

Article

Spatial–Temporal Correlation Considering Environmental Factor Fusion for Estimating Gross Primary Productivity in Tibetan Grasslands

Qinmeng Yang¹, Ningming Nie^{1,2}, Yangang Wang^{1,2,*}, Xiaojing Wu^{3,4}, Weihua Liu^{2,3,4}, Xiaoli Ren^{3,4}, Zijian Wang¹, Meng Wan¹ and Rongqiang Cao^{1,2}

¹ Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China; qmyang@cnic.cn (Q.Y.); nienm@sccas.cn (N.N.); wangzj@cnic.cn (Z.W.); wanmengdamon@cnic.cn (M.W.); caorq@sccas.cn (R.C.)

² University of Chinese Academy of Sciences, Beijing 100049, China; liuw.h.20b@igsnr.ac.cn

³ Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; wuxj@igsnr.ac.cn (X.W.); renxl@igsnr.ac.cn (X.R.)

⁴ National Ecosystem Science Data Center, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

* Correspondence: wangyg@sccas.cn; Tel.: +86-13718268975

Abstract: Gross primary productivity (GPP) is an important indicator in research on carbon cycling in terrestrial ecosystems. High-accuracy GPP prediction is crucial for ecosystem health and climate change assessments. We developed a site-level GPP prediction method based on the GeoMAN model, which was able to extract spatiotemporal features and fuse external environmental factors to predict GPP on the Tibetan Plateau. We evaluated four models' behavior—Random Forest (RF), Support Vector Machine (SVM), Deep Belief Network (DBN), and GeoMAN—in predicting GPP at nine flux observation sites on the Tibetan Plateau. The GeoMAN model achieved the best results ($R^2 = 0.870$, $RMSE = 0.788 \text{ g Cm}^{-2} \text{ d}^{-1}$, $MAE = 0.440 \text{ g Cm}^{-2} \text{ d}^{-1}$). Distance and vegetation type of the flux sites influenced GPP prediction, with the latter being more significant. The different grassland vegetation types exhibited different sensitivity to environmental factors (Ta, PAR, EVI, NDVI, and LSWI) for GPP prediction. Among them, the site located in the alpine swamp meadow was insensitive to changes in environmental factors; the GPP prediction accuracy of the site located in the alpine meadow steppe decreased significantly with the changes in environmental factors; and the GPP prediction accuracy of the site located in the alpine Kobresia meadow also varied with environmental factor changes, but to a lesser extent than the former. This study provides a good reference that deep learning model is able to achieve good accuracy in GPP simulation when considers spatial, temporal, and environmental factors, and the judgement made by deep learning model conforms to basic knowledge in the relevant field.

Keywords: deep learning; GeoMAN model; gross primary productivity; attention mechanism; interdisciplinary



Citation: Yang, Q.; Nie, N.; Wang, Y.; Wu, X.; Liu, W.; Ren, X.; Wang, Z.; Wan, M.; Cao, R. Spatial–Temporal Correlation Considering Environmental Factor Fusion for Estimating Gross Primary Productivity in Tibetan Grasslands. *Appl. Sci.* **2023**, *13*, 6290. <https://doi.org/10.3390/app13106290>

Academic Editor: Nathan J Moore

Received: 25 April 2023

Revised: 14 May 2023

Accepted: 18 May 2023

Published: 21 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Gross primary productivity (GPP) is the cumulative sum of organisms produced by plants absorbing CO_2 during photosynthesis [1,2]. It drives the seasonal and annual variations in atmospheric CO_2 concentration, which reflect the production capacity of terrestrial ecosystems under natural conditions [2,3], and is an important indicator for assessing ecosystem health and climate change [4]. Therefore, accurate prediction of GPP is crucial for ecosystem function evaluation and carbon balance research [2].

The common methods for quantifying and predicting GPP are based on processing observational data and process-based model simulation [2,5]. Observational data include

data obtained using the eddy covariance (EC) technique, in which GPP values are obtained by calculating net ecosystem exchange (NEE) from vertical turbulent transport in the atmosphere under meteorological conditions [5,6], and satellite data, which are commonly used for GPP estimation due to their stability and sustainability, such as the MODIS GPP standard product, the VPM model, and the EC-LUE model [2,7–9]. However, satellite GPP products cannot fully guarantee the reliability of data [10], which affects the accuracy of the prediction data and introduces uncertainties to related research. Process-based models mainly investigate and simulate ecological processes occurring in plants and have extensive theoretical foundations in related fields. However, process-based models have complex structures and often simulate ideal ecological processes that deviate from actual conditions [2,11], which affects model accuracy. In addition, plant organisms involve complex and nonlinear biological and chemical mechanisms [2,12,13], which pose a great challenge for process-based models to simulate these mechanisms.

Currently, artificial intelligence (AI) algorithms have been widely applied in various fields because they can fit complex nonlinear mapping relationships between predictive and driving factors without requiring as many complicated prior assumptions as traditional models do [14]. Commonly applied machine learning models include Random Forest (RF), Support Vector Machine (SVM), and neural networks such as Long-Short Term Memory (LSTM). Tramontana et al. [15], Ichii et al. [16], Wang et al. [17], and other researchers used AI methods for tree species classification and carbon flux prediction, demonstrating the potential of AI in ecology. In recent years, Zhang et al. [18], Yuan et al. [19], Yu et al. [20], Sarkar et al. [4], and others used RF, Convolutional Neural Network (CNN), and Deep Belief Network (DBN) to predict GPP and achieved good results.

In this work, we constructed a model with spatial–temporal correlation while considering environmental factor fusion based on the GeoMAN model, a network with a multi-level attention mechanism developed by Liang et al. [21]. We trained and parameterized the algorithm with observational data on GPP from various sites and environmental driving data to extract nonlinear mapping relationships between GPP and multiple environmental factors. We designed a series of case studies to assess the performance of this deep learning model, which was based on the attention mechanism, by examining the impacts of distance, vegetation, and environmental factors on the prediction results across various flux sites. Compared with the previous applications of AI models in cross-disciplinary fields, our method not only fully utilizes the high precision of AI in prediction but also considers the prior knowledge within the relevant field to ensure that the results are both more accurate and consistent with domain knowledge.

2. Materials and Methods

2.1. Study Area

The flux sites used in this study were distributed in the Tibetan Plateau region, which is located in the alpine climate zone and has the climatic characteristics of long sunshine hours, intense sunlight, low temperatures, and scant rainfall. The regional average elevation exceeds 4000 m, the annual average temperature ranges from -5.75 to 2.57 °C, and the annual average amount of precipitation is 200–600 mm. Alpine grassland covers more than 60% of the surface area of this region [14,22], and it is a distinctive grassland ecosystem within all alpine areas in the world [23,24].

According to the *Atlas of Grassland Resources in China* (1:1,000,000) [25], alpine grasslands are subdivided into four sub-categories: alpine Kobresia meadow (KO), alpine shrub meadow (SH), alpine swamp meadow (SW), and alpine meadow steppe (AS) (Figure 1).

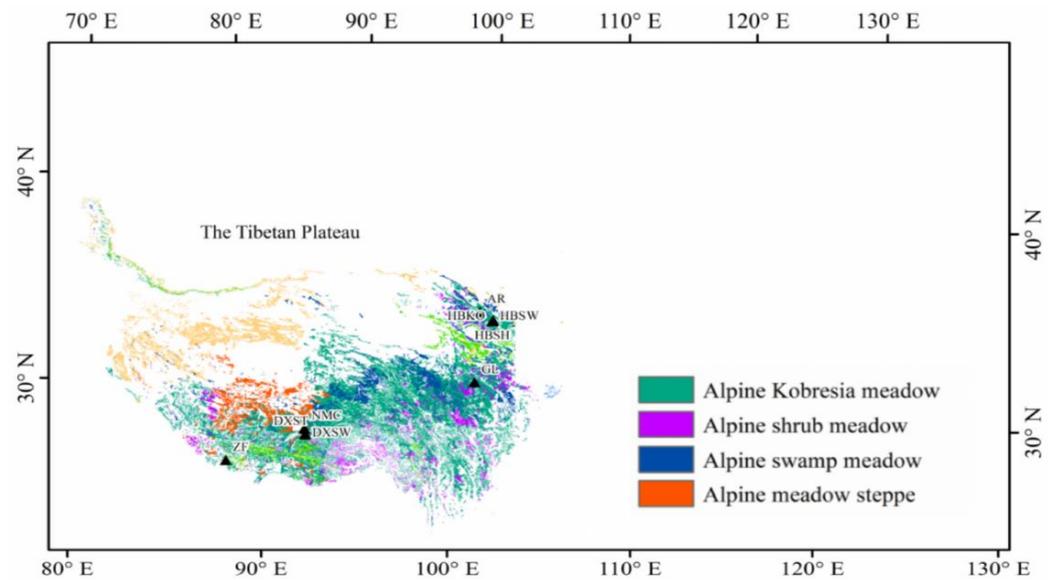


Figure 1. Geographical spread of alpine grasslands across China [14,23]. Black triangles represent the nine flux sites.

2.2. Data

2.2.1. Flux and Meteorological Data

The flux and meteorological data used in this study were collected from the China Terrestrial Ecosystem Flux Observation and Research Network (ChinaFLUX) [14,23,26], the Coordinated Observations and Integrated Research over Arid and Semi-arid China (COIRAS) [27], and the Heihe Watershed Allied Telemetry Experimental Research (HiWATER) [28], which were observed by nine flux stations distributed in the Tibetan Plateau region. The data spans from 2003 to 2014. These flux sites exemplify the broadest grassland ecosystem types, encompassing an extensive range of spatial, ecological, and weather-related circumstances [23].

Carbon flux data were processed using various methods, including triple coordinate rotation, Webb–Pearman–Leuning (WPL) correction, and outlier removal. The temporal resolution of the MODIS data was eight days, while that of temperature and photosynthetically active radiation was half an hour. The eddy covariance system was used to concurrently record these meteorological data, and any missing values were supplemented using the technique proposed by Schwalm et al. [29]. The data were then averaged and summed over eight days [14,23]. Finally, a total of 1421 site observation data points with a temporal resolution of eight days were obtained. The primary attributes of the nine flux sites in northern China’s grasslands were shown in Table 1.

Table 1. Primary attributes of the nine flux sites in northern China’s grasslands [14,23].

Site	Grassland Type	Latitude	Longitude	Elevation (m)	Operation Period
AR	Alpine Kobresia Meadows	38.04° N	100.46° E	3033	2014
GL		34.35° N	100.56° E	3980	2007, 2010–2011, and 2013
HBKO		37.61° N	101.31° E	3148	2003–2004
HBSH	Alpine Shrub Meadows	37.67° N	101.33° E	3293	2003–2012
DXSW	Alpine Swamp Meadows	30.47° N	91.06° E	4286	2009–2010
HBSW		37.61° N	101.33° E	3160	2004–2008 and 2010–2012
DXST	Alpine Meadow Steppes	30.5° N	91.06° E	4333	2004–2005, 2007, and 2009–2010
NMC		30.77° N	90.96° E	4730	2009
ZF		28.36° N	86.95° E	4293	2009

2.2.2. Remote Sensing Data

In this research, the remote sensing data utilized comprised the following MODIS products: normalized difference vegetation index (NDVI), enhanced vegetation index (EVI) (MOD13A2) [14,30], and surface reflectance (MOD09A1) [14,31]. The spatial resolution of the NDVI and EVI products was 1000 m and the temporal resolution was 16 days, while the spatial resolution of the surface reflectance was 500 m and the temporal resolution was 8 days. In order to acquire data with consistent spatial and temporal resolution, the quality control and data completion approaches proposed by Ma et al. [32] and Xiao et al. [33] were applied. The surface reflectance data were used to calculate the land surface water index (LSWI) [34].

2.3. Model

2.3.1. Deep Learning Model

In this study, we constructed our model based on the GeoMAN algorithm developed by Liang et al. in 2018, which was originally applied to predict air quality [21]. The GeoMAN algorithm can extract the spatial correlation of input variables and consider the influence of neighboring sites on the target site's GPP, which can help estimate the GPP of grasslands more accurately. The GeoMAN algorithm consists of an encoder and a decoder. The encoder contains a mechanism to consider the features within a site, a mechanism to consider the spatial features between sites, and an LSTM model to extract the local features of the site to be predicted and the spatial features of relevant surrounding sites. The decoder includes a temporal attention mechanism and an LSTM model, which decode the feature vector output by the encoder to predict grassland GPP.

There are complex correlations between environmental variables and GPP at each flux site. The inter-site feature attention mechanism of the GeoMAN algorithm dynamically captures the association between environmental variables and GPP within the site targeted for prediction. The inter-site feature attention mechanism for the target flux site is estimated as follows:

$$e_{k,t} = v_0^T \tanh(W_0 [h_{t-1}; s_{t-1}] + U_0 [I_k^0]) + b_0 \quad (1)$$

In Equation (1), $[h_{t-1}; s_{t-1}]$ denotes the concatenation operation in Tensorflow between h_{t-1} and s_{t-1} , as they are the hidden state and the cell state of the LSTM network at time $t - 1$, respectively, which contain the information of the previous $t - 1$ time steps, thereby forming the long-term and short-term memory of the LSTM network. I_k^0 means the k -th time series at the flux site to be predicted. v_0 , W_0 , U_0 , and b_0 are the learnable parameters: during the learning and training process of the model, they are continuously updated according to the loss function via backpropagation. The environmental factors selected for predicting GPP in this study include temperature (Ta), photosynthetically active radiation (PAR), enhanced vegetation index (EVI), normalized difference vegetation index (NDVI), and land surface water index (LSWI). The formula for calculating the weighting values of each factor based on the Geoman model's inter-site feature attention mechanism is as follows:

$$\alpha_{k,t} = \frac{\exp(e_{k,t})}{\sum_{j=1}^T \exp(e_{j,t})} \quad (2)$$

In Equation (2), $\alpha_{k,t}$ is the weighting value of the k -th feature (one each for Ta, PAR, EVI, NDVI, and LSWI) at time t . The sum of the weighting values of all features is 1. The calculated weighting values are multiplied by the corresponding feature values to distinguish the importance of different features according to the GeoMAN model.

As shown in Figure 2, the LSTM unit is an important computational unit in the GeoMAN model. Figure 3 shows the schematic diagram of the LSTM unit. The input at time t is x_t , and c_{t-1} and h_{t-1} are the cell state and the hidden state at time $t - 1$, respectively. They go through three main stages inside the LSTM unit: first, the forgetting stage, which selectively forgets the input from the previous time step; second, the selective memory stage, which selectively remembers the input from the current time step, emphasizing

important parts while remembering less important parts; and third, the output stage, which outputs the new hidden state h_t and the cell state c_t and inputs them to the next time step.

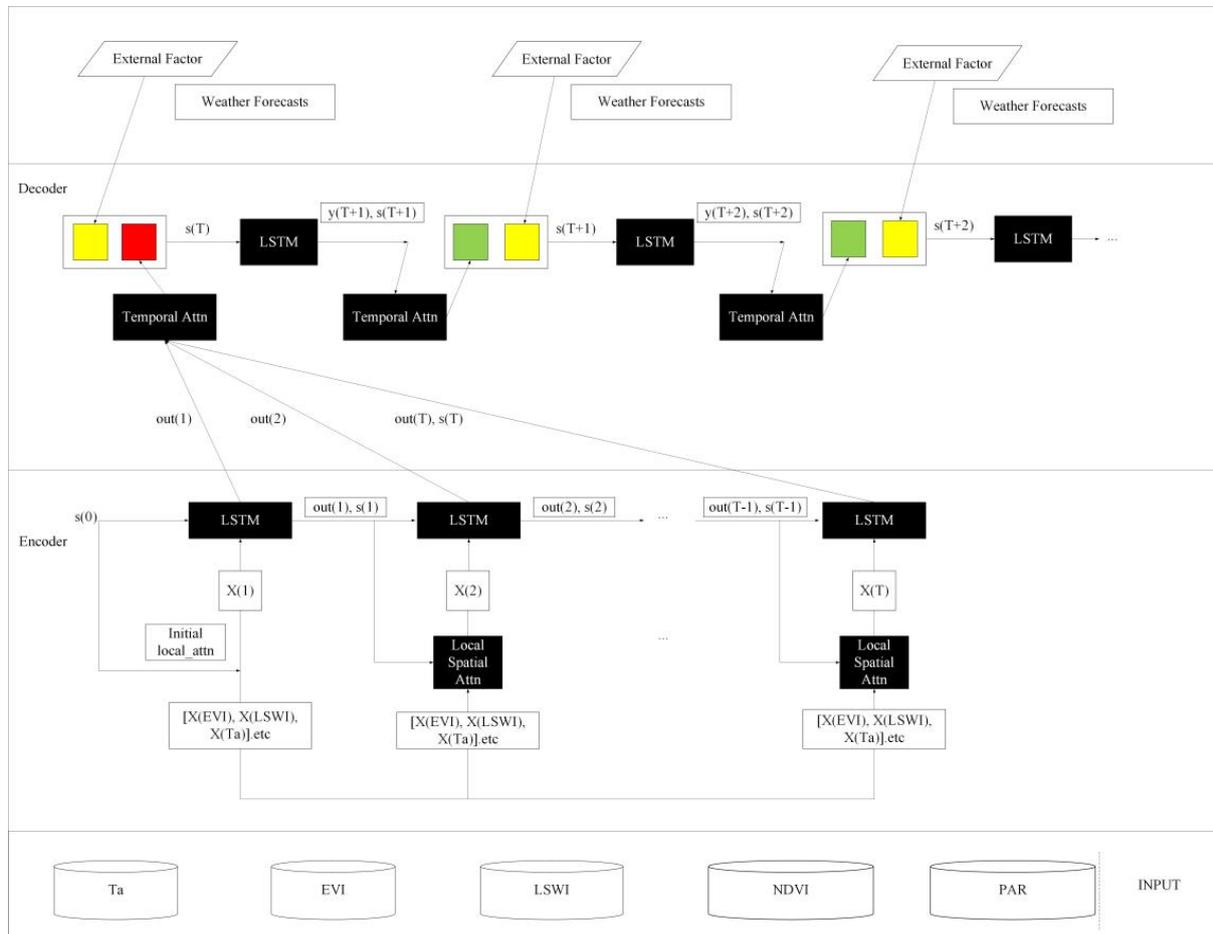


Figure 2. The structure of GeoMAN adapted from Liang et al. [21].

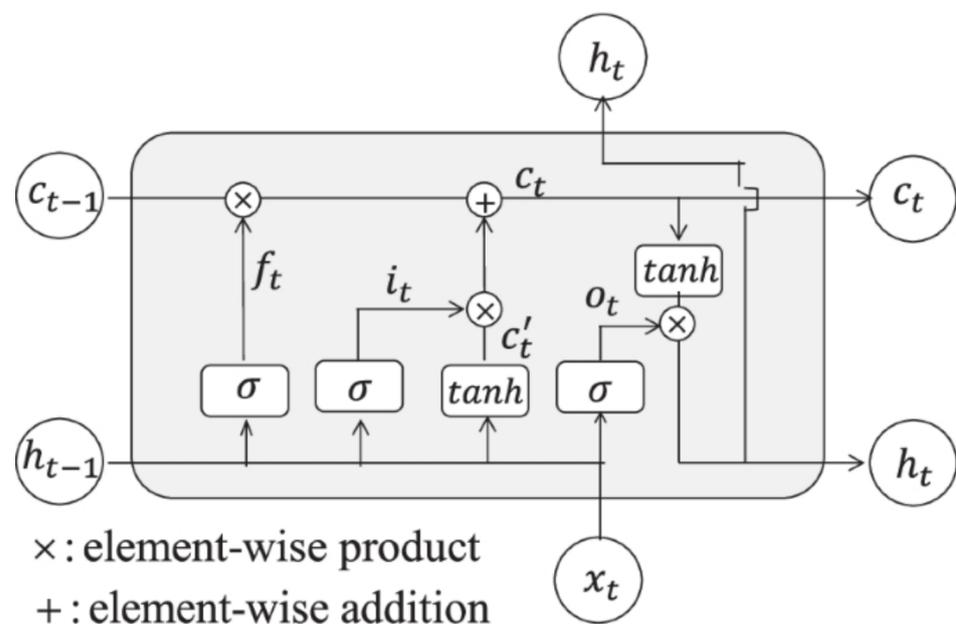


Figure 3. The structure of LSTM.

After the input data are assigned different weights by the inter-station feature attention mechanism and updated by the LSTM unit, it is essential to select pertinent time periods for GPP forecasting. It further enhances the precision of the prediction outcomes. Therefore, the GeoMAN model introduces a temporal attention mechanism. The formula for calculating the attention weight of each hidden state at a historical time step is as follows [35]:

$$u_{t,\tau} = v_d^T \tanh(W_d [h'_{t-1}; s'_{t-1}] + W'_d h_t + b_d) \tag{3}$$

$$\gamma_{t,\tau} = \frac{\exp(u_{t,\tau})}{\sum_{j=1}^T \exp(u_{j,\tau})} \tag{4}$$

In Equation (3), h'_{t-1} and s'_{t-1} are the hidden state and the cell state of the LSTM at time step $\tau - 1$. τ is the output prediction time step. v_d , W_d , U_d , and b_d are the learnable parameters. The output vector of the time attention mechanism at this time step is as follows [35]:

$$c_\tau = \sum_{t=1}^T \gamma_{t,\tau} h_t \tag{5}$$

2.3.2. Model Training and Evaluation

The values of each element of the flux site data used in this study have large differences. To ensure that model learning and training are not affected by this issue, each element of the data is standardized and normalized using the following formula:

$$x' = \frac{x - \mu}{\sigma} \tag{6}$$

In Equation (6), μ and σ are the mean and standard deviation of the corresponding element, respectively. The processed element data have a mean of 0 and a variance of 1, which prevents the model from being biased toward elements with large value ranges and ensures the accuracy of the model. After data pre-processing, the learning and training steps start. Since the observation data from the nine flux sites are not large in scale (a total of 1421 records), ten-fold cross-validation was applied to make full use of the data and to ensure the model's prediction performance on the whole data set. That is, for each fold, 10% of the data were taken as the test set, and the remaining 90% were used for learning and training. Then, the data used for learning and training were divided into training and validation sets at a ratio of 9:1, and the data order was randomly shuffled to avoid over-fitting. The model uses Mean Squared Error (MSE) as the loss function and the Adam optimizer to update model parameters. This process was repeated ten times to complete the predictions on all the data and evaluate the results.

This study used three common statistical indicators to evaluate model prediction accuracy: mean squared error, mean absolute error, and R-squared. The relevant formulas are as follows:

$$R^2 = \left(\frac{\sum_{i=1}^n (y_i - \bar{y})(y'_i - \bar{y}')}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (y'_i - \bar{y}')^2}} \right)^2 \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y'_i - y_i)^2} \tag{8}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y'_i - y_i| \tag{9}$$

where y'_i and y_i are the predicted and observed values of GPP, respectively; \bar{y}' and \bar{y} are the mean values of y'_i and y_i , respectively; and n is the number of observation samples.

3. Case Analysis

In this section, we introduced the basic idea of our experimental design. We tested the performance of different models on the aggregate data to prove that deep learning models have highly accurate prediction capabilities. Moreover, considering that the model predictions must conform to basic ecological knowledge, we were required to conduct multiple experiments by controlling the spatial distance of flux sites, vegetation types, and environmental factors.

3.1. Prediction Accuracy with All Factors

3.1.1. Comparison of Model Performance with the Use of All Data

In this analysis, we used the data from all flux sites to train the Random Forest (RF), Support Vector Machine (SVM), Deep Belief Network (DBN), and GeoMAN models. According to the ten-fold cross-validation results of all the models, the relationship between predicted GPP and observed GPP is shown in Figure 4 and Table 2. It is obvious that there are different training effects between the four models. The Random Forest model has the lowest prediction accuracy, with a prediction RMSE of $0.954 \text{ g Cm}^{-2} \text{ d}^{-1}$, MAE of $0.553 \text{ g Cm}^{-2} \text{ d}^{-1}$, and R^2 of 0.810; the Deep Belief Network model achieves a certain improvement in prediction accuracy compared to the Random Forest model, with a prediction RMSE of $0.912 \text{ g Cm}^{-2} \text{ d}^{-1}$, MAE of $0.559 \text{ g Cm}^{-2} \text{ d}^{-1}$, and R^2 of 0.827; the Support Vector Machine model has similar prediction accuracy to the Deep Belief Network model, with a prediction RMSE of $0.910 \text{ g Cm}^{-2} \text{ d}^{-1}$, MAE of $0.571 \text{ g Cm}^{-2} \text{ d}^{-1}$, and R^2 of 0.827; and the GeoMAN model has the highest prediction accuracy with a prediction RMSE of $0.788 \text{ g Cm}^{-2} \text{ d}^{-1}$, MAE of $0.440 \text{ g Cm}^{-2} \text{ d}^{-1}$, and R^2 of 0.870, which indicates that the GeoMAN model could fit the GPP values of the nine flux sites in the Tibetan Plateau better than the other three models.

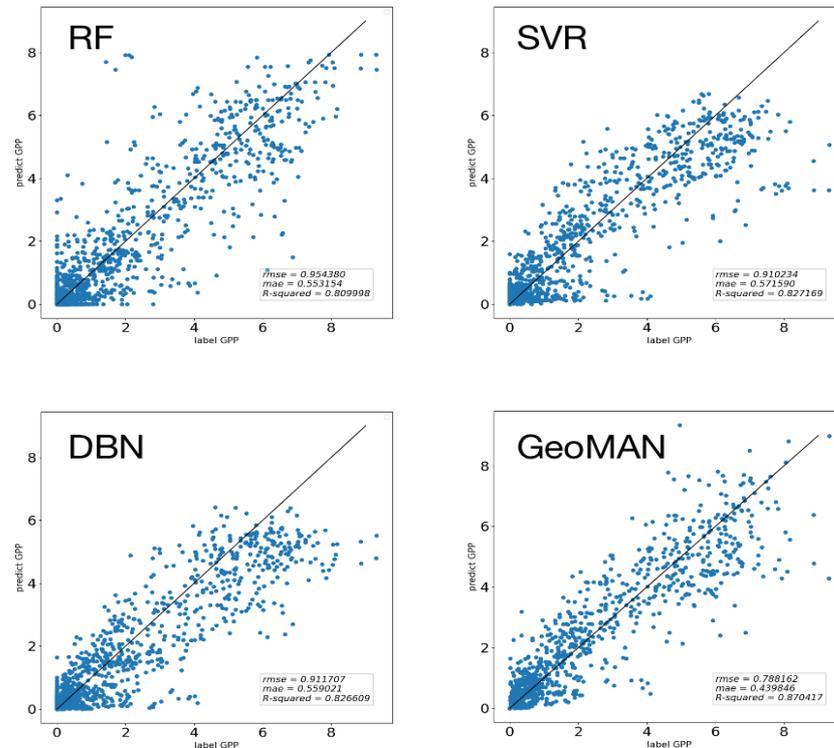


Figure 4. Performance of RF, SVR, DBN, and GeoMAN models in predicting GPP.

Table 2. Results of the four models with the use of all data.

Model	RF	SVR	DBN	GeoMAN
RMSE	0.954	0.910	0.912	0.788
MAE	0.553	0.571	0.559	0.440
R ²	0.810	0.827	0.827	0.870

3.1.2. Performance of Single Flux Site

As shown in Figure 1, the distribution map of the nine flux sites indicates that different sites have different vegetation and climate conditions. Therefore, it is necessary to test the GPP prediction accuracy of the GeoMAN model at different flux sites.

1. Test site GPP against remaining sites

We took the target flux site as the test data and the remaining sites as the training data, for a total of nine sites being tested. As shown in Figure 5, there is a large difference in performance among the different flux sites. The possible reasons for this difference are (1) some sites have different vegetation types from the target site, and (2) some sites have different climate conditions from the target site due to their long distance.

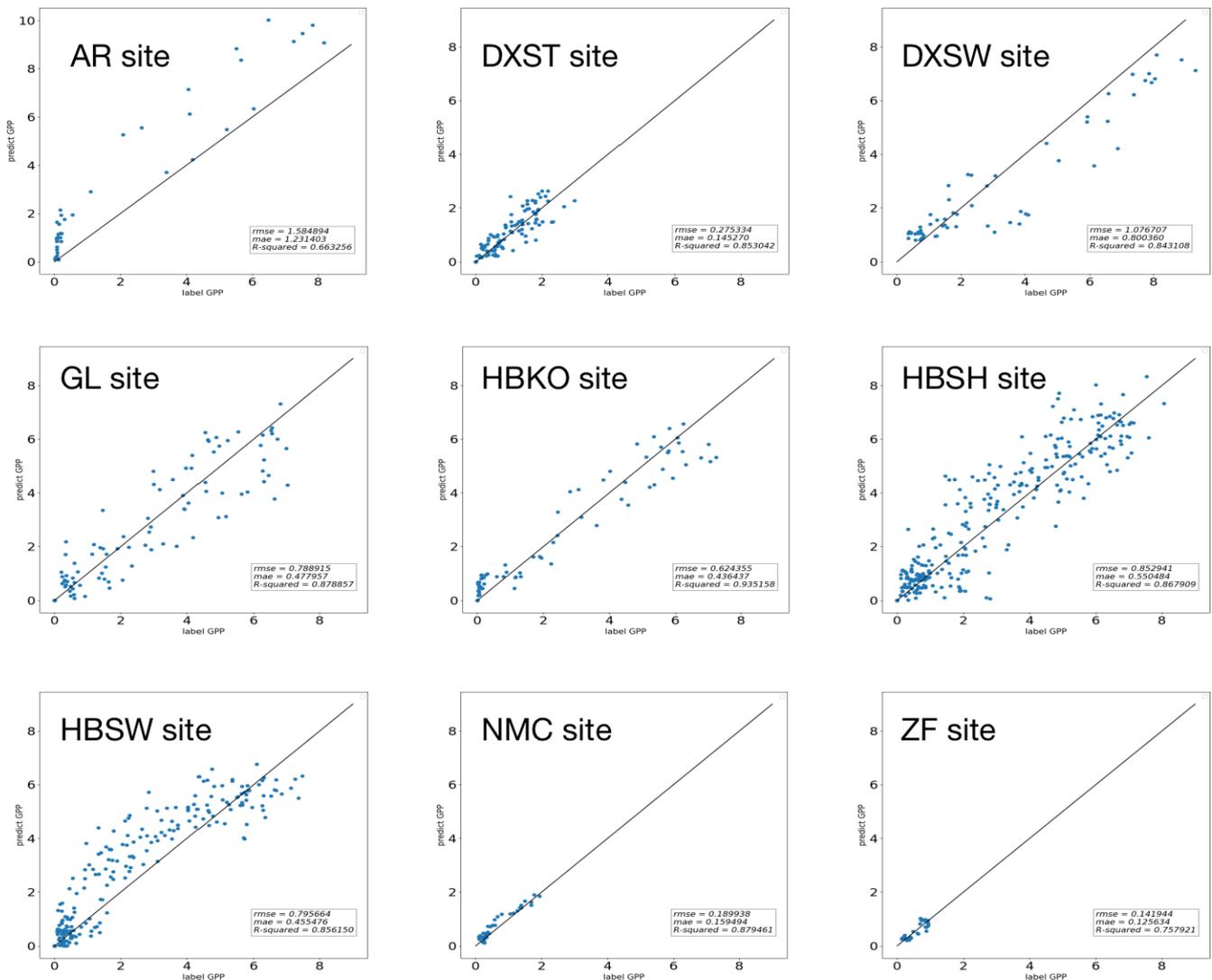


Figure 5. Predicted GPP vs. labeled GPP at a single site.

2. Test site GPP against the other sites at distance of 500 km and 100 km

The distances between each flux site were calculated using the Haversine formula based on their latitude and longitude. The calculation results are shown in Table 3.

Table 3. Distances between the flux sites (km).

	AR	DXST	DXSW	GL	HBKO	HBSH	HBSW	NMC	ZF
AR	/	1202.5	1204.9	410.4	88.7	86.8	90.1	1187.5	1651.6
DXST	1202.5	/	3.3	988.5	1230.1	1235.6	1231.6	31.5	463.7
DXSW	1204.9	3.3	/	990.1	1232.4	1237.8	1233.8	34.7	462.0
GL	410.4	988.5	990.1	/	368.7	375.6	369.1	983.3	1452.1
HBKO	88.7	1230.1	1232.4	368.7	/	6.9	1.8	1217.1	1685.3
HBSH	86.8	1235.6	1237.8	375.6	6.9	/	6.7	1222.44	1690.5
HBSW	90.1	1231.6	1233.8	369.1	1.8	6.7	/	1218.6	1686.8
NMC	1187.5	31.5	34.7	983.3	1217.1	1222.4	1218.6	/	471.4
ZF	1651.6	463.7	462.0	1452.1	1685.3	1690.5	1686.8	471.4	/

Based on the results in Table 3, each target flux site was predicted using the flux sites within 500 km and 100 km as the training data. There are no flux sites within 100 km of the GL and ZF sites, so they were not included in the prediction results for sites within 100 km. The prediction results are shown in Figures 6 and 7. According to the results, the accuracy of the AR site increases as the range of selected sites decreases, while the DXST, NMC, and ZF sites show a decreasing trend in accuracy as the range of selected sites decreases. The overall accuracy of the DXSW site is lower than when using all sites for prediction, but it shows a rebound trend as the range decreases. According to Table 1, the AR site with an increasing trend has alpine Kobresia meadow as its vegetation type, while the DXST, NMC, and ZF sites with a decreasing trend have alpine meadow steppe as their vegetation type. It can be speculated that the prediction effect of each site is related to the vegetation type of the other selected sites.

3. Selecting training data according to vegetation type

As shown in Table 1, the training data for each site to be predicted comes from the other flux sites with the same vegetation type. The final predictions are shown in Figure 8. It is obvious that selecting the training data based on vegetation type has higher overall prediction accuracy compared to selecting the training data based on distance (no corresponding prediction results are available for the HBSH site because it has a different vegetation type compared to the other flux sites). Next, we combined the results of the previous three experiments to examine the effect of selecting training data under different conditions on prediction accuracy, and the combined results are shown in Table 4.

As shown in Table 4, using vegetation type as the training data in the screening mechanism is better than using site distance as the training data from an overall perspective. However, from a single-site perspective, the GL and HBKO sites show a decreasing trend in accuracy. This is because for the AR and HBKO sites, which have significantly less data than other sites and share the same vegetation type as the GL site, the training data are insufficient, leading to a decrease in the prediction accuracy at the GL site. The HBKO site, which has always maintained an R-squared value above 0.9 from an overall perspective, does not have much room for accuracy improvement. At the same time, we compared the prediction results of the training data without screening and with screening for vegetation type. Although some sites have lower accuracy, the AR site shows a significant improvement in accuracy; as a result, the overall prediction accuracy of the latter method is not lower than that of the former method. This result shows that increasing the amount of training data with the same vegetation type can achieve equally good results as increasing the overall amount of training data.

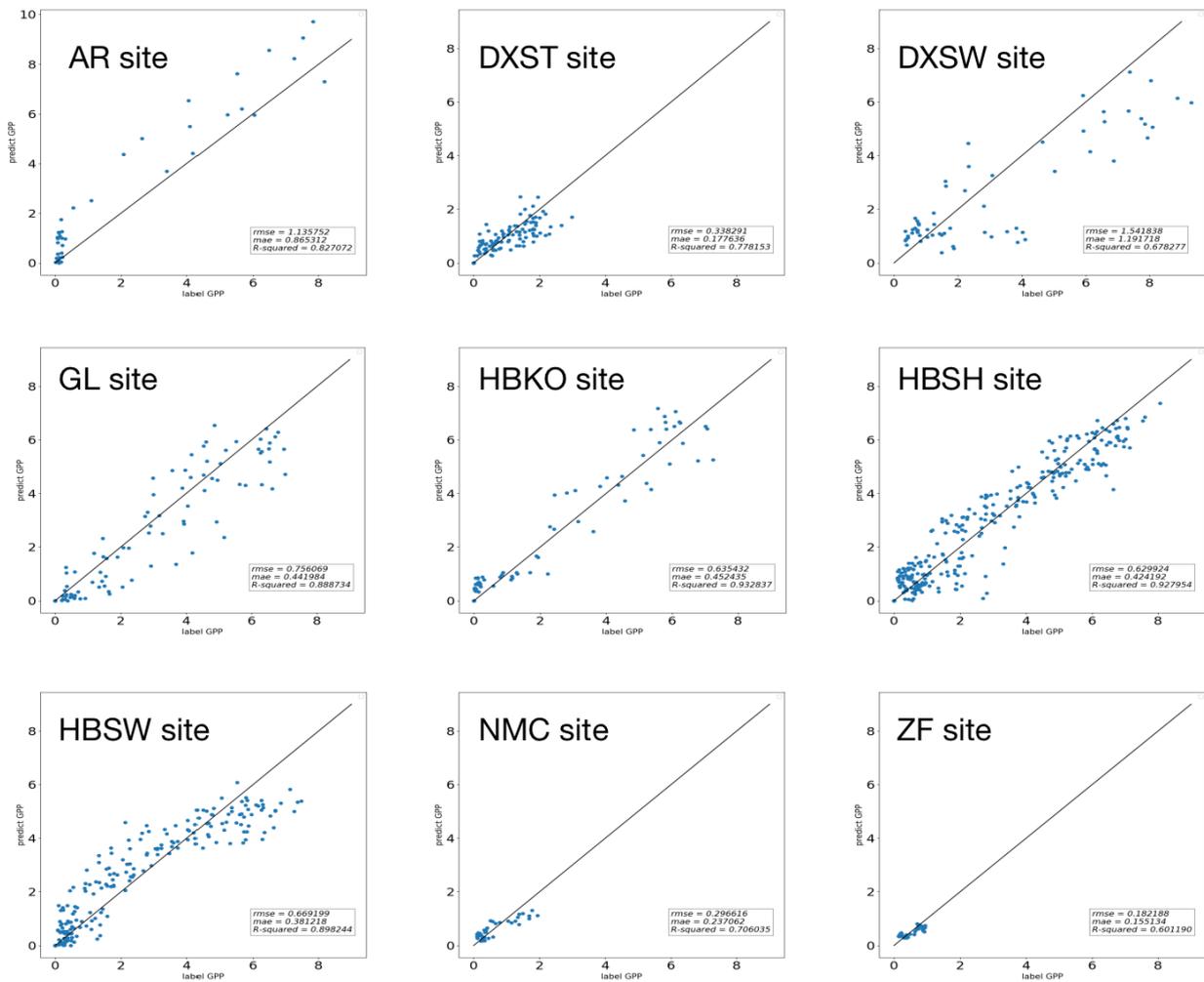


Figure 6. Predicted GPP vs. the labeled GPP at a single site (500 km).

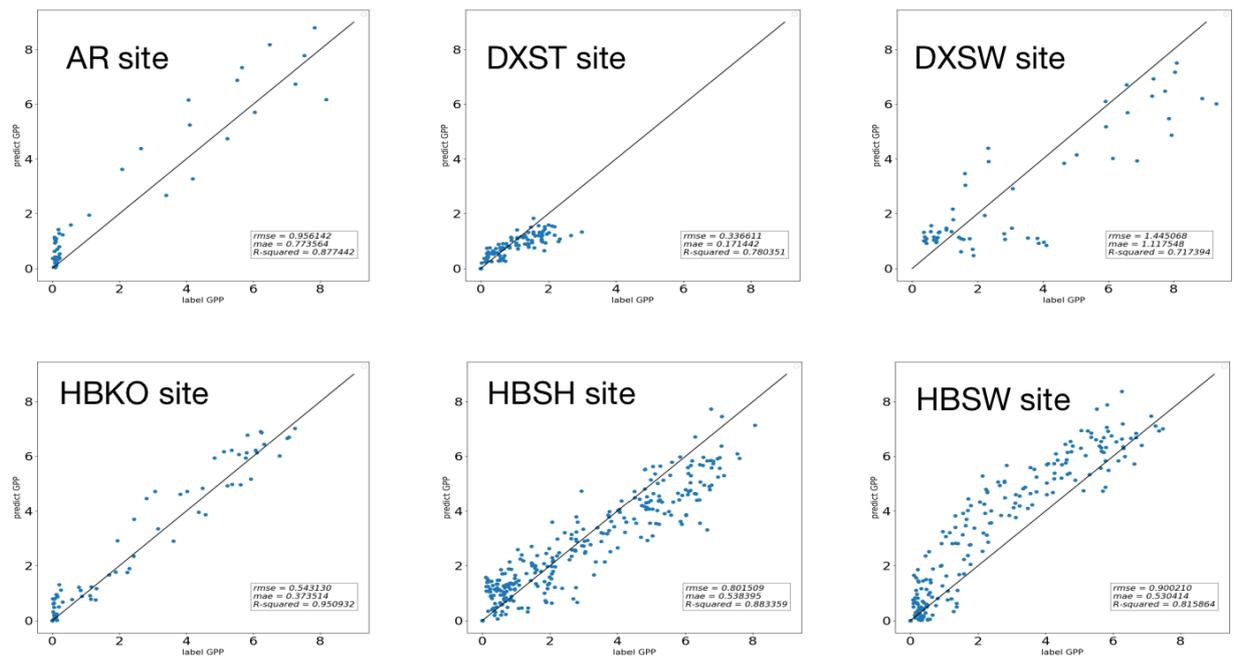


Figure 7. Cont.

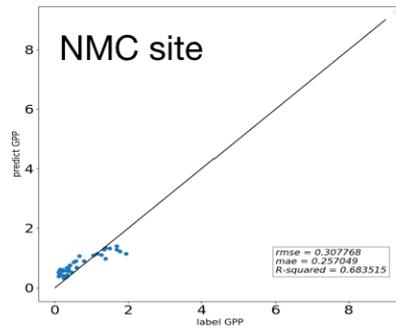


Figure 7. Predicted GPP vs. labeled GPP at a single site (100 km).

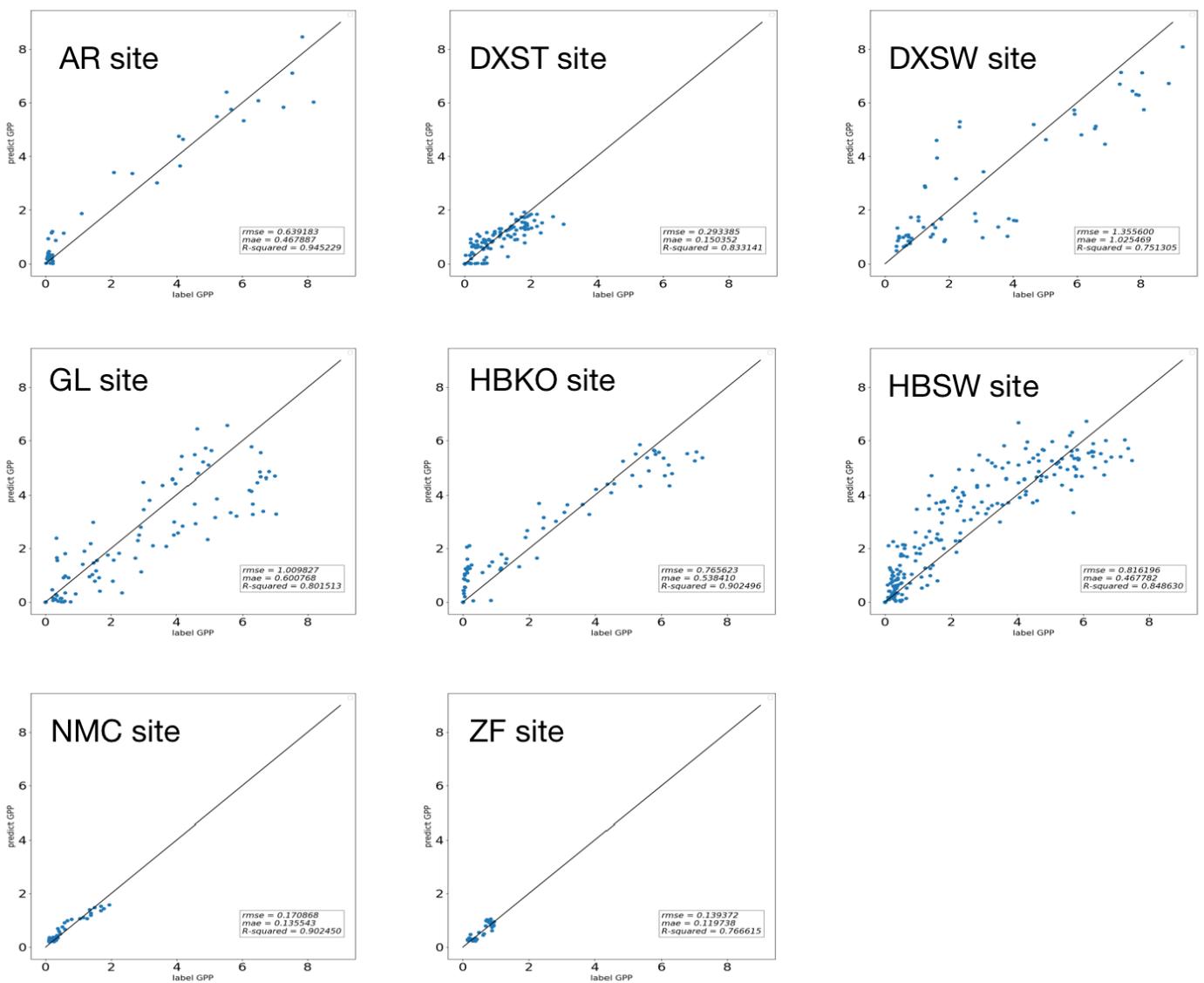


Figure 8. Predicted GPP vs. labeled GPP at a single site based on vegetation type.

Table 4. Results comparing all previous experiments.

Site	AR	DXST	DXSW	GL	HBKO	HBSH	HBSW	NMC	ZF
R ²	0.663	0.853	0.843	0.879	0.935	0.868	0.856	0.879	0.758
R ² (500 km)	0.827	0.778	0.678	0.889	0.933	0.928	0.898	0.706	0.601
R ² (100 km)	0.877	0.780	0.717	/	0.951	0.883	0.816	0.683	/
R ² (vt)	0.945	0.833	0.751	0.801	0.902	/	0.849	0.902	0.767

Numbers in red indicate that the prediction results for the corresponding site have increased in accuracy compared to the previous results without setting a distance range; numbers in blue indicate that the results have decreased in accuracy; and numbers in black indicate that the results have no significant changes.

3.2. Prediction Accuracy with Factor Ablation

In Section 3.1, the effects of vegetation type and distance between the flux sites on GPP prediction accuracy were then investigated. The training data included temperature (Ta), photosynthetically active radiation (PAR), enhanced vegetation index (EVI), normalized difference vegetation index (NDVI), and land surface water index (LSWI). In this analysis, a feature ablation experiment was conducted to explore the influence of each factor on GPP prediction accuracy.

3.2.1. Test Site GPP without Ta

In this experiment, all Ta data were deleted from the training data, which were then trained for each site flux. The final prediction results are shown in Figure 9. Compared to the prediction results without any feature ablation of the training data, the prediction accuracy of the AR site is significantly improved, while that of the DXST, DXSW, NMC, and ZF sites is greatly reduced. The prediction accuracy of the GL, HBKO, HBSH, and HBSW sites does not change noticeably.

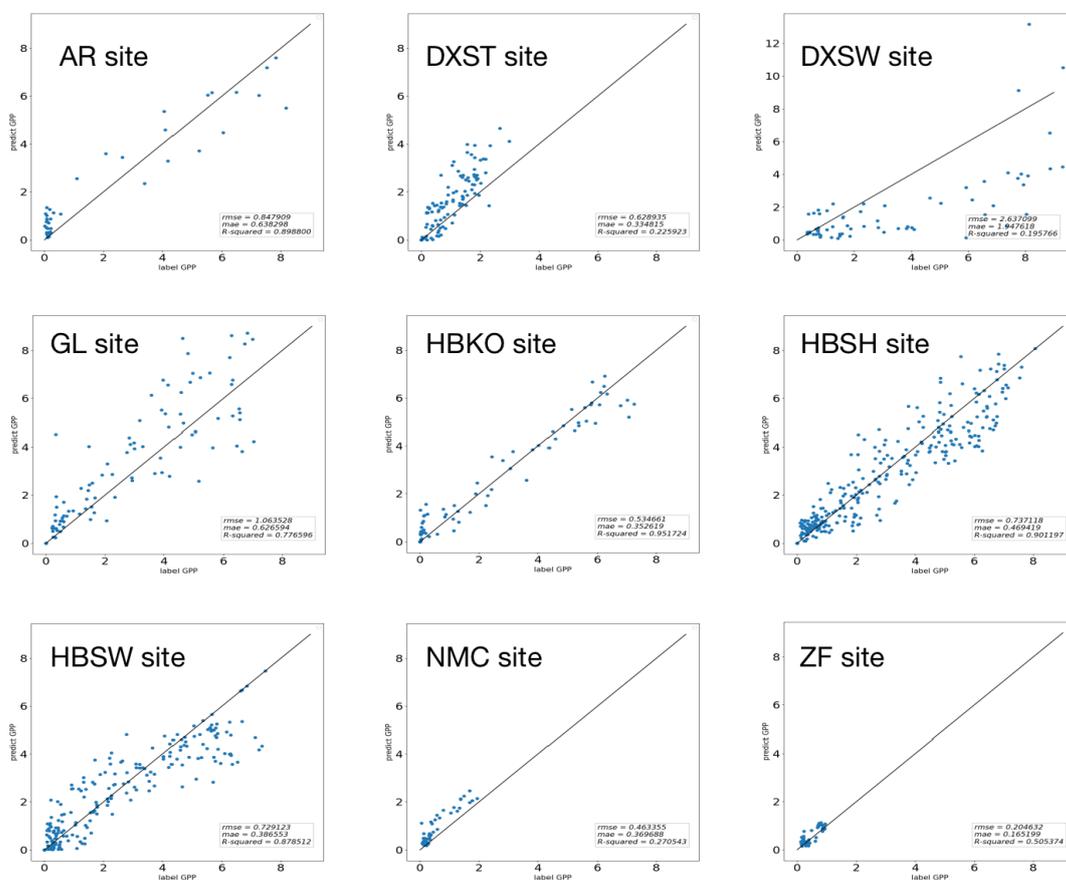


Figure 9. Predicted GPP vs. labeled GPP at a single site (no Ta).

3.2.2. Test Site GPP without PAR

In this experiment, all Par data were deleted from the training data, which were then trained for each site flux. The final prediction results are shown in Figure 10. Compared to the prediction results without any feature ablation of the training data, the prediction accuracy trends of the sites are similar to those obtained after removing Ta. The prediction accuracy of the AR site has significantly improved, but not as much as after the removal of Ta. The prediction accuracy of the DXST, DXSW, NMC, and ZF sites is sharply reduced. For the DXST site, the decrease is greater than that after removing Ta, whereas the DXSW, NMC, and ZF sites show some recovery but still perform worse than without any feature ablation. The prediction accuracy of the GL, HBKO, HBSH, and HBSW sites as a whole is lower than that after removing Ta, though it does not change appreciably.

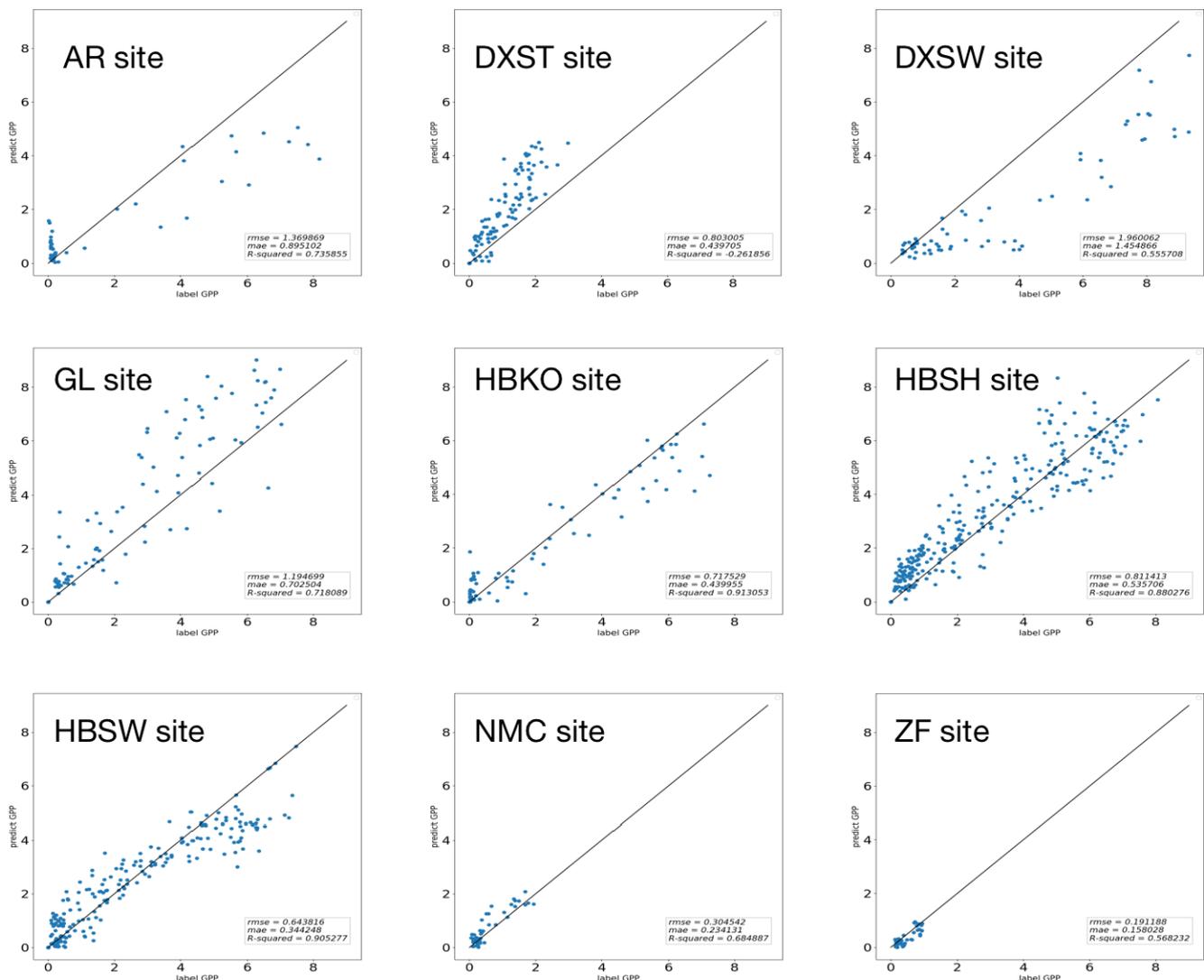


Figure 10. Predicted GPP vs. labeled GPP at a single site (no PAR).

3.2.3. Test Site GPP without EVI

In this experiment, all EVI data were deleted from the training data, which were then trained for each site flux. The final prediction results are shown in Figure 11. Compared to the prediction results without any feature ablation of the training data, the prediction accuracy of the AR site is still significantly improved and higher than that after removing Par but lower than that after removing Ta. The prediction accuracy of the DXST, DXSW, NMC, and ZF sites is similar to that after the removal of Ta and Par, with a sharp decrease.

The decrease is larger for the NMC and ZF sites than for the others. The prediction accuracy of the GL site shows a continuous decline compared to that after the removal of Ta and Par. The prediction accuracy of the HBKO, HBSH, and HBSW sites does not change noticeably.

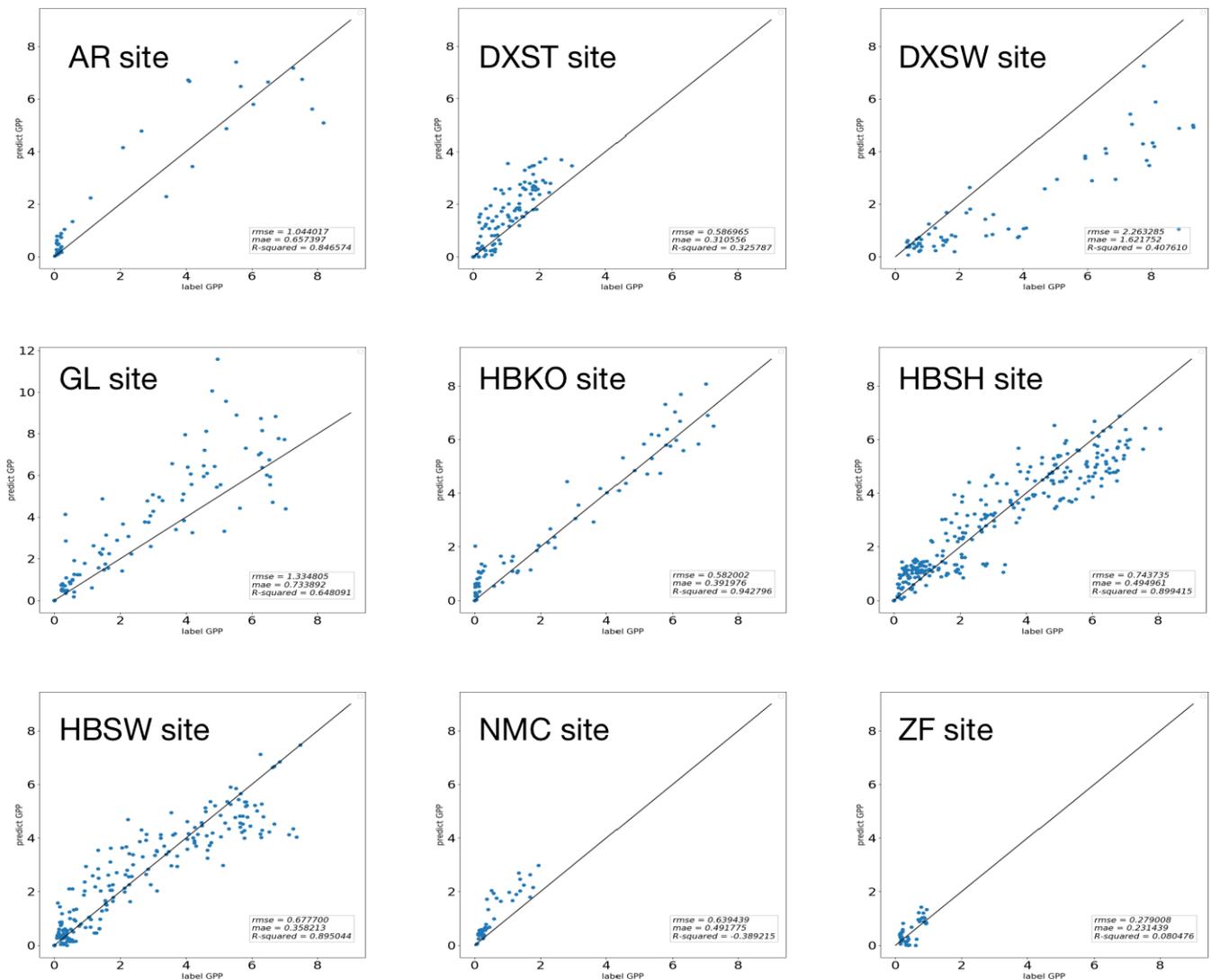


Figure 11. Predicted GPP vs. labeled GPP at a single site (no EVI).

3.2.4. Test Site GPP without NDVI

In this experiment, all NDVI data were deleted from the training data, which were then trained for each site flux. The final prediction results are shown in Figure 12. Compared to the prediction results without any feature ablation of the training data, the prediction accuracy of the AR site is significantly improved and higher than that after removing Ta, Par, and EVI. The prediction accuracy of the DXST, DXSW, and ZF sites decreases, which is consistent with the results of the previous ablation experiments. For the ZF site, the prediction accuracy is only better than that after removing Ta and EVI. The NMC site has an abnormal increase in accuracy. The prediction accuracy of the GL site is higher than that after removing Ta, Par, and EVI. The prediction accuracy of the HBKO, HBSH, and HBSW sites does not change noticeably.

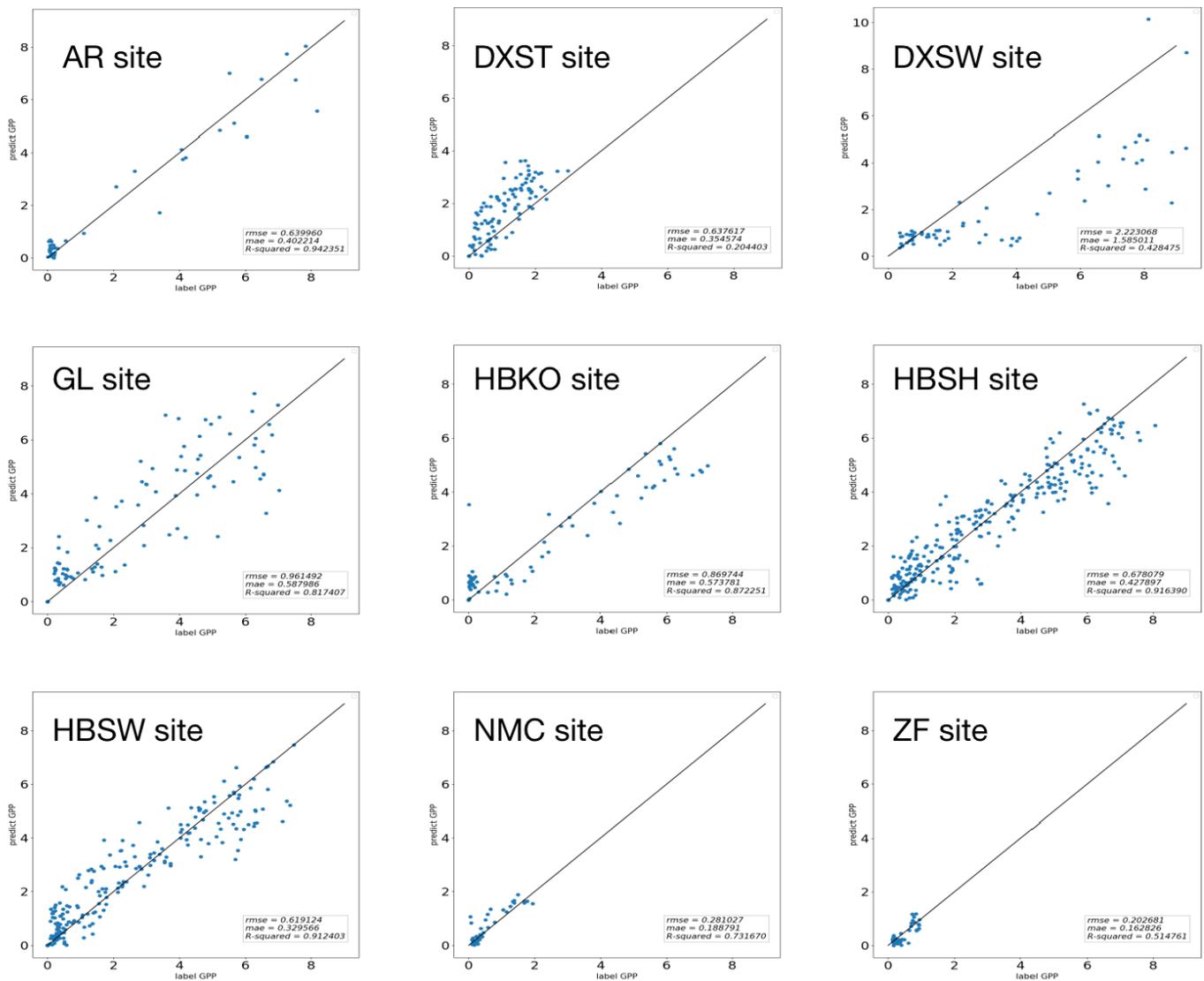


Figure 12. Predicted GPP vs. labeled GPP at a single site (no NDVI).

3.2.5. Test Site GPP without LSWI

In this experiment, all LSWI data were deleted from the training data, which were then trained for each site flux. The final prediction results are shown in Figure 13. Compared to the prediction results without any feature ablation of the training data, the prediction accuracy of the AR site improves significantly and is only lower than that after removing NDVI. The DXST, DXSW, NMC, and ZF sites have similar prediction accuracy as in the previous ablation experiments, with a large decrease in accuracy, and the prediction accuracy of the ZF site is only higher than that after removing EVI. The GL site shows a slight decrease in accuracy, while the HBKO, HBSH, and HBSW sites show no obvious changes in prediction accuracy.

3.2.6. Summary of Factor Ablation Experiments

To observe the influence of different features on the prediction accuracy of each site more intuitively, we summarized all the feature ablation experiment results, as shown in Table 5.

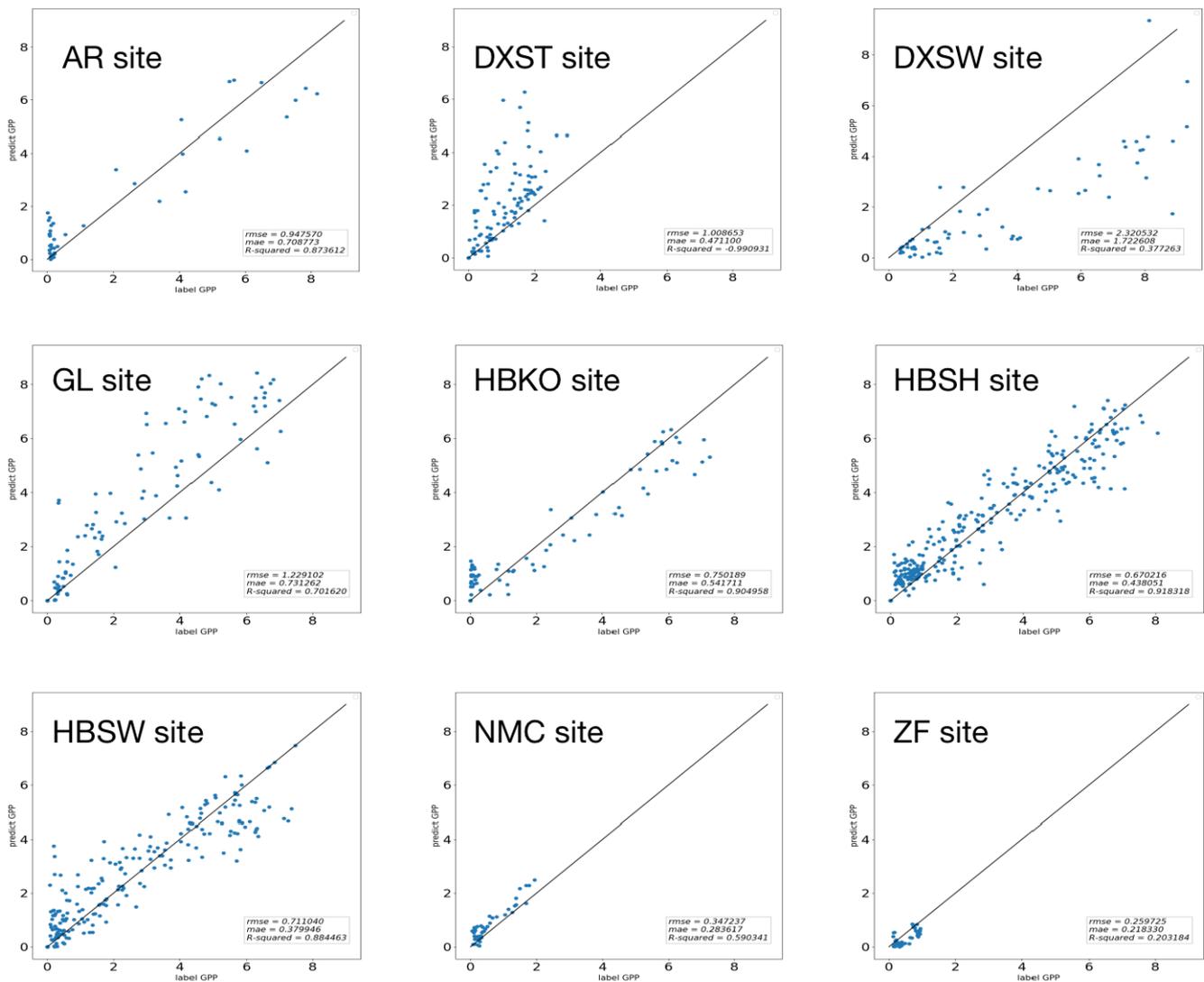


Figure 13. Predicted GPP vs. labeled GPP at a single site (no LSWI).

Table 5. Results of all factor ablation experiments.

Site	AR	DXST	DXSW	GL	HBKO	HBSH	HBSW	NMC	ZF
R ²	0.663	0.853	0.843	0.879	0.935	0.868	0.856	0.879	0.758
R ² (no-Ta)	0.899	0.226	0.196	0.777	0.952	0.901	0.879	0.271	0.505
R ² (no-PAR)	0.736	-0.262	0.556	0.718	0.913	0.880	0.905	0.685	0.568
R ² (no-EVI)	0.847	0.326	0.408	0.648	0.943	0.899	0.895	-0.385	0.080
R ² (no-NDVI)	0.942	0.204	0.428	0.817	0.872	0.916	0.912	0.732	0.515
R ² (no-LSWI)	0.874	-0.991	0.377	0.702	0.905	0.918	0.884	0.590	0.203

Numbers in red indicate that the prediction results for the corresponding site have increased in accuracy compared to the previous results without setting a distance range; numbers in blue indicate that the results have decreased in accuracy; and numbers in black indicate that the results have no significant changes.

Table 5 indicates that removing any feature from the AR site would result in a significant improvement in accuracy, with the largest improvement obtained after removing NDVI. The removal of any feature for the DXST, DXSW, and ZF sites would lead to a degradation of accuracy, with the DXST site showing a large accuracy decline and the

lowest accuracy after removing LSWI. The DXSW site also shows a decline in accuracy, although it is smaller than that of the DXST site. The ZF site has a noticeable decline in accuracy after removing EVI. The NMC site has an abnormal increase in accuracy after removing NDVI and a decline after removing other features except NDVI. The GL site is insensitive to the removal of Ta or NDVI and shows slight decreases in accuracy after removing other features in addition to Ta and NDVI. The HBKO, HBSH, and HBSW sites are insensitive to the removal of any feature and have no obvious changes in accuracy.

4. Conclusions

In this work, we used satellite remote sensing data and flux site observation data to introduce the GeoMAN model based on an encoder–decoder framework with an attention mechanism for site features, and we obtained good results. According to the experiments on training data selection based on distance and vegetation type, we found that both distance and vegetation type had an impact on GPP prediction results, with vegetation type having a larger impact. Through the feature ablation experiments, we found that different sites showed sensitivity to different factors, with the site located in the alpine swamp meadow being insensitive to changes in environmental factors, while the site located in the alpine meadow steppe showed a different trend since the GPP prediction accuracy decreased sharply with the changes in environmental factors. The GPP prediction accuracy of the site located in the alpine Kobresia meadow also varied with environmental factor changes but was more stable than the other sites. The results of this work show that deep learning models have high accuracy when simulating site-scale GPP and, to some extent, reflect the correlation between a target site's GPP and other sites' distances, vegetation types, and meteorological factors. Our work could be used in the prediction of other factors, for example, AGB (Above Ground Biomass) and RE (Ecosystem Respiration). However, this work has some limitations. Firstly, we do not consider some factors that have an influence on productivity in Tibetan grasslands, such as soil development and drought regimes. Secondly, the data we used in this work only cover a partial area of the Tibetan Plateau region, and this introduces constraints to regional GPP assessment. In our future work, we will add more factors to the training data, for example, soil pH, soil fertility, and soil organic matter (SOM), since high soil pH and a lack of soil fertility limit plant productivity [36], and SOM is able to enhance alpine grassland productivity by improving the soil structure, aggregates, and cation-exchange capacity (CEC) under high aridity conditions [37]. Moreover, ecological factors, such as growing and non-growing seasons, will be considered, and larger regional-scale data will be used in future training and learning processes. These improvements will help us perform more accurate and larger-scale GPP simulations.

Author Contributions: Conceptualization, Q.Y. and Y.W.; methodology, Q.Y. and N.N.; software, Q.Y.; validation, Q.Y., Z.W. and M.W.; formal analysis, R.C.; investigation, Q.Y.; resources, X.R.; data curation, Q.Y. and X.W.; writing—original draft preparation, Q.Y.; writing—review and editing, N.N. and X.W.; visualization, Q.Y. and W.L.; supervision, Y.W.; project administration, Q.Y.; funding acquisition, N.N. and Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (Grant No. 2021YFF0703902).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Williams, M.; Rastetter, E.B.; Fernandes, D.N.; Goulden, M.L.; Shaver, G.R.; Johnson, L.C. Predicting gross primary productivity in terrestrial ecosystems. *Ecol. Appl.* **1997**, *7*, 882–894. [[CrossRef](#)]
2. Zhang, X.; Wang, H.; Yan, H.; Ai, J. Analysis of spatio-temporal changes of gross primary productivity in China from 2001 to 2018 based on Remote Sensing. *Acta Ecol. Sin.* **2021**, *41*, 6351–6362.
3. Piao, S.L.; Sitch, S.; Ciais, P.; Friedlingstein, P.; Peylin, P.; Wang, X.H.; Ahlström, A.; Anav, A.; Canadell, J.G.; Cong, N.; et al. Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO₂ trends. *Glob. Chang. Biol.* **2013**, *19*, 2117–2132. [[CrossRef](#)]
4. Sarkar, D.P.; Shankar, B.U.; Parida, B.R. Machine Learning Approach to Predict Terrestrial Gross Primary Productivity using Topographical and Remote Sensing Data. *Ecol. Inform.* **2022**, *70*, 101697. [[CrossRef](#)]
5. Lee, B.; Kim, N.; Kim, E.-S.; Jang, K.; Kang, M.; Lim, J.-H.; Cho, J.; Lee, Y. An Artificial Intelligence Approach to Predict Gross Primary Productivity in the Forests of South Korea Using Satellite Remote Sensing Data. *Forests* **2020**, *11*, 1000. [[CrossRef](#)]
6. Kang, M.; Kim, J.; Kim, H.S.; Thakuri, B.M.; Chun, J.H. On the nighttime correction of CO₂ flux measured by eddy covariance over temperate forests in complex terrain. *Korean J. Agric. For. Meteorol.* **2014**, *16*, 233–245. (In Korean with English abstract) [[CrossRef](#)]
7. Running, S.W.; Nemani, R.R.; Heinsch, F.A.; Zhao, M.S.; Reeves, M.; Hashimoto, H. A continuous satellite-derived measure of global terrestrial primary production. *Bioscience* **2004**, *54*, 547–560. [[CrossRef](#)]
8. Xiao, X.M.; Zhang, Q.Y.; Braswell, B.; Urbanski, S.; Boles, S.; Wofsy, S.; Moore, B., III; Ojima, D. Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data. *Remote Sens. Environ.* **2004**, *91*, 256–270. [[CrossRef](#)]
9. Yuan, W.P.; Liu, S.G.; Zhou, G.S.; Zhou, G.Y.; Tieszen, L.L.; Baldocchi, D.; Bernhofer, C.; Gholz, H.; Goldstein, A.H.; Goulden, M.L.; et al. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. *Agric. For. Meteorol.* **2007**, *143*, 189–207. [[CrossRef](#)]
10. Reeves, M.C.; Zhao, M.; Running, S.W. Usefulness and limits of MODIS GPP for estimating wheat yield. *Int. J. Remote Sens.* **2007**, *26*, 1403–1421. [[CrossRef](#)]
11. Beer, C.; Reichstein, M.; Tomelleri, E.; Ciais, P.; Jung, M.; Carvalhais, N.; Rödenbeck, C.; Arain, M.A.; Baldocchi, D.; Bonan, G.B.; et al. Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate. *Science* **2010**, *329*, 834–838. [[CrossRef](#)]
12. Schindler, D.E.; Hilborn, R. Prediction, precaution, and policy under global change. *Science* **2015**, *347*, 953–954. [[CrossRef](#)]
13. Ye, H.; Beamish, R.J.; Glaser, S.M.; Grant, S.C.H.; Hsieh, C.H.; Richards, L.J.; Schnute, J.T.; Sugihara, G. Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E1569–E1576. [[CrossRef](#)]
14. Zhu, X.; He, H.; Ma, M.; Ren, X.; Zhang, L.; Zhang, F.; Li, Y.; Shi, P.; Chen, S.; Wang, Y.; et al. Estimating Ecosystem Respiration in the Grasslands of Northern China Using Machine Learning: Model Evaluation and Comparison. *Sustainability* **2020**, *12*, 2099. [[CrossRef](#)]
15. Tramontana, G.; Ichii, K.; Camps-Valls, G.; Tomelleri, E.; Papale, D. Uncertainty analysis of gross primary production upscaling using random forests, remote sensing and eddy covariance data. *Remote Sens. Environ.* **2015**, *168*, 360–373. [[CrossRef](#)]
16. Ichii, K.; Ueyama, M.; Kondo, M.; Saigusa, N.; Kim, J.; Alberto, M.C.; Ardö, J.; Euskirchen, E.S.; Kang, M.; Hirano, T.; et al. New data-driven estimation of terrestrial CO₂ fluxes in Asia using a standardized database of eddy covariance measurements, remote sensing data, and support vector regression. *J. Geophys. Res. Biogeosci.* **2017**, *122*, 767–795. [[CrossRef](#)]
17. Wang, X.; Yao, Y.; Zhao, S.; Jia, K.; Zhang, X.; Zhang, Y.; Zhang, L.; Xu, J.; Chen, X. MODIS-based estimation of terrestrial latent heat flux over North America using three machine learning algorithms. *Remote Sens.* **2017**, *9*, 1326. [[CrossRef](#)]
18. Zhang, K.; Liu, N.; Gao, S.; Zhao, S. Data-Driven Estimation of Gross Primary Production. *Remote Sens. Technol. Appl.* **2020**, *35*, 943–949. [[CrossRef](#)]
19. Yuan, D.; Zhang, S.; Li, H.; Zhang, J.; Yang, S.; Bai, Y. Improving the Gross Primary Productivity Estimate by Simulating the Maximum Carboxylation Rate of the Crop Using Machine Learning Algorithms. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4413115. [[CrossRef](#)]
20. Yu, T.; Zhang, Q.; Sun, R. Comparison of Machine Learning Methods to Up-Scale Gross Primary Production. *Remote Sens.* **2021**, *13*, 2448. [[CrossRef](#)]
21. Liang, Y.; Ke, S.; Zhang, J.; Yi, X.; Zheng, Y. Geoman: Multi-level attention networks for geo-sensory time series prediction. In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, Stockholm, Sweden, 13–19 July 2018.
22. Zhang, J.-W. *Vegetation of Xizang (Tibet)*; Science Press: Beijing, China, 1988.
23. Liu, W.; He, H.; Wu, X.; Ren, X.; Zhang, L.; Zhu, X.; Feng, L.; Lv, Y.; Chang, Q.; Xu, Q.; et al. Spatiotemporal Changes and Driver Analysis of Ecosystem Respiration in the Tibetan and Inner Mongolian Grasslands. *Remote Sens.* **2022**, *14*, 3563. [[CrossRef](#)]
24. Ge, R.; He, H.; Ren, X.; Zhang, L.; Li, P.; Zeng, N.; Yu, G.; Zhang, L.; Yu, S.-Y.; Zhang, F.; et al. A Satellite-Based Model for Simulating Ecosystem Respiration in the Tibetan and Inner Mongolian Grasslands. *Remote Sens.* **2018**, *10*, 149. [[CrossRef](#)]
25. Su, D. *The Atlas of Grassland Resources of China (1:1,000,000)*; Press of Map: Beijing, China, 1993. (In Chinese)
26. Yu, G.-R.; Wen, X.-F.; Sun, X.-M.; Tanner, B.D.; Lee, X.; Chen, J.-Y. Overview of ChinaFLUX and evaluation of its eddy covariance measurement. *Agric. For. Meteorol.* **2006**, *137*, 125–137. [[CrossRef](#)]

27. Wang, H.; Jia, G.; Fu, C.; Feng, J.; Zhao, T.; Ma, Z. Deriving maximal light use efficiency from coordinated flux measurements and satellite data for regional gross primary production modeling. *Remote Sens. Environ.* **2010**, *114*, 2248–2258. [[CrossRef](#)]
28. Li, X.; Cheng, G.D.; Liu, S.M.; Xiao, Q.; Ma, M.G.; Jin, R.; Che, T.; Liu, Q.H.; Wang, W.Z.; Qi, Y.; et al. Heihe watershed allied telemetry experimental research (hiwater): Scientific objectives and experimental design. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1145–1160. [[CrossRef](#)]
29. Schwalm, C.R.; Williams, C.A.; Schaefer, K.; Anderson, R.; Arain, M.A.; Baker, I.; Barr, A.; Black, T.A.; Chen, G.; Chen, J.M.; et al. A model-data intercomparison of CO₂ exchange across North America: Results from the North American Carbon Program site synthesis. *J. Geophys. Res. Biogeosci.* **2010**, *115*, G3. [[CrossRef](#)]
30. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [[CrossRef](#)]
31. Vermote, E.; Vermeulen, A. *MODIS Algorithm Technical Background Document, Atmospheric Correction Algorithm: Spectral Reflectances (MOD09)*; NASA Contract NAS5-96062; University of Maryland: College Park, MD, USA, 1999.
32. Ma, M.G.; Veroustraete, F. Reconstructing pathfinder AVHRR land NDVI time-series data for the Northwest of China. *Adv. Space Res. Ser.* **2006**, *37*, 835–840. [[CrossRef](#)]
33. Xiao, J.F.; Zhuang, Q.L.; Baldocchi, D.D.; Law, B.E.; Richardson, A.D.; Chen, J.Q.; Oren, R.; Starr, G.; Noormets, A.; Ma, S.Y.; et al. Estimation of net ecosystem carbon exchange for the conterminous United States by combining MODIS and AmeriFlux data. *Agric. For. Meteorol.* **2008**, *148*, 1827–1847. [[CrossRef](#)]
34. Xiao, X.; Hollinger, D.; Aber, J.; Goltz, M.; Davidson, E.A.; Zhang, Q.; Moore, B. Satellite-based modeling of gross primary production in an evergreen needleleaf forest. *Remote Sens. Environ.* **2004**, *89*, 519–534. [[CrossRef](#)]
35. Shi, M.; Wang, J.; Yin, R.; Zhang, P. Short-Term Photovoltaic Power Forecast Based on Grey Relational Analysis and GeoMAN Model. *Trans. China Electrotech. Soc.* **2021**, *36*, 2298–2305.
36. Zhao, Y.; Wang, X.; Jiang, S.; Xiao, J.; Li, J.; Zhou, X.; Liu, H.; Hao, Z.; Wang, K. Soil development mediates precipitation control on plant productivity and diversity in alpine grasslands. *Geoderma* **2022**, *412*, 115721. [[CrossRef](#)]
37. Zhao, Y.; Wang, X.; Chen, F.; Li, J.; Wu, J.; Sun, Y.; Zhang, Y.; Deng, T.; Jiang, S.; Zhou, X.; et al. Soil organic matter enhances aboveground biomass in alpine grassland under drought. *Geoderma* **2023**, *433*, 116430. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.