land

MDPI

Article

Determining Dominant Factors of Vegetation Change with Machine Learning and Multisource Data in the Ganjiang River Basin, China

Zhiming Xia [1,†], Kaitao Liao [2,3,†], Liping Guo [1,*], Bin Wang [4,5,*], Hongsheng Huang [6], Xiulong Chen [6], Xiangmin Fang [1], Kuiling Zu [1], Zhijun Luo [6], Faxing Shen [2] and Fusheng Chen [1]

1 Key Laboratory of National Forestry and Grassland Administration on Forest Ecosystem Protection and Restoration of Poyang Lake Watershed, College of Forestry, Jiangxi Agricultural University, Nanchang 330045, China; xiazming@126.com (Z.X.); fangxiangmin@jxau.edu.cn (X.F.); zukuiling@jxau.edu.cn (K.Z.); chenfusheng@jxau.edu.cn (F.C.)
2 Jiangxi Key Laboratory of Watershed Soil and Water Conservation, Jiangxi Provincial Technology Innovation Center for Ecological Water Engineering in Poyang Lake Basin, Jiangxi Academy of Water Science and Engineering, Nanchang 330029, China; liaokaitao@jxnu.edu.cn (K.L.); faixing2007@126.com (F.S.)
3 Key Laboratory of Poyang Lake Wetland and Watershed Research (Ministry of Education), School of Geography and Environment, Jiangxi Normal University, Nanchang 330022, China
4 NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW 2650, Australia
5 Gulbali Institute for Agriculture, Water and Environment, Charles Sturt University, Wagga Wagga, NSW 2650, Australia
6 College of Land Resources and Environment, Jiangxi Agricultural University, Nanchang 330045, China; huanghongsheng@jxau.edu.cn (H.H.); chenxiulong@jxau.edu.cn (X.C.); luozhijun@jxau.edu.cn (Z.L.)
* Correspondence: guoliping@jxau.edu.cn (L.G.); bin.a.wang@dpi.nsw.gov.au (B.W.)
† Theses authors contributed equally to this work.

Abstract: Vegetation is a fundamental component of terrestrial ecosystems, and accurately assessing the effects of seasonal climate variations, extreme weather events, and land use changes on vegetation dynamics is crucial. The Ganjiang River Basin (GRB), a key region for water conservation and recharge in southeastern China, has experienced significant land use changes and variable climate in the past. However, comprehensive evaluations of how these changes have impacted vegetation remain limited. To address this gap, we used machine learning models (random forest and XGBoost) to assess the impact of seasonal and extreme climate variables, land cover, topography, soil properties, atmospheric CO_2, and night-time light intensity on vegetation dynamics. We found that the annual mean NDVI showed a slight increase from 1990 to 1999 but has decreased significantly over the last 8 years. XGBoost was better than the RF model in simulating the NDVI when using all five types of data source (R^2 = 0.85; RMSE = 0.04). The most critical factors influencing the NDVI were forest and cropland ratio, followed by soil organic carbon content, elevation, cation exchange capacity, night-time light intensity, and CO_2 concentration. Spring minimum temperature was the most important seasonal climate variable. Both linear and nonlinear relationships were identified between these variables and the NDVI, with most variables exhibiting threshold effects. These findings underscore the need to develop and implement effective land management strategies to enhance vegetation health and promote ecological balance in the region.

Keywords: NDVI; machine learning; driving factors; climate variable; land cover; Ganjiang River Basin

Academic Editors: Dong Liang, Barjeece Bashir and Min Xu

Received: 31 October 2024
Revised: 26 December 2024
Accepted: 31 December 2024
Published: 3 January 2025

Citation: Xia, Z.; Liao, K.; Guo, L.; Wang, B.; Huang, H.; Chen, X.; Fang, X.; Zu, K.; Luo, Z.; Shen, F.; et al. Determining Dominant Factors of Vegetation Change with Machine Learning and Multisource Data in the Ganjiang River Basin, China. *Land* **2025**, *14*, 76. https://doi.org/10.3390/land14010076

1. Introduction

In recent years, global vegetation has been greening, leading to debates about the reasons and mechanisms behind this phenomenon, especially when considering different regions [1]. Areas like China and India, known for intensive cultivation and afforestation efforts, have shown significant greening trends [2]. There is a growing interest in unraveling the driving mechanisms behind this greening phenomenon [3]. Understanding these forces is crucial for comprehending vegetation dynamics in specific areas, evaluating ecosystem health and resilience, and guiding land management and conservation efforts in the face of climate change [4].

Numerous studies have established that human activities and climate change, especially variations in seasonal temperatures and precipitation patterns, are key drivers of vegetation dynamics [5,6]. As manifestations of climate change become more pronounced [7], the way that vegetation responds is also evolving. For instance, Liu et al. [8] observed that rising minimum temperatures could enhance vegetation greening in subtropical areas. Moreover, seasonal climate changes warrant attention, as previous studies have shown that seasonal patterns of rainfall and temperature significantly influence vegetation health and productivity. However, their roles in long-term vegetation dynamics are often underexplored. The climate–vegetation relationship varies across seasons [9], for example, temperature and NDVI may exhibit a negative correlation in autumn but shift to a positive correlation in winter. These shifts profoundly affect vegetation growth dynamics, highlighting the importance of incorporating seasonality to better capture the temporal variability of vegetation changes. Conversely, human actions, such as urbanization, significantly impact vegetation by converting green spaces into impermeable surfaces, thereby diminishing plant productivity [10]. Additionally, other elements like CO_2 fertilization [11], night-time light [12], soil properties [13], and elevation [14] also play crucial roles in vegetation dynamics. Despite these insights, ongoing debates about the factors affecting vegetation changes over the past decades continue, fueled by climate variability, diverse research methodologies, and inconsistencies in data availability. Insufficient knowledge of these factors may interfere with the understanding of ecosystem interactions, and an accurate method is needed to assess the impact of these factors on vegetation.

In recent years, many methods have emerged to study the dynamic drivers of vegetation in different regions, including regression analysis and geographic detectors [15–17]. However, there are still drawbacks associated with the identification of drivers of vegetation change. Concerning residual analysis, the response of vegetation to climate and human behavior is often non-linear, making it challenging to describe the relationship with a single regression equation [18]. As for geographic detectors, it is difficult to show the mechanism between the NDVI and the different influencing factors [17]. In recent years, the development of explainable machine learning methods has created certain conditions for understanding the response mechanism of vegetation to different influencing factors [19,20]. This method has achieved good results in the study of the influencing factors of vegetation [20,21]. For example, Wang et al. [22] used machine learning methods and found that higher water availability and a wetter environment are more likely to promote positive changes in plant productivity on the Loess Plateau, while negative changes are related to night-time light. The application of machine learning methods in these studies demonstrates their advantages in identifying drivers of vegetation change and quantifying their relative importance.

The Ganjiang River Basin (GRB) plays a critical role in conserving and recharging water sources in the Yangtze River [23], which is of great importance in regulating the hydrological balance of the Poyang Lake Basin and the flow of the Yangtze River. Since 1990, ecological protection projects including the protected forest system projects in the middle

and upper reaches of the Yangtze River, desertification control and plain afforestation, and key projects of water and soil conservation [24], have been implemented, resulting in an increase in forest coverage from 34.73% in 1983 to 63.10% in 2010 [25]. In recent years, the basin has experienced pronounced climate changes, with an increasing trend in precipitation extremes [26]. Rapid human development activities have also exerted a comprehensive impact on vegetation [27]. However, there are still knowledge gaps in identifying dominant drivers of vegetation dynamic in the GRB. Apart from the effects of climate factors, the impacts of human activities (e.g., land cover changes), extreme climate events, night-time light variations, and CO_2 fertilization remain unclear and underexplored in the GRB region.

In this study, we employed two widely used machine learning models, extreme gradient boosting (XGBoost) and random forest (RF), to explore the factors influencing vegetation dynamics in the GRB. Specifically, we analyzed the impacts of climate change, extreme climate events, atmospheric CO_2 concentration, soil properties, topography, land cover, and night-time light intensity on the NDVI within the GRB. The objectives of this study were to (1) identify the trend in vegetation change within the GRB from 1990 to 2018, (2) develop machine learning methods with different predictors to estimate the NDVI, and (3) identify dominant driving factors contributing to the NDVI change. We expect the results will provide insights for effective environmental management, climate change adaptation, and informed evidence-based policies for sustainable land use and environmental conservation.

2. Materials and Methods

2.1. Study Area

The Ganjiang River is the largest and most significant waterway in Jiangxi Province, with a mainstream length of 758 km [28]. It is also the seventh largest tributary of the Yangtze River, covering an area of about 83,500 km^2 and accounting for 51% of the territory in Jiangxi Province (Figure 1) [29]. The GRB experiences a humid subtropical monsoon climate, influenced by monsoon climates and typhoons, and interacts with a complex terrain. The basin has an average annual precipitation of 1400–1800 mm and an average temperature of around 18 °C. Precipitation is unevenly distributed throughout the year, with 78.6% of the total annual precipitation occurring during the wet season (April–September) [29]. The main soil types found in the region are red, yellow, purple, and paddy soil [24]. The GRB is characterized by superior natural conditions, including abundant forest resources and diverse plant species nationwide. The basin is primarily composed of warm coniferous forests and evergreen broadleaved forests [28]. In terms of land use types, the northern part of the basin is mostly mountainous, while the southern and central regions are predominantly farmland [30]. The boundary of the Poyang Lake basin was recorded in 2000 and it was downloaded from National Earth System Science Data Center, National Science & Technology Infrastructure of China (http://www.geodata.cn) (accessed on 1 July 2023).

2.2. Data

2.2.1. NDVI Data

The normalized difference vegetation index (NDVI), defined as the ratio of the difference between near-infrared reflectance and red visible reflectance to their sum, serves as a crucial indicator of vegetation greenness [31,32]. To investigate the long-term variations of the NDVI, an extended temporal scale was employed in this study. We integrated two NDVI datasets representing different periods. The initial dataset comprised GIMMS NDVI 3g data spanning the years 1982 to 2015. Derived from a global dataset reliant on NOAA weather satellite data and disseminated by NASA (https://data.nasa.gov/) (accessed on

1 July 2023), this dataset was developed by The Global Inventory Modeling and Mapping Studies (GIMMS). The second dataset encompassed SPOT/VGT NDVI data covering the period 1998–2018. This dataset was obtained from the National Cryosphere Desert Data Center, China. (http://www.ncdc.ac.cn) (accessed on 3 July 2023). The amalgamation of these datasets facilitates a comprehensive examination of NDVI dynamics, offering a robust foundation for a nuanced exploration of vegetation trends over the specified temporal range [33]. We used a pixel-scale linear regression method to fuse the two datasets. The pixel scale linear regression fusion method tended to be more accurate in the short term and small regions than the global model [34]. For example, Li et al. [35] used this method to fuse the GIMMS and MODIS NDVI data to produce the PKU GIMMS NDVI product.

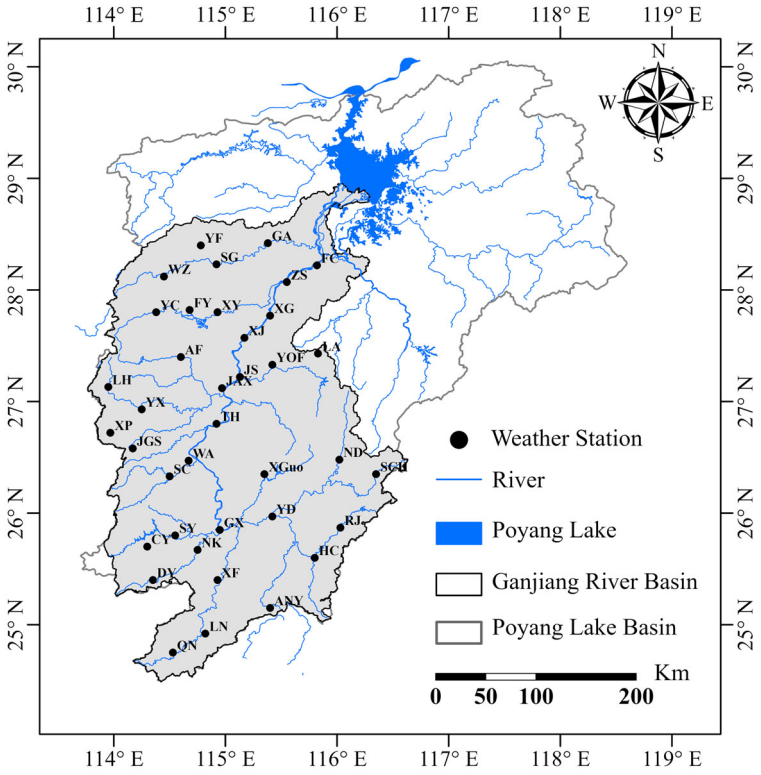

Figure 1. The location of Ganjiang River Basin and the weather stations used in this study. The following is the corresponding name of each weather station: AF—anfu, ANY—anyuan, CY—chongyi, DY—dingnan, FC—dayu, FY—fenyi, GA—gaoan, GX—ganxian, HC—huichang, JAX—jianxian, JGS—jinggangshan, JS—jishui, LA—lean, LH—lianhua, LN—longnan, ND—ningdu, NK—nankang, QN—quannan, RJ—ruijin, SC—suichuan, SCH—shicheng, SG—shanggao, SY—shangyou, TH—taihe, WA—wanan, WZ—wanzai, XF—xinfeng, XG—xingan, XGuo—xingguo, XJ—xiajiang, XP—xiaping, XY—xinyu, YC—yichun, YD—yudu, YF—yifeng, YOF—yongfeng, YX—yongxin, ZS—zhangshu.

The MVC (maximum value composites) method was employed to amalgamate GIMMS and SPOT/VGT data at a monthly scale [36]. Subsequently, the data underwent projection onto the WGS 1984 UTM Zone 50N coordinate system, with a resolution resampled to 1 km. The GIMMS and SPOT/VGT data for different months in the overlapping period 1999–2015 were disaggregated into 12-month segments, and pixel-wise conversion coefficients were calculated through the application of the linear regression method. Ultimately, we obtained the new monthly GIMMS NDVI data for the period 1982–1998. Subsequently, the regenerated monthly NDVI data for the same period were integrated with the SPOT/VGT NDVI monthly-scale data spanning 1998–2018, yielding a comprehensive and uninterrupted time series of NDVI data spanning 1982–2018. The NDVI data at an annual scale were also acquired utilizing the maximum value composites method. For 38 meteorological stations,

the average of nine pixels around each meteorological site was chosen as the NDVI value, representing the NDVI value for the corresponding sites at the corresponding times. Owing to constraints posed by the availability of land cover and NDVI data, we opted for a data period spanning from 1990 to 2018.

2.2.2. Climate and Atmospheric CO_2 Data

Climatic data spanning the years 1960 to 2020 from 38 meteorological stations within the GRB were acquired from the China Meteorological Data Service Center (http://data.cma.cn/) (accessed on 15 July 2023). The selected climatic variables encompassed daily precipitation, air mean temperature, maximum and minimum temperatures, air humidity, wind speed, and hours of solar insolation (Table 1). The dataset was stratified into four distinct sections corresponding to seasons, with each season's data determined by computing the average or sum values of the corresponding months: 1 for spring (March–May), 2 for summer (June–August), 3 for autumn (September–November), and 4 for winter (December–February); if no number is present, it indicates the entire year.

Table 1. The different types of data used in this study and their respective origins.

Type	Item	Temporal Scale	Indices	Definition	Data Source
Climate	Climate data	Seasonal and annual	Rain	Total precipitation	http://data.cma.cn/
			Tem	Mean temperature	
			Max_Tem	Max temperature	
			Min_Tem	Min temperature	
			Hum	Air humidity	
			WS	Wind speed	
			Sun	Sun hour	
	Climate extreme indices	Annual	SU25	Count of days for daily max temperature > 25 °C	
			CWD	Max number of consecutive days with daily precipitation ≥ 1 mm	
			FD0	Count of days for daily min temperature < 0 °C	
			R95P	Computes the annual sum of precipitation on days where daily precipitation exceeds the 95th percentile	
			GSL	Annual count between the first span of at least 6 days with daily mean temperature > 5 °C and first span after July 1st of 6 days with <5 °C	
			TR20	Count of days for daily min temperature > 20 °C	
Atmosphere	CO_2	Annual	CO_2	Annual atmospheric CO_2 concentration	https://gml.noaa.gov/
Soil	Soil properties	\	T_OC	0–30 cm organic carbon content	http://data.tpdc.ac.cn
			T_CEC_CLAY	The cation exchange capacity of clay layer soil	
Topography	DEM	\	Elevation	Elevation of station	http://www.gscloud.cn/
Human activities	Land cover type	Annual	Cropland	Area change of cropland	https://zenodo.org/records/4417810
			Forest	Area change of forest	
	Night-time light	Annual	Light	Light intensity at night	http://data.tpdc.ac.cn

Climate extreme indices had significant impacts on the growth of vegetation [37]. Here, we considered 27 climate extreme indices (http://etccdi.pacificclimate.org) (accessed on 19 July 2023), including 16 annual climate extreme indices (TR20, FD0, R95P, SU25, CWD, GSL, PRCPTOT, R99P, WSDI, CSDI, CDD, SDII, R25MM, R20MM, R10MM, and ID0) (Table 1), while the remaining 11 climate extreme indices were not considered due to their monthly scales.

Monthly CO_2 concentration from January 1979 to July 2023 was obtained from the Global Monitoring Laboratory (https://gml.noaa.gov/) (accessed on 20 July 2023). The average CO_2 concentration of the 12 months of each year was taken as the CO_2 concentration for that year.

2.2.3. Land Cover Data

In previous studies, the temporal resolution of the exploited land use data was often in the range of 5 years [38]. To account for the changes in vegetation conditions under continuous land use patterns, we utilized a land cover dataset released from a recent study with a temporal resolution of 1 year [39], representing the vegetation conditions under continuous land cover changes. This data included a complete image of available land cover for each year from 1990–2018 (https://zenodo.org/records/4417810) (accessed on 30 July 2023). The land cover types were classified into nine categories: cropland, forest, shrub, grassland, water, snow and ice, barren, impervious, and wetland. As the spatial resolution of the land cover data was 30 m, we resampled it to 1 km. We extracted the land cover types of the nine pixels around each weather station site as the land cover status at the corresponding time and location. The ratio distribution of land cover types within each station was determined for the period 1990–2018. Notably, the main land cover we extracted was forest and cropland for the study sites (Figure 2).

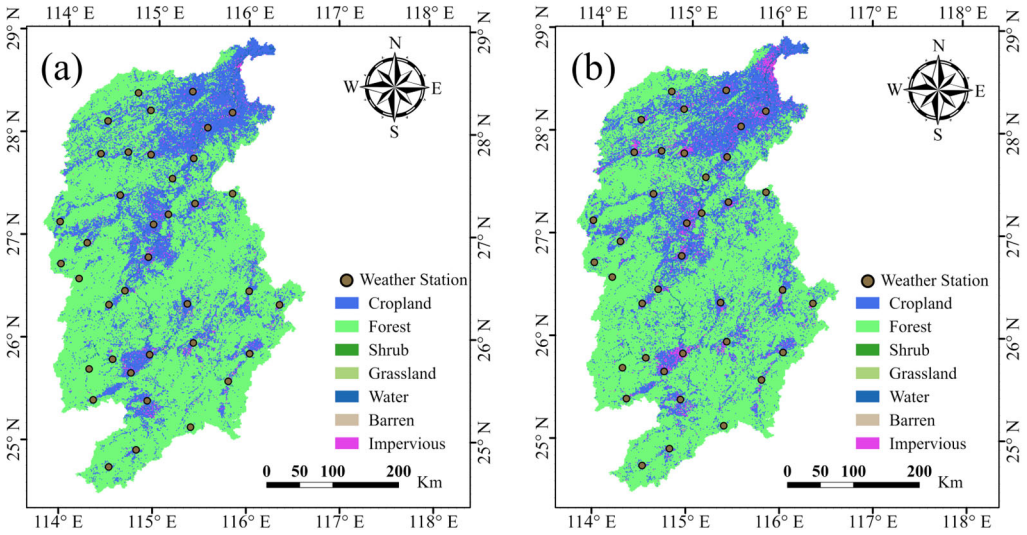

Figure 2. Land cover of Ganjiang River Basin for: (**a**) 1990 and (**b**) 2018.

2.2.4. Other Data

DEM data were obtained from the Geospatial Data Cloud (http://www.gscloud.cn/) (accessed on 25 July 2023) and resampled to a 1 km resolution. We considered elevation because its variation was thought to have a dramatic effect on vegetation growth [14,40]. The intensity of night-time light has been considered in numerous studies investigating the driving factors of vegetation change [12,41]. The Prolonged Artificial Nighttime-light Dataset of China (1984–2020) was provided by the National Tibetan Plateau/Third Pole Environment Data Center (http://data.tpdc.ac.cn) (accessed on 19 July 2023) [42].

The Harmonised World Soil Database (HWSD) version 1.1, a collaborative effort between the Food and Agriculture Organization of the United Nations (FAO) and the International Institute for Applied Systems Analysis (IIASA) in Vienna, was utilized. These datasets were accessible through the National Tibetan Plateau/Third Pole Environment Data Center [43]. We selected T_OC (0–30 cm organic carbon content) and T_CEC_CLAY (0–30 cm cation exchange capacity of clay layer soil) as soil property indicators because they were commonly used and more relevant to vegetation growth [44,45].

2.3. Methods

2.3.1. Feature Selection

In the study, a total of 58 diverse indicators were considered to estimate NDVI values. These included 35 climatic mean variables, 16 climate extreme indices, 1 atmospheric indicator, 2 land cover indicators, 1 elevation indicator, 1 night-time light indicator, and 2 soil indicators. Data anomalies were excluded using the principle of three times the standard deviation for NDVI data. We first conducted correlation analysis for the NDVI and 58 predictors. Then 38 indicators with a *p*-value of less than 0.1 (Table 2) were selected for following analysis.

Table 2. Correlation between NDVI and input variables. Selected features with a *p*-value less than 0.1 are shown in *italics*.

ID	Variable	Pearson	*p*-Value	ID	Variable	Pearson	*p*-Value	ID	Variable	Pearson	*p*-Value
1	*Forest*	0.57	<0.001	*21*	*Min_tem_4*	−0.125	<0.001	41	*PRCPTOT*	0.038	0.209
2	*Light*	−0.538	<0.001	*22*	*sun_1*	−0.121	<0.001	42	*rain*	0.037	0.223
3	*Min_tem_2*	−0.381	<0.001	*23*	*WS_2*	−0.121	<0.001	43	*R99P*	0.036	0.24
4	*Min_tem*	−0.34	<0.001	*24*	*T_CEC_CLAY*	−0.114	<0.001	44	*Max_tem*	−0.033	0.274
5	*TR20*	−0.309	<0.001	*25*	*Cropland*	0.111	<0.001	45	*Max_tem_4*	0.031	0.304
6	*Min_tem_3*	−0.305	<0.001	*26*	*sun_3*	−0.105	0.001	46	*rain_4*	−0.031	0.318
7	*DEM*	0.297	<0.001	*27*	*Max_tem_2*	−0.099	0.001	47	*WS_4*	−0.028	0.355
8	*tem_3*	−0.29	<0.001	*28*	*sun_4*	−0.094	0.002	48	*WSDI*	0.027	0.379
9	*Min_tem_1*	−0.257	<0.001	*29*	*R95P*	0.088	0.004	49	*CSDI*	0.027	0.384
10	*Hum*	0.256	<0.001	*30*	*WS_3*	−0.081	0.008	50	*rain_1*	0.023	0.46
11	*FD0*	0.241	<0.001	*31*	*WS*	−0.073	0.016	51	*CDD*	−0.018	0.547
12	*Hum_3*	0.237	<0.001	*32*	*tem_4*	−0.073	0.017	52	*SDII*	0.017	0.586
13	*Hum_4*	0.224	<0.001	*33*	*SU25*	−0.072	0.019	53	*R25MM*	0.016	0.611
14	*tem*	−0.221	<0.001	*34*	*CWD*	−0.065	0.033	54	*R20MM*	0.015	0.62
15	*Hum_2*	0.219	<0.001	*35*	*Max_tem_1*	−0.061	0.046	55	*sun*	−0.012	0.694
16	*sun_2*	−0.201	<0.001	*36*	*GSL*	−0.06	0.051	56	*rain_3*	0.009	0.765
17	*T_OC*	0.193	<0.001	*37*	*CO_2*	−0.058	0.057	57	*R10MM*	0.006	0.845
18	*tem_2*	−0.188	<0.001	*38*	*Max_tem_3*	−0.058	0.058	58	*ID0*	0.004	0.89
19	*Hum_1*	0.179	<0.001	*39*	rain_2	0.045	0.138				
20	*tem_1*	−0.134	<0.001	40	WS_1	−0.043	0.156				

2.3.2. Trend Analysis

In this study, we analyzed the NDVI trends in the GRB from 1990 to 2018 using the Theil–Sen slope test and evaluated their statistical significance through the Mann–Kendall significance test [46,47]. The Theil–Sen slope estimation was more robust than traditional slope analysis because of its resistance to noise and capability to mitigate the effects of outliers, measurement errors, and discrete data [48]. A positive slope (>0) denoted an upward trend in the study period for the time series data, while a negative slope indicated the opposite. The Mann–Kendall (MK) significance test was used to evaluate the significance of the observed change trend. Integrating these statistical tests not only mitigates the impact of anomalous NDVI values but also visually represents the spatial distribution of NDVI variability through NDVI variability partitioning, a method

recognized and widely used in assessing NDVI trends [49,50]. The calculation process is as follows:

$$\text{Slope} = \text{Median}\left[\frac{(x_j - x_i)}{(j - i)}\right] \quad 1990 \le i < j \le 2018 \tag{1}$$

where Slope is the value of the slope estimated by the Theil–Sen median; x is the value of the NDVI of each year in this study; and i and j are the different years between 1990 and 2018.

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{Sgn}(x_j - x_i) \tag{2}$$

$$\text{Sgn}(x_j - x_i) = \begin{cases} 1 & (x_j - x_i) > 0 \\ 0 & (x_j - x_i) = 0 \\ -1 & (x_j - x_i) < 0 \end{cases} \tag{3}$$

$$\text{Var}(s) = \frac{n(n-1)(2n+5)}{18} \tag{4}$$

$$Z = \begin{cases} \frac{(S-1)}{\sqrt{\text{Var}(S)}} & S > 0 \\ 0 & S = 0 \\ \frac{(S+1)}{\sqrt{\text{Var}(S)}} & S < 0 \end{cases} \tag{5}$$

where Z is the standardized test statistic; n is the number of time series data, which equals 29 in this study; and Sgn is the function symbol.

2.3.3. Random Forest

Random forest (RF), developed by Breiman in 2001, is a widely adopted ensemble learning technique capable of performing both classification and regression tasks [51]. This approach improves upon the classification and regression trees (CART) method by combining multiple decision trees. In the RF algorithm, at each node of the forest, the optimal feature for splitting is determined using a random subset of available features. This random feature selection strategy enhances the model's accuracy and robustness, ensuring precise and resilient results [52].

2.3.4. XGBoost

The extreme gradient boosting (XGBoost) algorithm, initially proposed by Chen and Guestrin in 2016 [53], is an efficient implementation of the gradient boosting decision tree (GBDT). The XGBoost model utilizes the decision tree as its base classifier. This technique involves a series of gradient-enhanced decision trees [54], where each tree learns from its predecessor and influences the subsequent tree, thereby enhancing the model and constructing a robust learner. This iterative process improves model generalization, addressing decision trees' tendency to overfit. XGBoost is known for its high accuracy, fast processing speed, and low computational cost and complexity. Moreover, as a decision tree algorithm, XGBoost is not negatively impacted by multicollinearity [55]. This further contributes to its robustness in handling complex datasets.

2.3.5. Model Performance Assessment

Both the RF and XGBoost model were run 100 times. In each run, 70% of the total data was randomly chosen for training the model, and the remaining 30% was used to evaluate the model's performance. Model performance was evaluated using the coefficient of determination (R^2) and root-mean-squared error (RMSE) [56]. Higher R^2 values and lower RMSE values generally indicate superior model performance [57]. Data processing

and machine learning was carried out using Python 3.9, mainly utilizing computational packages like "sklearn", "xgboost", "pandas, "numpy", "matplotlib", and "seaborn".

Initially, we trained the model using climate and atmospheric CO_2 data. Then, we incorporated elevation and soil data into the training. Lastly, we included land cover and night-time light data in the final round of training (Figure 3).

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(NDVI - NDVI_{pre})^2}{\sum_{i=1}^{N}(NDVI - \overline{NDVI})^2} \quad (6)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(NDVI - NDVI_{pre})^2}{N}} \quad (7)$$

where $NDVI$ is the measured value, $NDVI_{pre}$ is the predicted value, $\overline{NDVI}$ is the average value of measured $NDVI$, and N is the number of measured $NDVI$ samples.

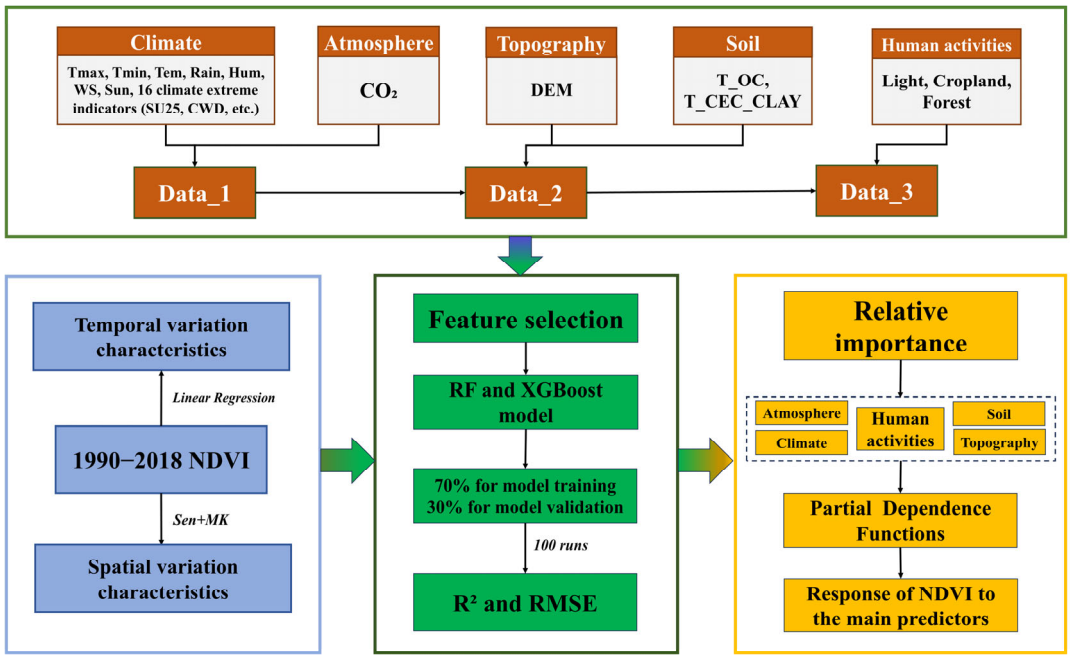

Figure 3. Framework of NDVI estimation integrating different sources of data with RF and XGBoost models. Rain, rainfall; Tem, mean temperature; Tmax, max temperature; Tmin, min temperature; Hum, air humidity; WS, wind speed; Sun, sun hour; SU25, summer days; CWD, consecutive wet days; DEM, elevation; T_OC, soil organic carbon content at 0–30 cm; T_CEC_CLAY, 0–30 cm cation exchange capacity of clay layer soil; Light, night-time light.

3. Results

3.1. Spatial and Temporal Change of the NDVI in the Study Sites

In this study, the years 1990 to 2018 were segmented into three distinct periods: 1990–1999, 2000–2009, and 2010–2018 (Figure 4). For each period, linear regression analysis was conducted on the NDVI data. The findings revealed trends in the NDVI across the 38 sites in the GRB, with specific trends varying across these time intervals. During 1990 to 1999, NDVI values showed an increasing trend, with an R^2 of 0.44 and $p < 0.05$, indicating a clear upward trend. From 2000 to 2009, NDVI values exhibited a slight upward trend, with the slope decreasing from 0.0021 to 0.0003, and the R^2 decreasing to 0.01 with $p > 0.05$, indicating no clear upward or downward trend during this period. Subsequently, from 2010 to 2018, the trend of the NDVI became negative (−0.009), with the R^2 increasing to 0.75 and $p < 0.005$, indicating a strong linear relationship. These trend changes reflected variations

in environmental conditions or vegetation dynamics over time. The NDVI initially showed a clear upward trend in the early 1990s but began to gradually weaken from the early 2000s onwards (Figure 4). Additionally, compared to the 1990–1999 and 2000–2009 periods, the NDVI exhibited a downward trend during the 2010–2018 period, suggesting the presence of significant interference during these times.

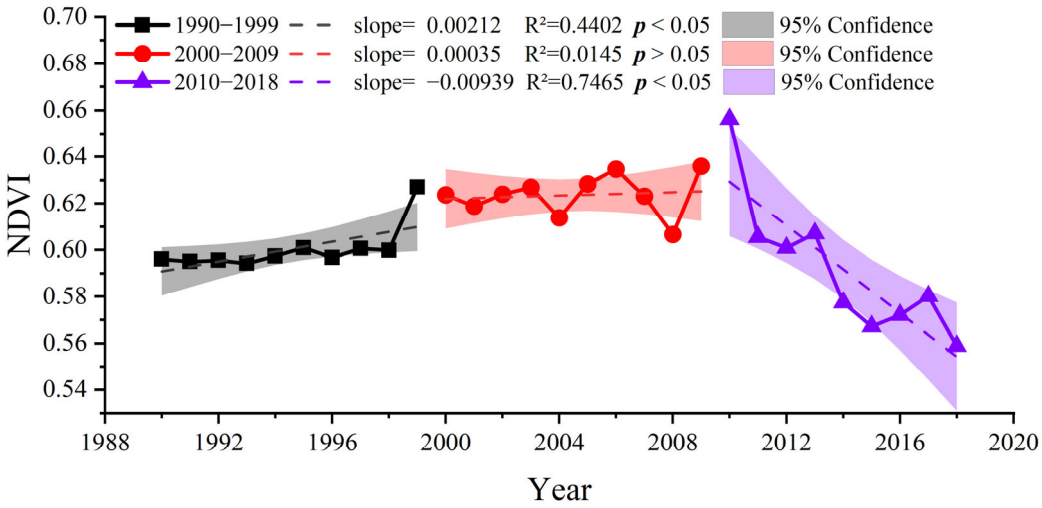

Figure 4. Temporal change of annual mean NDVI across the 38 sites in the GRB between 1990 and 2018, where the black solid-line box, the red solid-line circle, and the blue solid-line triangle represent three time spans: 1990–1999, 2000–2009, and 2010–2018, respectively.

Among the 38 weather stations analyzed, the majority exhibited a unsignificant trend (Figure 5). A total of 15 stations displayed an increasing trend, with 6 of these stations showing statistically significant results. The station with the strongest upward trend was ANY station, with a slope of 0.0036. Conversely, 23 stations demonstrated a declining trend, with 6 stations indicating a statistically significant decrease. NK station showed the steepest downward trend, with a slope of −0.0065. Analysis in conjunction with Figure 2 suggested that stations surrounded by impervious surfaces predominantly showed negative trends, while those near cropland and forests exhibited positive trends. This implied that land use around weather stations has a significant impact on NDVI variations.

3.2. NDVI Estimation with Machine Learning

Both RF and XGBoost models were utilized to assess the impact of different data sources on NDVI variations. The results clearly demonstrated that the performance of both the RF and XGBoost models consistently enhanced with the inclusion of additional input data. When only climate and atmospheric data (CO_2 concentration) were used as input, the RF model attained an R^2 value of 0.41 and an RMSE of 0.079 (Figure 6). Upon retraining the model for the second time with the addition of DEM and soil data, the R^2 value increased to 0.744 and the RMSE decreased to 0.052. Subsequently, after incorporating land cover and night-time light data into the second model training, the accuracy of the RF model improved, resulting in an increased R^2 value of 0.814 and a decreased RMSE to 0.044. It was observed that the performance of the XGBoost model and the RF model varied when subjected to the same data source. Although both models initially exhibited comparable performance using only climate and atmospheric data, the XGBoost model consistently outperformed the RF model as additional data were introduced during the second and third training rounds. Ultimately, the XGBoost model trained with the Data_3 data source (Climate + CO_2 + DEM + Soil + Light + Land cover) was selected as the optimal model,

demonstrating the highest degree of accuracy with an R^2 value of 0.848 and an RMSE of 0.04.

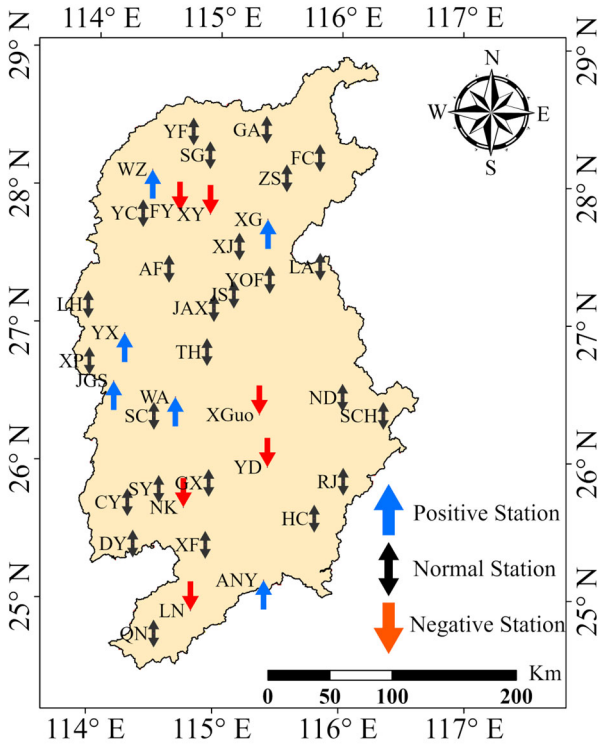

Figure 5. The change of NDVI slope trends and significance levels at 38 sites across the GRB during 1990–2018. The ascending arrows are for significant positive trends, descending arrows for significant negative trends, and double-headed arrows for no significant trend. The following is the corresponding name of each weather station: AF—anfu, ANY—anyuan, CY—chongyi, DY—dingnan, FC—dayu, FY—fenyi, GA—gaoan, GX—ganxian, HC—huichang, JAX—jianxian, JGS—jinggangshan, JS—jishui, LA—lean, LH—lianhua, LN—longnan, ND—ningdu, NK—nankang, QN—quannan, RJ—ruijin, SC—suichuan, SCH—shicheng, SG—shanggao, SY—shangyou, TH—taihe, WA—wanan, WZ—wanzai, XF—xinfeng, XG—xingan, XGuo—xingguo, XJ—xiajiang, XP—xiaping, XY—xinyu, YC—yichun, YD—yudu, YF—yifeng, YOF—yongfeng, YX—yongxin, ZS—zhangshu.

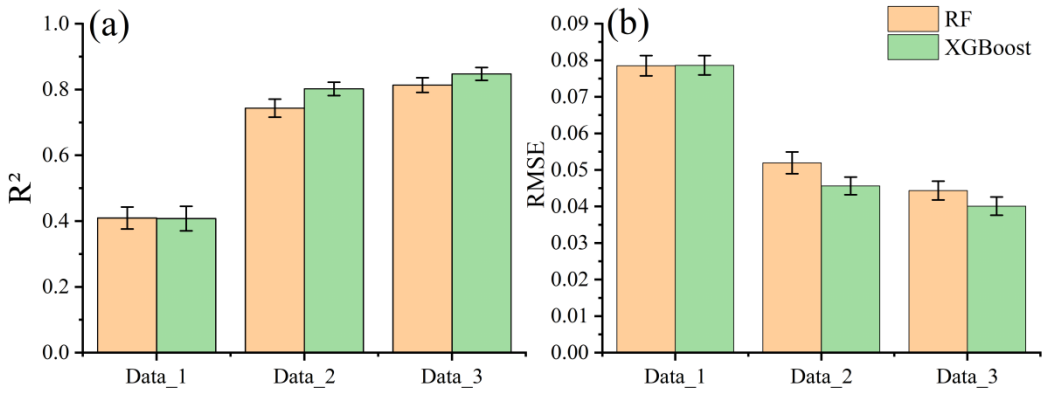

Figure 6. Evaluation of model performance using different sources of data based on the random forest (RF) and extreme gradient boosting (XGBoost) model. Data_1: Climate + CO_2 data; Data_2: Data_1 + DEM + Soil data; Data_3: Data_2 + Light + Land cover data; (**a**) coefficient of determination; (**b**) root mean squared error.

3.3. The Response of the NDVI to Predictors

The top 20 indicators of model importance are shown in Figure 7a and the results indicate that the area of forest was the most important factor influencing the NDVI. The change

in cropland area ranked as the second most influential factor, followed by soil organic carbon content as the third most significant influencer. Regarding climate, temperature was more significant than rainfall. The minimum temperature in spring emerged as the most influential climate factor, with minimum temperature generally outweighing maximum temperature in importance, except in summer.

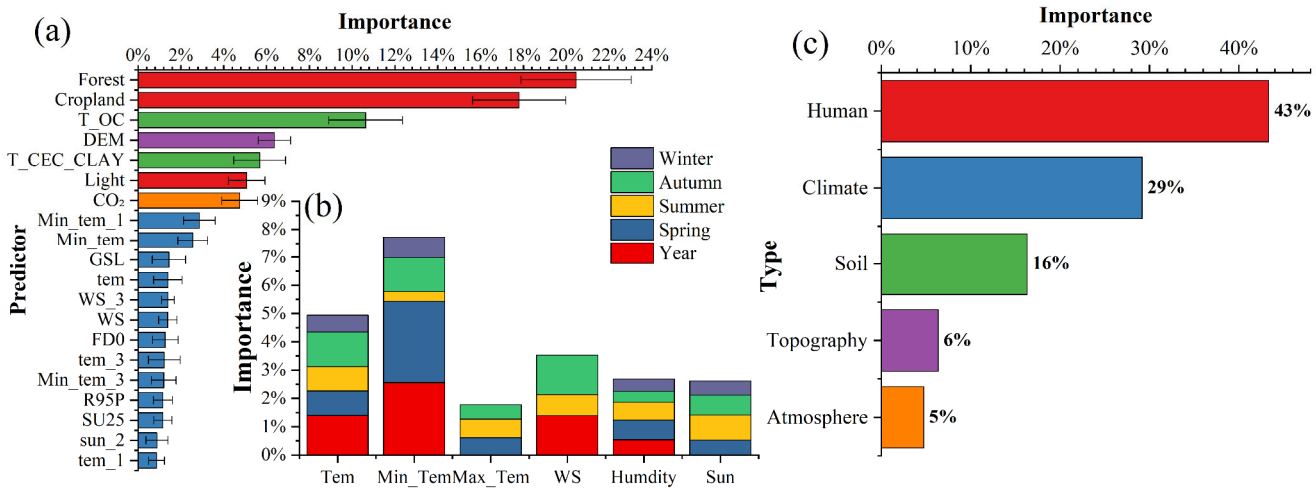

Figure 7. Variable importance of input features for the XGBoost model predicting the NDVI: (**a**) the relative importance of the first 20 predictor variables, (**b**) the relative importance of climate variables at different time scales, (**c**) the relative importance aggregated by feature type. For each category, the relative importance shown is the sum of that calculated for all features in each category. Note that 1, 2, 3, and 4 represent the different seasons (1 for spring, 2 for summer, 3 for autumn, and 4 for winter). For example, WS_1 indicates spring wind speed and tem_2 indicates summer temperature. Climate: climate variables and climate extreme index; Human: land cover and night-time light. Error bar is based on 100 runs for the model.

The relative importance of climate indicators for each season is shown in Figure 7b. Minimum temperature emerged as the most influential factor on both an annual and spring scale, exerting the greatest impact during spring. The importance of humidity remained consistent across seasonal scales. Sunshine hours held greater importance in summer compared to other seasons, serving as the primary forcing factor during this period. Wind speed was highest in autumn and was the main influence factor in autumn. Overall, minimum and mean temperatures wielded greater influence compared to other factors. Climatic factors in winter exhibited a weaker influence on the NDVI, while those in spring and autumn demonstrated a more significant impact. When categorizing each indicator (Figure 7c), it was revealed that climatic factors contributed approximately 29.19%. Human activities accounted for 43.36% of the relative importance, with changes in forest area exerting the greatest influence, followed by alterations in cropland area. Elevation and soil factors contributed 22.7% of the total importance, with soil organic carbon content being more significant than other soil factors.

Land cover strongly influenced vegetation growth. Changes in forest and cropland exerted a significant impact on the NDVI (Figure 8). An increase in the forest and cropland area ratio had a positive effect, and this relationship was found to be near-linear. However, the influence diminished when the proportion of forest exceeded 60%. Additionally, increased night-time lighting due to human development negatively affected the NDVI, and this effect was also near-linear (Figure 8f). Regarding soil properties, there was a positive linear relationship between the soil organic carbon content and the NDVI within the range of 1–1.5%, with the NDVI reaching its maximum value at approximately 1.5%. Conversely, there was a negative near-linear relationship between the cation exchange

capacity and the NDVI in the range of 10–40 cmol/kg. Elevation affected the NDVI by about 6.4%, with a gradual increase in the NDVI observed in the 100–350 m range. NDVI values steadily increased with rising CO_2 concentrations. However, at high CO_2 concentrations, particularly in the range of 380–390 ppm, there was minimal to no increase in the NDVI despite rising CO_2 levels (Figure 8g). Beyond this range, the NDVI showed a rapid increase with further CO_2 elevation. An increase in spring minimum temperature (Figure 8h) negatively impacted the NDVI, higher spring minimum temperatures correlated with decreased NDVI values. Additionally, it was worth noting that the adverse effect of spring minimum temperature on the NDVI diminished once it surpassed 16 °C. Similarly, the annual minimum temperature began to demonstrate a decreasing effect only after exceeding 14.5 °C.

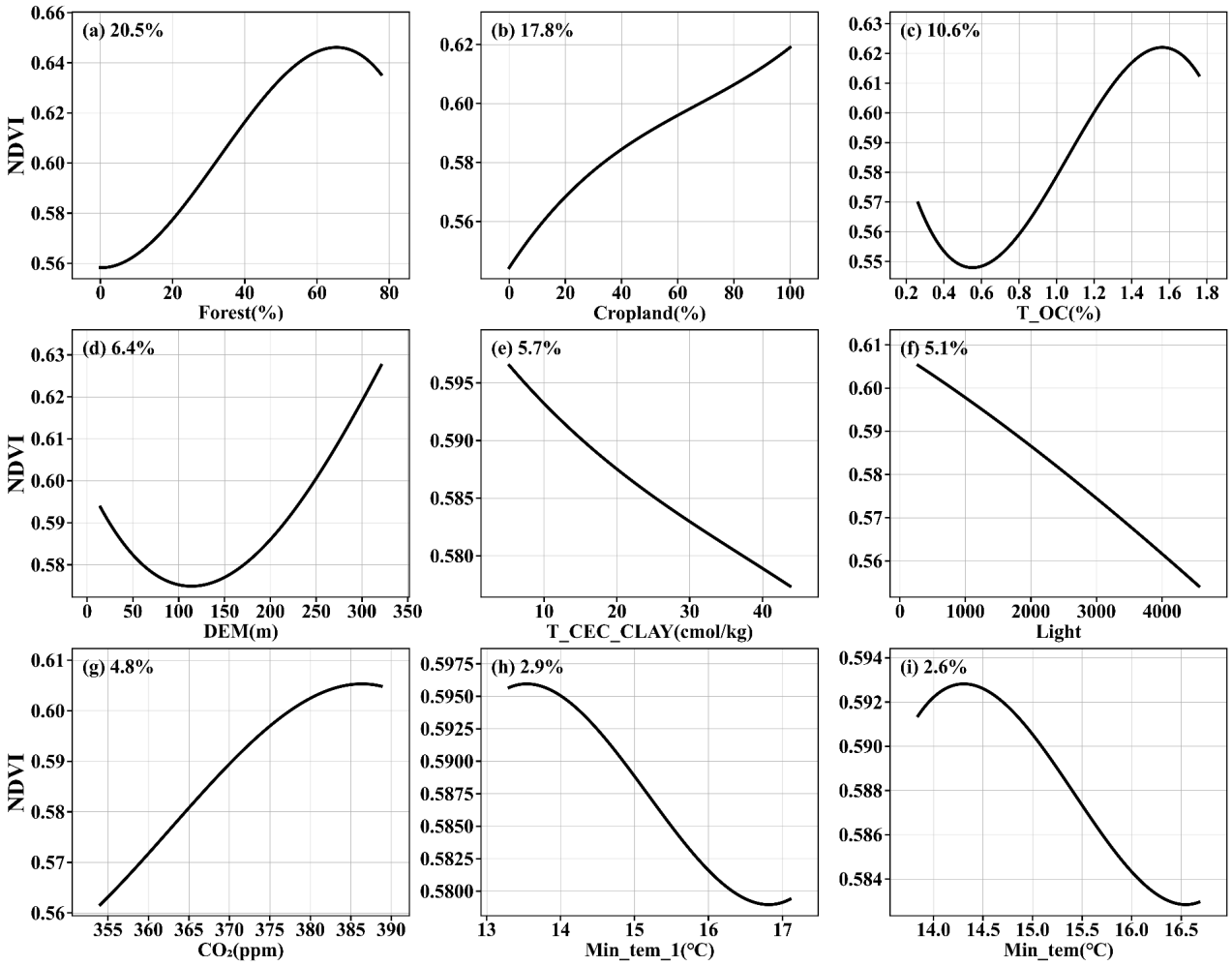

Figure 8. Partial dependence plots for the most important nine features based on the XGBoost model: (**a**) forest ratio; (**b**) cropland ration; (**c**) organic carbon content; (**d**) DEM; (**e**) cation exchange capacity of clay layer; (**f**) night −time light; (**g**) CO_2 concentration; (**h**) spring min temperature; (**i**) min temperature. The black lines are smoothed representations of the response, with fitted values (model predictions) for the training data. The trend of the line, rather than the actual values, describe the nature of the dependence of the NDVI on the predictors.

At the 38 sites, the proportion of cropland area had consistently been higher than that of forest area (Figure 9), indicating that cropland was the primary land cover type near the weather stations. Notably, both the proportions of forest and cropland area showed a significant declining trend from 1990 to 2018. Forest area decreased at a rate of approximately 0.2% per year, while cropland area decreased at a faster rate of 0.8% per

year. This indicated that while the overall trend in the past 30 years has been a decline in forest and farmland area, there were exceptions, such as in 1993, 2000, and 2010, when the trend shifted to an increase.

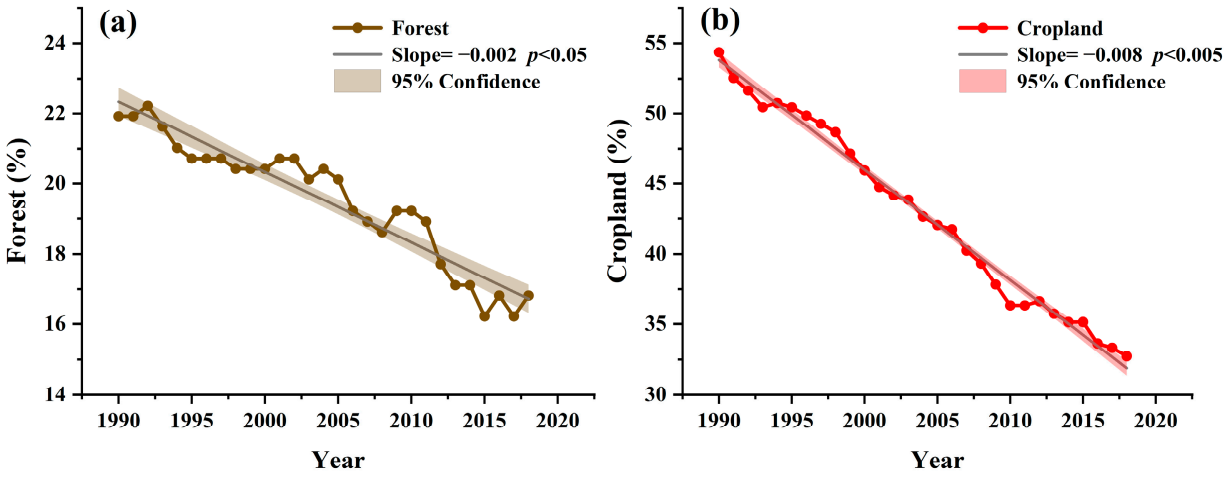

Figure 9. Temporal changes in the average ratio of (**a**) forest and (**b**) cropland across the 38 sites in the GRB between 1990 and 2018.

4. Discussion

4.1. The Change Trend of the NDVI in the GRB

From the analysis of 38 weather stations, we observed that the greening trend at most stations was either negligible or had deteriorated. Between 1990 and 1999, the NDVI exhibited an upward trend, likely due to the relatively small decrease in forest and farmland areas near the stations and minimal temperature variation during this period. These factors had limited negative effects on the NDVI, while the rising CO_2 concentration exerted a stronger positive influence, leading to a modest increase in vegetation. However, from 2000 to 2018, intensified human activities caused a more significant reduction in nearby forest and farmland areas. Concurrently, the rise in minimum temperatures exacerbated the negative impacts on vegetation, outweighing the positive effects and resulting in a noticeable decline in the NDVI during this period. These findings suggest that the NDVI changes were driven by the combined interactions of multiple factors rather than by individual influences, which aligns with prior research [58,59].

4.2. The Application of the Machine Learning Approach in the GRB

In this study, we employed two widely used machine learning models, RF and XGBoost, to evaluate the driving factors of the NDVI. This approach allowed us to reveal complex variable interactions influencing vegetation changes, yielding accurate and interpretable results [20,21]. Our results demonstrated the feasibility of developing machine learning models at the site scale. Previous studies also investigated NDVI driving factors. For instance, Zeng et al. [60] used RF to explain browning in the Dosso Reserve, Niger, achieving an R^2 value greater than 0.71. Similarly, a study in the Amazon region combined XGBoost with SHAP, resulting in an R^2 of 0.92 [61]. Roy [62] applied SVR, RF, linear, and polynomial models for NDVI prediction, with R^2 values exceeding 0.8. On a larger scale, research on NDVI prediction across China also reported an R^2 above 0.8 [63]. Compared to these previous studies, our prediction results were similar, with an R^2 of 0.85. The slight variations may be due to differences in scale and geographic region across the studies.

4.3. The Dominant Driving Factors on the NDVI in the GRB

We found that human activities emerged as the dominant driving factor of NDVI change, constituting 43.36% of the total importance. This finding aligns with previous studies suggesting that land cover changes in southeastern China exert a significant influence on vegetation dynamics [2,64]. Cropland and forest were the primary contributors to NDVI variation in the GRB, with their fluctuations having a substantial impact on vegetation. Previous studies showed that NDVI values for cropland and forest were higher than for other land use types [65], indicating that if the ratio of forest and cropland decreases, the NDVI value will also decrease. From a spatial perspective, land retirement and afforestation convert farmland into forest, increasing vegetation cover. Since the NDVI measures vegetation density and health, this increase in cover leads to more sunlight absorption and reflection, thus raising NDVI values. In contrast, reclamation often transforms forests into farmland, reducing vegetation cover and lowering the NDVI. Future research should consider spatial variability, such as the conversion between land use types like forests, farmland, urban areas, grasslands, and water bodies, which requires more detailed spatial datasets and research methods. Additionally, while natural disasters like mudslides, earthquakes, and floods can alter land use, their frequency determines that their impact is small. We found a negative correlation between night-time lights and the NDVI, as night-time lights reflected regional economic development, human activity intensity, and the expansion of built-up land. Higher levels of night-time light typically indicated stronger human interference [41]. The impact of human activities is mainly related to the development policies adopted by the government. If human behavior is allowed to affect the environment without regulation, further reductions in cropland and forest areas (Figure 9) may impede vegetation greening trends and contribute to environmental degradation.

We identified CO_2 as a major influencing factor due to its role as a crucial substrate for photosynthesis. The rise in atmospheric CO_2 concentration can enhance photosynthesis by increasing the rate of carboxylation [66], a phenomenon known as the "CO_2 fertilization effect", which has been confirmed in previous work [1]. Regarding soil factors, soil organic carbon and clay cation exchange capacity contributed approximately 16% to NDVI changes. Higher organic carbon content enriched the soil, providing essential nutrients for vegetation growth [67], thus promoting plant development. Additionally, vegetation luxuriance was often linked to cation exchange capacity, as noted in previous research [68]. It was important to note that human activities impact SOC and CEC [69]. The application of fertilizers can increase organic carbon in soil [70], while land use changes, such as the conversion of forests to farmland, reduce SOC [71]. Fertilization and land management practices can raise soil pH, thereby increasing the amount of variable negative charges and subsequently enhancing CEC [69]. Future studies could quantify the extent of these indirect effects to better understand these interactions. Elevation also emerged as an important factor, particularly at altitudes below 400 m, where the NDVI and altitude exhibited a nearly linear relationship. This can be attributed to increased human activity at lower elevations. As altitude rises, human influence on vegetation diminishes [40]. However, if altitude continues to rise, the growth of vegetation will be inhibited due to the drop in ambient temperature.

We found that climate is a secondary driver of vegetation dynamics, with temperature playing a more significant role than rainfall. In the subtropical humid monsoon climate of the GRB, abundant rainfall ensures that water is not a limiting factor for vegetation growth, making temperature the primary constraint [72]. However, during the rainy season, cloud cover reduces the solar radiation available to vegetation, which may weaken the positive impact of rainfall on growth [37]. This does not mean that rainfall is a limiting factor, but rather that its benefits for vegetation growth may be less pronounced under specific

conditions. These findings align with previous studies that highlight the greater importance of temperature over rainfall in driving vegetation changes in the GRB [28,37].

4.4. Limitations

Our study has certain limitations that should be acknowledged. Firstly, we relied on data from only 38 weather stations in the GRB. Although this approach ensured the accuracy of the climate data, it limited the overall scope of the study. Future research could benefit from using gridded high-resolution meteorological data to facilitate large-scale basin analyses. Second, other factors such as population density, economic development, and grazing conditions have been shown to influence the NDVI [41]. However, we did not consider them due to their mismatched conditions, low impact levels, and difficulty in downscaling at the site level. Third, it is important to acknowledge that statistical models are data-driven, and their efficacy heavily relies on data quality [56]. Machine learning models may be constrained by data and algorithms [73]. In addition, the relationships identified between environmental factors and vegetation growth through machine learning may be statistical rather than causal [74], which limits our comprehensive understanding of the mechanisms driving NDVI changes. Finally, vegetation dynamics are shaped by the interaction of multiple factors, including climate conditions, land use changes, and human activities. While our study identifies the primary drivers, the potential interactions among these factors remain underexplored. For instance, the synergistic effects of deforestation and climate change may lead to a compounded decline in vegetation health [75], while afforestation can mitigate soil degradation [76]. To better understand these combined effects, future analyses could employ multivariate statistical models to assess interactions among predictive factors, providing more reliable insights into the drivers of vegetation change.

5. Conclusions

In this study, we investigated the spatiotemporal dynamics of the annual mean NDVI in the GRB and applied machine learning techniques to estimate the NDVI and identify the key factors driving vegetation changes. These factors included climate variables, atmospheric conditions, soil characteristics, topography, and land cover patterns. The main findings of the study are as follows:

1. The annual mean NDVI experienced a slight increase from 1990 to 1999 but has significantly declined over the last 8 years.
2. XGBoost outperformed the RF model in simulating the NDVI, yielding the best performance when incorporating all data sources (R^2 = 0.85; RMSE = 0.04).
3. The most influential variables affecting the NDVI were forest and cropland ratio, followed by organic carbon content, elevation, cation exchange capacity, night-time light intensity, CO_2 concentration, and spring minimum temperature.
4. Partial dependence plots revealed both linear and nonlinear relationships between the NDVI and the variables, with most variables exhibiting threshold effects on NDVI.

Our results underscored the large impact of land cover changes on NDVI variations in the GRB over the past 30 years. The expansion of human settlements has notably altered forest and crop cover, contributing to the observed NDVI decline. This highlights the urgent need to develop and implement effective land management strategies that can regulate and mitigate the adverse effects of these changes. By prioritizing sustainable land use practices, such as protecting forested areas and optimizing agricultural land management, it is possible to enhance vegetation health and support ecological balance in our study region.

Author Contributions: Conceptualization, L.G., B.W., X.C., X.F. and K.Z.; Methodology, Z.X., K.L., B.W. and H.H.; Software, Z.X. and H.H.; Formal analysis, K.L. and X.F.; Resources, L.G. and X.C.;

Data curation, Z.L. and F.S.; Writing—original draft, Z.X. and K.L.; Writing—review & editing, L.G. and B.W.; Supervision, L.G. and F.C.; Project administration, L.G. and F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (42161073, 41905104, 42367032).

Data Availability Statement: Dataset available on request from the authors.

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

References

1. Piao, S.; Wang, X.; Park, T.; Chen, C.; Lian, X.; He, Y.; Bjerke, J.W.; Chen, A.; Ciais, P.; Tømmervik, H.; et al. Characteristics, Drivers and Feedbacks of Global Greening. *Nat. Rev. Earth Environ.* **2019**, *1*, 14–27. [CrossRef]
2. Chen, C.; Park, T.; Wang, X.; Piao, S.; Xu, B.; Chaturvedi, R.K.; Fuchs, R.; Brovkin, V.; Ciais, P.; Fensholt, R.; et al. China and India Lead in Greening of the World Through Land-Use Management. *Nat. Sustain.* **2019**, *2*, 122–129. [CrossRef] [PubMed]
3. Corlett, R.T. Impacts of Warming on Tropical Lowland Rainforests. *Trends Ecol. Evol.* **2011**, *26*, 606–613. [CrossRef] [PubMed]
4. Ukkola, A.M.; De Kauwe, M.G.; Roderick, M.L.; Burrell, A.; Lehmann, P.; Pitman, A.J. Annual Precipitation Explains Variability in Dryland Vegetation Greenness Globally but Not Locally. *Glob. Chang. Biol.* **2021**, *27*, 4367–4380. [CrossRef] [PubMed]
5. Berdimbetov, T.; Ilyas, S.; Ma, Z.; Bilal, M.; Nietullaeva, S. Climatic Change and Human Activities Link to Vegetation Dynamics in the Aral Sea Basin Using NDVI. *Earth Syst. Environ.* **2021**, *5*, 303–318. [CrossRef]
6. Castro Sardiña, L.; Irisarri, G.; Texeira, M. Climate Factors Rather than Human Activities Controlled NDVI Trends across Wet Meadow Areas in the Andes Centrales of Argentina. *J. Arid. Environ.* **2023**, *214*, 104983. [CrossRef]
7. Ghebrezgabher, M.G.; Yang, T.; Yang, X.; Eyassu Sereke, T. Assessment of NDVI Variations in Responses to Climate Change in the Horn of Africa. *Egypt. J. Remote Sens. Space Sci.* **2020**, *23*, 249–261. [CrossRef]
8. Liu, H.; Li, X.; Mao, F.; Zhang, M.; Zhu, D.; He, S.; Huang, Z.; Du, H. Spatiotemporal Evolution of Fractional Vegetation Cover and Its Response to Climate Change Based on MODIS Data in the Subtropical Region of China. *Remote Sens.* **2021**, *13*, 913. [CrossRef]
9. Li, C.; Jia, X.; Zhu, R.; Mei, X.; Wang, D.; Zhang, X. Seasonal Spatiotemporal Changes in the NDVI and Its Driving Forces in Wuliangsu Lake Basin, Northern China from 1990 to 2020. *Remote Sens.* **2023**, *15*, 2965. [CrossRef]
10. Van Vliet, J. Direct and Indirect Loss of Natural Area from Urban Expansion. *Nat. Sustain.* **2019**, *2*, 755–763. [CrossRef]
11. Winkler, A.J.; Myneni, R.B.; Hannart, A.; Sitch, S.; Haverd, V.; Lombardozzi, D.; Arora, V.K.; Pongratz, J.; Nabel, J.E.M.S.; Goll, D.S.; et al. Slowdown of the Greening Trend in Natural Vegetation with Further Rise in Atmospheric CO_2. *Biogeosciences* **2021**, *18*, 4985–5010. [CrossRef]
12. Jiang, L.; Liu, Y.; Wu, S.; Yang, C. Analyzing Ecological Environment Change and Associated Driving Factors in China Based on NDVI Time Series Data. *Ecol. Indic.* **2021**, *129*, 107933. [CrossRef]
13. Zhang, S.; Bai, X.; Zhao, C.; Tan, Q.; Luo, G.; Cao, Y.; Deng, Y.; Li, Q.; Li, C.; Wu, L.; et al. Limitations of Soil Moisture and Formation Rate on Vegetation Growth in Karst Areas. *Sci. Total Environ.* **2022**, *810*, 151209. [CrossRef]
14. Mishra, N.B.; Mainali, K.P. Greening and Browning of the Himalaya: Spatial Patterns and the Role of Climatic Change and Human Drivers. *Sci. Total Environ.* **2017**, *587–588*, 326–339. [CrossRef]
15. Ibrahim, Y.; Balzter, H.; Kaduk, J.; Tucker, C. Land Degradation Assessment Using Residual Trend Analysis of GIMMS NDVI3g, Soil Moisture and Rainfall in Sub-Saharan West Africa from 1982 to 2012. *Remote Sens.* **2015**, *7*, 5471–5494. [CrossRef]
16. Eze, E.; Siegmund, A. Identifying Disaster Risk Factors and Hotspots in Africa from Spatiotemporal Decadal Analyses Using INFORM Data for Risk Reduction and Sustainable Development. *Sustain. Dev.* **2024**, *32*, 4020–4041. [CrossRef]
17. Venkatesh, K.; John, R.; Chen, J.; Xiao, J.; Amirkhiz, R.G.; Giannico, V.; Kussainova, M. Optimal Ranges of Social-Environmental Drivers and Their Impacts on Vegetation Dynamics in Kazakhstan. *Sci. Total Environ.* **2022**, *847*, 157562. [CrossRef] [PubMed]
18. Bao, Z.; Zhang, J.; Wang, G.; Guan, T.; Jin, J.; Liu, Y.; Li, M.; Ma, T. The Sensitivity of Vegetation Cover to Climate Change in Multiple Climatic Zones Using Machine Learning Algorithms. *Ecol. Indic.* **2021**, *124*, 107443. [CrossRef]
19. Shi, Y.; Jin, N.; Ma, X.; Wu, B.; He, Q.; Yue, C.; Yu, Q. Attribution of Climate and Human Activities to Vegetation Change in China Using Machine Learning Techniques. *Agric. For. Meteorol.* **2020**, *294*, 108146. [CrossRef]
20. Berdugo, M.; Gaitán, J.J.; Delgado-Baquerizo, M.; Crowther, T.W.; Dakos, V. Prevalence and Drivers of Abrupt Vegetation Shifts in Global Drylands. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2123393119. [CrossRef]
21. Nguyen, K.A.; Seeboonruang, U.; Chen, W. Projected Climate Change Effects on Global Vegetation Growth: A Machine Learning Approach. *Environments* **2023**, *10*, 204. [CrossRef]
22. Wang, Z.; Fu, B.; Wu, X.; Wang, S.; Li, Y.; Feng, Y.; Zhang, L.; Hu, Y.; Cheng, L.; Li, B. Distinguishing Trajectories and Drivers of Vegetated Ecosystems in China's Loess Plateau. *Earth's Future* **2024**, *12*, e2023EF003769. [CrossRef]

23. Liu, X.; Zhang, L.; She, D.; Chen, J.; Xia, J.; Chen, X.; Zhao, T. Postprocessing of Hydrometeorological Ensemble Forecasts Based on Multisource Precipitation in Ganjiang River Basin, China. *J. Hydrol.* **2022**, *605*, 127323. [CrossRef]
24. Zheng, B.; Wu, S.; Song, X.; Huang, Y.; Wu, H.; Liu, Z.; Zhu, J.; Wan, W. Impacts of Landscape Pattern Evolution on Typical Ecosystem Services in Ganjiang River Basin, China. *Environ. Sci. Pollut. Res.* **2023**, *30*, 110562–110578. [CrossRef]
25. Huang, L.; Ye, A.; Tang, C.; Duan, Q.; Zhang, Y. Impact of Rural Depopulation and Climate Change on Vegetation, Runoff and Sediment Load in the Gan River Basin, China. *Hydrol. Res.* **2020**, *51*, 768–780. [CrossRef]
26. Li, H.; Zou, A.; Kong, D.; Ma, Z. Spatiotemporal Characteristics of Precipitation Extremes Based on Reanalysis Precipitation Data during 1950–2020 over the Ganjiang River Basin and Its Surroundings, China. *Atmos. Sci. Lett.* **2023**, *24*, e1149. [CrossRef]
27. Liu, H.; Zheng, L.; Liao, M. Dynamics of Vegetation Change and Its Relationship with Nature and Human Activities—A Case Study of Poyang Lake Basin, China. *J. Sustain. For.* **2021**, *40*, 47–67. [CrossRef]
28. Fu, B.; Yang, W.; Yao, H.; He, H.; Lan, G.; Gao, E.; Qin, J.; Fan, D.; Chen, Z. Evaluation of Spatio-Temporal Variations of FVC and Its Relationship with Climate Change Using GEE and Landsat Images in Ganjiang River Basin. *Geocarto Int.* **2022**, *37*, 13658–13688. [CrossRef]
29. Huang, Y.; Huang, B.; Qin, T.; Nie, H.; Wang, J.; Li, X.; Shen, Z. Assessment of Hydrological Changes and Their Influence on the Aquatic Ecology over the Last 58 Years in Ganjiang Basin, China. *Sustainability* **2019**, *11*, 4882. [CrossRef]
30. Tang, J.; Song, P.; Hu, X.; Chen, C.; Wei, B.; Zhao, S. Coupled Effects of Land Use and Climate Change on Water Supply in SSP–RCP Scenarios: A Case Study of the Ganjiang River Basin, China. *Ecol. Indic.* **2023**, *154*, 110745. [CrossRef]
31. Myneni, R.B.; Keeling, C.D.; Tucker, C.J.; Asrar, G.; Nemani, R.R. Increased Plant Growth in the Northern High Latitudes from 1981 to 1991. *Nature* **1997**, *386*, 698–702. [CrossRef]
32. Huang, S.; Tang, L.; Hupy, J.P.; Wang, Y.; Shao, G. A Commentary Review on the Use of Normalized Difference Vegetation Index (NDVI) in the Era of Popular Remote Sensing. *J. For. Res.* **2021**, *32*, 1–6. [CrossRef]
33. Mao, D.; Wang, Z.; Luo, L.; Ren, C. Integrating AVHRR and MODIS Data to Monitor NDVI Changes and Their Relationships with Climatic Parameters in Northeast China. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *18*, 528–536. [CrossRef]
34. Yang, W.; Kogan, F.; Guo, W.; Chen, Y. A Novel Re-Compositing Approach to Create Continuous and Consistent Cross-Sensor/Cross-Production Global NDVI Datasets. *Int. J. Remote Sens.* **2021**, *42*, 6023–6047. [CrossRef]
35. Li, M.; Cao, S.; Zhu, Z. Spatiotemporally Consistent Global Dataset of the GIMMS Normalized Difference Vegetation Index (PKU GIMMS NDVI) from 1982 to 2020. *Earth Syst. Sci. Data* **2023**, *15*, 4181–4203. [CrossRef]
36. Holben, B.N. Characteristics of Maximum-Value Composite Images from Temporal AVHRR Data. *Int. J. Remote Sens.* **1986**, *7*, 1417–1434. [CrossRef]
37. Tan, Z.; Tao, H.; Jiang, J.; Zhang, Q. Influences of Climate Extremes on NDVI (Normalized Difference Vegetation Index) in the Poyang Lake Basin, China. *Wetlands* **2015**, *35*, 1033–1042. [CrossRef]
38. Xiao, X.; Wang, Q.; Guan, Q.; Zhang, Z.; Yan, Y.; Mi, J.; Yang, E. Quantifying the Nonlinear Response of Vegetation Greening to Driving Factors in Longnan of China Based on Machine Learning Algorithm. *Ecol. Indic.* **2023**, *151*, 110277. [CrossRef]
39. Yang, J.; Huang, X. The 30m Annual Land Cover Dataset and Its Dynamics in China from 1990 to 2019. *Earth Syst. Sci. Data* **2021**, *13*, 3907–3925. [CrossRef]
40. Liu, Y.; Huang, H.; Meng, L.; Liu, M.; Wu, Z.; Liu, T.; Labat, D. Spatial-Temporal Evolution of Vegetation Coverage and Its Relationship with Terrain and Human Factors in the Upper Reaches of Ganjiang River Basin, China. *Front. Earth Sci.* **2023**, *10*, 1043403. [CrossRef]
41. Zhang, Y.; Zhang, L.; Wang, J.; Dong, G.; Wei, Y. Quantitative Analysis of NDVI Driving Factors Based on the Geographical Detector Model in the Chengdu-Chongqing Region, China. *Ecol. Indic.* **2023**, *155*, 110978. [CrossRef]
42. Zhang, L.; Chen, B.; Fu, H.; Xu, B.; Ren, Z.; Gao, P. A Prolonged Artificial Nighttime-Light Dataset of China (1984–2020). *Sci. Data* **2024**, *11*, 414. [CrossRef] [PubMed]
43. International Institute for Applied Systems Analysis. *Food and Agriculture Organization of the United Nations (FAO) China Soil Map Based Harmonized World Soil Database (HWSD) (v1.1) (2009)*; National Tibetan Plateau Data Center: Tibet, China, 2021.
44. Qiu, K.; Xie, Y.; Xu, D.; Pott, R. Ecosystem Functions Including Soil Organic Carbon, Total Nitrogen and Available Potassium Are Crucial for Vegetation Recovery. *Sci. Rep.* **2018**, *8*, 7607. [CrossRef]
45. Yan, J.-H.; Zhou, G.-Y.; Zhang, D.-Q.; Chu, G.-W. Changes of Soil Water, Organic Matter, and Exchangeable Cations Along a Forest Successional Gradient in Southern China. *Pedosphere* **2007**, *17*, 397–405. [CrossRef]
46. Martínez, B.; Gilabert, M.A. Vegetation Dynamics from NDVI Time Series Analysis Using the Wavelet Transform. *Remote Sens. Environ.* **2009**, *113*, 1823–1842. [CrossRef]
47. Hou, W.; Gao, J.; Wu, S.; Dai, E. Interannual Variations in Growing-Season NDVI and Its Correlation with Climate Variables in the Southwestern Karst Region of China. *Remote Sens.* **2015**, *7*, 11105–11124. [CrossRef]
48. Wang, Z.; Wang, Y.; Liu, Y.; Wang, F.; Deng, W.; Rao, P. Spatiotemporal Characteristics and Natural Forces of Grassland NDVI Changes in Qilian Mountains from a Sub-Basin Perspective. *Ecol. Indic.* **2023**, *157*, 111186. [CrossRef]

49. Rhif, M.; Ben Abbes, A.; Martinez, B.; Farah, I.R. An Improved Trend Vegetation Analysis for Non-Stationary NDVI Time Series Based on Wavelet Transform. *Env. Sci. Pollut. Res.* **2021**, *28*, 46603–46613. [CrossRef]
50. Lamchin, M.; Wang, S.W.; Lim, C.-H.; Ochir, A.; Pavel, U.; Gebru, B.M.; Choi, Y.; Jeon, S.W.; Lee, W.-K. Understanding Global Spatio-Temporal Trends and the Relationship between Vegetation Greenness and Climate Factors by Land Cover during 1982–2014. *Glob. Ecol. Conserv.* **2020**, *24*, e01299. [CrossRef]
51. Breiman, L. Random Forest. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
52. Liaw, A.; Wiener, M. Classification and Regression by randomForest. *R News* **2002**, *2*, 18–22.
53. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; ACM: San Francisco, CA, USA, 2016; pp. 785–794.
54. Friedman, J.H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Statist.* **2001**, *29*, 1189–1232. [CrossRef]
55. Chen, M.; Liu, Q.; Chen, S.; Liu, Y.; Zhang, C.-H.; Liu, R. XGBoost-Based Algorithm Interpretation and Application on Post-Fault Transient Stability Status Prediction of Power System. *IEEE Access* **2019**, *7*, 13149–13158. [CrossRef]
56. Li, L.; Wang, B.; Feng, P.; Li Liu, D.; He, Q.; Zhang, Y.; Wang, Y.; Li, S.; Lu, X.; Yue, C.; et al. Developing Machine Learning Models with Multi-Source Environmental Data to Predict Wheat Yield in China. *Comput. Electron. Agric.* **2022**, *194*, 106790. [CrossRef]
57. Jia, Z.; Zhang, Z.; Cheng, Y.; Buhebaoyin; Borjigin, S.; Quan, Z. Grassland Biomass Spatiotemporal Patterns and Response to Climate Change in Eastern Inner Mongolia Based on XGBoost Model Estimates. *Ecol. Indic.* **2024**, *158*, 111554. [CrossRef]
58. Li, H.; Li, K.; Zhao, X.; Zhao, J. Changes in Vegetation Greenness and Their Influencing Factors in Southern China. *Remote Sens.* **2022**, *14*, 3291. [CrossRef]
59. Liu, H.; Lin, N.; Zhang, H.; Liu, Y.; Bai, C.; Sun, D.; Feng, J. Driving Force Analysis of Natural Wetland in Northeast Plain Based on SSA-XGBoost Model. *Sensors* **2023**, *23*, 7513. [CrossRef] [PubMed]
60. Zeng, Y.; Jia, L.; Jiang, M.; Zheng, C.; Menenti, M.; Bennour, A.; Lv, Y. Hydrological Factor and Land Use/Land Cover Change Explain the Vegetation Browning in the Dosso Reserve, Niger. *Remote Sens.* **2024**, *16*, 1728. [CrossRef]
61. Li, L.; Zeng, Z.; Zhang, G.; Duan, K.; Liu, B.; Cai, X. Exploring the Individualized Effect of Climatic Drivers on MODIS Net Primary Productivity through an Explainable Machine Learning Framework. *Remote Sens.* **2022**, *14*, 4401. [CrossRef]
62. Roy, B. Optimum Machine Learning Algorithm Selection for Forecasting Vegetation Indices: MODIS NDVI & EVI. *Remote Sens. Appl. Soc. Environ.* **2021**, *23*, 100582. [CrossRef]
63. Li, X.; Yuan, W.; Dong, W. A Machine Learning Method for Predicting Vegetation Indices in China. *Remote Sens.* **2021**, *13*, 1147. [CrossRef]
64. Zhu, Z.; Piao, S.; Myneni, R.B.; Huang, M.; Zeng, Z.; Canadell, J.G.; Ciais, P.; Sitch, S.; Friedlingstein, P.; Arneth, A.; et al. Greening of the Earth and Its Drivers. *Nat. Clim. Chang.* **2016**, *6*, 791–795. [CrossRef]
65. Xue, S.-Y.; Xu, H.-Y.; Mu, C.-C.; Wu, T.-H.; Li, W.-P.; Zhang, W.-X.; Streletskaya, I.; Grebenets, V.; Sokratov, S.; Kizyakov, A.; et al. Changes in Different Land Cover Areas and NDVI Values in Northern Latitudes from 1982 to 2015. *Adv. Clim. Chang. Res.* **2021**, *12*, 456–465. [CrossRef]
66. Farquhar, G.D.; Sharkey, T.D. Stomatal Conductance and Photosynthesis. *Annu. Rev. Plant. Physiol.* **1982**, *33*, 317–345. [CrossRef]
67. Bationo, A.; Kihara, J.; Vanlauwe, B.; Waswa, B.; Kimetu, J. Soil Organic Carbon Dynamics, Functions and Management in West African Agro-Ecosystems. *Agric. Syst.* **2007**, *94*, 13–25. [CrossRef]
68. Eni, D.D.; Iwara, A.I.; Offiong, R.A. Analysis of Soil-Vegetation Interrelationships in a South-Southern Secondary Forest of Nigeria. *Int. J. For. Res.* **2012**, *2012*, 469326. [CrossRef]
69. Bi, X.; Chu, H.; Fu, M.; Xu, D.; Zhao, W.; Zhong, Y.; Wang, M.; Li, K.; Zhang, Y. Distribution Characteristics of Organic Carbon (Nitrogen) Content, Cation Exchange Capacity, and Specific Surface Area in Different Soil Particle Sizes. *Sci. Rep.* **2023**, *13*, 12242. [CrossRef] [PubMed]
70. Dror, I.; Yaron, B.; Berkowitz, B. The Human Impact on All Soil-Forming Factors during the Anthropocene. *ACS Environ. Au* **2022**, *2*, 11–19. [CrossRef]
71. Awoonor, J.K.; Dogbey, B.F.; Salis, I. Human-Induced Land Use Changes and Phosphorus Limitation Affect Soil Microbial Biomass and Ecosystem Stoichiometry. *PLoS ONE* **2023**, *18*, e0290687. [CrossRef]
72. Zhang, L.; Chen, X.; Cai, X.; Salim, H. Spatial-Temporal Changes of NDVI and Their Relations with Precipitation and Temperature in Yangtze River Basin from 1981 to 2001. *Geo-Spat. Inf. Sci.* **2010**, *13*, 186–190. [CrossRef]
73. Silva, S.J.; Keller, C.A.; Hardin, J. Using an Explainable Machine Learning Approach to Characterize Earth System Model Errors: Application of SHAP Analysis to Modeling Lightning Flash Occurrence. *J. Adv. Model. Earth Syst.* **2022**, *14*, e2021MS002881. [CrossRef]
74. Meng, X.; Gao, X.; Li, S.; Lei, J. Spatial and Temporal Characteristics of Vegetation NDVI Changes and the Driving Forces in Mongolia during 1982–2015. *Remote Sens.* **2020**, *12*, 603. [CrossRef]

75. Hartmann, H.; Bastos, A.; Das, A.J.; Esquivel-Muelbert, A.; Hammond, W.M.; Martínez-Vilalta, J.; McDowell, N.G.; Powers, J.S.; Pugh, T.A.M.; Ruthrof, K.X.; et al. Climate Change Risks to Global Forest Health: Emergence of Unexpected Events of Elevated Tree Mortality Worldwide. *Annu. Rev. Plant Biol.* **2022**, *73*, 673–702. [CrossRef]
76. Deng, L.; Shangguan, Z. Afforestation Drives Soil Carbon and Nitrogen Changes in China. *Land Degrad. Dev.* **2017**, *28*, 151–165. [CrossRef]