Zeng *et al.*
Carbon Balance and Management (2023) 18:19
https://doi.org/10.1186/s13021-023-00239-9

Carbon Balance and Management

RESEARCH

Open Access

The fusion of multiple scale data indicates that the carbon sink function of the Qinghai-Tibet Plateau is substantial

Jingyu Zeng[1,2], Tao Zhou[1,2*], Yixin Xu[1,2], Qiaoyu Lin[1,2], E. Tan[1,2], Yajie Zhang[1,2], Xuemei Wu[1,2], Jingzhou Zhang[1,2] and Xia Liu[1,2]

Abstract

Background The Qinghai-Tibet Plateau is the "sensitive area" of climate change, and also the "driver" and "amplifier" of global change. The response and feedback of its carbon dynamics to climate change will significantly affect the content of greenhouse gases in the atmosphere. However, due to the unique geographical environment characteristics of the Qinghai-Tibet Plateau, there is still much controversy about its carbon source and sink estimation results. This study designed a new algorithm based on machine learning to improve the accuracy of carbon source and sink estimation by integrating multiple scale carbon input (net primary productivity, NPP) and output (soil heterotrophic respiration, Rh) information from remote sensing and ground observations. Then, we compared spatial patterns of NPP and Rh derived from the fusion of multiple scale data with other widely used products and tried to quantify the differences and uncertainties of carbon sink simulation at a regional scale.

Results Our results indicate that although global warming has potentially increased the Rh of the Qinghai-Tibet Plateau, it will also increase its NPP, and its current performance is a net carbon sink area (carbon sink amount is 22.3 Tg C/year). Comparative analysis with other data products shows that CASA, GLOPEM, and MODIS products based on remote sensing underestimate the carbon input of the Qinghai-Tibet Plateau (30–70%), which is the main reason for the severe underestimation of the carbon sink level of the Qinghai-Tibet Plateau (even considered as a carbon source).

Conclusions The estimation of the carbon sink in the Qinghai-Tibet Plateau is of great significance for ensuring its ecological barrier function. It can deepen the community's understanding of the response to climate change in sensitive areas of the plateau. This study can provide an essential basis for assessing the uncertainty of carbon sources and sinks in the Qinghai-Tibet Plateau, and also provide a scientific reference for helping China achieve "carbon neutrality" by 2060.

Highlights

1. This study designed a carbon sink estimation method for Qinghai-Tibet Plateau by integrating machine learning and multiple scale ground- and remote sensed-based data.

*Correspondence:
Tao Zhou
tzhou@bnu.edu.cn
Full list of author information is available at the end of the article

2. The estimated total carbon sink of the Qinghai-Tibet Plateau is 22.3 Tg C/year, accounting for about 10% of China's total carbon sink.
3. The carbon sink of former estimation maybe greatly underestimated due to the underestimation of carbon input in the Qinghai-Tibet Plateau.

Keywords Qinghai-Tibet Plateau, Carbon source and sink, Multiple scale information fusion, Machine learning, Uncertainty assessment

Background

Terrestrial ecosystems have always been the largest natural carbon sink. Especially since the 1960s, more than a quarter of anthropogenic emissions have been stored on average [1]. Considering its importance in combating climate warming, accurate estimation of carbon sinks is crucial for assessing mitigation policies [2–4]. With the development of technology, significant progress has been made in the annual updated global carbon estimate [1, 5]. However, in terms of its response, such as the trend of atmospheric CO_2 and the increase of interannual variability, climate variability and other environmental factors, the terrestrial carbon sink is still highly uncertain [6–10]. These uncertainties are the main obstacle to our ability to predict climate feedback and its potential impact on the human-natural system [11].

As a unique geomorphic unit on the earth, the Qinghai-Tibet Plateau (QTP) is very sensitive to global warming [12], and its warming rate in the past 50 years is twice the global average [13–15]. It is observed that the permafrost on the Qinghai-Tibet Plateau has been significantly degraded due to the rapid warming of the climate [12, 16]. These changes have a significant impact on the ecology, hydrology, biogeochemistry and engineering infrastructure of the third pole and surrounding areas [17, 18]. In particular, the melting of permafrost will cause the release of soil organic carbon (SOC) into the atmosphere, which will affect climate warming [19–21]. At the same time, the Qinghai-Tibet Plateau is considered to have a profound impact on the regional and global climate system and ecological economy [22]. It is a sensitive indicator of climate change and a key component of the global carbon (C) cycle [23]. The response of the Qinghai-Tibet Plateau C dynamics to climate change may significantly affect the concentrations of carbon dioxide (CO_2) and methane (CH_4) in the atmosphere [24, 25], thus affecting the pace of future climate change [26, 27].

At present, whether the Qinghai-Tibet Plateau is a carbon source or a carbon sink is controversial [28]. For example, based on ORCHIDE and TEM models, the Qinghai-Tibet Plateau (QTP) is estimated to be a carbon sink [29, 30]. The results of CASA model and empirical statistical model show that the grassland ecosystem in the central and southern parts of the Qinghai-Tibet Plateau shows weak net carbon absorption, and its net ecosystem productivity (NEP) is 42.03 g C/m^2 year [31]. The research based on the traditional empirical model believes that the Qinghai-Tibet Plateau is a carbon source [32, 33]. Studies also found that the Qinghai-Tibet Plateau is a carbon sink at present, but in the case of climate warming at the end of the twenty-first century, it will change from carbon sink to carbon source [34]. Koven et al. [35] revealed that by the end of the twenty-first century, the climate was warming, and the permafrost ecosystem in high latitudes could shift from carbon sink to carbon source. Although the warming of the Qinghai-Tibet Plateau will increase the carbon sink capacity of the ecosystem in the growing season, it will also increase the soil carbon decomposition in the non-growing season, and the degradation of the permafrost region will increase the carbon release of the ecosystem. This will also destroy the surface vegetation, causing the ecosystem to change from carbon sink to carbon source in the growing season [12]. However, due to the high spatiotemporal heterogeneity of the Qinghai-Tibet Plateau, it is challenging to simulate, evaluate and predict the spatiotemporal dynamics of carbon flux [15, 36].

The significant uncertainty in estimating carbon sinks in the Qinghai Tibet Plateau is due to differences in methods. The estimation methods for carbon sources and sinks in terrestrial ecosystems include "top-down" atmospheric inversion method, "bottom-up" inventory method, vorticity correlation method, and ecosystem process model simulation method [19, 37, 38]. These methods may be limited to varying degrees by model parameters, number and distribution of stations, scale conversion and simulation accuracy, especially in the Qinghai Tibet Plateau region, where the differences in estimation results between different methods may exceed 10 times [12, 39]. Due to the complex underlying surface conditions and high spatial heterogeneity of the Qinghai Tibet Plateau, it is difficult to localize the parameters of the process model [38]. Remote sensing methods are also not suitable for most areas of the sparsely vegetated Qinghai Tibet Plateau. The carbon source and sink estimation method that combines remote sensing products

with ground observation data to ensure spatial continuity and quantitative accuracy is very promising [40, 41]. This data-driven method can retain the effective information of remote sensing products and sample observation data, capture the complex nonlinear relationship between input and output variables, and achieve the goal of unifying different data scales, so it has high flexibility and data adaptability [42, 43]. Based on the data-driven method, new information can also be mined from data to promote a new understanding of the new mechanism [44]. Therefore, the multiple scale information fusion method of carbon sink estimation using machine learning as a bridge to connect multiple-source data such as remote sensing and ground sample observation data is an effective solution to reduce the uncertainty of the results.

In order to more accurately simulate the size and spatial pattern of carbon sources and sinks in the Qinghai-Tibet Plateau, we designed a multiple scale information fusion method for carbon sink estimation that uses machine learning as a bridge to connect multiple-source data such as remote sensing and ground sample observation data. Thus, a more accurate and spatially continuous carbon source and sink and its spatial pattern from 2000 to 2018, which can reflect the regional characteristics of the Qinghai-Tibet Plateau, are obtained. Then, we compared the ecosystem carbon input and carbon output with other widely used similar products, and discussed the differences and uncertainties of carbon sink simulation in the Qinghai-Tibet Plateau.

Data sources and methods

Study area

The Qinghai-Tibet Plateau locates in western China (73.50–104.67° E, 25.99–39.83° N), with an area of more than 2.56×10^6 km^2 (Fig. 1). As an Asian water tower, about 40% of the world's population depends on or is affected by it [45]. Grassland is the main vegetation type in the Qinghai-Tibet Plateau. Since the twenty-first century, the annual average temperature in the coldest region of the Qinghai-Tibet Plateau has been lower than − 20 °C, while that in the warmest region is higher than 20 °C [46]. Since the mid-1950s, the Qinghai-Tibet Plateau has experienced significant warming, with the annual average temperature increasing by 0.3 °C every decade [30]. The spatial heterogeneity of precipitation on the Qinghai-Tibet Plateau is extremely high. The southeast Qinghai-Tibet Plateau is the wettest region with an annual precipitation of more than 1000 mm. The annual precipitation in the driest area in the northwest is less than 50 mm [47].

Data sources

Ground observation data

NPP ground observation data are from relevant literature, and soil respiration and heterotrophic respiration ground observation data are from the fifth edition of soil respiration database [48]. Ground observation data are used for model construction and precision evaluation. The details for these ground observation list in appendix.

Spatial data

Three NPP product data sets were used in this study. MOD17A3H has a spatial resolution of 500 m, a time resolution of years and a time span of 2000–2018(NPP_{MODIS}). Chen [49] estimated China's 1 km NPP product (NPP_{CASA}) from 1985 to 2015 based on CASA model. Institute of Geographic Sciences and Natural Resources, Chinese Academy of Sciences based on light energy utilization model GLO_PEM estimated China's 1 km NPP product (NPP_{GLO_PEM}) from 2000 to 2010. The random forest model uses the high-resolution meteorological driven data set developed by He et al. [50] for input,

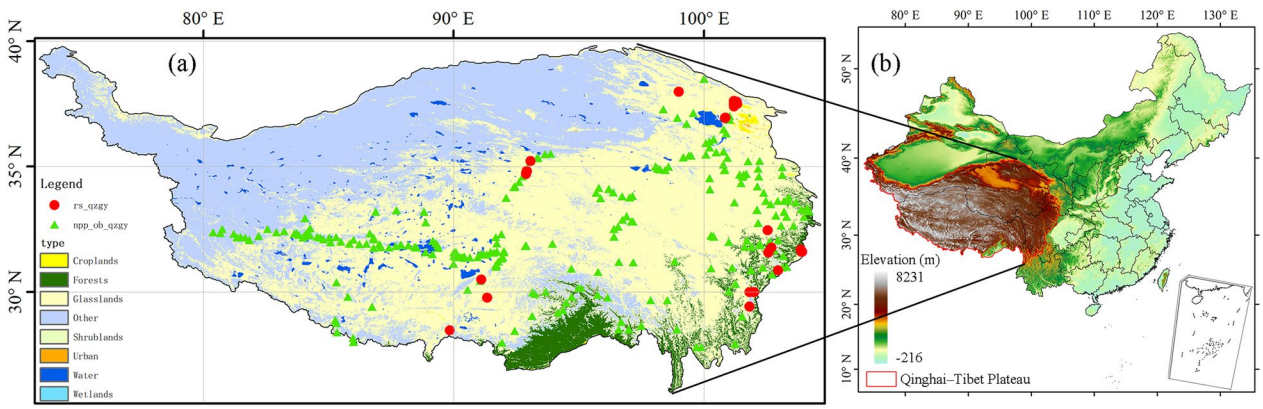

Fig. 1 Overview of the study area. **a** Land use types and location of site observations of ground sample in the Qinghai-Tibet Plateau; **b** The elevation of the Qinghai-Tibet Plateau and its position in China

including air temperature, rainfall and downward short-wave radiation. The spatial resolution is 0.1°, the temporal resolution is years, and the time span is 1979–2018. The product data of land cover type includes MCD12Q1, with a spatial resolution of 500 m, a temporal resolution of years, and a time span of 2001–2018. The leaf area index (LAI) data is from Liu et al. [51], with a spatial resolution of 8 km and a temporal resolution of 8d, and a time span of 1981–2019. Digital Elevation Model (DEM) includes SRTMDEMUTM products with a spatial resolution of 90 m. The global soil organic carbon (SOC) comes from the soil geographic database with a spatial resolution of 250 m. The global 1 km spatial resolution soil respiration data (Rs_{QRFM}) estimated by quantile regression forest model is from Stell et al. [52]. The above data were resampled by the nearest neighbor at a spatial resolution of 1 km, and the annual data were synthesized by the average method.

Methods

NPP estimate

The random forest (RF) model includes many binary decision trees, which are grown independently using a two-stage randomization program [53]. Using a bootstrap resampling method, RF creates a tree structure classifier through independent and identically distributed random vectors, which randomly extracts quantitative samples as training samples. RF then constructs a decision tree model for each training sample set. The model obtains the final regression result by calculating the average value predicted by multiple decision trees [54]. A large number of theoretical and practical studies have proved that the RF algorithm has high accuracy and stability and has a good tolerance to outlier and noise [55]. Based on the NPP data observed on the ground, we built a random forest model by integrating the NPP data estimated by remote sensing and environmental variables, and simulated the NPP of China from 2000 to 2018. The model structure is as follows:

$$NPP = f(Precipitation, Temperature, SRAD, LULC, DEM, LAI, NPP_{modis}) \quad (1)$$

where NPP is the net primary productivity of vegetation (g C/m^2 year) after integrating ground observation and remote sensing information, SRAD is short-wave radiation, LULC is land cover type data, DEM is digital elevation model, LAI is leaf area index, NPP_{modis} is the NPP remote sensing product of MODIS, which has been widely used and has high spatial resolution [56, 57]. Using NPP_{modis} and LAI to replace the traditional vegetation index is a new idea in this study. It can effectively estimate the photosynthetic capacity and growth status of vegetation in complex environments [58, 59]. We estimate the relationship between NPP and independent variables through the random forest model established by R language. The ntree parameter in the model is set to 1500. Since there are seven input variables in the model, the mrty parameter is set to six.

Rs and Rh estimate

Considering that soil respiration (Rs) has more observation data than soil heterotrophic respiration (Rh), the estimation of Rh is carried out in two steps. First, soil respiration (Rs) is estimated by random forest model, and then Rh is further estimated by Rs. The random forest model of Rs is as follows:

$$Rs = f(Precipitation, Temperature, SRAD, LULC, DEM, LAI, NPP_{modis}, SOC) \quad (2)$$

where Rs is soil respiration (g C/m^2 year), SOC is soil organic carbon, and other variables are consistent with NPP model (Eq. 1). The parameter of ntree in the model is set to 1500. Since there are eight input variables in the model, the parameter of mrty is set to seven. After obtaining the simulated value of Chinese Rs, we further estimate the Rh of the Qinghai-Tibet Plateau based on the statistical model of Chinese Rs and Rh (Eq. 3) (see the Additional file 1: Fig. S1 for details).

$$Rh = 0.62 * Rs + 9.01 R^2 = 0.77 \quad (3)$$

NEP estimate

Based on the estimated NPP and Rh, we calculate the net ecosystem productivity (NEP) of the Qinghai-Tibet Plateau:

$$NEP = NPP - Rh \quad (4)$$

where NEP is the net productivity of the ecosystem (g C/m^2yr), which is used to characterize the carbon source/sink level of the ecosystem. When NEP is positive, it indicates that the ecosystem is a carbon sink; When NEP is negative, it indicates that the ecosystem is a carbon source.

Evaluation indicators

Multiple indicators were used to evaluate the reliability of random forest models and the accuracy of different models. These indicators include decision coefficient, correlation coefficient (r), root mean square error (RMSE), percent bias (PBIAS), mean absolute percentage error (MAPE).

$$\mathrm{PBIAS} = \frac{\sum_{i=1}^{n} (\mathrm{pi} - \mathrm{Oi}) * 100}{\sum_{i=1}^{n} \mathrm{Oi}} \quad (5)$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\mathrm{pi} - \mathrm{Oi})^2}{N}} \quad (6)$$

$$\mathrm{MAPE} = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{Oi - Pi}{Oi} \right| \quad (7)$$

$$\mathrm{r} = \frac{\sum_{i=1}^{n} (\mathrm{x_i} - \overline{\mathrm{x}})(\mathrm{y_i} - \overline{\mathrm{y}})}{\sqrt{\sum_{i=1}^{n} (\mathrm{x_i} - \overline{\mathrm{x}})^2} \sqrt{\sum_{i=1}^{n} (\mathrm{y_i} - \overline{\mathrm{y}})^2}} \quad (8)$$

where Pi is the i-th predicted value of the evaluated element, O_i is the i-th observed value, P is the average of the predicted values, O is the average of the observed values, and n is the total number of observations.

The flow framework of this study is shown in Fig. 2. The idea of this study is to first build a database based on multi-source spatial and remote sensing data. Next, a random forest model is used as a bridge to connect ground sample observation data and spatial data to address the system bias caused by scale differences between different data. Simultaneously establish two sets of random forest models to simulate NPP and Rs, so that the scale of simulated carbon input and carbon output data matches. Finally, multiple indicators are used to evaluate the simulation accuracy of different models.

Results

NPP estimation of Qinghai-Tibet Plateau

The carbon flux estimation model of NPP and Rs is constructed by using the random forest model. The simulation results of the model are reliable and accurate (Additional file 1: Figs. S2 and S3). The R^2 value of the random forest model based on 394 NPP ground samples in the Qinghai-Tibet Plateau is 0.88. The observed and simulated data are basically near the 1:1 line, and the slope of the trend line is 0.81 ($P < 0.001$). The R^2 value of the random forest model based on 62 Rs ground samples in the Qinghai-Tibet Plateau is 0.83, the observed and simulated data are basically near the 1:1 line, and the slope of the trend line is 0.74 ($P < 0.001$). The results show that the random forest constructed by combining multiple-source data such as remote sensing

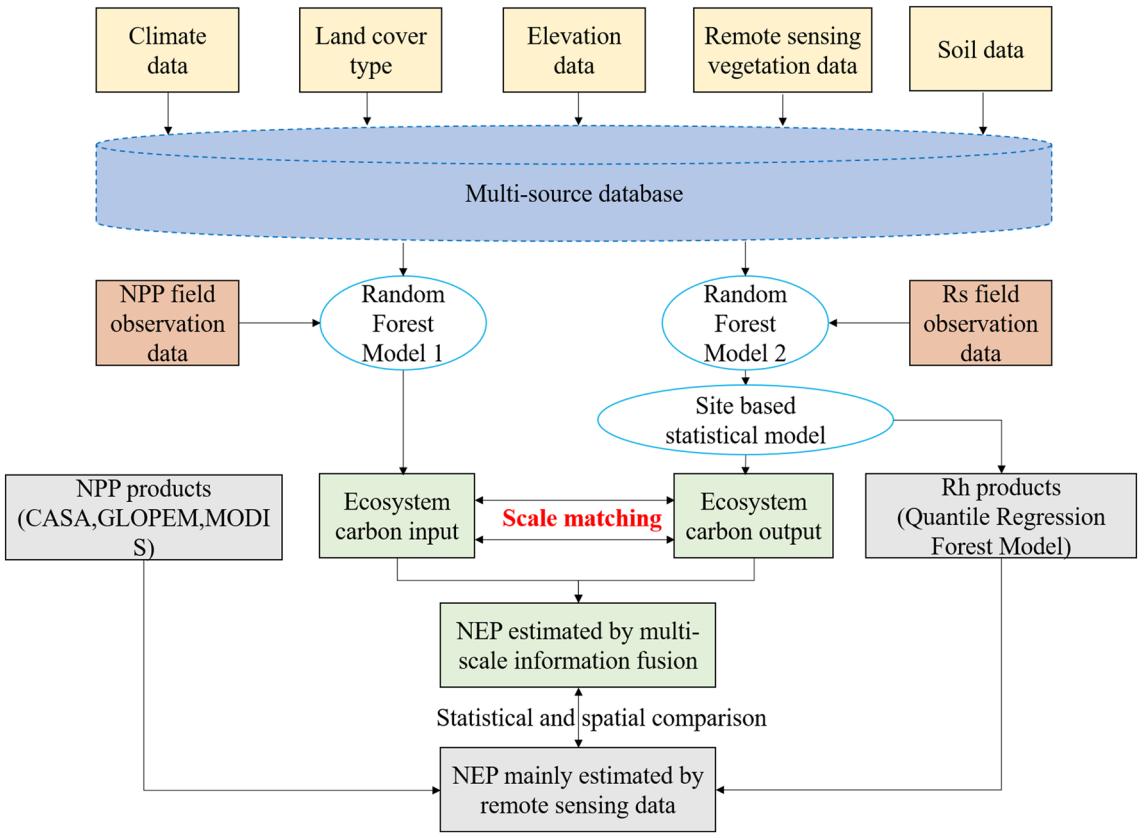

Fig. 2 The flow framework of this study

and ground observation can effectively estimate the key carbon flux of the Qinghai-Tibet Plateau.

Table 1 compares the NPP simulation results of the four models on the Qinghai-Tibet Plateau. In order to fairly compare the simulation results of each model, we extracted the NPP values of each model based on the spatial location of the observation data. The NPP results of the Qinghai-Tibet Plateau simulated by the random forest model are superior to the other three models. The correlation coefficient between the NPP simulated by the random forest model and the observed data of the ground sample points is 0.42, while the correlation coefficient between the MODIS product, GLOPEM model and CASA model and the observed NPP is 0.38, 0.24 and 0.23, respectively. From the root mean square error, the results of random forest simulation improved 18.32% on average compared with the other three models. From the percentage of deviation and the average absolute percentage error, the random forest model is also more suitable for the Qinghai-Tibet Plateau than the other three models.

By comparing the NPP spatial pattern of the Qinghai-Tibet Plateau simulated by different models, we find that MODIS NPP products, CASA models and GLOPEM have a large underestimation of the NPP of the Qinghai-Tibet Plateau (Fig. 3). The NPP results of random forest simulation are generally higher than 100 g C/m^2 year, while the NPP results of the other three models are generally lower than 100 g C/m^2 year in the vast central and western regions of the Qinghai-Tibet Plateau, especially the GLOPEM model is lower than 50 g C/m^2 year. We also compared the spatial differences between the random forest and the other three models. The NPP of the random forest simulation is higher than the other three models in the central and western regions of the Qinghai-Tibet Plateau, while there are differences in a small part of the eastern regions with relatively high vegetation coverage. In general, the NPP products of MODIS, the NPP simulated by CASA model and GLOPEM model are underestimated to a large extent, which may seriously affect the accuracy of carbon sink estimation in the Qinghai-Tibet Plateau.

Rs estimation of Qinghai-Tibet Plateau

Table 2 shows the simulation results of the two models for the Rs of the Qinghai-Tibet Plateau. In order to compare the simulation results of the two models fairly, we still extract the spatial results of the Rs of the two models based on the coordinates of the observed data. The results of Rs simulated by random forest model in Qinghai-Tibet Plateau are better than that of quantile regression forest model. The correlation coefficient between the Rs simulated by the random forest model and the ground sample observation data is 0.70, while the correlation coefficient between the quantile regression forest model and the observed Rs is 0.65. From the root mean square error, the results of random forest simulation improved 7.31% on average compared with the quantile regression forest model. From the percentage of deviation and the average absolute percentage error, the random forest model is also more suitable for the Qinghai-Tibet Plateau than the quantile regression forest model, but the difference is not as obvious as NPP.

By comparing the spatial pattern of Rs in the Qinghai-Tibet Plateau simulated by the two models (Fig. 4), we find that the quantile regression forest model has a large underestimation of Rs in the Qinghai-Tibet Plateau. The Rs simulated by the random forest model have obvious east–west differentiation characteristics in the Qinghai-Tibet Plateau, and generally higher than 500 g C/m^2 year in the east. The Rs simulated by quantile regression forest model also have east–west differentiation characteristics in the Qinghai-Tibet Plateau, but the difference is smaller than that of the random forest model. The Rs in most areas of the east are between 300 and 500 g C/m^2 year. Comparing the spatial difference between the two models, it was found that the quantile regression forest models in most areas of the Qinghai-Tibet Plateau underestimated Rs.

NEP estimation of Qinghai-Tibet Plateau

Based on NPP and Rs simulated by various models, we compared the spatial pattern of carbon source and sink (NEP) in the Qinghai-Tibet Plateau estimated by four pairs of models and their differences (Fig. 5). The NEP simulated by the random forest model shows that the

Table 1 Comparison of four NPP models

NPP	MODIS	GLOPEM	CASA	Mean	Random forest	Difference (%)
r	0.38	0.24	0.23	0.28	0.42	+49.00
RMSE	258.62	323.99	264.37	282.33	230.60	+18.32
PBIAS	− 28.09	− 54.62	− 35.33	− 39.35	− 25.74	+34.58
MAPE	88.61	105.08	81.18	91.62	68.41	+25.33

r is the correlation coefficient, RMSE is the root mean square error, PBIAS is the percentage of deviation, MAPE is the average absolute percentage error, Mean is the average value of NPP simulated by MODIS, GLOPEM and CASA models, and Difference is the percentage of difference between Random Forest and Average

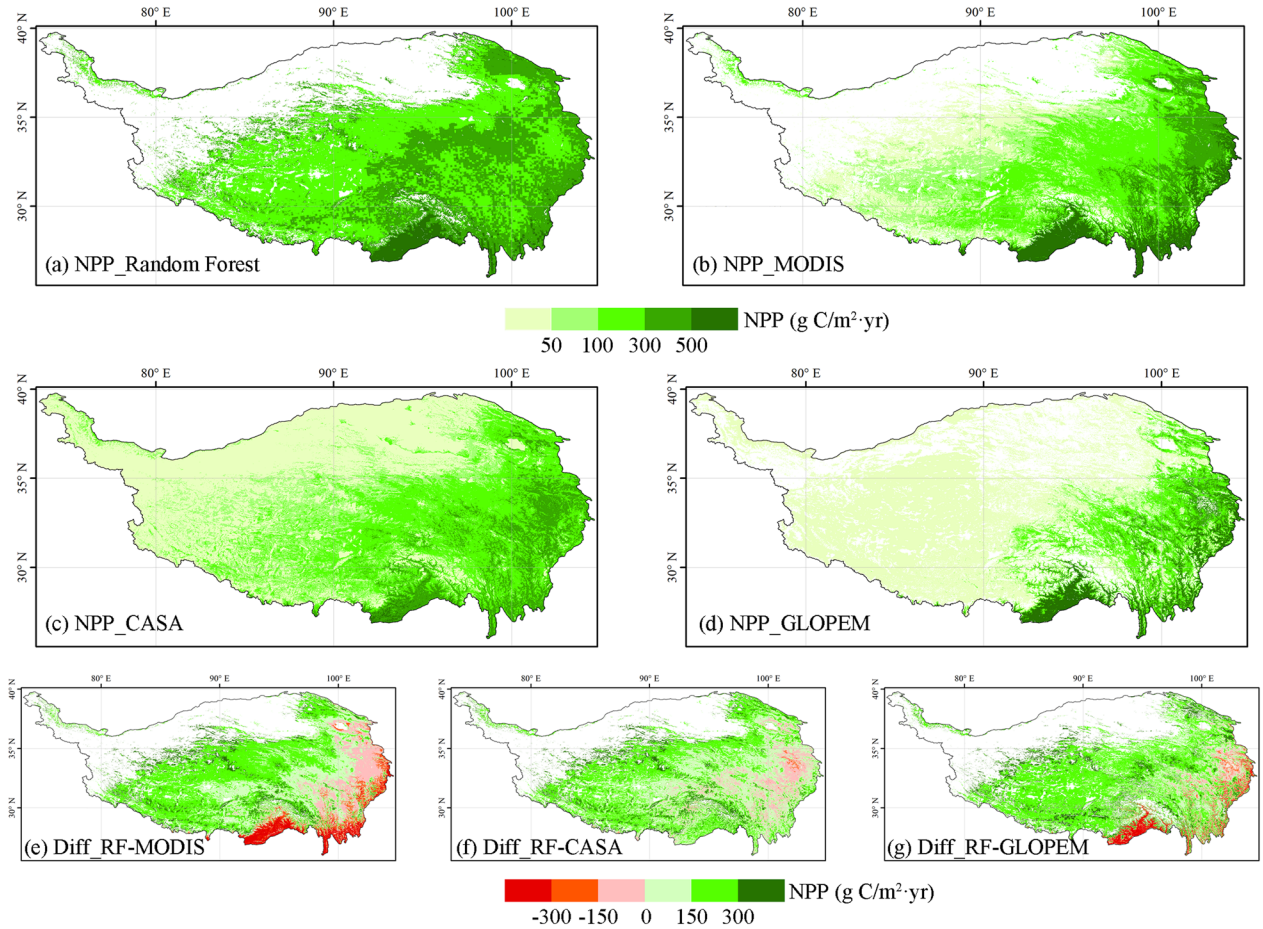

Fig. 3 Four models simulate the spatial pattern of NPP and its differences. **a** NPP of random forest simulation; **b** NPP products of MODIS; **c** NPP simulated by CASA; **d** NPP simulated by GLOPEM; **e** NPP difference between random forest and MODIS; **f** NPP difference between random forest and CASA; **g** NPP difference between random forest and GLOEPM

Table 2 Comparison of two Rs models

Rs	Quantile regression forest model	Random forest	Difference (%)
r	0.65	0.70	7.31
RMSE	0.73	0.59	18.92
PBIAS	− 5.30	− 0.31	94.17
MAPE	15.98	10.96	31.43

r is the correlation coefficient, RMSE is the root mean square error, PBIAS is the percentage of deviation, MAPE is the average absolute percentage error, and Difference is the percentage of difference between Random Forest and Quantitative Regression Forest Model

spatial pattern of carbon sinks in the Qinghai-Tibet Plateau presents a law of east–west differentiation, and the spatial heterogeneity is high. The west is generally a carbon sink area, while the east shows a carbon source. The other three pairs of model results show that the Qinghai-Tibet Plateau is a carbon source in general, and only the area with high vegetation coverage in the eastern part of the Qinghai-Tibet Plateau is a carbon sink. The results of spatial differences show that the carbon sink of the Qinghai-Tibet Plateau is underestimated to a large extent by the other three pairs of models except the random forest model. Most regions underestimate more than 50 g C/m^2 year, which may lead to the misjudgment of whether the Qinghai-Tibet Plateau is a carbon sink or a carbon source.

Figure 6 shows that the difference in NPP estimates of the four models is greater than Rs, and there are differences in the amount and direction of carbon sink estimates of the Qinghai-Tibet Plateau. For NPP, the random forest model estimates an average of 321.4 g/Cm^2 year, while other models underestimate it between 30 and 70%. For Rs, the average estimate of random forest model is about 255.3 g/Cm^2 year, and the estimate of quantile regression forest model is relatively low, with an underestimation of about 5%. The random forest model estimates

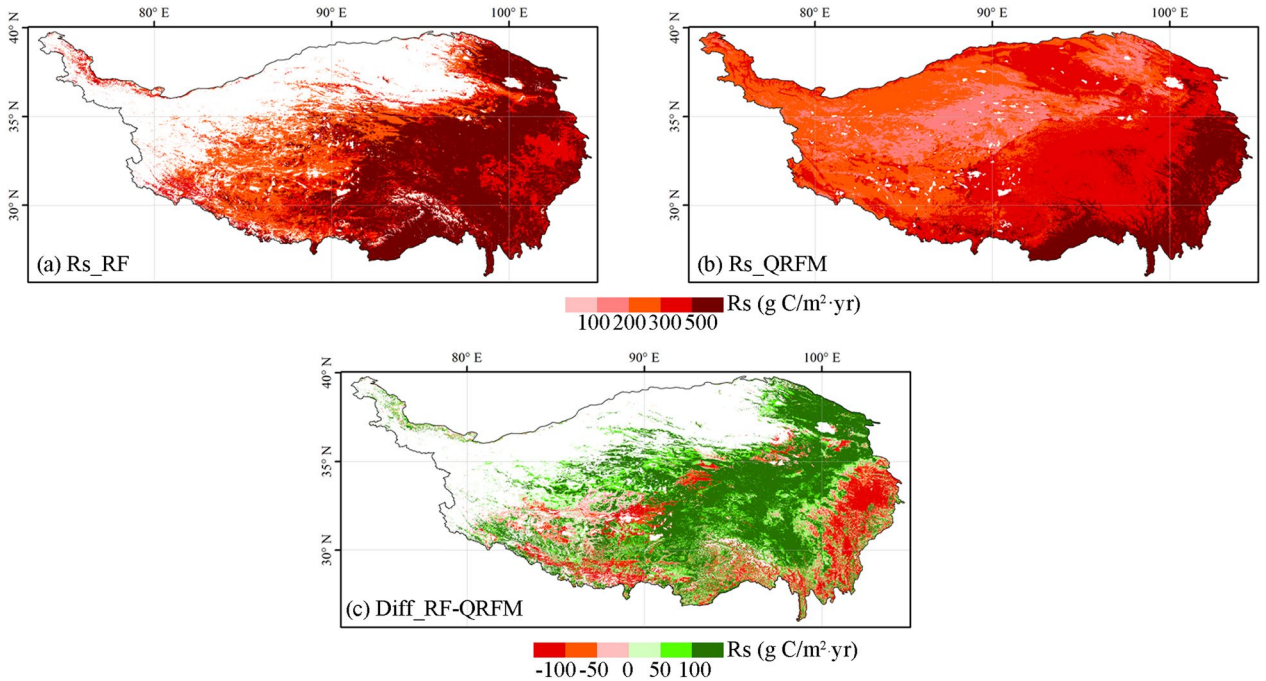

Fig. 4 The two models simulate the spatial distribution pattern of Rs and their differences. **a** Rs simulated by random forest model; **b** Rs of QRFM; **c** Rs difference between random forest and QRFM

that the Qinghai-Tibet Plateau is a carbon sink in general, and will absorb 22.3 Tg of C per year (the average carbon sink intensity is 68.06 g/Cm^2 year). The other three models believe that the Qinghai-Tibet Plateau is a carbon source area in general. The GLOPEM estimation results show that the Qinghai-Tibet Plateau may emit 144.1 Tg C (average carbon source intensity is 128.16 g/Cm^2 year) per year, the CASA model estimates that the Qinghai-Tibet Plateau will emit about 120 Tg C (average carbon source intensity is 119.65 g/Cm^2 year) per year, and the MODIS product estimates that the Qinghai-Tibet Plateau will emit about 36.7 Tg C (average carbon source intensity is 25.75 g/Cm^2 year) per year. In general, the accuracy of carbon flux estimation of the Qinghai-Tibet Plateau by different models is quite different.

Discussion

Comparison of carbon source and sink simulation methods

The top-down atmospheric retrieval method can estimate the real-time changes of carbon sources and sinks worldwide [60], but its spatial resolution is low, and the retrieval accuracy is limited by the number and distribution of atmospheric observation stations, which cannot distinguish between man-made carbon emissions and natural process carbon emissions [61]. The bottom-up ecosystem process model method can simulate all aspects of the carbon cycle process, and has strong physical interpretation ability [62], but the structure and parameters of the model have great uncertainty [63]. The assumption of ecosystem model is usually based on carbon balance, while in the actual natural environment, the ecosystem is usually in a state of non-equilibrium disturbance, especially under the dual impact of human activities and climate change [64–66]. In addition, the widely used carbon source and sink estimation method based on carbon flux sites or ecological sample sites has high accuracy and can realize long-term continuous positioning observation of ecosystem carbon flux on a fine time scale [67, 68]. However, this method is greatly affected by the terrain and meteorological conditions, especially in the Qinghai-Tibet Plateau region. In the process of scaling up, there are problems such as insufficient sample representation, high cost and large spatial heterogeneity [39]. Satellite remote sensing products are increasingly used in terrestrial ecosystem carbon sink estimation because they can identify different underlying surface conditions, are suitable for large-scale carbon cycle monitoring, have strong real-time updating capability, high accuracy and spatial continuity, and have unparalleled cost-effectiveness [69–71]. However, there is usually a scale difference between the ecosystem carbon flux estimated by remote sensing data and the carbon flux estimated based on sample sites, and there is a systematic deviation between the estimates based on remote sensing products and the results estimated by sample sites [72]. Direct comparison of carbon input and output of ecosystems with scale

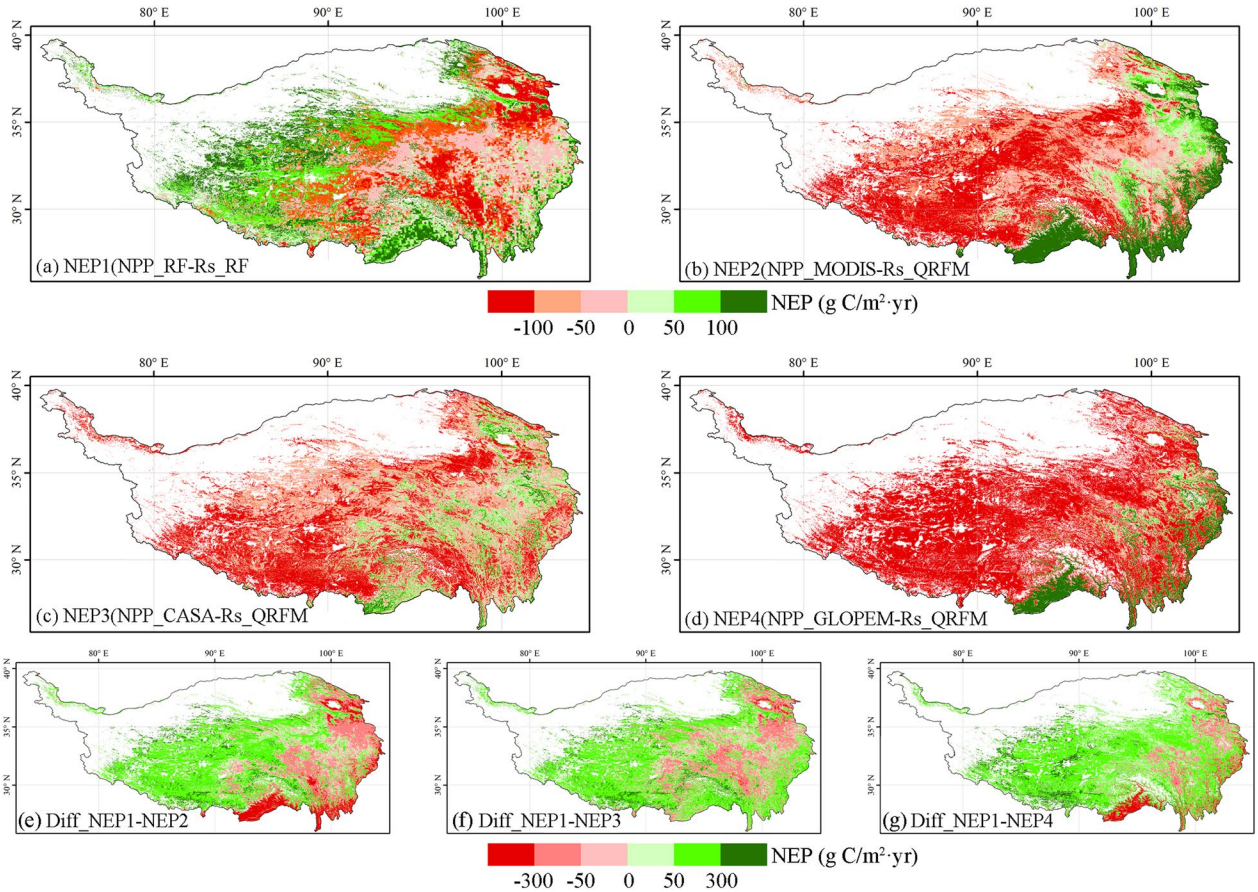

Fig. 5 Various models simulate the spatial pattern of NEP and its differences. **a** NEP1, NPP simulated by random forest model minus Rh simulated by random forest and linear regression model; **b** NEP2, NPP of MODIS minus Rh simulated by QRFM and linear regression model; **c** NEP3, NPP simulated by CASA minus Rh simulated by QRFM and linear regression model; **d** NEP4, NPP simulated by GLOPEM minus Rh simulated by QRFM and linear regression model; **e** Difference between NEP1 and NEP2; **f** Difference between NEP1 and NEP3; **g** Difference between NEP1 and NEP4

differences has great uncertainty in the estimation of ecosystem carbon sink, which is one of the main reasons for the large difference in the estimation results of different methods [39, 72].

Figures 3 and 4 show that other models have different degrees of underestimation of NPP and Rs compared with random forest models. One reason is that most of the other models use remote sensing data as the main input source. Although remote sensing data can ensure spatial continuity, the ability to capture details of the Qinghai-Tibet Plateau, a region with complex terrain and sparse vegetation, is insufficient [32, 73, 74]. Therefore, remote sensing data in the Qinghai-Tibet Plateau is more suitable as a supplementary source of information than the main data [75]. Another important reason is that compared with the random forest model, these models lack the key ground sample observation information as a supplement. The Qinghai-Tibet Plateau has a complex terrain and is extremely vulnerable to the impact of climate change. In recent years, the intensity of human activities has gradually increased [12]. The traditional empirical model can no longer meet the requirements for the simulation and assessment of the carbon source and sink of the ecosystem in the region, and it is more necessary to supplement the observation information of the ground sample points to improve the reliability and applicability of the model [75]. Therefore, we propose a multiple scale information fusion method for carbon sink estimation that uses machine learning as a bridge to connect multiple-source data such as remote sensing and ground sample observation data. This method can not only achieve a wide range of carbon source and sink spatial simulation, but also ensure a high accuracy, so it is more suitable for the Qinghai-Tibet Plateau than other models. Machine learning can explore the nonlinear relationships between multi-source data, provide more accurate simulation results, and have stronger inclusiveness towards the data itself. It also has the characteristics of easy operation, economy, and applicability to multiple scales [76]. Therefore, compared to traditional carbon

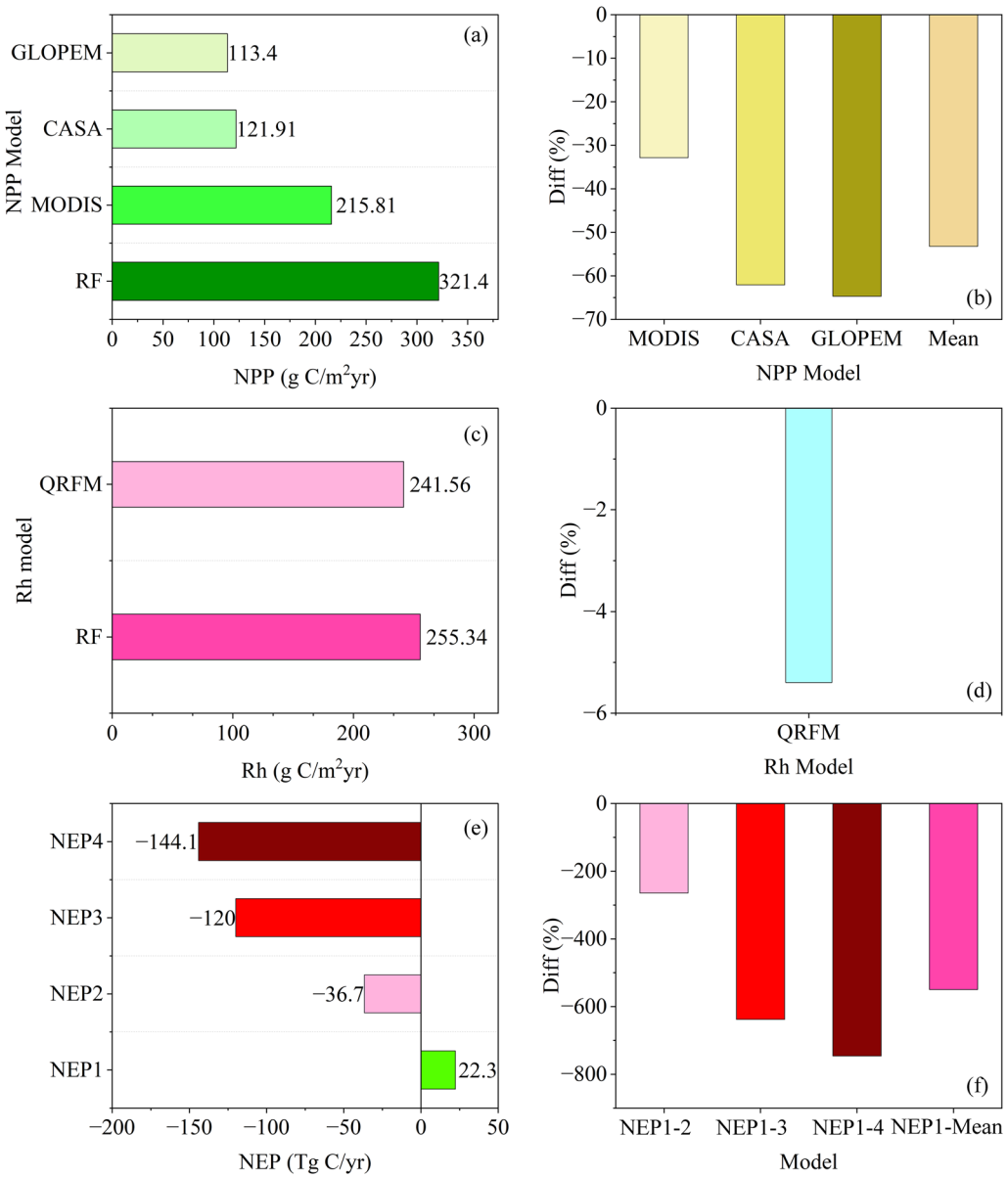

Fig. 6 Various models simulate the carbon sink and its component differences in the Qinghai-Tibet Plateau. **a** NPP simulated by four models; **b** The percentage difference between the mean NPP of random forest simulation and that of other three models; **c** Rh simulated by random forest model and quantile regression forest model; **d** The percentage difference of Rh mean simulated by random forest model and quantile regression forest model; **e** NEP simulated by multiple combination models; **f** Percentage of difference between NEP1 and NEP2, NEP3, NEP4 and their mean values

sink estimation methods, it is more suitable for carbon sink estimation in the Qinghai Tibet Plateau.

Estimation difference of carbon sink in Qinghai-Tibet Plateau

The carbon sink estimates of the Qinghai-Tibet Plateau vary greatly among different models (Fig. 5). It is estimated that the Qinghai-Tibet Plateau is a carbon source or weak carbon sink through remote sensing products and simple linear models [77, 78]. The carbon source level of the Qinghai-Tibet Plateau is estimated to be 39.05 g/Cm^2 year by using the model estimated by remote sensing products [32]. Due to the lack of ground sample data to optimize and constrain the model, these methods have produced a large deviation in the estimation of carbon sink results, especially in the Qinghai-Tibet Plateau, a region with complex terrain [79]. The study of combining the ground sample observation data with the model

to constrain it shows that the Qinghai-Tibet Plateau is a carbon sink in general, with a carbon sink level of about 43.16 Tg C/year, and its carbon sink capacity may weaken in the future with the impact of climate change [36]. At the same time, we also emphasize that for the region with complex conditions such as the Qinghai-Tibet Plateau, more ground sample data are needed to improve the model of carbon sink estimation on the Qinghai-Tibet Plateau, otherwise the results may have a large deviation [80]. For example, only 46 observation data were used to estimate the terrestrial ecological carbon sink in China, and only 4 observations were made in the Qinghai-Tibet Plateau area, and the results obtained were more than 5 times higher than those of other scholars [39, 81].

Our results show that the Qinghai-Tibet Plateau is a carbon sink, about 22.3 Tg C/year (Fig. 6), which is also supported by relevant studies [30, 34, 82]. For example, a large number of studies based on field observations show that the carbon sink of China's terrestrial ecosystem is 0.2–0.3 Pg C/year [1, 39, 68], and the Qinghai-Tibet Plateau is a carbon sink in general, accounting for about 10% of China's carbon sink [62, 83]. The main reason for the difference in the estimation of carbon sinks in the Qinghai-Tibet Plateau is the lack of sufficient ground sample observation data to optimize and constrain the model. The lack of observation data or a small amount of observation data may cause a large deviation in the estimation results [79, 80]. Second, the applicability of remote sensing data in the Qinghai-Tibet Plateau is insufficient. Due to the complex underlying surface conditions of the Qinghai-Tibet Plateau itself, the remote sensing data may not be able to completely retrieve the vegetation level of the Qinghai-Tibet Plateau, resulting in a serious underestimate of NPP [32, 84]. Therefore, the multiple scale information fusion carbon sink estimation method designed by us, which uses machine learning as a bridge to connect multiple-source data such as remote sensing and ground sample observation data, can not only preserve the spatial continuity of remote sensing data, capture spatial details, but also combine the ground sample observation data to optimize and constrain the model, making the simulation results more accurate.

Impacts, limitations, and prospects

In this study, we designed a multiple scale information fusion method for carbon sink estimation that uses machine learning as a bridge to connect multiple-source data such as remote sensing and ground sample observation data. This method can not only preserve the spatial continuity of remote sensing data, capture spatial details, but also combine the ground sample observation data to optimize and constrain the model to obtain a set of spatially continuous, statistically reliable carbon input and output data, Make the simulation results more accurate. Then, based on the carbon input and output data of scale matching, the carbon sink level and its spatial pattern of the Qinghai-Tibet Plateau from 2000 to 2018 are estimated. In addition, we quantitatively evaluated the carbon sink results simulated by various models, and discussed the differences between the models. This study answers the controversial question of whether the Qinghai-Tibet Plateau is a carbon source or a carbon sink, and emphasizes the importance of carbon process simulation in areas with complex terrain and vulnerable to climate change, such as the Qinghai-Tibet Plateau, combined with ground sample observation data. The research results provide an important basis for accurately estimating the carbon sink and its spatial distribution in the Qinghai-Tibet Plateau, and provide a scientific reference for helping China achieve "carbon neutrality" and sustainable development by 2060.

This study uses multiple-source data, including climate-driven data sets, MODIS and related vegetation remote sensing products, and ground sample observations. Remote sensing products are affected by mixed pixels, and the fusion process of spatial data at different scales may affect the carbon sink estimation results [69]. In addition, although machine learning can explore nonlinear relationships between data, it has high requirements for the data itself, and abnormal data may have a significant impact on the results. Therefore, in the future, more and more representative samples are needed to further improve the reliability of the model [76]. Scaling the site data to the regional level places higher demands on the homogeneity of the samples, especially in the complex terrain of the Qinghai Tibet Plateau, where the uneven distribution of spatial sample points may lead to uncertainty in carbon si60nk estimation results in certain regions [41].

This study compares the key carbon fluxes of the Qinghai-Tibet Plateau simulated by various models, revealing that the Qinghai-Tibet Plateau is an important carbon sink. In the future, in-depth evaluation of the response mechanism of carbon sinks and related environmental factors (climate and land use change) may be a promising work [85, 86]. Especially in the context of global warming, global warming may cause the melting of permafrost in the Qinghai Tibet Plateau to release more soil carbon, while increasing water content may also promote vegetation photosynthesis, resulting in greater uncertainty in the estimation of carbon sources and sinks in the Qinghai Tibet Plateau region [87, 88]. In addition, it is necessary to combine the CMIP6 model to estimate carbon sink changes in different climate change and human activity scenarios on the Qinghai Tibet Plateau, especially to conduct

research on the impact assessment of carbon sinks after a series of ecological projects, which can further provide basic support for related research on how ecosystems respond to climate change and has important practical significance [37, 73, 89].

Conclusions

This study designed a multiple scale information fusion method for carbon sink estimation, which uses machine learning as a bridge to connect multiple-source data such as remote sensing and ground sample observation data. This method can not only preserve the spatial continuity of remote sensing data, but also combine the ground sample observation data to optimize and constrain the model, improve the simulation accuracy, so as to obtain a group of spatially continuous, statistically reliable carbon input and carbon output data. The results derived by the fusion of multiple scale data revealed that the Qinghai-Tibet Plateau, as an important ecological security barrier for China and even East Asia, has a carbon sink of about 22.3 Tg C/year. Compared with the random forest model, the traditional CASA, GLOPEM and MODIS products underestimated the carbon input of the Qinghai-Tibet Plateau by 30–70%, while the quantile regression forest model underestimated the carbon output of the Qinghai-Tibet Plateau by about 5%. These models underestimate carbon input more than carbon output, and may mistake the Qinghai-Tibet Plateau as a carbon source. This study provides an important basis for accurately estimating the carbon sink and its spatial distribution in the Qinghai-Tibet Plateau, and provides a scientific reference for helping China achieve "carbon neutrality" and sustainable development by 2060.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13021-023-00239-9.

Additional file 1: Figure S1. Scatter fitting of Rs and Rh based on observation data. **Figure S2.** Scatter fitting of NPP of random forest simulation and ground sample observation data. The green scatter points in the figure represent the training set of the random forest model, the red scatter points represent the verification set of the random forest model, the black dotted line represents the 1:1 line, and the blue color is realized as the fitting line. **Figure S3.** Scatter fitting of Ln (Rs) of random forest simulation and ground sample observation data. The green scatter points in the figure represent the training set of the random forest model, the red scatter points represent the verification set of the random forest model, the black dotted line represents the 1:1 line, and the blue line is the fitting line

Additional file 2. NPP/Rs/Rh ground observation sample information.

Acknowledgements
We thank all the data providers for their field observations, environmental datasets and climate datasets. Special thanks to the editor and several anonymous reviewers for their help and support.

Author contributions
JYZ and TZ conceived and designed the study. JYZ performed material preparation, interpreted the results and wrote the original manuscript. TZ supervised this project. YXX, QYL, and ET collected the data. YJZ, XMW, JZZ and XL edited the manuscript. All authors read and approved the final manuscript.

Funding
This work was supported by the Second Tibetan Plateau Scientific Expedition and Research Program (2019QZKK0405-02), the National Natural Science Foundation of China (42277206), and the Key Laboratory of Environmental Change and Natural Disasters of Ministry of Education, Beijing Normal University (2022-KF-15).

Availability of data and materials
The data that support the findings of this study were derived from the following resources available in the public domain: NPPMOSID dataset is available from https://e4ftl01.cr.usgs.gov/MOLT/MOD17A3HGF.006/. NPPCASA dataset can be downloaded from https://doi.org/10.6084/m9.figshare.20387889.v1. NPPGLO_PEM dataset can be downloaded from https://doi.org/10.6084/m9.figshare.20387874.v1. MOD12Q1 dataset can be downloaded from https://e4ftl01.cr.usgs.gov/MOTA/MCD12Q1.006/. Meteorological-driven dataset can be downloaded from https://doi.org/10.6084/m9.figshare.c.4557599.v1. The leaf area index is accessed from https://doi.org/10.6084/m9.figshare.20387928.v1. SRTMDEMUTM DEM is accessed from http://data.tpdc.ac.cn/en/data/acb49ce8-2bfe-4ab4-97ff-e6e727110703/. Global soil organic carbon dataset can be downloaded from https://soilgrids.org/. VHI can be downloaded from https://doi.org/10.6084/m9.figshare.19811854.v3. Soil respiration and Heterotrophic respiration dataset https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1736.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Key Laboratory of Environmental Change and Natural Disasters of Chinese Ministry of Education, Beijing Normal University, Beijing 100875, China. [2]State Key Laboratory of Earth Surface Processes and Resource Ecology (ESPRE), Beijing Normal University, Beijing 100875, China.

Received: 20 February 2023 Accepted: 3 September 2023
Published online: 11 September 2023

References

1. Friedlingstein P, et al. Global carbon budget 2019. Earth Syst Sci Data. 2019;11:1783–838. https://doi.org/10.5194/essd-11-1783-2019.
2. Nishina K, et al. Quantifying uncertainties in soil carbon responses to changes in global mean temperature and precipitation. Earth Syst Dyn. 2014;5:197–209. https://doi.org/10.5194/esd-5-197-2014.
3. Schimel D, Stephens BB, Fisher JB. Effect of increasing CO2 on the terrestrial carbon cycle. Proc Natl Acad Sci USA. 2015;112:436–41. https://doi.org/10.1073/pnas.1407302112.
4. Zhu L, et al. Decadal variability in land carbon sink efficiency. Carbon Balance Manage. 2021;16:15. https://doi.org/10.1186/s13021-021-00178-3.
5. Le Quere C, et al. Global carbon budget 2018. Earth Syst Sci Data. 2018;10:2141–94. https://doi.org/10.5194/essd-10-2141-2018.
6. Bastos A, et al. Direct and seasonal legacy effects of the 2018 heat wave and drought on European ecosystem productivity. Sci Adv. 2020. https://doi.org/10.1126/sciadv.aba2724.

7. Green JK, et al. Large influence of soil moisture on long-term terrestrial carbon uptake. Nature. 2019;565:476. https://doi.org/10.1038/s41586-018-0848-x.
8. O'Sullivan M, et al. Have synergies between nitrogen deposition and atmospheric CO2 driven the recent enhancement of the terrestrial carbon sink? Glob Biogeochem Cycles. 2019;33:163–80. https://doi.org/10.1029/2018gb005922.
9. Piao SL, et al. Characteristics, drivers and feedbacks of global greening. Nat Rev Earth Environ. 2020;1:14–27. https://doi.org/10.1038/s43017-019-0001-x.
10. Trugman AT, Medvigy D, Mankin JS, Anderegg WRL. Soil moisture stress as a major driver of carbon cycle uncertainty. Geophys Res Lett. 2018;45:6495–503. https://doi.org/10.1029/2018gl078131.
11. Huntzinger DN, et al. Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions. Sci Rep. 2017. https://doi.org/10.1038/s41598-017-03818-2.
12. Cheng GD, et al. Characteristic, changes and impacts of permafrost on Qinghai-Tibet Plateau. Chin Sci Bull-Chin. 2019;64:2783–95. https://doi.org/10.1360/tb-2019-0191.
13. Piao S, et al. Detection and attribution of vegetation greening trend in China over the last 30 years. Glob Change Biol. 2015;21:1601–9. https://doi.org/10.1111/gcb.12795.
14. Yao T, et al. Recent Third pole's rapid warming accompanies cryospheric melt and water cycle intensification and interactions between monsoon and environment: multidisciplinary approach with observations, modeling, and analysis. Bull Am Meteor Soc. 2019;100:423–44. https://doi.org/10.1175/bams-d-17-0057.1.
15. Liu X, et al. Uncertainties of soil organic carbon stock estimation caused by paleoclimate and human footprint on the Qinghai Plateau. Carbon Balance Manage. 2022;17:8. https://doi.org/10.1186/s13021-022-00203-z.
16. Biskaborn BK, et al. Permafrost is warming at a global scale. Nat Commun. 2019;10:264. https://doi.org/10.1038/s41467-018-08240-4.
17. Song X-D, et al. Significant loss of soil inorganic carbon at the continental scale. Natl Sci Rev. 2021. https://doi.org/10.1093/nsr/nwab120.
18. Wang W, et al. Exploring the ground ice recharge near permafrost table on the central Qinghai-Tibet Plateau using chemical and isotopic data. J Hydrol. 2018;560:220–9. https://doi.org/10.1016/j.jhydrol.2018.03.032.
19. Cheng F, et al. Alpine permafrost could account for a quarter of thawed carbon based on Plio-Pleistocene paleoclimate analogue. Nat Commun. 2022;13:1329. https://doi.org/10.1038/s41467-022-29011-2.
20. DeConto RM, et al. Past extreme warming events linked to massive carbon release from thawing permafrost. Nature. 2012;484:87–91. https://doi.org/10.1038/nature10929.
21. Mu C, et al. The status and stability of permafrost carbon on the Tibetan Plateau. Earth-Sci Rev. 2020;211:103433. https://doi.org/10.1016/j.earscirev.2020.103433.
22. Kato T, et al. Carbon dioxide exchange between the atmosphere and an alpine meadow ecosystem on the Qinghai-Tibetan Plateau, China. Agric For Meteorol. 2004;124:121–34. https://doi.org/10.1016/j.agrformet.2003.12.008.
23. Chen H, et al. The impacts of climate change and human activities on biogeochemical cycles on the Qinghai-Tibetan Plateau. Glob Change Biol. 2013;19:2940–55. https://doi.org/10.1111/gcb.12277.
24. Fu Y, et al. Quantification of year-round methane and nitrous oxide fluxes in a typical alpine shrub meadow on the Qinghai-Tibetan Plateau. Agr Ecosyst Environ. 2018;255:27–36. https://doi.org/10.1016/j.agee.2017.12.003.
25. Li Y, et al. Seasonal changes of CO2, CH4 and N2O fluxes in different types of alpine grassland in the Qinghai-Tibetan Plateau of China. Soil Biol Biochem. 2015;80:306–14. https://doi.org/10.1016/j.soilbio.2014.10.026.
26. Guo JH, Huang GH, Wang XQ, Li YP, Yang L. Future changes in precipitation extremes over China projected by a regional climate model ensemble. Atmos Environ. 2018;188:142–56. https://doi.org/10.1016/j.atmosenv.2018.06.026.
27. Jiang ZY, et al. High-resolution projections of mean and extreme precipitation over china by two regional climate models. J Meteorol Res. 2020;34:965–85. https://doi.org/10.1007/s13351-020-9208-5.
28. Song CL, et al. Net ecosystem carbon budget of a grassland ecosystem in central Qinghai-Tibet Plateau: integrating terrestrial and aquatic carbon fluxes at catchment scale. Agric Forest Meteorol. 2020. https://doi.org/10.1016/j.agrformet.2020.108021.
29. Zhuang Q, et al. Carbon dynamics of terrestrial ecosystems on the Tibetan Plateau during the 20th century: an analysis with a process-based biogeochemical model. Glob Ecol Biogeogr. 2010;19:649–62. https://doi.org/10.1111/j.1466-8238.2010.00559.x.
30. Piao S, et al. Impacts of climate and CO2 changes on the vegetation growth and carbon balance of Qinghai-Tibetan grasslands over the past five decades. Global Planet Change. 2012;98–99:73–80. https://doi.org/10.1016/j.gloplacha.2012.08.009.
31. Guo D, et al. Large-scale analysis of the spatiotemporal changes of Net Ecosystem Production in Hindu Kush Himalayan Region. Remote Sens. 2021. https://doi.org/10.3390/rs13061180.
32. Liang L, et al. Remote sensing estimation and spatiotemporal pattern analysis of terrestrial net ecosystem productivity in China. Remote Sens. 2022. https://doi.org/10.3390/rs14081902.
33. Zhao L, et al. Seasonal variations in carbon dioxide exchange in an alpine wetland meadow on the Qinghai-Tibetan Plateau. Biogeosciences. 2010;7:1207–21. https://doi.org/10.5194/bg-7-1207-2010.
34. Ma W, et al. Warming could shift steppes to carbon sinks and meadows to carbon sources in permafrost regions: evidence from the improved IBIS model. CATENA. 2021;200:105168. https://doi.org/10.1016/j.catena.2021.105168.
35. Koven CD, et al. Permafrost carbon-climate feedbacks accelerate global warming. Proc Natl Acad Sci USA. 2011;108:14769–74. https://doi.org/10.1073/pnas.1103910108.
36. Wu TH, et al. Weakening of carbon sink on the Qinghai–Tibet Plateau. Geoderma. 2022. https://doi.org/10.1016/j.geoderma.2022.115707.
37. Selmants PC, et al. Ecosystem carbon balance in the Hawaiian Islands under different scenarios of future climate and land use change. Environ Res Lett. 2021. https://doi.org/10.1088/1748-9326/ac2347.
38. Lu Y, et al. China's terrestrial ecosystem carbon balance during the 20th century: an analysis with a process-based biogeochemistry model. Carbon Balance Manage. 2022;17:16. https://doi.org/10.1186/s13021-022-00215-9.
39. Piao SL, He Y, Wang XH, Chen FH. Estimation of China's terrestrial ecosystem carbon sink: methods, progress and prospects. Sci China-Earth Sci. 2022;65:641–51. https://doi.org/10.1007/s11430-021-9892-6.
40. Zhang Y, Yao YT, Wang XH, Liu YW, Piao SL. Mapping spatial distribution of forest age in China. Earth Space Sci. 2017;4:108–16. https://doi.org/10.1002/2016ea000177.
41. Du H, et al. Responses of autumn vegetation phenology to climate change and urbanization at northern middle and high latitudes. Int J Appl Earth Observ Geoinf. 2022;115:103086. https://doi.org/10.1016/j.jag.2022.103086.
42. Liu ZL, et al. Application of machine-learning methods in forest ecology: recent progress and future challenges. Environ Rev. 2018;26:339–50. https://doi.org/10.1139/er-2018-0034.
43. Ali I, Greifeneder F, Stamenkovic J, Neumann M, Notarnicola C. Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. Remote Sens. 2015;7:16398–421. https://doi.org/10.3390/rs71215841.
44. Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science. 2015;349:255–60. https://doi.org/10.1126/science.aaa8415.
45. Foggin JM. Depopulating the Tibetan grasslands—National policies and perspectives for the future of Tibetan herders in Qinghai Province, China. Mt Res Devt. 2008;28:26–31. https://doi.org/10.1659/mrd.0972.
46. Piao S, et al. Altitude and temperature dependence of change in the spring vegetation green-up date from 1982 to 2006 in the Qinghai-Xizang Plateau. Agric For Meteorol. 2011;151:1599–608. https://doi.org/10.1016/j.agrformet.2011.06.016.
47. Zheng D. The system of physico-geographical regions of the Qinghai-Xizang (Tibet) plateau. Sci China Ser D-Earth Sci. 1996;39:410–7.
48. Jian J, Vargas R, Anderson-Teixeira KJ, Stell E, Herrmann V, Horn M, Kholod N, Manzon J, Marchesi R, Paredes D, Bond-Lamberty BP. A Global Database of Soil Respiration Data, Version 5.0. ORNL DAAC, Oak Ridge, Tennessee, USA. 2021. https://doi.org/10.3334/ORNLDAAC/1827.
49. Chen PF. Monthly NPP dataset covering China's terrestrial ecosystems at North of 18°N (1985–2015). J Glob Change Data Discov. 2019;3:34–41. https://doi.org/10.3974/geodp.2019.01.05.

50. He J, et al. The first high-resolution meteorological forcing dataset for land process studies over China. Sci Data. 2020;7:25. https://doi.org/10.1038/s41597-020-0369-y.
51. Liu Y, Liu R, Chen JM. Retrospective retrieval of long-term consistent global leaf area index (1981–2011) from combined AVHRR and MODIS data. J Geophys Res: Biogeosci. 2012. https://doi.org/10.1029/2012JG002084.
52. Stell E, Warner DL, Jian J, Bond-Lamberty BP, Vargas R. Global Gridded 1-km Soil and Soil Heterotrophic Respiration Derived from SRDB v5. ORNL DAAC, Oak Ridge, Tennessee, USA. 2021. https://doi.org/10.3334/ORNLDAAC/1928.
53. Zeng J, et al. Global terrestrial carbon fluxes of 1999–2019 estimated by upscaling eddy covariance data with a random forest. Sci Data. 2020;7:313. https://doi.org/10.1038/s41597-020-00653-5.
54. Li H, et al. Spatiotemporal distribution and national measurement of the global carbonate carbon sink. Sci Total Environ. 2018;643:157–70. https://doi.org/10.1016/j.scitotenv.2018.06.196.
55. Rodriguez-Galiano V, Mendes MP, Garcia-Soldado MJ, Chica-Olmo M, Ribeiro L. Predictive modeling of groundwater nitrate pollution using Random Forest and multisource variables related to intrinsic and specific vulnerability: a case study in an agricultural setting (Southern Spain). Sci Total Environ. 2014;476–477:189–206. https://doi.org/10.1016/j.scitotenv.2014.01.001.
56. Gulbeyaz O, Bond-Lamberty B, Akyurek Z, West TO. A new approach to evaluate the MODIS annual NPP product (MOD17A3) using forest field data from Turkey. Int J Remote Sens. 2018;39:2560–78. https://doi.org/10.1080/01431161.2018.1430913.
57. Son S, Wang MH, Harding LW. Satellite-measured net primary production in the Chesapeake Bay. Remote Sens Environ. 2014;144:109–19. https://doi.org/10.1016/j.rse.2014.01.018.
58. Potter C, Klooster S, Huete A, Genovese V. Terrestrial carbon sinks for the United States predicted from MODIS satellite data and ecosystem modeling. Earth Interact. 2007;11:1–21. https://doi.org/10.1175/ei228.1.
59. Pan N, Wang S, Wei FL, Shen MG, Fu BJ. Inconsistent changes in NPP and LAI determined from the parabolic LAI versus NPP relationship. Ecol Indic. 2021;131:108134. https://doi.org/10.1016/j.ecolind.2021.108134.
60. Fernandez-Martinez M, et al. Global trends in carbon sinks and their relationships with CO2 and temperature. Nat Clim Change. 2019;9:73. https://doi.org/10.1038/s41558-018-0367-7.
61. Jiang F, et al. A comprehensive estimate of recent carbon sinks in China using both top-down and bottom-up approaches. Sci Rep. 2016. https://doi.org/10.1038/srep22130.
62. Piao SL, et al. The carbon balance of terrestrial ecosystems in China. Nature. 2009;458:1009-U1082. https://doi.org/10.1038/nature07944.
63. Zhou T, et al. Age-dependent forest carbon sink: estimation via inverse modeling. J Geophys Res-Biogeosci. 2015;120:2473–92. https://doi.org/10.1002/2015jg002943.
64. Zhou T, Shi PJ, Jia GS, Luo YQ. Nonsteady state carbon sequestration in forest ecosystems of China estimated by data assimilation. J Geophys Res-Biogeosci. 2013;118:1369–84. https://doi.org/10.1002/jgrg.20114.
65. Zhu K, Zhang J, Niu S, Chu C, Luo Y. Limits to growth of forest biomass carbon sink under climate change. Nat Commun. 2018;9:2709. https://doi.org/10.1038/s41467-018-05132-5.
66. Ge R, et al. Underestimated ecosystem carbon turnover time and sequestration under the steady state assumption: a perspective from long-term data assimilation. Glob Change Biol. 2019;25:938–53. https://doi.org/10.1111/gcb.14547.
67. Heiskanen J, et al. The integrated carbon observation system in Europe. Bull Am Meteor Soc. 2022;103:E855–72. https://doi.org/10.1175/bams-d-19-0364.1.
68. Fang JY, Yu GR, Liu LL, Hu SJ, Chapin FS. Climate change, human impacts, and carbon sequestration in China. Proc Natl Acad Sci USA. 2018;115:4015–20. https://doi.org/10.1073/pnas.1700304115.
69. Smith WK, et al. Remote sensing of dryland ecosystem structure and function: progress, challenges, and opportunities. Remote Sens Environ. 2019. https://doi.org/10.1016/j.rse.2019.111401.
70. Espirito-Santo FDB, et al. Size and frequency of natural forest disturbances and the Amazon forest carbon balance. Nat Commun. 2014. https://doi.org/10.1038/ncomms4434.
71. Zeng J, et al. Improving the drought monitoring capability of VHI at the global scale via ensemble indices for various vegetation types from 2001 to 2018. Weather Clim Extremes. 2022;35:100412. https://doi.org/10.1016/j.wace.2022.100412.
72. Yang YH, et al. Terrestrial carbon sinks in China and around the world and their contribution to carbon neutrality. Sci China-Life Sci. 2022;65:861–95. https://doi.org/10.1007/s11427-021-2045-5.
73. Guo P, et al. The influence of temperature and precipitation on the vegetation dynamics of the tropical island of Hainan. Theoret Appl Climatol. 2021;143:429–45. https://doi.org/10.1007/s00704-020-03430-x.
74. Kong R, et al. Projected global warming-induced terrestrial ecosystem carbon across China under SSP scenarios. Ecol Indic. 2022. https://doi.org/10.1016/j.ecolind.2022.108963.
75. Kang XM, et al. Precipitation and temperature regulate the carbon allocation process in alpine wetlands: quantitative simulation. J Soils Sediments. 2020;20:3300–15. https://doi.org/10.1007/s11368-020-02643-x.
76. Zeng J, et al. Spatial patterns of China's carbon sinks estimated from the fusion of remote sensing and field-observed net primary productivity and heterotrophic respiration. Ecol Inf. 2023;76:102152. https://doi.org/10.1016/j.ecoinf.2023.102152.
77. Chuai X, et al. Land degradation monitoring using terrestrial ecosystem carbon sinks/sources and their response to climate change in China. Land Degrad Dev. 2018;29:3489–502. https://doi.org/10.1002/ldr.3117.
78. Ye X, Chuai X. Carbon sinks/sources? Spatiotemporal evolution in China and its response to built-up land expansion. J Environ Manag. 2022. https://doi.org/10.1016/j.jenvman.2022.115863.
79. Zhou DC, et al. Potential impacts of climate change on vegetation dynamics and ecosystem function in a mountain watershed on the Qinghai-Tibet Plateau. Clim Change. 2019;156:31–50. https://doi.org/10.1007/s10584-019-02524-4.
80. Wang Y, et al. The size of the land carbon sink in China. Nature. 2022;603:E7–9. https://doi.org/10.1038/s41586-021-04255-y.
81. Yao YT, et al. A new estimation of China's net ecosystem productivity based on eddy covariance measurements and a model tree ensemble approach. Agric For Meteorol. 2018;253:84–93. https://doi.org/10.1016/j.agrformet.2018.02.007.
82. Song C, Wang G. Land carbon sink of the Tibetan Plateau may be overestimated without accounting for the aquatic carbon export. Proc Natl Acad Sci. 2021;118:e2114694118. https://doi.org/10.1073/pnas.2114694118.
83. Piao SL, et al. Responses and feedback of the Tibetan Plateau's alpine ecosystem to climate change. Chin Sci Bull-Chin. 2019;64:2842–55. https://doi.org/10.1360/tb-2019-0074.
84. Filippa G, et al. On the distribution and productivity of mountain grasslands in the Gran Paradiso National Park, NW Italy: a remote sensing approach. Int J Appl Earth Observ Geoinf. 2022. https://doi.org/10.1016/j.jag.2022.102718.
85. Huang BB, et al. Ecological restoration and rising CO2 enhance the carbon sink, counteracting climate change in northeastern China. Environ Res Lett. 2022. https://doi.org/10.1088/1748-9326/ac3871.
86. He H, et al. Large-scale estimation and uncertainty analysis of gross primary production in Tibetan alpine grasslands. J Geophys Res Biogeosci. 2014;119:466–86. https://doi.org/10.1002/2013JG002449.
87. Wu ZT, et al. Recent changes in the drought of China from 1960 to 2014. Int J Climatol. 2020;40:3281–96. https://doi.org/10.1002/joc.6397.
88. Zeng J, et al. Drought frequency characteristics of China, 1981–2019, based on the vegetation health index. Climate Res. 2020;81:131–47. https://doi.org/10.3354/cr01616.
89. Koch A, Hubau W, Lewis SL. Earth system models are not capturing present-day tropical forest carbon dynamics. Earths Future. 2021;9:e2020EF001874. https://doi.org/10.1029/2020EF001874.

Publisher's Note