901 \text{ g C m}^{-2} \text{ yr}^{-1}$), in long-term trend strength, and in their treatment of high-productivity regions
689 and wetlands. About half of the products systematically underestimate monthly GPP exceeding 300 g
690 C m^{-2} , with the smallest underestimation in SIF-derived GOSIF. TCH-based uncertainty analysis
691 reveals that products sharing satellite-derived canopy structure as input show stronger inter-product
692 consistency, while DGVMs and products driven by alternative greenness proxies (e.g., VOD) display
693 larger spatial heterogeneity, particularly in eastern and southern China. Tower-level validation and
694 TCH analysis are complementary rather than redundant because they reveal different facets of
695 product reliability.

696 Three fusion approaches were tested on multiple product subsets. The arithmetic mean of five
697 products (GLASS, VPM, TL-LUE, PML-V2, and GOSIF), hereafter MEAN-5, achieved the best
698 overall agreement with flux-tower observations ($r = 0.87$; $\text{RMSE} = 53.9 \text{ g C m}^{-2} \text{ month}^{-1}$), preserved
699 both magnitude (predicted mean of 110.99 vs. observed $110.63 \text{ g C m}^{-2} \text{ month}^{-1}$) and variability
700 across ecosystems, and yielded a national multi-year mean GPP of $7.72 \pm 2.03 \text{ Pg C yr}^{-1}$, in line with



701 recent independent estimates. Notably, RF fusion underperformed simple averaging in capturing
702 long-term trends, indicating that for high-dimensional GPP fusion across heterogeneous landscapes
703 data-driven weighting does not automatically outperform parsimonious averaging when training
704 samples are limited and trend signals are subtle.

705 Joint analysis of the MEAN-5 GPP ensemble and the satellite-based LAI ensemble shows that
706 57.68% of China's land area exhibits a statistically significant increase in GPP, broadly aligned with
707 the 67.49% area showing significant LAI greening. In most regions, structural and functional change
708 moved in the same direction, supporting the view that China's greening is functionally meaningful.
709 Where the two diverged, LAI tended to increase faster than GPP across dryland restoration regions
710 (e.g., Loess Plateau, Hexi Corridor) where water availability limits photosynthetic gains, while GPP
711 increased faster than LAI on the Tibetan Plateau and in Northeast China, consistent with warming-
712 driven productivity enhancement. SHAP-based attribution identifies long-term mean CO₂,
713 precipitation, and soil moisture as the dominant controls on this LAI-GPP decoupling. These results
714 support continued investment in afforestation and ecological-restoration programs as effective
715 contributors to China's "dual-carbon" targets, while underscoring that carbon-sink effectiveness
716 depends on regional water and energy constraints and should guide spatially differentiated planting
717 and management strategies.

718 **CRedit authorship contribution statement**

719 Bingxiao Liu: Conceptualization, Methodology, Data curation, Formal analysis, Software,
720 Writing – original draft, Visualization, Writing – review & editing. Wen Wang: Conceptualization,
721 Supervision, Writing – review & editing, Project administration, Funding acquisition. Wei Wu: Data



722 curation, Formal analysis, Software, Visualization, Writing – review & editing. Yujie Liu: Data

723 curation, Software, Visualization, Writing – review & editing.

724 **Declaration of competing interest**

725 The authors declare that they have no known competing financial interests or personal

726 relationships that could have appeared to influence the work reported in this paper.

727 **Data availability**

728 The GPP fusion product generated in this study has been deposited publicly available at

729 <https://doi.org/10.5281/zenodo.21827542>. The original GPP products and ChinaFLUX observations

730 were obtained from publicly available sources listed in Table 1 and Table S1.

731 **Acknowledgements**

732 This work was supported by the National Natural Science Foundation of China (No.42471027).

733

734



735 Reference

- 736 Badgley, G., Field, C. B., and Berry, J. A.: Canopy near-infrared reflectance and terrestrial photosynthesis,
 737 Sci. Adv., 3, e1602244, <https://doi.org/10.1126/sciadv.1602244>, 2017.
- 738 Bao, S., Alonso, L., Wang, S., Gensheimer, J., De, R., and Carvalhais, N.: Toward Robust Parameterizations in
 739 Ecosystem-Level Photosynthesis Models, J. Adv. Model. Earth Syst., 15, e2022MS003464,
 740 <https://doi.org/10.1029/2022MS003464>, 2023.
- 741 Bi, W., He, W., Zhou, Y., Ju, W., Liu, Y., Liu, Y., Zhang, X., Wei, X., and Cheng, N.: A global 0.05° dataset for
 742 gross primary production of sunlit and shaded vegetation canopies from 1992 to 2020, Sci. Data, 9, 213,
 743 <https://doi.org/10.1038/s41597-022-01309-2>, 2022.
- 744 Breiman, L.: Random Forests, Machine Learning, 45, 5–32, <https://doi.org/10.1023/A:1010933404324>, 2001.
- 745 Cao, D., Huang, X., Liu, G., Tian, L., Xin, Q., and Yang, Y.: Evaluating the Performance of Satellite-Derived
 746 Vegetation Indices in Gross Primary Productivity (GPP) Estimation at 30 m and 500 m Spatial Resolution,
 747 Remote Sensing, 17, 3291, <https://doi.org/10.3390/rs17193291>, 2025.
- 748 Cao, S., Li, M., Zhu, Z., Wang, Z., Zha, J., Zhao, W., Duanmu, Z., Chen, J., Zheng, Y., Chen, Y., Myneni, R.
 749 B., and Piao, S.: Spatiotemporally consistent global dataset of the GIMMS leaf area index (GIMMS LAI4g)
 750 from 1982 to 2020, Earth Syst. Sci. Data, 15, 4877–4899, <https://doi.org/10.5194/essd-15-4877-2023>, 2023.
- 751 Cappelli, F., Castronuovo, G., Grimaldi, S., and Telesca, V.: Random Forest and Feature Importance Measures
 752 for Discriminating the Most Influential Environmental Factors in Predicting Cardiovascular and Respiratory
 753 Diseases, IJERPH, 21, 867, <https://doi.org/10.3390/ijerph21070867>, 2024.
- 754 Chen, C., Park, T., Wang, X., Piao, S., Xu, B., Chaturvedi, R. K., Fuchs, R., Brovkin, V., Ciais, P., Fensholt,
 755 R., Tømmervik, H., Bala, G., Zhu, Z., Nemani, R. R., and Myneni, R. B.: China and India lead in greening of
 756 the world through land-use management, Nat. Sustain., 2, 122–129, [https://doi.org/10.1038/s41893-019-0220-](https://doi.org/10.1038/s41893-019-0220-7)
 757 7, 2019a.
- 758 Chen, J. M., Ju, W., Ciais, P., Viovy, N., Liu, R., Liu, Y., and Lu, X.: Vegetation structural change since 1981
 759 significantly enhanced the terrestrial carbon sink, Nat. Commun., 10, 4259, [https://doi.org/10.1038/s41467-](https://doi.org/10.1038/s41467-019-12257-8)
 760 019-12257-8, 2019b.
- 761 Chen, X., Cai, A., Guo, R., Liang, C., and Li, Y.: Variation of gross primary productivity dominated by leaf
 762 area index in significantly greening area, J. Geogr. Sci., 33, 1747–1764, [https://doi.org/10.1007/s11442-023-](https://doi.org/10.1007/s11442-023-2151-5)
 763 2151-5, 2023.
- 764 Chen, Y., Feng, X., Tian, H., Wu, X., Gao, Z., Feng, Y., Piao, S., Lv, N., Pan, N., and Fu, B.: Accelerated
 765 increase in vegetation carbon sequestration in China after 2010: A turning point resulting from climate and
 766 human interaction, Global Change Biology, 27, 5848–5864, <https://doi.org/10.1111/gcb.15854>, 2021.
- 767 Cheng, K., Yang, H., Tao, S., Su, Y., Guan, H., Ren, Y., Hu, T., Li, W., Xu, G., Chen, M., Lu, X., Yang, Z.,



- 768 Tang, Y., Ma, K., Fang, J., and Guo, Q.: Carbon storage through China's planted forest expansion, *Nat.*
 769 *Commun.*, 15, 4106, <https://doi.org/10.1038/s41467-024-48546-0>, 2024.
- 770 Dąbrowska-Zielińska, K., Misiura, K., Malińska, A., Gurdak, R., Grzybowski, P., Bartold, M., and Kluczek,
 771 M.: Spatiotemporal estimation of gross primary production for terrestrial wetlands using satellite and field
 772 data, *Remote Sens. Appl.: Soc. Environ.*, 27, 100786, <https://doi.org/10.1016/j.rsase.2022.100786>, 2022.
- 773 Fang, H., Baret, F., Plummer, S., and Schaepman-Strub, G.: An Overview of Global Leaf Area Index (LAI):
 774 Methods, Products, Validation, and Applications, *Reviews of Geophysics*, 57, 739–799,
 775 <https://doi.org/10.1029/2018RG000608>, 2019.
- 776 Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P.,
 777 Le Quéré, C., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G.,
 778 Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Barbero, L., Bates, N. R., Becker, M., Bellouin, N.,
 779 Decharme, B., Bopp, L., Brasika, I. B. M., Cadule, P., Chamberlain, M. A., Chandra, N., Chau, T.-T.-T.,
 780 Chevallier, F., Chini, L. P., Cronin, M., Dou, X., Enyo, K., Evans, W., Falk, S., Feely, R. A., Feng, L., Ford, D.
 781 J., Gasser, T., Ghattas, J., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Hefner, M.,
 782 Heinke, J., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jacobson, A. R., Jain, A., Jarníková, T., Jersild,
 783 A., Jiang, F., Jin, Z., Joos, F., Kato, E., Keeling, R. F., Kennedy, D., Klein Goldewijk, K., Knauer, J.,
 784 Korsbakken, J. I., Körtzinger, A., Lan, X., Lefèvre, N., Li, H., Liu, J., Liu, Z., Ma, L., Marland, G., Mayot, N.,
 785 McGuire, P. C., McKinley, G. A., Meyer, G., Morgan, E. J., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O'Brien,
 786 K. M., Olsen, A., Omar, A. M., Ono, T., Paulsen, M., Pierrot, D., Pocock, K., Poulter, B., Powis, C. M.,
 787 Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T. M., Schwinger, J., Séférian, R.,
 788 Smallman, T. L., Smith, S. M., Sospedra-Alfonso, R., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tans, P.
 789 P., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., van Ooijen, E., Wanninkhof, R.,
 790 Watanabe, M., Wimart-Rousseau, C., Yang, D., Yang, X., Yuan, W., Yue, X., Zaehle, S., Zeng, J., and Zheng,
 791 B.: Global Carbon Budget 2023, *Earth System Science Data*, 15, 5301–5369, [https://doi.org/10.5194/essd-15-](https://doi.org/10.5194/essd-15-5301-2023)
 792 5301-2023, 2023.
- 793 Gao, M., Piao, S., Chen, A., Yang, H., Liu, Q., Fu, Y. H., and Janssens, I. A.: Divergent changes in the
 794 elevational gradient of vegetation activities over the last 30 years, *Nat. Commun.*, 10, 2970,
 795 <https://doi.org/10.1038/s41467-019-11035-w>, 2019.
- 796 Graham, J. D., Ricciuto, D. M., Glenn, N. F., and Hanson, P. J.: Incorporating Microtopography in a Land
 797 Surface Model and Quantifying the Effect on the Carbon Cycle, *J. Adv. Model. Earth Syst.*, 14,
 798 e2021MS002721, <https://doi.org/10.1029/2021MS002721>, 2022.
- 799 Guan, X., Li, Y., Chen, J. M., Ma, Y., and Shen, H.: A process model-guided transfer learning framework for
 800 mapping global gross primary production, *Agricultural and Forest Meteorology*, 372, 110678,
 801 <https://doi.org/10.1016/j.agrformet.2025.110678>, 2025.
- 802 He, J., Yang, K., Tang, W., Lu, H., Qin, J., Chen, Y., and Li, X.: The first high-resolution meteorological
 803 forcing dataset for land process studies over China, *Sci Data*, 7, 25, [https://doi.org/10.1038/s41597-020-0369-](https://doi.org/10.1038/s41597-020-0369-y)
 804 y, 2020.



- 805 He, P., Ma, X., Han, Z., Meng, X., and Sun, Z.: Uncertainties of gross primary productivity of Chinese
 806 grasslands based on multi-source estimation, *Front. Environ. Sci.*, 10,
 807 <https://doi.org/10.3389/fenvs.2022.928351>, 2022.
- 808 He, Z., Yang, Y., Fang, R., Zhou, S., Zhao, W., Bai, Y., Li, J., and Wang, B.: Integration of shapley additive
 809 explanations with random forest model for quantitative precipitation estimation of mesoscale convective
 810 systems, *Front. Environ. Sci.*, 10, 1057081, <https://doi.org/10.3389/fenvs.2022.1057081>, 2023.
- 811 Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T.: Bayesian model averaging: a tutorial (with
 812 comments by M. Clyde, David Draper and E. I. George, and a rejoinder by the authors, *Statist. Sci.*, 14,
 813 <https://doi.org/10.1214/ss/1009212519>, 1999.
- 814 Hu, Z., Piao, S., Knapp, A. K., Wang, X., Peng, S., Yuan, W., Running, S., Mao, J., Shi, X., Ciais, P.,
 815 Huntzinger, D. N., Yang, J., and Yu, G.: Decoupling of greenness and gross primary productivity as aridity
 816 decreases, *Remote Sens. Environ.*, 279, 113120, <https://doi.org/10.1016/j.rse.2022.113120>, 2022.
- 817 Huang, R., Ma, C., Zhang, W., Liu, X., Wu, S., Pang, H., and Hou, S.: Significant greening in the western
 818 Tibetan Plateau and surroundings occurred after the 1970s, *Commun. Earth Environ.*, 6, 82,
 819 <https://doi.org/10.1038/s43247-025-02061-2>, 2025.
- 820 Huang, Y., Xin, Z., Dor-ji, T., and Wang, Y.: Tibetan Plateau greening driven by warming-wetting climate
 821 change and ecological restoration in the 21st century, *Land Degrad Dev*, 33, 2407–2422,
 822 <https://doi.org/10.1002/ldr.4319>, 2022.
- 823 Jacobson, A. R., Schuldt, K. N., Tans, P., Arlyn Andrews, Miller, J. B., Oda, T., Mund, J., Weir, B., Ott, L.,
 824 Aalto, T., Abshire, J. B., Aikin, K., Aoki, S., Apadula, F., Arnold, S., Baier, B., Bartyzel, J., Beyersdorf, A.,
 825 Biermann, T., Biraud, S. C., Boenisch, H., Brailsford, G., Brand, W. A., Chen, G., Huilin Chen, Lukasz
 826 Chmura, Clark, S., Colomb, A., Commane, R., Conil, S., Couret, C., Cox, A., Cristofanelli, P., Cuevas, E.,
 827 Curcoll, R., Daube, B., Davis, K. J., De Wekker, S., Coletta, J. D., Delmotte, M., DiGangi, E., DiGangi, J. P.,
 828 Di Sarra, A. G., Dlugokencky, E., Elkins, J. W., Emmenegger, L., Shuangxi Fang, Fischer, M. L., Forster, G.,
 829 Frumau, A., Galkowski, M., Gatti, L. V., Gehrlein, T., Gerbig, C., Francois Gheusi, Gloor, E., Gomez-Trueba,
 830 V., Goto, D., Griffis, T., Hammer, S., Hanson, C., Haszpra, L., Hatakka, J., Heimann, M., Heliasz, M., Hensen,
 831 A., Hermansen, O., Hintsa, E., Holst, J., Ivakhov, V., Jaffe, D. A., Jordan, A., Joubert, W., Karion, A., Kawa,
 832 S. R., Kazan, V., Keeling, R. F., Keronen, P., Kneuer, T., Kolari, P., Kateřina Komínková, Kort, E., Kozlova,
 833 E., Krummel, P., Kubistin, D., Labuschagne, C., Lam, D. H. Y., Lan, X., Langenfelds, R. L., Laurent, O.,
 834 Laurila, T., Lauvaux, T., Lavric, J., Law, B. E., Lee, J., Lee, O. S. M., Lehner, I., Lehtinen, K., Leppert, R.,
 835 Leskinen, A., Leuenberger, M., Levin, I., Levula, J., Lin, J., Lindauer, M., Loh, Z., Lopez, M., Luijkx, I. T.,
 836 Lunder, C. R., Machida, T., Mammarella, I., Manca, G., Manning, A., Manning, A., Marek, M. V., Martin, M.
 837 Y., Matsueda, H., McKain, K., Meijer, H., Meinhardt, F., Merchant, L., N. Mihalopoulos, Miles, N. L., Miller,
 838 C. E., Mitchell, L., Mölder, M., Montzka, S., Moore, F., Moossen, H., Morgan, E., Josep-Anton Morgui,
 839 Morimoto, S., Müller-Williams, J., J. William Munger, Munro, D., Myhre, C. L., Shin-Ichiro Nakaoka,
 840 Jaroslaw Necki, Newman, S., Nichol, S., Niwa, Y., Obersteiner, F., O'Doherty, S., Paplawsky, B., Peischl, J.,
 841 Peltola, O., Piacentino, S., Jean-Marc Pichon, Pickers, P., Piper, S., Pitt, J., Plass-Dülmer, C., Platt, S. M.,
 842 Prinzivalli, S., Ramonet, M., Ramos, R., Reyes-Sanchez, E., Richardson, S. J., Riris, H., Rivas, P. P., Ryerson,



- 843 T., Saito, K., Sargent, M., Sasakawa, M., Scheeren, B., Schuck, T., Schumacher, M., Seifert, T., Sha, M. K.,
 844 Shepson, P., Shook, M., Sloop, C. D., Smith, P., Stanley, K., Steinbacher, M., Stephens, B., Sweeney, C.,
 845 Thoning, K., Timas, H., Torn, M., Tørseth, K., Trisolino, P., Turnbull, J., Van Den Bulk, P., Van Dinter, D.,
 846 Vermeulen, A., Viner, B., Vitkova, G., Walker, S., Watson, A., Wofsy, S. C., Worsey, J., Worthy, D., Dickon
 847 Young, Zaehle, S., Zahn, A., and Mirosław Zimnoch: CarbonTracker CT2022, [https://doi.org/10.25925/Z1GJ-](https://doi.org/10.25925/Z1GJ-3254)
 848 3254, 2023., 2023.
- 849 Jian, J., Bailey, V., Dorheim, K., Konings, A. G., Hao, D., Shiklomanov, A. N., Snyder, A., Steele, M.,
 850 Teramoto, M., Vargas, R., and Bond-Lamberty, B.: Historically inconsistent productivity and respiration
 851 fluxes in the global terrestrial carbon cycle, *Nat. Commun.*, 13, 1733, [https://doi.org/10.1038/s41467-022-](https://doi.org/10.1038/s41467-022-29391-5)
 852 29391-5, 2022.
- 853 Jiao, W., Wang, L., Smith, W. K., Chang, Q., Wang, H., and D'Odorico, P.: Observed increasing water
 854 constraint on vegetation growth over the last three decades, *Nat. Commun.*, 12, 3777,
 855 <https://doi.org/10.1038/s41467-021-24016-9>, 2021.
- 856 Ju, W. and Zhou, Y.: Global Daily GPP Simulation Data Products from 1981 to 2019,
 857 <https://doi.org/10.12199/nesdc.ecodb.2016YFA0600200.02.001>, 2021.
- 858 Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., Anthoni, P., Besnard, S.,
 859 Bodesheim, P., Carvalhais, N., Chevallier, F., Gans, F., Goll, D. S., Haverd, V., Köhler, P., Ichii, K., Jain, A.
 860 K., Liu, J., Lombardozzi, D., Nabel, J. E. M. S., Nelson, J. A., O'Sullivan, M., Pallandt, M., Papale, D.,
 861 Peters, W., Pongratz, J., Rödenbeck, C., Sitch, S., Tramontana, G., Walker, A., Weber, U., and Reichstein, M.:
 862 Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM
 863 approach, *Biogeosciences*, 17, 1343–1365, <https://doi.org/10.5194/bg-17-1343-2020>, 2020.
- 864 Kannenberg, S. A., Cabon, A., Babst, F., Belmecheri, S., Delpierre, N., Guerrieri, R., Maxwell, J. T., Meinzer,
 865 F. C., Moore, D. J. P., Pappas, C., Ueyama, M., Ulrich, D. E. M., Voelker, S. L., Woodruff, D. R., and
 866 Anderegg, W. R. L.: Drought-induced decoupling between carbon uptake and tree growth impacts forest
 867 carbon turnover time, *Agric. For. Meteorol.*, 322, 108996, <https://doi.org/10.1016/j.agrformet.2022.108996>,
 868 2022.
- 869 Konings, A. G., Holtzman, N. M., Rao, K., Xu, L., and Saatchi, S. S.: Interannual Variations of Vegetation
 870 Optical Depth are Due to Both Water Stress and Biomass Changes, *Geophys. Res. Lett.*, 48, e2021GL095267,
 871 <https://doi.org/10.1029/2021GL095267>, 2021.
- 872 Konkathi, P. and Karthikeyan, L.: Error and uncertainty characterization of soil moisture and VOD retrievals
 873 obtained from L-band SMAP radiometer, *Remote Sens. Environ.*, 280, 113146,
 874 <https://doi.org/10.1016/j.rse.2022.113146>, 2022.
- 875 Li, C., Fu, B., Wang, S., Stringer, L. C., Wang, Y., Li, Z., Liu, Y., and Zhou, W.: Drivers and impacts of
 876 changes in China's drylands, *Nat. Rev. Earth Environ.*, 2, 858–873, [https://doi.org/10.1038/s43017-021-](https://doi.org/10.1038/s43017-021-00226-z)
 877 00226-z, 2021.
- 878 Li, D., Wu, S., Liang, Z., and Li, S.: The impacts of urbanization and climate change on urban vegetation



- 879 dynamics in China, *Urban For. Urban Green.*, 54, 126764, <https://doi.org/10.1016/j.ufug.2020.126764>, 2020.
- 880 Li, X. and Xiao, J.: A Global, 0.05-Degree Product of Solar-Induced Chlorophyll Fluorescence Derived from
 881 OCO-2, MODIS, and Reanalysis Data, *Remote Sens.*, 11, 517, <https://doi.org/10.3390/rs11050517>, 2019.
- 882 Li, X., Wigneron, J.-P., Fan, L., Frappart, F., Yueh, S. H., Colliander, A., Ebtehaj, A., Gao, L., Fernandez-
 883 Moran, R., Liu, X., Wang, M., Ma, H., Moisy, C., and Ciais, P.: A new SMAP soil moisture and vegetation
 884 optical depth product (SMAP-IB): Algorithm, assessment and inter-comparison, *Remote Sens. Environ.*, 271,
 885 112921, <https://doi.org/10.1016/j.rse.2022.112921>, 2022.
- 886 Li, Y., Li, Z.-L., Wu, H., Zhou, C., Liu, X., Leng, P., Yang, P., Wu, W., Tang, R., Shang, G.-F., and Ma, L.:
 887 Biophysical impacts of earth greening can substantially mitigate regional land surface temperature warming,
 888 *Nat. Commun.*, 14, 121, <https://doi.org/10.1038/s41467-023-35799-4>, 2023.
- 889 Liang, S., Cheng, J., Jia, K., Jiang, B., Liu, Q., Xiao, Z., Yao, Y., Yuan, W., Zhang, X., Zhao, X., and Zhou, J.:
 890 The Global Land Surface Satellite (GLASS) Product Suite, *Bulletin of the American Meteorological Society*,
 891 102, E323–E337, <https://doi.org/10.1175/BAMS-D-18-0341.1>, 2021.
- 892 Liao, Z., Zhou, B., Zhu, J., Jia, H., and Fei, X.: A critical review of methods, principles and progress for
 893 estimating the gross primary productivity of terrestrial ecosystems, *Front. Environ. Sci.*, 11,
 894 <https://doi.org/10.3389/fenvs.2023.1093095>, 2023.
- 895 Lin, S., Huang, X., Wang, C., He, T., Zhang, X., Shen, R., Peng, Q., Chen, X., Zheng, Y., Dong, J., Liang, S.,
 896 and Yuan, W.: A 30-m gross primary production dataset from 2016 to 2020 in China, *Sci. Data*, 11, 1065,
 897 <https://doi.org/10.1038/s41597-024-03893-x>, 2024.
- 898 Liu, L., Wang, Y., Wang, Z., Li, D., Zhang, Y., Qin, D., and Li, S.: Elevation-dependent decline in vegetation
 899 greening rate driven by increasing dryness based on three satellite NDVI datasets on the Tibetan Plateau, *Ecol.*
 900 *Indic.*, 107, 105569, <https://doi.org/10.1016/j.ecolind.2019.105569>, 2019.
- 901 Lu, C., Zhao, T., Shi, X., and Cao, S.: Ecological restoration by afforestation may increase groundwater depth
 902 and create potentially large ecological and water opportunity costs in arid and semiarid China, *J. Clean. Prod.*,
 903 176, 1213–1222, <https://doi.org/10.1016/j.jclepro.2016.03.046>, 2018.
- 904 Maina, F. Z., Kumar, S. V., Albergel, C., and Mahanama, S. P.: Warming, increase in precipitation, and
 905 irrigation enhance greening in High Mountain Asia, *Commun Earth Environ*, 3, 43,
 906 <https://doi.org/10.1038/s43247-022-00374-0>, 2022.
- 907 McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., and Stoffelen, A.: Extended triple
 908 collocation: Estimating errors and correlation coefficients with respect to an unknown target, *Geophys. Res.*
 909 *Lett.*, 41, 6229–6236, <https://doi.org/10.1002/2014GL061322>, 2014.
- 910 Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S.,
 911 Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J.,
 912 Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for



- land applications, *Earth Syst. Sci. Data*, 13, 4349–4383, <https://doi.org/10.5194/essd-13-4349-2021>, 2021.
- Nelson, J. A., Walther, S., Gans, F., Kraft, B., Weber, U., Novick, K., Buchmann, N., Migliavacca, M., Wohlfahrt, G., Šigut, L., Ibrom, A., Papale, D., Göckede, M., Duveiller, G., Knohl, A., Hörtnagl, L., Scott, R. L., Dušek, J., Zhang, W., Hamdi, Z. M., Reichstein, M., Aranda-Barranco, S., Ardö, J., Op de Beeck, M., Billesbach, D., Bowling, D., Bracho, R., Brümmer, C., Camps-Valls, G., Chen, S., Cleverly, J. R., Desai, A., Dong, G., El-Madany, T. S., Euskirchen, E. S., Feigenwinter, I., Galvagno, M., Gerosa, G. A., Gielen, B., Godec, I., Goslee, S., Gough, C. M., Heinesch, B., Ichii, K., Jackowicz-Korczynski, M. A., Klosterhalfen, A., Knox, S., Kobayashi, H., Kohonen, K.-M., Korkiakoski, M., Mammarella, I., Gharun, M., Marzuoli, R., Matamala, R., Metzger, S., Montagnani, L., Nicolini, G., O’Halloran, T., Ourcival, J.-M., Peichl, M., Pendall, E., Ruiz Reverter, B., Roland, M., Sabbatini, S., Sachs, T., Schmidt, M., Schwalm, C. R., Shekhar, A., Silberstein, R., Silveira, M. L., Spano, D., Tagesson, T., Tramontana, G., Trotta, C., Turco, F., Vesala, T., Vincke, C., Vitale, D., Vivoni, E. R., Wang, Y., Woodgate, W., Yopez, E. A., Zhang, J., Zona, D., and Jung, M.: X-BASE: the first terrestrial carbon and water flux products from an extended data-driven scaling framework, *FLUXCOM-X, Biogeosciences*, 21, 5079–5115, <https://doi.org/10.5194/bg-21-5079-2024>, 2024.
- O’Sullivan, M., Smith, W. K., Sitch, S., Friedlingstein, P., Arora, V. K., Haverd, V., Jain, A. K., Kato, E., Kautz, M., Lombardozzi, D., Nabel, J. E. M. S., Tian, H., Vuichard, N., Wiltshire, A., Zhu, D., and Buermann, W.: Climate-Driven Variability and Trends in Plant Productivity Over Recent Decades Based on Three Global Products, *Global Biogeochemical Cycles*, 34, e2020GB006613, <https://doi.org/10.1029/2020GB006613>, 2020.
- Piao, S., Yin, G., Tan, J., Cheng, L., Huang, M., Li, Y., Liu, R., Mao, J., Myneni, R. B., Peng, S., Poulter, B., Shi, X., Xiao, Z., Zeng, N., Zeng, Z., and Wang, Y.: Detection and attribution of vegetation greening trend in China over the last 30 years, *Global Change Biology*, 21, 1601–1609, <https://doi.org/10.1111/gcb.12795>, 2015.
- Pu, G., Han, L., Chen, L., Wan, D., and Teng, H.: Elevational dynamics of vegetation changes in response to climate change on the Tibetan plateau, *Sci. Rep.*, 15, 9813, <https://doi.org/10.1038/s41598-025-94896-0>, 2025.
- Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M.: Using Bayesian Model Averaging to Calibrate Forecast Ensembles, *Monthly Weather Review*, 133, 1155–1174, <https://doi.org/10.1175/MWR2906.1>, 2005.
- Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T., Havránková, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D., and Valentini, R.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm, *Global Change Biology*, 11, 1424–1439, <https://doi.org/10.1111/j.1365-2486.2005.001002.x>, 2005.
- Ruehr, S., Gerlein-Safdi, C., Falco, N., Seibert, P. O., Chou, C., Albert, L., and Keenan, T. F.: Quantifying Seasonal and Diurnal Cycles of Solar-Induced Fluorescence With a Novel Hyperspectral Imager, *Geophys. Res. Lett.*, 51, e2023GL107429, <https://doi.org/10.1029/2023GL107429>, 2024.



- 951 Running, S., Mu, Q., and Zhao, M.: MODIS/Terra Gross Primary Productivity (MOD17A2H) 8-Day L4
952 Global 500 m, Version 6.1 (Version 6.1), <https://doi.org/10.5067/MODIS/MOD17A2H.061>, 2021.
- 953 Shao, Q., Liu, S., Ning, J., Liu, G., Yang, F., Zhang, X., Niu, L., Huang, H., Fan, J., and Liu, J.: Assessment of
954 ecological benefits of key national ecological projects in China in 2000–2019 using remote sensing [in
955 Chinese], *Acta Geogr. Sin.*, 77, 2133–2153, <https://doi.org/10.11821/dlxb202209001>, 2022.
- 956 Shekhar, A., Buchmann, N., and Gharun, M.: How well do recently reconstructed solar-induced fluorescence
957 datasets model gross primary productivity?, *Remote Sens. Environ.*, 283, 113282,
958 <https://doi.org/10.1016/j.rse.2022.113282>, 2022.
- 959 Sitch, S., O’Sullivan, M., Robertson, E., Friedlingstein, P., Albergel, C., Anthoni, P., Arneth, A., Arora, V. K.,
960 Bastos, A., Bastrikov, V., Bellouin, N., Canadell, J. G., Chini, L., Ciais, P., Falk, S., Harris, I., Hurtt, G., Ito,
961 A., Jain, A. K., Jones, M. W., Joos, F., Kato, E., Kennedy, D., Klein Goldewijk, K., Kluzek, E., Knauer, J.,
962 Lawrence, P. J., Lombardozzi, D., Melton, J. R., Nabel, J. E. M. S., Pan, N., Peylin, P., Pongratz, J., Poulter,
963 B., Rosan, T. M., Sun, Q., Tian, H., Walker, A. P., Weber, U., Yuan, W., Yue, X., and Zaehle, S.: Trends and
964 Drivers of Terrestrial Sources and Sinks of Carbon Dioxide: An Overview of the TRENDY Project, *Glob.*
965 *Biogeochem. Cycles*, 38, e2024GB008102, <https://doi.org/10.1029/2024GB008102>, 2024.
- 966 Sjoberg, J. P., Anthes, R. A., and Rieckh, T.: The Three-Cornered Hat Method for Estimating Error Variances
967 of Three or More Atmospheric Datasets. Part I: Overview and Evaluation, *Journal of Atmospheric and*
968 *Oceanic Technology*, 38, 555–572, <https://doi.org/10.1175/JTECH-D-19-0217.1>, 2021.
- 969 Stoffelen, A.: Toward the true near-surface wind speed: Error modeling and calibration using triple
970 collocation, *J. Geophys. Res.: Oceans*, 103, 7755–7766, <https://doi.org/10.1029/97JC03180>, 1998.
- 971 Sun, Y., Peng, D., Guan, X., Chu, D., Ma, Y., and Shen, H.: Impacts of the data quality of remote sensing
972 vegetation index on gross primary productivity estimation, *GIScience & Remote Sensing*, 60, 2275421,
973 <https://doi.org/10.1080/15481603.2023.2275421>, 2023.
- 974 Tian, Z., Yi, C., Fu, Y., Kutter, E., Krakauer, N. Y., Fang, W., Zhang, Q., and Luo, H.: Fusion of Multiple
975 Models for Improving Gross Primary Production Estimation With Eddy Covariance Data Based on Machine
976 Learning, *Journal of Geophysical Research: Biogeosciences*, 128, e2022JG007122,
977 <https://doi.org/10.1029/2022JG007122>, 2023.
- 978 Van Jaarsveld, B., Hauswirth, S. M., and Wanders, N.: Machine learning and global vegetation: random
979 forests for downscaling and gap filling, *Hydrol. Earth Syst. Sci.*, 28, 2357–2374, [https://doi.org/10.5194/hess-](https://doi.org/10.5194/hess-28-2357-2024)
980 [28-2357-2024](https://doi.org/10.5194/hess-28-2357-2024), 2024.
- 981 Verma, M., Friedl, M. A., Richardson, A. D., Kiely, G., Cescatti, A., Law, B. E., Wohlfahrt, G., Gielen, B.,
982 Roupsard, O., Moors, E. J., Toscano, P., Vaccari, F. P., Gianelle, D., Bohrer, G., Varlagin, A., Buchmann, N.,
983 Van Gorsel, E., Montagnani, L., and Propastin, P.: Remote sensing of annual terrestrial gross primary
984 productivity from MODIS: an assessment using the FLUXNET La Thuile data set, *Biogeosciences*, 11, 2185–
985 2200, <https://doi.org/10.5194/bg-11-2185-2014>, 2014.



- 986 Wang, H., Yakir, D., and Rotenberg, E.: Assessing the effectiveness of a central flux tower in representing the
 987 spatial variations in gross primary productivity in a semi-arid pine forest, *Agricultural and Forest*
 988 *Meteorology*, 333, 109415, <https://doi.org/10.1016/j.agrformet.2023.109415>, 2023.
- 989 Wang, L., Zhu, H., Lin, A., Zou, L., Qin, W., and Du, Q.: Evaluation of the Latest MODIS GPP Products
 990 across Multiple Biomes Using Global Eddy Covariance Flux Data, *Remote Sens.*, 9, 418,
 991 <https://doi.org/10.3390/rs9050418>, 2017.
- 992 Wild, B., Teubner, I., Moesinger, L., Zotta, R.-M., Forkel, M., van der Schalie, R., Sitch, S., and Dorigo, W.:
 993 VODCA2GPP – a new, global, long-term (1988–2020) gross primary production dataset from microwave
 994 remote sensing, *Earth Syst. Sci. Data*, 14, 1063–1085, <https://doi.org/10.5194/essd-14-1063-2022>, 2022.
- 995 Xiao, J., Baldocchi, D., Ichii, K., Li, F., and Papale, D.: Insights into terrestrial carbon and water cycling from
 996 the global eddy covariance network, *Nat Rev Earth Environ*, <https://doi.org/10.1038/s43017-025-00743-1>,
 997 2025.
- 998 Xiao, Z., Liang, S., Wang, J., Xiang, Y., Zhao, X., and Song, J.: Long-Time-Series Global Land Surface
 999 Satellite Leaf Area Index Product Derived From MODIS and AVHRR Surface Reflectance, *IEEE Trans.*
 1000 *Geosci. Remote Sensing*, 54, 5301–5318, <https://doi.org/10.1109/TGRS.2016.2560522>, 2016.
- 1001 Xie, X., Li, A., Jin, H., Yin, G., and Bian, J.: Spatial Downscaling of Gross Primary Productivity Using
 1002 Topographic and Vegetation Heterogeneity Information: A Case Study in the Gongga Mountain Region of
 1003 China, *Remote Sensing*, 10, 647, <https://doi.org/10.3390/rs10040647>, 2018.
- 1004 Xie, X., Li, A., Jin, H., Tan, J., Wang, C., Lei, G., Zhang, Z., Bian, J., and Nan, X.: Assessment of five
 1005 satellite-derived LAI datasets for GPP estimations through ecosystem models, *Science of The Total*
 1006 *Environment*, 690, 1120–1130, <https://doi.org/10.1016/j.scitotenv.2019.06.516>, 2019.
- 1007 Xu, J., Chen, T., Chen, X., Zhou, S., Gu, Z., Li, W., Cui, Y., Wang, S., and Liu, S.: Estimating GPP in China
 1008 using different site-level datasets, vegetation classification and vegetation indices, *Ecol Process*, 14, 65,
 1009 <https://doi.org/10.1186/s13717-025-00617-w>, 2025.
- 1010 Yang, R., Wang, J., Zeng, N., Sitch, S., Tang, W., McGrath, M. J., Cai, Q., Liu, D., Lombardozzi, D., Tian, H.,
 1011 Jain, A. K., and Han, P.: Divergent historical GPP trends among state-of-the-art multi-model simulations and
 1012 satellite-based products, *Earth Syst. Dyn.*, 13, 833–849, <https://doi.org/10.5194/esd-13-833-2022>, 2022.
- 1013 Yao, Y., Humphrey, V., Konings, A. G., Wang, Y., Yin, Y., Holtzman, N., Wood, J. D., Bar-On, Y., and
 1014 Frankenberg, C.: Investigating Diurnal and Seasonal Cycles of Vegetation Optical Depth Retrieved From
 1015 GNSS Signals in a Broadleaf Forest, *Geophys. Res. Lett.*, 51, e2023GL107121,
 1016 <https://doi.org/10.1029/2023GL107121>, 2024.
- 1017 Yu, G., Wen, X., Sun, X., Bertrand D. Tanner, Lee, X., and Chen, J.: Overview of ChinaFLUX and evaluation
 1018 of its eddy covariance measurement, *Agric. For. Meteorol.*, 137, 125–137,
 1019 <https://doi.org/10.1016/j.agrformet.2006.02.011>, 2006.



- 1020 Zaehle, S., Friend, A. D., Friedlingstein, P., Dentener, F., Peylin, P., and Schulz, M.: Carbon and nitrogen cycle
1021 dynamics in the O-CN land surface model: 2. Role of the nitrogen cycle in the historical terrestrial carbon
1022 balance, *Glob. Biogeochem. Cycles*, 24, 2009GB003522, <https://doi.org/10.1029/2009GB003522>, 2010.
- 1023 Zhan, Y., Deng, Y., Wang, J., Liu, Q., and Wang, W.: Environmental response characteristics of the carbon and
1024 water fluxes above complex urban surfaces of a subtropical megacity in China, *Phys. Chem. Earth*, 135,
1025 103681, <https://doi.org/10.1016/j.pce.2024.103681>, 2024.
- 1026 Zhang, Y. and Ye, A.: Improving global gross primary productivity estimation by fusing multi-source data
1027 products, *Heliyon*, 8, e09153, <https://doi.org/10.1016/j.heliyon.2022.e09153>, 2022.
- 1028 Zhang, Y., Peña-Arancibia, J. L., McVicar, T. R., Chiew, F. H. S., Vaze, J., Liu, C., Lu, X., Zheng, H., Wang,
1029 Y., Liu, Y. Y., Miralles, D. G., and Pan, M.: Multi-decadal trends in global terrestrial evapotranspiration and its
1030 components, *Sci. Rep.*, 6, 19124, <https://doi.org/10.1038/srep19124>, 2016.
- 1031 Zhang, Y., Xiao, X., Wu, X., Zhou, S., Zhang, G., Qin, Y., and Dong, J.: A global moderate resolution dataset
1032 of gross primary production of vegetation for 2000-2016,
1033 <https://doi.org/10.6084/M9.FIGSHARE.C.3789814.V1>, 2017., 2017.
- 1034 Zhang, Y., Keenan, T. F., and Zhou, S.: Exacerbated drought impacts on global ecosystems due to structural
1035 overshoot, *Nat Ecol Evol*, 5, 1490–1498, <https://doi.org/10.1038/s41559-021-01551-8>, 2021.
- 1036 Zhang, Y., Fang, H., Hu, Z., Wang, Y., Li, S., and Wu, G.: Validation of Global Moderate-Resolution FAPAR
1037 Products over Boreal Forests in North America Using Harmonized Landsat and Sentinel-2 Data, *Remote*
1038 *Sens.*, 17, 2658, <https://doi.org/10.3390/rs17152658>, 2025.
- 1039 Zhang-Zheng, H., Deng, X., Aguirre-Gutiérrez, J., Stocker, B. D., Thomson, E., Ding, R., Adu-Bredu, S.,
1040 Duah-Gyamfi, A., Gvozdevaite, A., Moore, S., Oliveras Menor, I., Prentice, I. C., and Malhi, Y.: Why models
1041 underestimate West African tropical forest primary productivity, *Nat. Commun.*, 15, 53949,
1042 <https://doi.org/10.1038/s41467-024-53949-0>, 2024.
- 1043 Zhao, M., Heinsch, F. A., Nemani, R. R., and Running, S. W.: Improvements of the MODIS terrestrial gross
1044 and net primary production global data set, *Remote Sensing of Environment*, 95, 164–176,
1045 <https://doi.org/10.1016/j.rse.2004.12.011>, 2005.
- 1046 Zheng, Y., Shen, R., Wang, Y., Li, X., Liu, S., Liang, S., Chen, J. M., Ju, W., Zhang, L., and Yuan, W.:
1047 Improved estimate of global gross primary production for reproducing its long-term variation, 1982–2017,
1048 *Earth Syst. Sci. Data*, 12, 2725–2746, <https://doi.org/10.5194/essd-12-2725-2020>, 2020.
- 1049 Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., Ciais, P., Sitch, S., Friedlingstein, P.,
1050 Arneeth, A., Cao, C., Cheng, L., Kato, E., Koven, C., Li, Y., Lian, X., Liu, Y., Liu, R., Mao, J., Pan, Y., Peng,
1051 S., Peñuelas, J., Poulter, B., Pugh, T. A. M., Stocker, B. D., Viovy, N., Wang, X., Wang, Y., Xiao, Z., Yang, H.,
1052 Zaehle, S., and Zeng, N.: Greening of the Earth and its drivers, *Nat. Clim. Chang*, 6, 791–795,
1053 <https://doi.org/10.1038/nclimate3004>, 2016.



1054 Zuo, L., Zou, L., Xia, J., Zhang, L., Cao, H., and She, D.: Multi-scale analysis of six evapotranspiration
1055 products across China: Accuracy, uncertainty and spatiotemporal pattern, J. Hydrol., 650, 132516,
1056 <https://doi.org/10.1016/j.jhydrol.2024.132516>, 2025.

1057