https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

2001-2022 global gross primary productivity dataset using an ensemble model based on random forest

Xin Chen[1], Tiexi Chen[1,2,3*], Xiaodong Li[4], Yuanfang Chai[5], Shengjie Zhou[1], Renjie Guo[6], Jie Dai[1]

[1]School of Geographical Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China.
[2]Qinghai Provincial Key Laboratory of Plateau Climate Change and Corresponding Ecological and Environmental Effects, Qinghai University of Science and Technology, Xining 810016, China
[3]School of Geographical Sciences, Qinghai Normal University, Xining 810008, Qinghai, China.
[4]Qinghai Institute of Meteorological Science, Xining 810008, Qinghai, China.
[5]Department of Earth Sciences, Vrije Universiteit Amsterdam, Boelelaan 1085, 1081 HV, Amsterdam, the Netherlands
[6]Faculty of Geographical Science, Beijing Normal University, Beijing, China.

Correspondence to: Tiexi Chen (txchen@nuist.edu.cn)

Abstract. The continuous advancement of remote sensing technology has been instrumental in improving models for estimating terrestrial gross primary productivity (GPP). However, challenges arise from inconsistent spatial distributions and interannual variations in GPP datasets, impeding our comprehensive understanding of the entire terrestrial carbon cycle. In contrast to previous models relying on remote sensing and environmental variables, we developed a an ensemble model based on random forest named GPP_{ERF}. This model utilized the GPP outputs from established remote sensing-based models (EC-LUE, GPP-kNDVI, GPP-NIRv, Revised-EC-LUE) as inputs for GPP estimations. GPP_{ERF} demonstrated significant effectiveness by explaining 83.7% of the monthly variation in GPP across 171 sites. This performance surpassed the selected remote sensing models (72.4%-77.1%) and an independent random forest model using remote sensing and environmental variables (77.7%). Over the period from 2001 to 2022, the global estimated GPP value using the ensemble model based on random forest was 131.2 PgC yr^{-1}, exhibiting a trend of 0.45 PgC yr^{-2}. Furthermore, evaluation results utilizing flux sites from ChinaFlux indicated that the dateset exhibited good generalization. In summary, the machine learning-based ensemble method helps to reduce the uncertainty in the estimation of a single remote sensing model and provides a more reliable estimation of global GPP.

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

1 Introduction

Gross primary productivity (GPP) is the largest carbon flux in the global carbon cycle, and it is also the input of carbon during the carbon cycle. Uncertainties in the estimation of GPP will be further propagated to other carbon flux estimates, so it is important to clarify the spatio-temporal pattern of GPP (Xiao et al., 2019; Ruehr et al., 2023). However, different studies estimate global GPP to be between 90 PgC yr^{-1} and 160 PgC yr^{-1}, and this uncertainty may be more pronounced when extended to regional scales or specific ecosystem types, so it is necessary to develop some new methods to reduce the uncertainty of GPP estimates (Jung et al., 2019; Ryu et al., 2019; Anav et al., 2015).

Currently, there are several remote sensing data-driven methods to estimate GPP, including light use efficiency (LUE) models, vegetation index models, machine learning models, and process models (Sun et al., 2019; Mengistu et al., 2021). Direct validation of flux towers from FLUXNET shows that these models usually only explain about 70% of the monthly variation in GPP (Wang et al., 2021b; Badgley et al., 2019). One possible reason is that remote sensing models cannot fully characterize all the processes of photosynthesis. This is understandable, most of the existing models use linear or nonlinear mathematical formulas to express a certain process of photosynthesis. However, the ecosystem is highly complex, the bias introduced by such a numerical model in a process will increase the uncertainty in the final product (GPP) estimates. For example, in the LUE model, the difference in the meteorological constraints alone can lead to a difference of more than 10% in the explanatory power of the model (Yuan et al., 2014). As an important factor affecting photosynthesis, some models consider the effect of CO_2 fertilization. However, a study revealed that the effect of CO_2 fertilization showed a significant negative trend in the past 40 years, and this process may be missing in the model (Wang et al., 2020). Limited by the imperfection of the model mechanism, adjusting the model parameters is the most effective way to improve the simulation accuracy. The usual practice of the modeler is to divide the directly observed GPP data according to different vegetation types, and randomly select the testset through the cross-validation method to calibrate and validate the model parameters. However, this method is based on the assumption that the model parameters of the same vegetation type in different regions are roughly the same. In fact, the photosynthetic characteristics of the same vegetation type are also quite different in different regions. A typical example is the difference between C3 and C4 crops in the cropland, the GPP of C4 crops during the growing season may reach 600-800 gC m^{-2} $month^{-1}$, accounting for more than 60% of the annual GPP, in contrast, the GPP of C3 crops in the growing season is only 200-300 m^{-2} $month^{-1}$, or even lower (Chen et al., 2014). Some other studies have also found that the maximum carboxylation rate (Vcmax) that determines photosynthesis at the leaf scale not only varies with vegetation types, but also depends on environmental factors (Wang et al., 2021a). The same vegetation type also has a difference of 40umol m^{-2} s^{-1} in different geographical areas (Groenendijk et al., 2011), all of which may lead to uncertainties in GPP estimate. A widespread problem is that the deviation of model structure and model parameters may lead to poor estimation of GPP in the monthly extreme value, and the phenomenon of "high value underestimation and low value overestimate" occurs. Especially for extremely high values, which usually occur during the growing season and largely

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

determine the annual value and inter-annual variation of GPP, this underestimation may hinder our understanding of the entire carbon cycle process.

It is difficult for a single model to have good estimation in all regions of the globe. Previous studies have shown that ensemble model may perform better than a single model, which may improve some potential problems in model estimation (Chen et al., 2020; Yao et al., 2014). Traditional multi-model ensemble methods usually use multi-model simple average or bayesian weighted average. However, these methods usually only provide fixed weights for each model, and are essentially linear combinations between multiple models. Some recent studies apply machine learning methods to multi-model ensembles to establish nonlinear relationships between multiple simulated target variables and real target variables, improving the simulation performance (Bai et al., 2021; Yao et al., 2017; Tian et al., 2023). However, few studies have applied this method to the global GPP estimation, which provides a new idea for improving some common problems of a single remote sensing model (such as high value underestimation and ground value overestimation).

In this study, we attempt to use an ensemble model based on machine learning methods to improve the estimation of global GPP. Specifically, the work of this study includes the following points: (1) After re-calibrating the parameters of each model, the performance of five remote sensing models and the ensemble models was compared; (2) Focusing on the phenomenon of "high value underestimation and low value overestimation" in each model, and compared the performance of each model in each months, each vegetation types and different subvalues (high value, median value, low value); (3) Developing a global GPP dataset using an ensemble model based on machine learning methods, and using GPP observations from ChinaFlux as a complementary validation set to test the generalization of this dataset, i.e. the extent to which the dataset capture changes in GPP in regions where fewer sites are used in the modeling process.

2 Method

2.1 Data at the global scale

In this study, we selected remote sensing data from Moderate Resolution Imaging Spectroradiometer (MODIS) and meteorological data from EAR5-Land to estimate global GPP. For remote sensing data, surface reflectance, leaf area Index (LAI) and Fraction of Photosynthetically Active Radiation (FPAR) were used in this study. For meteorological data, we selected air temperature, dew point temperature, total solar radiation, and direct solar radiation. The dew point temperature and air temperature were used to calculate the saturated vapor pressure difference (VPD) (Yuan et al., 2019), and the diffuse solar radiation was calculated as the difference between the total solar radiation and the direct solar radiation. The CO_2 comes from the monthly average carbon dioxide levels measured by the Mauna Loa Observatory in Hawaii. Table 1 shows the details of these data.

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

Table 1. Overview of the datasets used in this study.

Variable	Dataset	Spatial resolution	Temporal resolution	Temporal coverage
Surface reflectance	MCD43C4	0.05°	daily	2001-2022
LAI	MOD15AH	500m	8d	2001-2022
FPAR	MOD15AH	500m	8d	2001-2022
Air temperature (AT)	ERA5	0.1°	Monthly	2001-2022
Dew point temperature (DPT)	ERA5	0.1°	Monthly	2001-2022
Total solar radiation (TSR)	ERA5	0.25°	Monthly	2001-2022
Direct solar radiation (DirSR)	ERA5	0.25°	Monthly	2001-2022
CO_2	NOAA's Earth System Research Laboratory	/	Monthly	2001-2022
Distribution map of C4 crops	Harvested Area and Yield for 175 Crops	1/12°	Annual	2000
Land use	MCD12C1	0.05°	Annual	2010

Previous studies have shown that the photosynthetic capacity of C4 crops is much higher than that of C3 crops (Chen et al., 2014; Chen et al., 2011), so it is necessary to divide the cropland into C3 crops and C4 crops. When estimating the global GPP, we used the "175 Crop harvested Area and yield" dataset, which describes the global harvested area and yield of 175 crops in 2000 (Monfreda et al., 2008). We extracted the sum of the area ratios of all C4 crops (corn, corn feed, sorghum, sorghum feed, sugarcane, millet) at each grid point as the coverage of C4 crops (Figure S1). Therefore, the estimated value of cropland GPP can be expressed as: coverage of C3 crops × GPP simulated value of C3 crops + coverage of C4 crops × GPP simulated value of C4 crops, which has been used in a previous study (Guo et al., 2023).

The land use map comes from the IGBP classification of MCD12Q1, and 2010 was selected as the reference year. In order to meet the need of subsequent research, the land cover types were combined into 9 categories: deciduous Broadleaf Forest (DBF), evergreen coniferous forest (ENF), Evergreen Broadleaf Forest (EBF), Mixed Forest (MF), Grassland (GRA), Cropland (including CRO-C3 and CRO-C4), Savannah (SAV), Shrub (SHR), Wetland (WET). Ultimately, all data were resampled to a spatial resolution of 0.05°, while data from MODIS were aggregated to a monthly scale to meet spatiotemporal consistency.

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

2.2 Observation data at the site scale

The modeling used GPP observations from the FLUXNET 2015 dataset, which includes carbon fluxes and meteorological variables from more than 200 flux sites around the world (Pastorello et al., 2020). GPP cannot be obtained directly from the flux site and usually needs to be obtained by dismantling the Net Ecosystem Exchange. We chose a month-scale GPP based on the nighttime partitioning method and retained only high quality data (NEE_VUT_REF_QC > 0.8) for every year, and finally selected 171 sites with 10824 monthly values for this study. In addition, temperature, radiation and VPD on the monthly scale were selected. Since part of the data required by the model is not directly available at the flux site, surface reflectance, LAI and FPAR on a scale of 500m were extracted, which are roughly similar to the footprint of the flux site and can represent the land surface of the site situation (Chu et al., 2021).

2.3 Remote sensing models and ensemble models for estimating GPP

In this study, five independent remote sensing models were selected to estimate GPP. The five models are EC-LUE, Revised-EC-LUE, NIRv-based linear model, kNDVI-based linear model and traditional random forest model. EC-LUE is a LUE model driven by remote sensing data. The model assumes that GPP is proportional to the photosynthetically active radiation absorbed by the canopy, and the seasonal variation of GPP is corrected by meteorological constraints (Yuan et al., 2007); Recently, Zheng et al. revised the EC-LUE model and proposed the Revised-EC-LUE model, which divides the canopy into sunlit and shaded leaves, and considers long-term changes in CO_2 to improve the estimation of global GPP (Zheng et al., 2020). NIRv and kNDVI are newly proposed vegetation indices, which are calculated from the red and near-infrared bands of the reflectance spectrum (Badgley et al., 2017; Camps-Valls et al., 2021). Similar to SIF, they exhibit a linear relationship with GPP and are considered to be effective proxies for GPP. Detailed descriptions of all models are presented in Text S1. Random forest method has been widely used in GPP estimation, which usually uses meteorological variables and vegetation index for modeling (Jung et al., 2019). In this study, we used air temperature, VPD, radiation and LAI to estimate GPP.

To reduce the uncertainty in estimating GPP from a single model, we also used a multi-modal ensemble method, the basic idea of which is to re-model the simulated values of multiple models. In this study, ensemble model based on the random forest method was used. Unlike traditional machine learning methods, we directly used random forests to establish the relationship between the GPP simulated by the above four models and the GPP observations. A summary of all models used is shown in Table 2.

Table 2. Overview of the models used in this study.

ID	Model	Input data	Output
1	EC-LUE	FPAR, VPD, AT, SRAD, CO_2	GPP_{EC}
2	Revised-EC-LUE	LAI, VPD, AT, DifSR, DirSR, CO_2	GPP_{REC}
3	kNDVI-GPP	Red band and near infrared band	GPP_{kNDVI}

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

4	NIRv-GPP	Red band and near infrared band	GPP_{NIRv}
5	Traditional random forest model	AT, VPD, SRAD, LAI	GPP_{RF}
6	Ensemble model based on random forest	GPP_{EC}, GPP_{REC}, GPP_{kNDVI}, GPP_{NIRv}	GPP_{ERF}

2.4 Model parameter calibration and Validation

Due to the difference between meteorological data and vegetation data, we did not use default parameters in the model, but carried out parameter calibration and model validation for all remote sensing models according to different vegetation types. For EC-LUE and Revised EC-LUE, the Markov chain Monte Carlo method (MCMC) was used to calibrate the model parameters. The traditional MCMC method usually takes the mean value of the posterior distribution of parameters as the optimal value, while previous studies have shown that some model parameters cannot be well constrained when calibrating multiple model parameters (Xu et al., 2006; Wang et al., 2017), so we use the parameter with the smallest root-mean-square error (RMSE) as the optimal parameter in each iteration. For each vegetation type, we randomly selected 70% of the sites for parameter calibration, and the process was repeated 200 times. In order to avoid overfitting, we took the mean of the 200 calibrated parameters as the final model parameters. Similarly, for the two vegetation index models, we randomly selected 70% of the sites in each vegetation type for parameter calibration. The process was repeated 200 times, and the mean of the 200 calibrated parameters was used as the final model parameters.

After obtaining GPP estimates from four remote sensing models, we tested the simulation performance of traditional random forest model and random forest-based ensemble model respectively. For both models, we tested model performance using 5-fold cross-validation, the process was repeated 200 times, and the mean of the GPP estimated 200 times as the final GPP estimate. Goodness of Fit (R^2) and RMSE were used to measure the simulation performance of all models. In addition, we used the ratio of GPP simulations to GPP observations (Sim/Obs) to measure whether the model was overestimated or underestimated.

2.5 Evaluation of the generalization of different GPP datasets

Most of the flux sites in Fluxnet2015 are concentrated in Europe and North America, it is not clear whether the different GPP estimation methods are suitable for some regions with sparse flux sites. Recently, ChinaFlux published GPP observations from multiple sites, providing an opportunity to test the generalization of different GPP datasets. However, the spatial resolution of most GPP datasets is 0.05°, and direct comparison with GPP observations at flux sites is challenging. Therefore, we extracted 0.05° MODIS land use covering the flux tower, and when the type of vegetation observed by the flux tower was consistent with the MODIS land use, the site was used for analysis. Finally, a total of 12 flux sites were selected (Figure S2), and Table S1 shows the information of these sites.

Based on site-scale models, we estimated the global GPP for 2001-2022 using an ensemble model based on random forest (ERF_GPP). We tested the generalization of ERF_GPP on 12 ChinaFlux sites. In addition, we selected a number of widely

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

used GPP datasets for comparison, including BESS (Li et al., 2023), GOSIF (Li and Xiao, 2019), ECGC (Guo et al., 2023), NIRv (Wang et al., 2021b), Revise-EC-LUE (Zheng et al., 2020), which are generated using different GPP estimation methods. These GPP products all have a spatial resolution of 0.05°, avoiding the uncertainty of GPP validation introduced due to resolution differences. The common time range for these products is 2001-2018, and the time resolution was unified to monthly to be consistent with GPP observations.

3 Result

3.1 Performance of six models at site scale

Table S2-S5 shows the optimization results of four remote sensing model parameters. Similar to the previous study, in the Revised EC-LUE model, the light use efficiency parameter of shade leaves was significantly higher than that of the sunlit leaves (Yuan et al., 2019; Zheng et al., 2020). It is necessary to divide the cropland into C3 crops and C4 crops. In all models, the parameters of C4 crops were significantly higher than those of C3 crops, which was especially reflected in the two vegetation index models of GPP_{kNDVI} and GPP_{NIRv}, the slope of the linear regression was a direct reflection of the difference in photosynthetic capacity of different crops.

Figure 1 shows the performance of all models on vegetation types. Overall, the performance of the ensemble model was better than that of the remote sensing model. GPP_{ERF} always had the highest accuracy among all models, with R^2 between 0.6-0.9 and RMSE between 0.8-3 gC m^{-2} d^{-1}. In contrast, the performance of the two vegetation index models was relatively poor, especially for evergreen forests, the R^2 of GPP_{kNDVI} and GPP_{NIRv} was significantly lower than other models. It is worth noting that compared to other vegetation types, the RMSE of cropland was the higher, with 5 out of 6 models in C4 Crop exceeding 3 gC m^{-2} d^{-1}, which suggested that these existing GPP models may not properly track seasonal changes in cropland. No significant estimation bias in vegetation type was found in four remote sensing model with calibration parameters and the ensemble model. However, GPP_{RF} was significantly underestimated in C4 crops and significantly overestimated in SHR and WET.

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

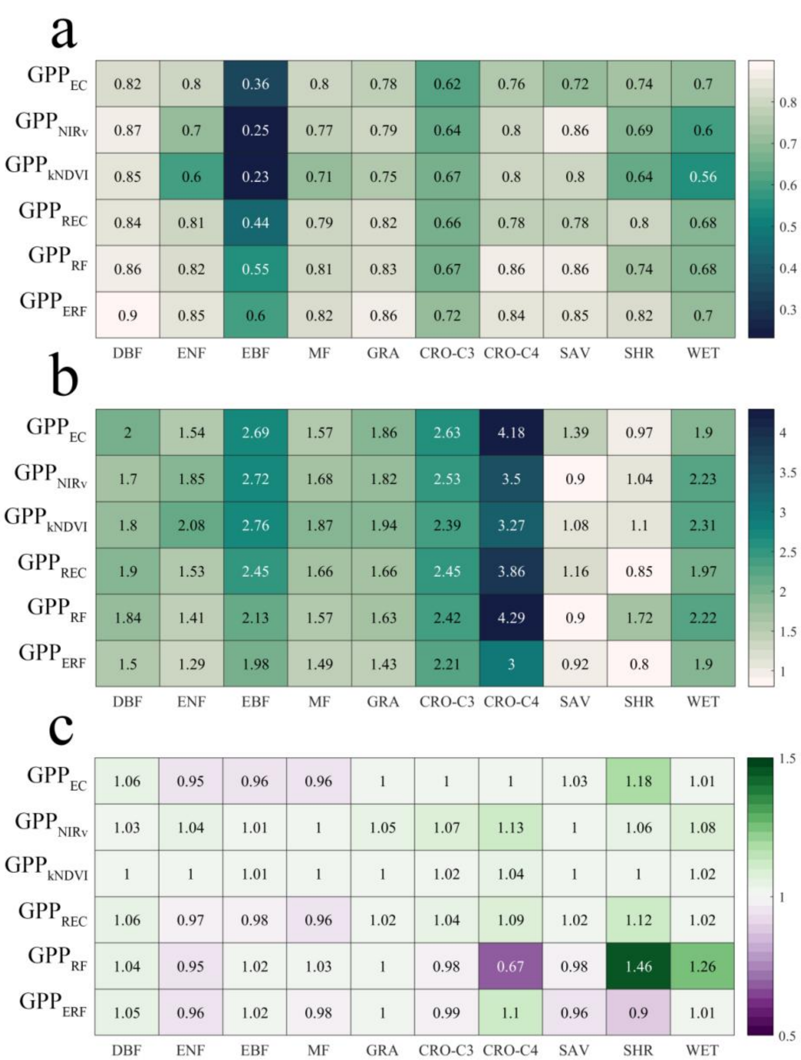

Figure 1. The performance of the six models on different vegetation types. a, b and c represent R^2, RMSE, and Sim/Obs respectively.

We further counted the simulation performance of different models at each site. As shown in Figure S3, we averaged the evaluation indicators of all sites and found that the accuracy of GPP_{ERF} was the highest, R^2 was 0.75, RMSE was 1.53 gC m^{-2} d^{-1}, Sim/Obs was also the closest to 1, which was 1.04. Combining the results of all flux sites, GPP_{ERF} could explain 83.7% of the monthly GPP variation, while the five remote sensing models only explained 72.4%-77.7% of the monthly variation in GPP (Figure 2).

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

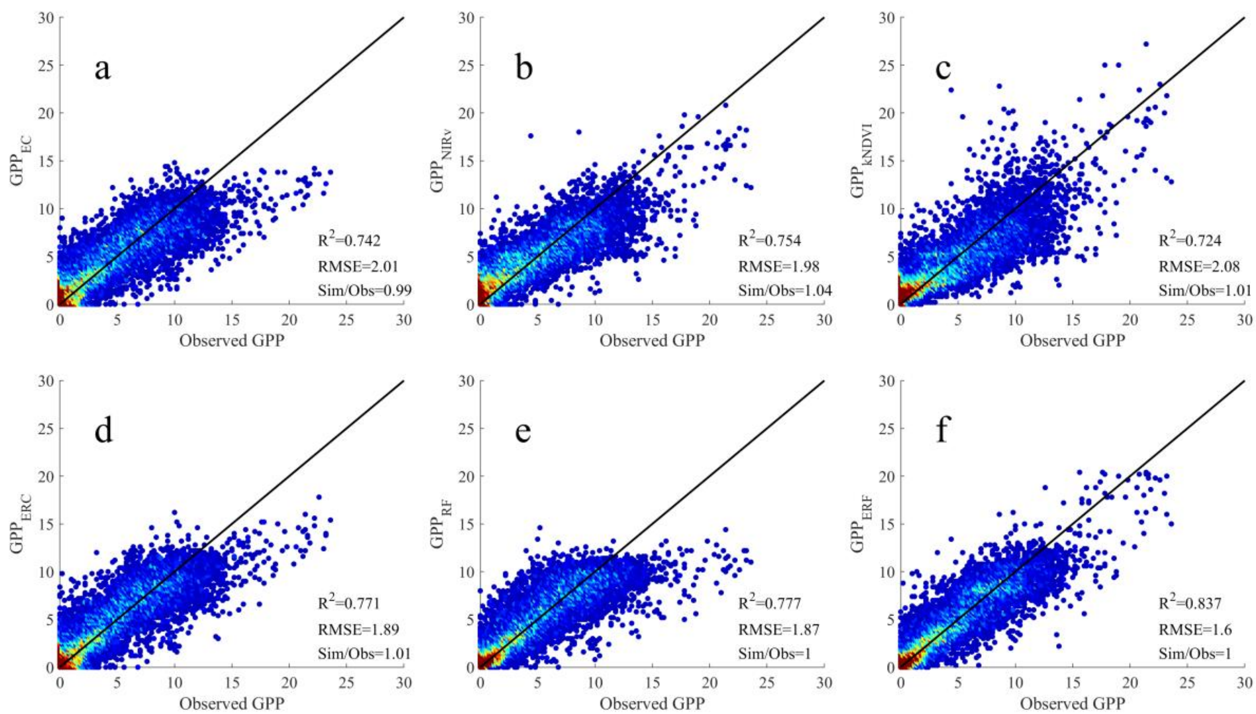

Figure 2. Comparison between the GPP simulations of the six models and the GPP observations. a-f represents GPP_{EC}, GPP_{NIRv}, GPP_{kNDVI}, GPP_{REC}, GPP_{RF}, GPP_{ERF}, respectively.

Overall, GPP_{ERF} exhibited high accuracy in terms of site scale, vegetation type, and the ability to interpret monthly variation in GPP, which also illustrated the potential of machine learning-based ensemble models in improving GPP estimation. However, we also found that most of the GPP simulations have the phenomenon of “high value underestimation and low value overestimate”. For example, GPP_{EC}, GPP_{REC} and GPP_{RF} showed obvious underestimation in the month when the monthly GPP value was greater than 10 gC m^{-2} d^{-1} (Figure 2), it is therefore necessary to evaluate the performance of different models in each month and in different subvalues.

3.2 Performance of six models in each month and different subvalues

Figure 3 shows the simulation accuracy of the six models in each month. The accuracy of the ensemble model was still higher than that of the remote sensing model. GPP_{ERF} maintained the higher R^2 and the lower RMSE in each month, and there was no obvious “high value underestimation and low value undervaluation”. In contrast, the accuracy of the remote sensing model was not satisfactory, especially for winter (most flux sites are concentrated in the northern hemisphere), the LUE models underestimated the GPP per month, and the Sim/Obs remained at 0.78-0.96, but R^2 was above 0.7, while the vegetation index models overestimated GPP, Sim/Obs remained at 1.34-1.73, and R^2 was relatively low, mostly around 0.6.

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

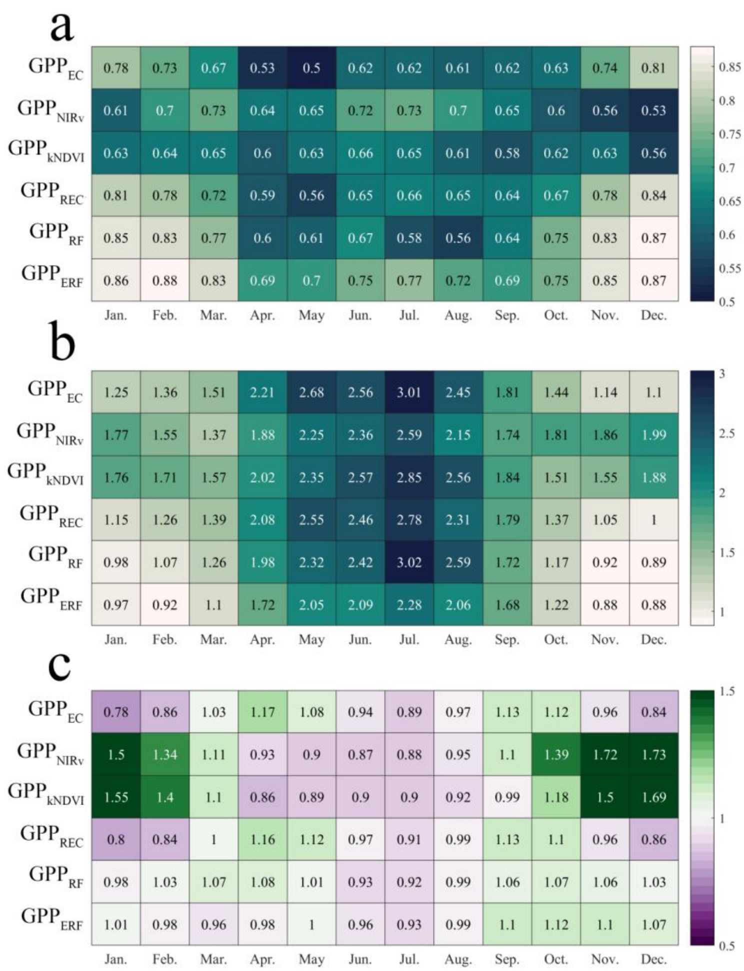

Figure 3. Performance of the six models in each month. a, b and c represent R^2, RMSE, and Sim/Obs respectively.

We compared the performance of all models in different subvalues, including high value (GPP > 10 gC m^{-2} d^{-1}), median value (10 gC m^{-2} d^{-1} > GPP > 2 gC m^{-2} d^{-1}), low value (GPP < 2 gC m^{-2} d^{-1}). In the extreme value, all models performed poorly (Figure 4), the R^2 of the remote sensing model was all below 0.3, while GPP_{ERF} showed a more obvious improvement in the high value, R^2 increased to 0.38, RMSE decreased to 3.03 gC m^{-2} d^{-1}, Sim/Obs also increased to 0.82, and only a slight improvement in the low value. In the median value, all models performed well without serious GPP estimation biases. The R^2 of the remote sensing model was between 0.43 and 0.6, and the RMSE remained between 1.71 and 2.1 gC m^{-2} d^{-1}. It could be seen that there was a large deviation in the estimation of the existing remote sensing model in the GPP extreme value, and the estimation in the median value was relatively good, while the ensemble model based on the machine learning method could improve the simulation accuracy of high value, which was of great significance for accurately estimating the annual values and inter-annual variation of GPP in terrestrial ecosystems.

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

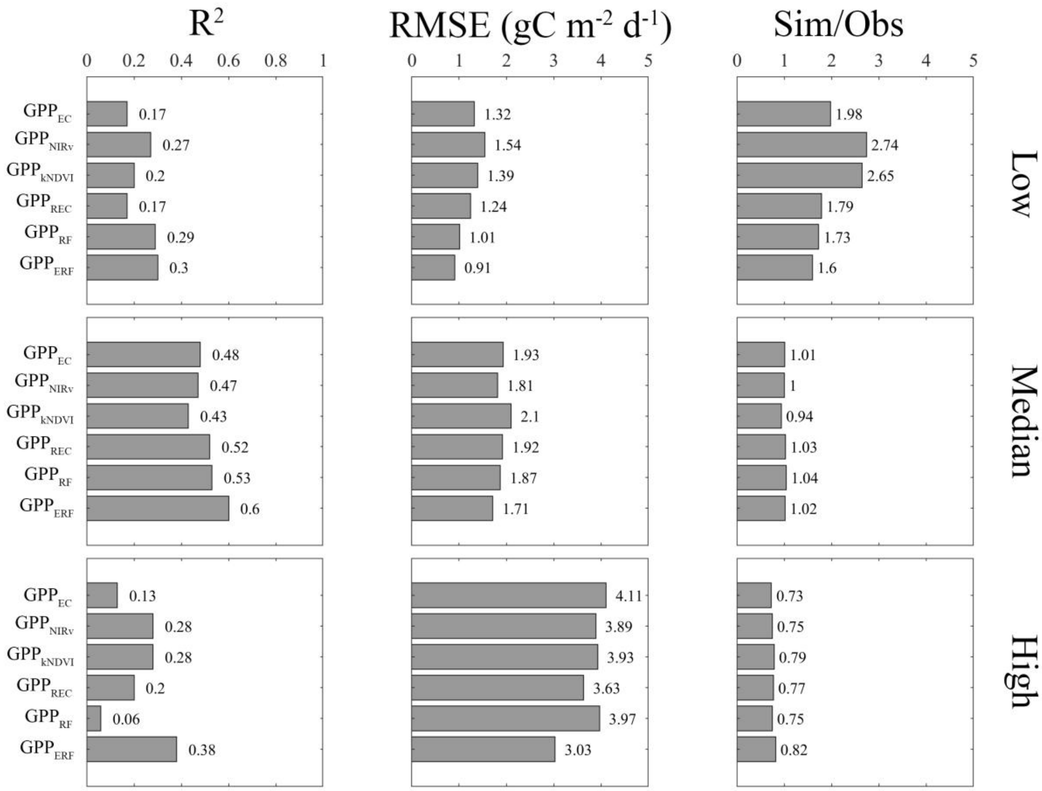

Figure 4. Performance of six models in different subvalues.

3.3 Global GPP estimation based on ensemble model and its generalization evaluation

Based on remote sensing data and meteorological data, we estimated the global GPP from 2001 to 2022 using the ensemble model based on random forest. Figure 5a shows the spatial distribution of ERF_GPP. The high value of GPP was mainly concentrated in tropical areas, exceeding 10 gC m^{-2} d^{-1}, and relatively high in southeastern North America, Europe and southern China, about 4-6 gC m^{-2} d^{-1}. From 2001-2022, China and India showed the fastest increase in GPP, mostly at 0.1 gC m^{-2} d^{-1} (Figure 5b), similar to a previous study that reported that China and India led the global greening (Chen et al., 2019). We further estimated the annual maximum GPP, as shown in Figure 5c, and the North American corn belt was by far the global leader in GPP at more than 15 gC m^{-2} d^{-1}, compared to only 10 gC m-2 d-1 in most tropical forests. In 2001-2022, the global GPP was 131.2 ± 3.1 PgC yr^{-1}, the trend was 0.45 PgC yr^{-2}, the lowest value was 126.4 PgC yr^{-1} in 2001, and the highest value was 135.9 PgC yr^{-1} in 2020 (Figure 5d).

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

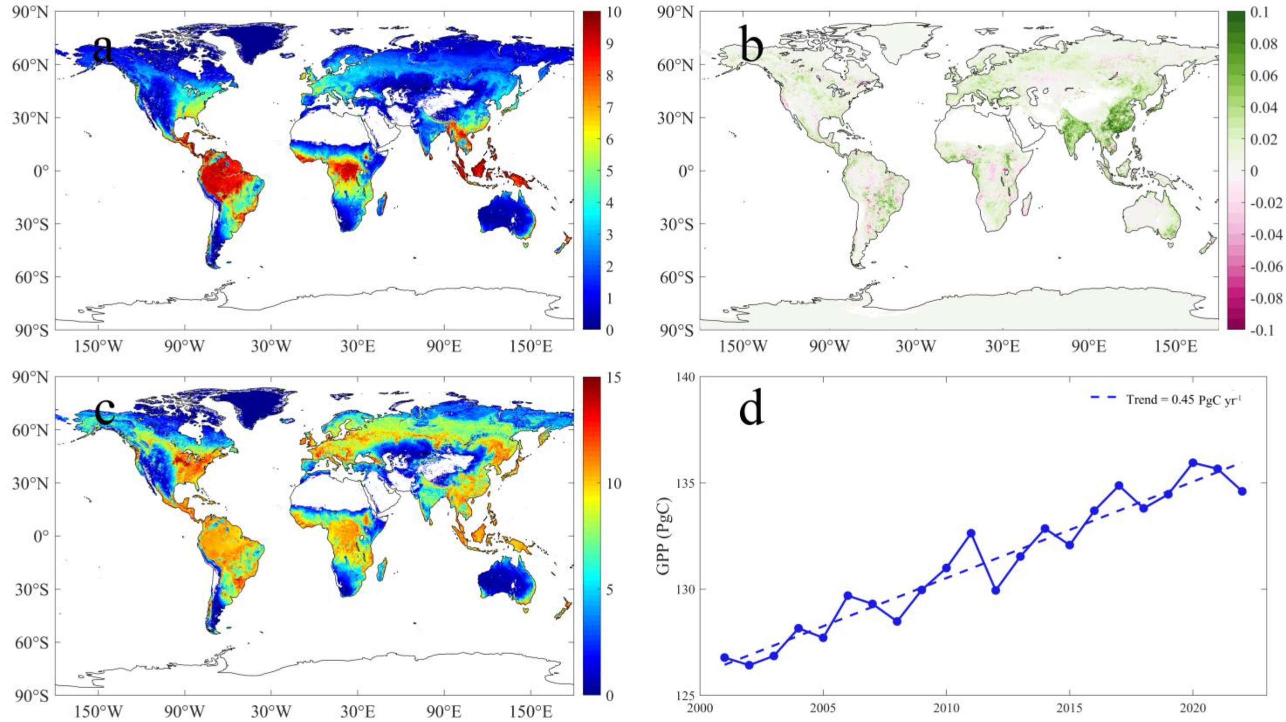

Figure 5. Spatial distribution and interannual change of ERF_GPP during 2001-2022. a represents the multi-year average, b represents the trend, c represents the annual maximum, and d represents the interannual change of GPP.

As shown in Figure 6, the generalizations of ERF_GPP and other GPP datasets were validated using GPP observations from ChinaFlux. Overall, in China, ERF_GPP has a high generalization, R^2 of 0.75, RMSE of 1.75 gC m^{-2} d^{-1}, there was no "high value underestimation and low value overestimation", which was comparable to the simulation accuracy of BESS, ECGC and GOSIF. However, the simulation accuracy of the other two GPP datasets in China was relatively poor, with the R^2 of NIRv being only 0.64, while the Revised EC-LUE was significantly underestimated, with the Sim/Obs being only 0.71. We further examined the different GPP datasets at each site, similar to the results at all sites, the ERF_GPP was relatively robust, with R^2 and RMSE of 0.77 and 1.49 gC m^{-2} d^{-1}, respectively (Figure S4). The R^2 of NIRv and Revised EC-LUE was 0.68 and 0.69, and Revised EC-LUE also showed a significant underestimate (Sim/Obs at 0.66). It should be noted that from the perspective of the average simulation accuracy of each site, BESS seemed to overestimate the GPP (Sim/Obs at 1.2).

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

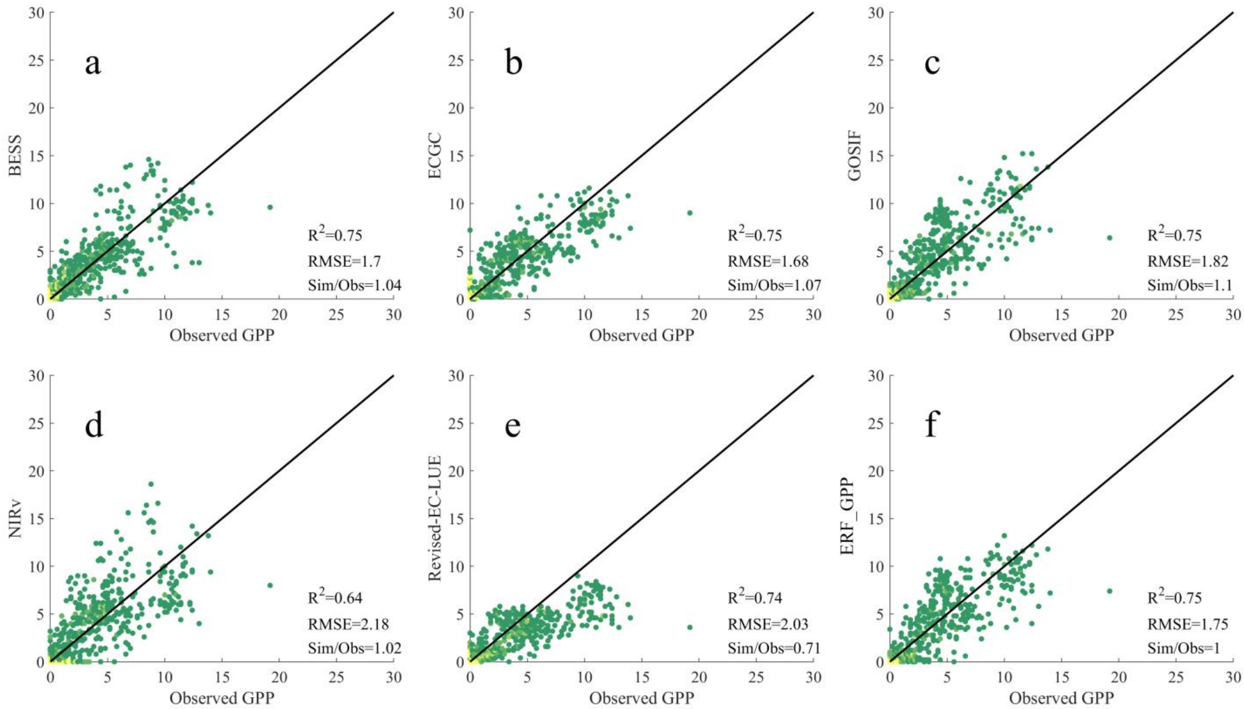

Figure 6. Comparison between the GPP datasets and the GPP observations from ChinaFlux. a-f represents BESS, GOSIF, ECGC, NIRv, Revise-EC-LUE, ERF_GPP, respectively.

4 Discussion

4.1 Performance analysis of different models

With the continuous development of remote sensing technology and carbon cycle models, the existing models for estimating GPP are gradually increasing, including LUE models, process models, machine learning models and the newly developed vegetation index models (such as SIF, NIRV, KNDVI), these "big class" models also include many "small classes". For example, the differences in the environmental restriction function in the LUE model are extended to CASA, VPM, EC-LUE and other models (Xiao et al., 2004; Potter et al., 1993; Yuan et al., 2007). A recent study collected the response functions of GPP to different environmental variables, and under the LUE theory, 5600 LUE models could be generated (Bao et al., 2022). These different model structures greatly increase the uncertainty of global GPP estimation, which make people still confused about the annual value and inter-annual trend of global GPP. All models can obtain a reliable model accuracy after calibrating the parameters, however none of the model accuracy is particularly outstanding, so it is urgent to provide a new method to further improve the accuracy of GPP estimation.

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

Multi-model ensemble may be a proven approach, and previous studies have shown that even simple multi-model average can reduce model estimation uncertainty. In this study, we used an ensemble models to improve the estimation of GPP. Compared with the remote sensing model, the ensemble model could indeed show higher accuracy, the R^2 reached 0.837, which is significantly higher than the accuracy of the machine learning model based on meteorological variables and remote sensing variables (R^2=0.777). Since there are no physical constraints, machine learning models need to find the relationship between explanatory variables and target variables from a large amount of training data (such as GPP=f (LAI,T,P, etc.)) (Tramontana et al., 2016; Jung et al., 2019), so the reliability of the model usually depends on the representativeness of training data, such as LAI can explain GPP to a large extent, however, due to the complexity of the surrounding environment of flux sites, it is difficult to guarantee consistent modeling relationships even for the same vegetation type. The difference between ensemble models based on machine learning lies in the differences in explanatory variables. These explanatory variables are the results of multiple model simulations, and these results are usually more representative and more consistent with the relationship between the target variables, which makes the GPP simulations more accurate.

The simulation results of different models in each months and different subvalues showed that the existing GPP estimation model widely existed the phenomenon of "high value underestimation and low value overestimate". For the LUE model, this phenomenon is most obvious in winter (Figure 3), and the GPP was underestimated by about 20%, which may be due to the deviation in the form of environmental factor. In the expression form of the temperature constraint adopted by the LUE model, the maximum temperature, minimum temperature and optimum temperature for limiting photosynthesis are all constants, however these values may not be fixed (Huang et al., 2019; Grossiord et al., 2020). The two vegetation index models were overestimated in winter, and even overestimated by 70% in December. The vegetation index model does not consider the constraints of environmental factors. They believe that all environmental impacts on vegetation have been included in the vegetation index (kNDVI, NIRv), however, this aspect is still controversial (Wu et al., 2020; Dechant et al., 2022), and the relationship between these vegetation indices and GPP is not robust, and the vegetation indices based on reflectance may have hysteresis (Wang et al., 2022), and our results also showed that only using vegetation indices modeling GPP should be carefully considered. In the low value and high value, the effects of all remote sensing models are not ideal, which may be caused by the model structure itself. Simple mathematical expressions cannot characterize the entire photosynthesis process, and these models are usually only empirical or semi-empirical. Although the ensemble model based on machine learning can improve this phenomenon to a certain extent, it still depends on the reliability of the remote sensing model in the extreme value. Therefore, we believe that in the future model development, it is necessary to focus on the simulation performance of GPP in the extreme value.

4.2 Robustness of global GPP estimation based on ensemble model

In this study, based on site-scale validation, we demonstrate the reliability of the random forest-based ensemble model in GPP estimation. However, what needs to be discussed further is whether the spatial distribution, spatial trends and global total of ERF_GPP are reliable. Since the current GPP datasets are generated based on remote sensing observation, all GPP

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

datasets are very similar in spatial distribution. Therefore, the validation of GPP observations independent of FLUXNET2015 are very important. Validation results from GPP observations from ChinaFlux show that GPP_{ERF} showed good generalization in China (R^2=0.75), which was slightly lower than the accuracy of the 5 fold cross validation during modeling, possibly due to the mismatch between the 0.05° GPP and the footprint of the flux tower (Chu et al., 2021). Overall, however, this is comparable to or slightly better than the simulation accuracy of current mainstream GPP datasets. In addition, we also found a clear improvement in the spatial maximum value of ERF_GPP in some corn growing regions, such as the North American Corn Belt (Figure 5c), which is supported by previous studies showing that C4 crops have much higher GPP peaks than other vegetation types (Yuan et al., 2015; Chen et al., 2011).

Due to the drought trend, the constraint effect of water on vegetation is gradually increasing, and some studies have reported the decoupling phenomenon of LAI and GPP under some specific conditions (Jiao et al., 2021; Huang et al., 2019). However, in China and India that two regions with significant greening, GPP is still increasing in most datasets, and ERF_GPP supports this view. This phenomenon may be due to the low drought pressure on farmland in China and India due to irrigation, which is less of a constraint on GPP (Ambika and Mishra, 2020; Ai et al., 2020). The global estimate of ERF_GPP was 131.2 PgC yr^{-1}, which is close to estimates from most previous studies (Wang et al., 2021b; Badgley et al., 2019). Some studies have suggested that the global GPP may reach 150-175 PgC yr^{-1} (Welp et al., 2011), however, there is no further evidence to support this view.

4.3 Limitations and uncertainties

In this study, we improved GPP estimates based on the ensemble model. However, there are still some limitations and uncertainties due to the availability of data and methods. First, C4 crop distribution maps were used in our study to improve estimates of cropland GPP. However, it is important to note that this dataset only represents the spatial distribution of crops around the year 2000, which may add uncertainty to GPP simulations of cropland in a few C3 and C4 alternating areas. Secondly, only the GPP simulations of four remote sensing models were considered in our model, and it is not clear whether adding more GPP simulations to the model can further improve the GPP estimation. Finally, our model did not consider the effect of soil moisture on GPP, and some previous studies have highlighted the importance of considering soil moisture in GPP estimates, especially for dry years (Stocker et al., 2019; Stocker et al., 2018).

5 Conclusion

In this study, we evaluated the performance of five remote sensing models and one ensemble model to simulate GPP. Overall, GPP_{ERF} had higher model accuracy, explaining 83.7% of the monthly variation in GPP, and showed good accuracy in different vegetation types, different months and different extreme regions. The global GPP of ERF_GPP for 2001-2022 is 131.2 PgC yr^{-1}. The results from ChinaFlux show that ERF_GPP has good generalization. For the current emerging GPP estimation models, machine learning-based ensemble models provide another method of GPP estimation, and this may lead to higher model accuracy and more reliable global GPP estimation.

Data and code availability

The global GPP dataset based on the ensemble model for 2001-2022 is available at https://doi.org/10.6084/m9.figshare.24417649 (Chen et al., 2023). The spatial resolution of ERF_GPP is 0.05° and the temporal resolution is monthly. Code is available from the author upon reasonable request.

Author contributions

X.C. and T.X.C. conceived the scientific ideas and designed this research. X.C. compiled the data, conducted analysis, prepared figures. X.C., T.X.C. and Y.F.C. wrote the manuscript. D.X.L., R.J.G., J.D., and S.J.Z. gave constructive suggestions for improving the manuscript.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (No. 42130506, 42161144003 and 31570464) and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (No. KYCX23_1322).

Declaration of interests

The authors have not disclosed any competing interests.

References

Ai, Z., Wang, Q., Yang, Y., Manevski, K., Yi, S., and Zhao, X.: Variation of gross primary production, evapotranspiration and water use efficiency for global croplands, Agricultural and Forest Meteorology, 287, 10.1016/j.agrformet.2020.107935, 2020.

Ambika, A. K. and Mishra, V.: Substantial decline in atmospheric aridity due to irrigation in India, Environmental Research Letters, 15, 10.1088/1748-9326/abc8bc, 2020.

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin, P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., and Zhao, M.: Spatiotemporal patterns of terrestrial gross primary production: A review, Reviews of Geophysics, 53, 785-818, 10.1002/2015rg000483, 2015.

Badgley, G., Field, C. B., and Berry, J. A.: Canopy near-infrared reflectance and terrestrial photosynthesis, Science advances, 3, e1602244, 2017.

Badgley, G., Anderegg, L. D., Berry, J. A., and Field, C. B.: Terrestrial gross primary production: Using NIRV to scale from site to globe, Global change biology, 25, 3731-3740, 2019.

Bai, Y., Zhang, S., Bhattarai, N., Mallick, K., Liu, Q., Tang, L., Im, J., Guo, L., and Zhang, J.: On the use of machine learning based ensemble approaches to improve evapotranspiration estimates from croplands across a wide environmental gradient, Agricultural and Forest Meteorology, 298, 108308, 2021.
Bao, S., Wutzler, T., Koirala, S., Cuntz, M., Ibrom, A., Besnard, S., Walther, S., Šigut, L., Moreno, A., and Weber, U.: Environment-sensitivity functions for gross primary productivity in light use efficiency models, Agricultural and Forest Meteorology, 312, 108708, 2022.
Camps-Valls, G., Campos-Taberner, M., Moreno-Martínez, Á., Walther, S., Duveiller, G., Cescatti, A., Mahecha, M. D., Muñoz-Marí, J., García-Haro, F. J., and Guanter, L.: A unified vegetation index for quantifying the terrestrial biosphere, Science Advances, 7, eabc7447, 2021.
Chen, C., Park, T., Wang, X., Piao, S., Xu, B., Chaturvedi, R. K., Fuchs, R., Brovkin, V., Ciais, P., Fensholt, R., Tommervik, H., Bala, G., Zhu, Z., Nemani, R. R., and Myneni, R. B.: China and India lead in greening of the world through land-use management, Nature Sustainability, 2, 122-129, 10.1038/s41893-019-0220-7, 2019.
Chen, T., van der Werf, G. R., Dolman, A. J., and Groenendijk, M.: Evaluation of cropland maximum light use efficiency using eddy flux measurements in North America and Europe, Geophysical Research Letters, 38, 10.1029/2011gl047533, 2011.
Chen, T., Van Der Werf, G., Gobron, N., Moors, E., and Dolman, A.: Global cropland monthly gross primary production in the year 2000, Biogeosciences, 11, 3871-3880, 2014.
Chen, Y., Yuan, H., Yang, Y., and Sun, R.: Sub-daily soil moisture estimate using dynamic Bayesian model averaging, Journal of Hydrology, 590, 125445, 2020.
Chu, H., Luo, X., Ouyang, Z., Chan, W. S., Dengel, S., Biraud, S. C., Torn, M. S., Metzger, S., Kumar, J., and Arain, M. A.: Representativeness of Eddy-Covariance flux footprints for areas surrounding AmeriFlux sites, Agricultural and Forest Meteorology, 301, 108350, 2021.
Dechant, B., Ryu, Y., Badgley, G., Köhler, P., Rascher, U., Migliavacca, M., Zhang, Y., Tagliabue, G., Guan, K., and Rossini, M.: NIRVP: A robust structural proxy for sun-induced chlorophyll fluorescence and photosynthesis across scales, Remote Sensing of Environment, 268, 112763, 2022.
Groenendijk, M., Dolman, A., Van der Molen, M., Leuning, R., Arneth, A., Delpierre, N., Gash, J., Lindroth, A., Richardson, A., and Verbeeck, H.: Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data, Agricultural and forest meteorology, 151, 22-38, 2011.
Grossiord, C., Buckley, T. N., Cernusak, L. A., Novick, K. A., Poulter, B., Siegwolf, R. T., Sperry, J. S., and McDowell, N. G.: Plant responses to rising vapor pressure deficit, New Phytologist, 226, 1550-1566, 2020.
Guo, R., Chen, T., Chen, X., Yuan, W., Liu, S., He, B., Li, L., Wang, S., Hu, T., Yan, Q., Wei, X., and Dai, J.: Estimating Global GPP From the Plant Functional Type Perspective Using a Machine Learning Approach, Journal of Geophysical Research-Biogeosciences, 128, 10.1029/2022jg007100, 2023.
Huang, M., Piao, S., Ciais, P., Peñuelas, J., Wang, X., Keenan, T. F., Peng, S., Berry, J. A., Wang, K., and Mao, J.: Air temperature optima of vegetation productivity across global biomes, Nature ecology & evolution, 3, 772-779, 2019.
Jiao, W., Wang, L., Smith, W. K., Chang, Q., Wang, H., and D'Odorico, P.: Observed increasing water constraint on vegetation growth over the last three decades, Nature Communications, 12, 10.1038/s41467-021-24016-9, 2021.
Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G., and Reichstein, M.: The FLUXCOM ensemble of global land-atmosphere energy fluxes, Scientific data, 6, 1-14, 2019.
Li, B., Ryu, Y., Jiang, C., Dechant, B., Liu, J., Yan, Y., and Li, X.: BESSv2.0: A satellite-based and coupled-process model for quantifying long-term global land-atmosphere fluxes, Remote Sensing of Environment, 295, 10.1016/j.rse.2023.113696, 2023.
Li, X. and Xiao, J.: A Global, 0.05-Degree Product of Solar-Induced Chlorophyll Fluorescence Derived from OCO-2, MODIS, and Reanalysis Data, Remote Sensing, 11, 10.3390/rs11050517, 2019.
Mengistu, A. G., Mengistu Tsidu, G., Koren, G., Kooreman, M. L., Boersma, K. F., Tagesson, T., Ardö, J., Nouvellon, Y., and Peters, W.: Sun-induced fluorescence and near-infrared reflectance of vegetation track the seasonal dynamics of gross primary production over Africa, Biogeosciences, 18, 2843-2857, 2021.
Monfreda, C., Ramankutty, N., and Foley, J. A.: Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000, Global Biogeochemical Cycles, 22, 10.1029/2007gb002947, 2008.

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., Poindexter, C., Chen, J., Elbashandy, A., and Humphrey, M.: The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data, Scientific data, 7, 1-27, 2020.

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., and Klooster, S. A.: Terrestrial ecosystem production: a process model based on global satellite and surface data, Global Biogeochemical Cycles, 7, 811-841, 1993.

Ruehr, S., Keenan, T. F., Williams, C., Zhou, Y., Lu, X., Bastos, A., Canadell, J. G., Prentice, I. C., Sitch, S., and Terrer, C.: Evidence and attribution of the enhanced land carbon sink, Nature Reviews Earth & Environment, 4, 518-534, 10.1038/s43017-023-00456-3, 2023.

Ryu, Y., Berry, J. A., and Baldocchi, D. D.: What is global photosynthesis? History, uncertainties and opportunities, Remote sensing of environment, 223, 95-114, 2019.

Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Penuelas, J., and Seneviratne, S. I.: Quantifying soil moisture impacts on light use efficiency across biomes, New Phytologist, 218, 1430-1449, 10.1111/nph.15123, 2018.

Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Seneviratne, S. I., and Penuelas, J.: Drought impacts on terrestrial primary production underestimated by satellite monitoring, Nature Geoscience, 12, 264-+, 10.1038/s41561-019-0318-6, 2019.

Sun, Z., Wang, X., Zhang, X., Tani, H., Guo, E., Yin, S., and Zhang, T.: Evaluating and comparing remote sensing terrestrial GPP models for their response to climate variability and CO2 trends, Science of the total environment, 668, 696-713, 2019.

Tian, Z., Yi, C., Fu, Y., Kutter, E., Krakauer, N. Y., Fang, W., Zhang, Q., and Luo, H.: Fusion of Multiple Models for Improving Gross Primary Production Estimation With Eddy Covariance Data Based on Machine Learning, Journal of Geophysical Research: Biogeosciences, 128, e2022JG007122, https://doi.org/10.1029/2022JG007122, 2023.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., Reichstein, M., Arain, M. A., Cescatti, A., and Kiely, G.: Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms, Biogeosciences, 13, 4291-4313, 2016.

Wang, J., Dong, J., Yi, Y., Lu, G., Oyler, J., Smith, W., Zhao, M., Liu, J., and Running, S.: Decreasing net primary production due to drought and slight decreases in solar radiation in China from 2000 to 2012, Journal of Geophysical Research: Biogeosciences, 122, 261-278, 2017.

Wang, J., Jiang, F., Wang, H., Qiu, B., Wu, M., He, W., Ju, W., Zhang, Y., Chen, J. M., and Zhou, Y.: Constraining global terrestrial gross primary productivity in a global carbon assimilation system with OCO-2 chlorophyll fluorescence data, Agricultural and Forest Meteorology, 304, 108424, 2021a.

Wang, S., Zhang, Y., Ju, W., Qiu, B., and Zhang, Z.: Tracking the seasonal and inter-annual variations of global gross primary production during last four decades using satellite near-infrared reflectance data, Science of the Total Environment, 755, 142569, 2021b.

Wang, S., Zhang, Y., Ju, W., Chen, J. M., Ciais, P., Cescatti, A., Sardans, J., Janssens, I. A., Wu, M., and Berry, J. A.: Recent global decline of CO2 fertilization effects on vegetation photosynthesis, Science, 370, 1295-1300, 2020.

Wang, X., Biederman, J. A., Knowles, J. F., Scott, R. L., Turner, A. J., Dannenberg, M. P., Köhler, P., Frankenberg, C., Litvak, M. E., and Flerchinger, G. N.: Satellite solar-induced chlorophyll fluorescence and near-infrared reflectance capture complementary aspects of dryland vegetation productivity dynamics, Remote Sensing of Environment, 270, 112858, 2022.

Welp, L. R., Keeling, R. F., Meijer, H. A. J., Bollenbacher, A. F., Piper, S. C., Yoshimura, K., Francey, R. J., Allison, C. E., and Wahlen, M.: Interannual variability in the oxygen isotopes of atmospheric CO₂ driven by El Nino, Nature, 477, 579-582, 10.1038/nature10421, 2011.

Wu, G., Guan, K., Jiang, C., Peng, B., Kimm, H., Chen, M., Yang, X., Wang, S., Suyker, A. E., and Bernacchi, C. J.: Radiance-based NIRv as a proxy for GPP of corn and soybean, Environmental Research Letters, 15, 034009, 2020.

Xiao, J., Chevallier, F., Gomez, C., Guanter, L., Hicke, J. A., Huete, A. R., Ichii, K., Ni, W., Pang, Y., and Rahman, A. F.: Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years, Remote Sensing of Environment, 233, 111383, 2019.

Xiao, X., Zhang, Q., Braswell, B., Urbanski, S., Boles, S., Wofsy, S., Moore III, B., and Ojima, D.: Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data, Remote sensing of environment, 91, 256-270, 2004.

https://doi.org/10.5194/egusphere-2024-114
Preprint. Discussion started: 9 February 2024

Xu, T., White, L., Hui, D., and Luo, Y.: Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction, Global Biogeochemical Cycles, 20, 2006.

Yao, Y., Liang, S., Li, X., Chen, J., Liu, S., Jia, K., Zhang, X., Xiao, Z., Fisher, J. B., and Mu, Q.: Improving global terrestrial evapotranspiration estimation using support vector machine by integrating three process-based algorithms, Agricultural and Forest Meteorology, 242, 55-74, 2017.

Yao, Y., Liang, S., Li, X., Hong, Y., Fisher, J. B., Zhang, N., Chen, J., Cheng, J., Zhao, S., and Zhang, X.: Bayesian multimodel estimation of global terrestrial latent heat flux from eddy covariance, meteorological, and satellite observations, Journal of Geophysical Research: Atmospheres, 119, 4521-4545, 2014.

Yuan, W., Cai, W., Nguy-Robertson, A. L., Fang, H., Suyker, A. E., Chen, Y., Dong, W., Liu, S., and Zhang, H.: Uncertainty in simulating gross primary production of cropland ecosystem from satellite-based models, Agricultural and Forest Meteorology, 207, 48-57, 10.1016/j.agrformet.2015.03.016, 2015.

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

Yuan, W., Liu, S., Zhou, G., Zhou, G., Tieszen, L. L., Baldocchi, D., Bernhofer, C., Gholz, H., Goldstein, A. H., and Goulden, M. L.: Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes, Agricultural and Forest Meteorology, 143, 189-207, 2007.

Yuan, W., Zheng, Y., Piao, S., Ciais, P., Lombardozzi, D., Wang, Y., Ryu, Y., Chen, G., Dong, W., and Hu, Z.: Increased atmospheric vapor pressure deficit reduces global vegetation growth, Science advances, 5, eaax1396, 2019.

Zheng, Y., Shen, R., Wang, Y., Li, X., Liu, S., Liang, S., Chen, J. M., Ju, W., Zhang, L., and Yuan, W.: Improved estimate of global gross primary production for reproducing its long-term variation, 1982–2017, Earth System Science Data, 12, 2725-2746, 2020.