remote sensing

MDPI

Article

Spatial-Temporal Characteristics and Driving Forces of Aboveground Biomass in Desert Steppes of Inner Mongolia, China in the Past 20 Years

Nitu Wu [1], Guixiang Liu [2], Deji Wuyun [3], Bole Yi [4], Wala Du [2] and Guodong Han [1,*]

[1] Key Laboratory of Grassland Resources of the Ministry of Education, College of Grassland, Resources and Environment, Inner Mongolia Agricultural University, Hohhot 010011, China
[2] Institute of Grassland Research, Chinese Academy of Agricultural Sciences, Hohhot 010010, China
[3] Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China
[4] College of Forestry, Inner Mongolia Agricultural University, Hohhot 010011, China
* Correspondence: hanguodong@imau.edu.cn

Abstract: The desert steppe serves as a transitional zone between grasslands and deserts, and long-term monitoring of aboveground biomass (AGB) in the desert steppe is essential for understanding grassland changes. While AGB observation techniques based on multisource remote-sensing data and machine-learning algorithms have been widely applied, research on monitoring methods specifically for the desert steppe remains limited. In this study, we focused on the desert steppe of Inner Mongolia, China, as the study area and used field sampling data, MODIS data, MODIS-based vegetation indices (VI), and environmental factors (topography, climate, and soil) to compare the performance of four commonly used machine-learning algorithms: multiple linear regression (MLR), partial least-squares regression (PLS), random forest (RF), and support vector machine (SVM) in AGB estimation. Based on the optimal model, the spatial–temporal characteristics of AGB from 2000 to 2020 were calculated, and the driving forces of climate change and human activities on AGB changes were quantitatively analyzed using the random forest algorithm. The results are as follows: (1) RF demonstrated outstanding performance in terms of prediction accuracy and model robustness, making it suitable for AGB estimation in the desert steppe of Inner Mongolia; (2) VI contributed the most to the model, and no significant difference was found between soil-adjusted VIs and traditional VIs. Elevation, slope, precipitation, and temperature all had positive effects on the model; (3) from 2000 to 2020, the multiyear average AGB in the study area was 58.34 g/m^2, exhibiting a gradually increasing distribution pattern from the inner region to the outer region (from north to south); (4) from 2000 to 2020, the proportions of grassland with AGB slightly and significantly increasing trend in the study area were 87.08% and 5.13%, respectively, while the proportions of grassland with AGB slightly and significantly decreasing trend were 7.76% and 0.05%, respectively; and (5) over the past 20 years, climate change, particularly precipitation, has been the primary driving force behind AGB changes of the study area. This research holds reference value for improving desert steppe monitoring capabilities and the rational planning of grassland resources.

Keywords: desert steppe; aboveground biomass; remote sensing; machine learning; random forest

Citation: Wu, N.; Liu, G.; Wuyun, D.; Yi, B.; Du, W.; Han, G. Spatial-Temporal Characteristics and Driving Forces of Aboveground Biomass in Desert Steppes of Inner Mongolia, China in the Past 20 Years. *Remote Sens.* **2023**, *15*, 3097. https://doi.org/10.3390/rs15123097

Academic Editors: Sadegh Jamali, Torbern Tagesson, Feng Tian, Meisam Amani, Per-Ola Olsson and Arsalan Ghorbanian

Received: 7 April 2023
Revised: 2 June 2023
Accepted: 5 June 2023
Published: 13 June 2023

1. Introduction

The Desert steppe in Inner Mongolia is an important component of the temperate grasslands in Central Asia and plays a crucial role in regulating the regional climate, preserving soil and water resources, and providing habitats for numerous species [1–3]. However, climate change and inappropriate land use have put these areas under severe pressure [4]. Aboveground biomass (AGB) is a valuable indicator for assessing terrestrial ecosystem health and understanding its spatial–temporal characteristics and underlying factors is

critical for effectively conserving and managing the desert steppe ecosystem [5,6]. Field observation and remote sensing are the two main methods for grassland AGB estimation. Compared with traditional field-based methods, remote-sensing-based techniques have emerged as a valuable tool for efficiently and accurately tracking AGB changes over large areas [5,7]. In recent decades, a multitude of remote-sensing-based AGB observation technologies has been developed and found to be suitable for vegetation observation problems at various scales [7–9]. However, constructing a fitting model that accurately captures changes in AGB depends heavily on the selection of both feature variables and appropriate modeling algorithms.

Feature selection is a critical step in remote-sensing-based AGB monitoring [10]. Currently, the main variables used to retrieve vegetation AGB are derived from spectral reflectance and vegetation indices (VI) [11,12]. As VIs are calculated based on the reflectance of multiple spectral bands, they exhibit high sensitivity to surface features, including vegetation, making them the preferred variables for biomass observation [13]. In addition to traditional VIs, such as the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI), modified indices, such as the soil-adjusted vegetation index (SAVI), the modified soil-adjusted vegetation index (MSAVI), the transformed soil-adjusted vegetation index (TSAVI), the optimized soil-adjusted vegetation index (OSAVI), and the adjusted transformed soil-adjusted vegetation index (ATSAVI), have been proposed and widely used in vegetation observation [14–20]. However, studies showed that some soil-adjusted VIs may not perform better than traditional indices in sparsely vegetated dry grasslands [21,22]. In addition, relying solely on VI in biomass monitoring models may result in regional differences in form and accuracy. To address this issue, some researchers integrated environmental variables such as geography, topography, and climate into single-factor models and developed multivariate regression models that significantly improve model-fitting accuracy [23–26]. Nevertheless, further research is necessary to determine which variables are suitable for estimating AGB in the desert steppe.

Modeling algorithm selection is a crucial aspect of the biomass estimation process [27]. Currently, there are numerous algorithms available for retrieving AGB, which can be broadly categorized into parametric and nonparametric algorithms based on whether they require assumptions about the model structure. Studies used VIs as the observed variable and constructed single-variable models, such as linear, exponential, and power to fit vegetation biomass [28–30]. However, the accuracy of these simple models is prone to be affected by soil and spectral saturation effects, particularly in heterogeneous regions [31,32]. To improve the generalization ability and accuracy, studies emphasized the establishment of multiple linear regression (MLR) models by increasing the feature variables [33–35]. In recent years, MLR models incorporating dimensionality reduction techniques such as the partial least squares regression (PLS) have also gained popularity [36,37]. However, as vegetation growth is a complex process, parametric models often lack the ability to fit this complex relationship, resulting in a lack of ecological significance in the entire operation process. In contrast, nonparametric models do not require assumptions about the form of the objective function and can thus fit complex relationships, making them the most popular machine-learning algorithms for estimating AGB [5,13]. Some studies showed that these algorithms, such as the random forest (RF), and support vector machine (SVM) perform better than traditional parametric models in estimating grassland AGB [38,39]. However, no indepth comparative study of AGB in Inner Mongolia's desert steppe has been conducted. The desert steppe in Inner Mongolia, China, is characterized by sparse vegetation and complex environmental conditions, leading to uncertainty in the performance of these algorithms in fitting vegetation AGB. Therefore, further research is needed to evaluate the performance of these models in estimating AGB in this region.

Human activities and climate change are widely considered to be the main factors affecting grassland health, but quantifying their dominance is challenging [40–43]. Some studies used linear regression coefficients and variance ratios to assess the impact of external drivers on grasslands [43–45]. However, these methods are prone to overlooking

ecological processes and cannot reflect complex ecological relationships [46]. Other studies attempted to establish quantitative analysis conditions from an ecological perspective by using ecological process models to explain driving factors [47–49]. However, this approach makes it difficult to provide data references for some specific driving forces, resulting in a lack of indepth analysis. Furthermore, some recent studies attempted to use nonparametric models such as RF [50] and boosted regression trees [51] to analyze the impacts of external driving forces on land changes. Similarly, we believe that this approach is also applicable to the topic of our research. The relationship between vegetation and external factors is complex and diverse, making nonparametric algorithms more advantageous in simulating this complex relationship. In addition, selecting appropriate algorithms and building grassland AGB inversion models that are suitable for the characteristics of the study area is key to accurately reflecting the spatial–temporal characteristics and driving forces of desert steppe in Inner Mongolia.

To fill the knowledge gaps mentioned above, this study focused on the desert steppe of Inner Mongolia, China, and carried out the following studies: (1) compared the performance of four commonly used machine-learning algorithms (MLR, PLS, RF, and SVM) on grassland AGB estimation, and generated a suitable modeling method; (2) extracted the spatial–temporal feature of grassland AGB in past 20 years (2000–2020); and (3) analyzed the driving forces of grassland AGB change in past 20 years.

2. Materials and Methods

2.1. Study Area

This research focuses on five regions (counties) in central Inner Mongolia, namely Sunite East Banner (SE), Sunite West Banner (SW), Erenhot City (ER), Siziwang Banner (SZW), and Damao Banner (DM) (Figure 1). The region is predominantly characterized by desert grasslands, covering an approximate area of 56,200 km^2, and represents the main body of the desert steppe in Inner Mongolia. The study area spans from 37°N to 46°N and 105°E to 115°E, with an elevation range of 800–2100 m and an average elevation of 1100 m. The climate in this area is characterized by a continental arid climate with an annual precipitation of 100–200 mm and an average annual temperature of 2–5 °C [2]. The grassland vegetation in this region is typified by a low diversity of species and a simple community structure. The drought-tolerant bunchgrasses are the dominant grass species, often accompanied by numerous drought-tolerant shrubs. The common herbaceous plants include *Stipa klemenzii*, *Stipa breviflora*, *Artemisia frigida*, and *Cleistogenes songorica* [2,52].

2.2. Data Used

2.2.1. Field Measurement Data

During the peak growing season (July and August) from 2018 to 2020, 97 field plots from various terrains and vegetation coverage were collected (Figure 1). The plot size was 500 m × 500 m, with each corner aligned as closely as possible with the corresponding MODIS pixel. Within each plot, 3–5 samples of 1 m × 1 m were selected along the diagonal, and the average value of these sample points was used as the AGB for that plot (Figure 1). The distance between adjacent plots was no less than 1 km, and the recorded information included fresh biomass, latitude, longitude, and elevation.

2.2.2. Remote-Sensing Data

- Multispectral Bands and Vegetation Indices

The MOD09GA product, which provides surface reflectance data, was acquired from the Google Earth Engine (GEE) platform. The time period covers from July 2000 to September 2020. The MOD09GA product records the surface spectral reflectance of MODIS bands 1 through 7 at a spatial resolution of 500 m, corrected for atmospheric conditions such as gases, aerosols, and Rayleigh scattering [53]. Specifically, the bands included in the MOD09GA are red (band 1), near infrared (band 2), blue (band 3), green (band 4), shortwave infrared 1 (band 6), and shortwave infrared 2 (band 7). In this study, in addition to

selecting the above-mentioned six bands, seven MODIS-based VIs were considered. These VIs include two traditional VIs (NDVI, EVI) and five soil-adjusted vegetation indices (SAVI, MSAVI, TSAVI, OSAVI, and ATSAVI). The specific calculation formulas for these indices are given in Table 1.

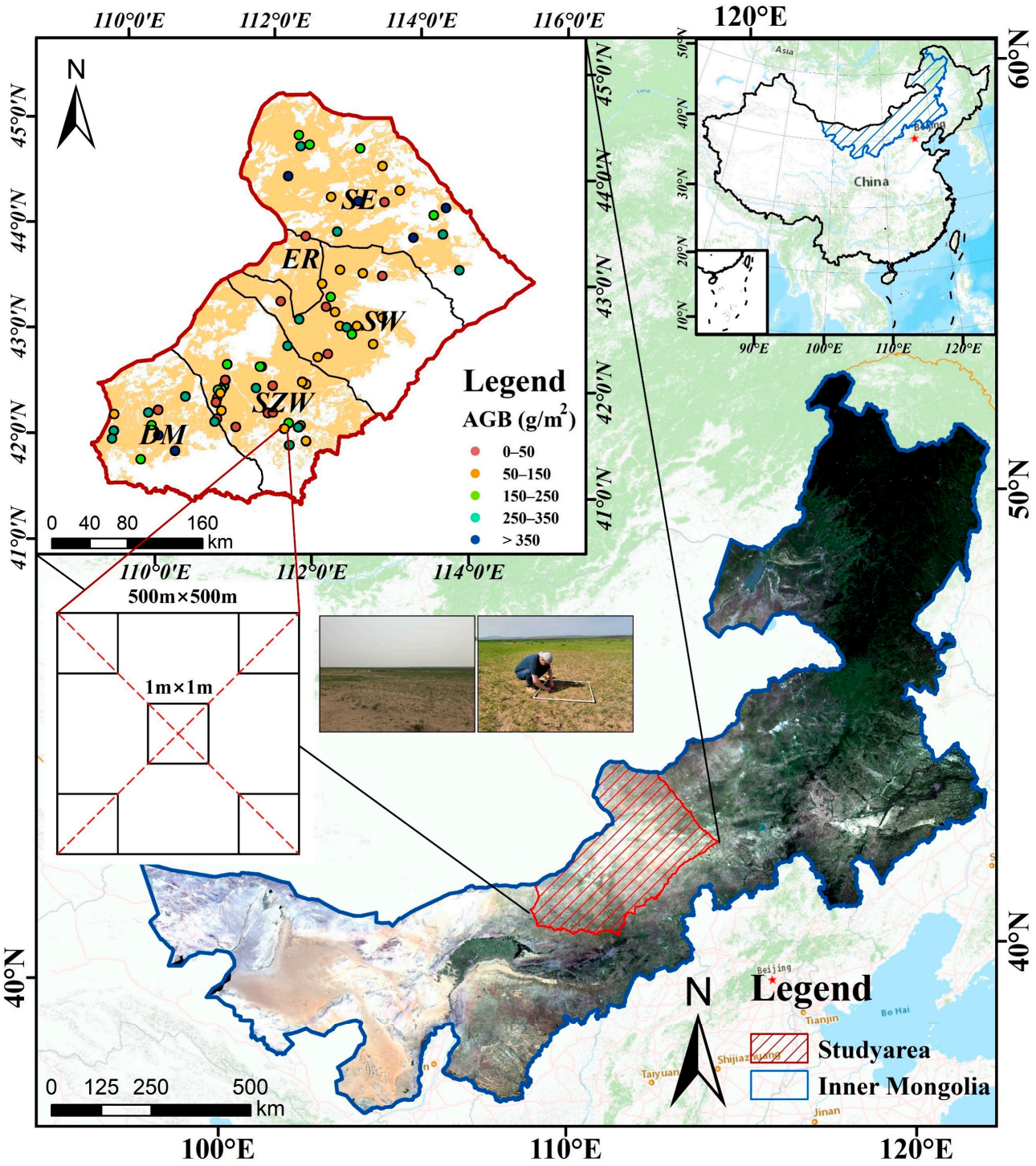

Figure 1. Study area and field-samples distribution.

As some of these indices require correction parameters, 30 barren sample sites were selected in the study area to establish a soil line using the reflectance of soil in the near-infrared and red spectral bands, and obtain the correction parameters *a* and *b*.

Table 1. Selected vegetation indices.

Vegetation Indices Type	Description
Traditional vegetation indices	$NDVI = \frac{\rho_{NIR}-\rho_{RED}}{\rho_{NIR}+\rho_{RED}}$ $EVI = 2.5 \times \frac{\rho_{NIR}-\rho_{RED}}{\rho_{NIR}+6\times\rho_{RED}-7.5\times\rho_{BLUE}+1}$
Soil-adjusted vegetation indices	$SAVI = (1+L) \times \frac{\rho_{NIR}-\rho_{RED}}{\rho_{NIR}+\rho_{RED}+L}$ $MSAVI = 0.5 \times [(2 \times \rho_{NIR}+1) - \sqrt{(2 \times \rho_{NIR}+1)^2 - 8 \times (\rho_{NIR}-\rho_{RED})}]$ $TSAVI = \frac{a\times(\rho_{NIR}-a\times\rho_{RED}-b)}{a\times\rho_{NIR}+\rho_{RED}+a\times b}$ $OSAVI = 1.16 \times \frac{\rho_{NIR}-\rho_{RED}}{\rho_{NIR}+\rho_{RED}+0.16}$ $ATSAVI = \frac{a\times(\rho_{NIR}-a\times\rho_{RED}-b)}{a\times\rho_{NIR}+\rho_{RED}-a\times b+0.08\times(1+a^2)}$

- Terrain data

The terrain variables, elevation, slope, and aspect were from ASTER GDEM version 2, which is a global DEM generated through a collaboration between NASA and Japan's Ministry of Economy, Trade, and Industry (METI). Despite variations in accuracy depending on terrain and other factors, the ASTER GDEM v2 is generally considered a high-quality DEM with a root mean square error (RMSE) of less than 10 m [54]. In this study, we applied the GEE platform to complete preprocessing operations such as projection transformation, resampling, and clipping.

- Meteorological data

Annual total precipitation data were obtained from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), which is a quasi-global rainfall dataset spanning over 30 years. CHIRPS combines 0.05° resolution satellite imagery with in situ station data to generate gridded rainfall time series suitable for trend analysis and seasonal drought monitoring [55].

Annual mean temperature and annual mean wind data were sourced from the ERA5-Land dataset, the latest global meteorological reanalysis dataset released by the European Centre for Medium-Range Weather Forecasts (ECMWF). The dataset provides continuous meteorological records since 1950 at a pixel resolution of up to 9 km and encompasses all conventional meteorological variables [56].

Annual total evapotranspiration data were extracted from the MOD16A2 Version 6 product, which is an 8-day composite dataset produced at a 500 m pixel resolution. The algorithm utilized for the MOD16 data product collection is based on the Penman-Monteith equation and incorporates inputs of daily meteorological reanalysis data, along with MODIS remotely sensed data products such as vegetation property dynamics, albedo, and land cover [57]. In this study, we employed the GEE platform to preprocess the aforementioned factors, including projection conversion, resampling, cropping, and unit conversion. The datasets cover the time period from 2000 to 2020.

- Soil data

The soil factors, including soil bulk density, soil pH, clay content, and sand content at 10 cm depth, were obtained from the OpenLandMap soil dataset, which is an environmental data portal that provides global-scale data on land cover, vegetation, soil, climate, terrain, and other related information to the public [58]. The soil dataset in OpenLandMap is established through the integration of global soil profile data points and machine-learning algorithms and provides multiple soil properties data at six standard depths (0, 10, 30, 60, 100, and 200 cm) with a resolution of 250 m [59]. In this study, we utilized the GEE platform to extract each factor and performed preprocessing operations such as projection transformation, resampling, and cropping.

- Population data

The population data was sourced from the WorldPop project, which is an interdisciplinary research group dedicated to improving the evidence base for spatial demographics and utilizing the data for health and development applications [60,61]. The project takes recent population counts from censuses and breaks them down into grid cells of approximately 100 m × 100 m using machine-learning techniques. These techniques exploit the relationships between population densities and different geospatial covariate layers. In this study, we employed the GEE platform to carry out preprocessing tasks. The datasets used encompass the period from 2000 to 2020.

2.2.3. Statistical Data

The statistical data involved in this study include annual livestock quantities, agricultural gross domestic product (GDP), industrial GDP, and consumer GDP for each county from 2000 to 2020. These data were obtained from local statistical yearbooks. To capture the multiyear trends of these statistical variables, we employed the least squares method and assigned the resulting trends to the corresponding administrative boundary polygons.

2.3. Analysis Methods

The technical route of this study can be divided into four parts: data preprocessing, optimal model construction, extraction of AGB spatial–temporal features, and analysis of driving forces (Figure 2). As mentioned in the previous section, the data used in this study mainly consist of field-sample data, remote-sensing data, and statistical data. First, extract the raster values corresponding to the coordinates of the sample plots for the year 2020 to build a sample dataset. Then, split the dataset into training and testing datasets using a 70:30 ratio. Next, the training dataset is used in combination with leave-one-out cross-validation (LOOCV) to train four different models: MLR, PLSR, RF, and SVM. The performance of these models is compared on the testing dataset, and the best model and variable combination are selected. Then, using the selected model, a spatial–temporal dataset of AGB in the desert steppe of the study area is generated for the period from 2000 to 2020. Following this, change detection and trend analysis methods are employed to characterize the changes in desert steppe AGB over the past 20 years. Lastly, the RF algorithm is utilized to quantitatively analyze the potential driving forces behind the AGB changes.

2.3.1. Modeling Algorithms

MLR is the most used parametric statistical algorithm for relating linear relationships between various predictors and independent variables [62]. Stepwise regression is a statistical approach used to build an MLR model by selecting the most significant predictors from a larger set. The algorithm adds or removes variables from the model in a stepwise manner based on predetermined criteria until the optimal subset of variables is achieved. The final model consists of the most significant predictors, and the coefficients represent the strength and direction of the relationship between predictors and the outcome variable [63].

PLS is a statistical method that is similar to principal components regression in that it finds a linear regression model by projecting the predicted variables and the observable variables to a new space rather than finding hyperplanes of maximum variance between the response and independent variables [64]. This algorithm is especially useful when the predictor matrix contains more variables than observations and there is multicollinearity among the variables.

RF is a powerful machine-learning algorithm used for regression analysis that falls under the category of ensemble learning algorithms. The algorithm works by creating multiple decision trees on randomly sampled subsets of the training data and averaging their predictions to obtain the final output [65]. Generally, RF involves two important hyperparameters: the number of decision trees (n_estimators) and the maximum number of features considered when splitting a node in each decision tree (max_features) [65]. The

n_estimators refer to the number of decision trees used to construct the random forest, and the max_features refer to the maximum number of features considered when splitting a node in each decision tree.

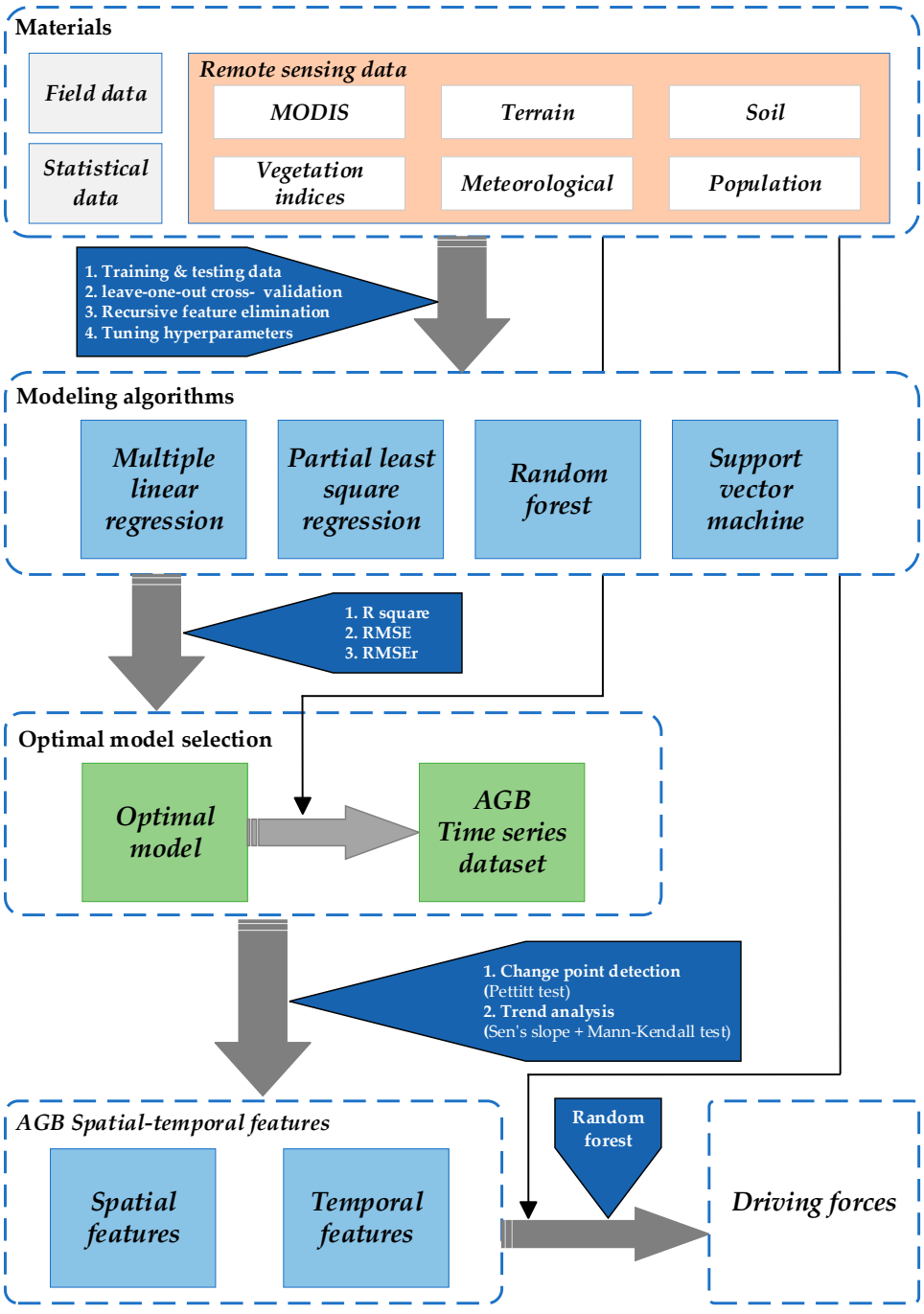

Figure 2. Technical route.

SVM is a popular machine-learning technique for predicting continuous output values. The algorithm aims to find the best hyperplane that minimizes the error between the predicted and true output values, with a penalty for the weight magnitude [66]. It can handle nonlinear data using kernel functions such as the radial basis function (RBF), which maps data into higher dimensional feature spaces. The SVM-RBF algorithm is capable of modeling nonlinear relationships between input and output variables and is robust to noise and outliers. Two hyperparameters need to be adjusted: penalty factor (c) and kernel parameter (gamma). The C parameter controls the trade-off between achieving a low

error on the training data and minimizing the weights' mean, while the gamma parameter controls the distance of influence of a single training point.

Feature selection has always been an important step in model training. In this study, the recursive feature elimination (RFE) method was employed to iteratively select feature variables for both RF and SVM. RFE is a feature selection technique that aims to iteratively eliminate features with lower predictive contributions to the model. In this method, the model is repeatedly trained, and less informative features are removed until a subset of the most predictive features is obtained [67].

2.3.2. Model Assessment

In this study, three validation standards, namely R^2, RMSE, and RMSEr, were selected to evaluate model accuracy, with the following specific formulas:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (1)$$

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

$$RMSEr = \frac{RMSE}{\bar{y}} \times 100 \quad (3)$$

Here, $\hat{y}_i$ represents the measured AGB value of the i-th sample, y_i represents the predicted AGB value of the i-th sample, $\bar{y}$ represents the average AGB value of the samples, and n represents the total number of samples.

2.3.3. Change Point Detection

The Pettitt test was used in this study to extract the abrupt change point of grassland AGB time-series data. The Pettitt test is a nonparametric test based on Mann–Whitney two independent samples and its hypothesis is that H_0 states that if two samples have no difference, then their central positions will be the same, while H_1 states that if there is a difference, then their central positions will not be the same [68].

The nonparametric statistic is defined as follows:

$$\widehat{U} = \max\left|U_k\right| \quad (4)$$

$$U_k = 2\sum_{i=1}^{k} r_i - k(n+1), k = 1, 2, \ldots, n \quad (5)$$

Here, $\widehat{U}$ is the test statistic, and its value is equal to the maximum absolute value of the vector U_k; r_i is the order of the series data, and k is the possible change year.

The formula for the p-value of Pettitt's test:

$$p = 2\exp\left(\frac{-6\widehat{U}^2}{n^3 + n^2}\right) \quad (6)$$

Here, if p is less than a given significance level β, reject the null hypothesis, namely series data has a change point; On the contrary, if p is greater than a given significance level β, accept the null hypothesis, namely the series data without change point.

2.3.4. Trend Analysis

In this study, the grassland AGB changing trend was analyzed using Sen's slope method. The Mann–Kendall test was employed to test the significance of the sequence

trend. The Mann–Kendall test was preferred as it does not make any assumptions about the distribution of the sequence and is robust to outliers [69].

Sen's slope formula is:

$$\beta = Median(\frac{AGB_j - AGB_i}{j - i}) \quad (7)$$

Here, Median(x) represents the median function, and β is the Sen's slope value. A positive value of β indicates an increasing trend of AGB, while a negative value of β indicates a decreasing trend.

The Mann–Kendall test involves testing two hypotheses: H_0 assumes that the sequence data are randomly arranged and there is no significant change trend, while H_1 assumes that the sequence data exhibit an upward or downward trend. The formula for the test statistic S is:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn(AGB_j - AGB_i) \quad (8)$$

$$sgn(AGB_j - AGB_i) = \begin{cases} 1 & AGB_j - AGB_i > 0 \\ 0 & AGB_j - AGB_i = 0 \\ -1 & AGB_j - AGB_i < 0 \end{cases} \quad (9)$$

Here, sgn(x) is the sign function that returns 1, −1, or 0, depending on whether x is positive, negative, or zero, respectively.

The values of Sen's slope (β) and the Mann–Kendall test statistic (p) were used to determine the levels of the AGB change trend, which included significantly increasing, slightly increasing, significantly decreasing, and slightly decreasing (Table 2).

Table 2. Levels of AGB change the trend of grassland.

AGB Change Trend and Test Results	Levels of AGB Change Trend
$\beta > 0, p < 0.05$	significantly increasing
$\beta > 0, p > 0.05$	slightly increasing
$\beta < 0, p < 0.05$	significantly decreasing
$\beta < 0, p > 0.05$	slightly decreasing

2.3.5. Driving Factors Analysis

The RF algorithm was employed in this study to analyze the driving factors behind changes in grassland AGB quantitatively. The following steps were taken in the analysis: (1) nine potential driving forces, from climate and human activity categories, were selected as independent variables to investigate their impact on grassland AGB changes in the study area (Table 3); (2) 300 sample points were randomly selected, and the corresponding independent variable and the AGB change trend (dependent variable) values were extracted for each point; (3) the sample points were divided into training and testing datasets in a 70:30 ratio, and an RF regression model was optimized using the training dataset and a 10-fold cross-validation method; (4) the feature importance was extracted to quantitatively evaluate the driving factors.

RF provides two metrics for measuring the importance of features: mean decrease impurity (MDI) and mean decrease accuracy (MDA). The MDI metric measures the number of times a feature is selected in the decision tree and its contribution to the impurity reduction, which can evaluate the contribution of a feature to classification. On the other hand, the MDA metric calculates the change in model accuracy on a test set by randomly shuffling the values of a feature, where a greater decrease in accuracy indicates a higher feature importance. It should be noted that although both metrics can be used to measure feature importance, they are not always consistent.

Table 3. Selected potential driving forces of AGB changes.

Factor Type	Factor Name	Description	No.	Unit
Climate change	Precipitation trends	Annual precipitation trends from 2000 to 2020	D1	mm
	Temperature trends	Annual temperature trends from 2000 to 2020	D2	°C
	Evapotranspiration trends	Annual evapotranspiration trends from 2000 to 2020	D3	mm
	Wind trends	Annual wind trends from 2000 to 2020	D4	m
Human activity	Population density trends	Annual population density trends from 2000 to 2020	D5	%
	Livestock growth rate	Annual average livestock growth rate	D6	%
	Primary industry GDP growth rate	Averaged annual growth rate of primary industry GDP from 2000 to 2020	D7	%
	Secondary industry GDP growth rate	Averaged annual growth rate of secondary industry GDP from 2000 to 2020	D8	%
	Tertiary industry GDP growth rate	Averaged annual growth rate of tertiary industry GDP from 2000 to 2020	D9	%

3. Results

3.1. The Performance of the Selected Algorithms

Figure 3 presents the relationship between algorithm performances and the number of variables (components). The cross-validation accuracy of these algorithms on the training dataset was unstable, indicating that an excessive number of variables does not improve the performance of the model. MLR retained eight variables for optimal performance after stepwise feature testing, indicating the presence of multicollinearity among the selected features. The optimal performance of PLS was achieved by retaining only one principal component. After evaluating each feature using the RFE method, RF and SVM chose 17 and 22 effective variables, respectively, to achieve the best performance.

Table 4 lists the detailed features selected by the modeling algorithms. MLR selected the feature variables MIR, SWIR, MSAVI, ATSAVI, soil bulk density, clay content, precipitation, and temperature to construct the AGB regression model, expressed as AGB = 35.14 × MIR − 33.69 × SWIR − 200.18 × MSAVI + 229.79 × ATSAVI + 18.44 × soil bulk density − 21.27 × clay content + 22.94 × precipitation − 12.82 × temperature + 73.41. PLS retained all variables since it integrates dimensionality reduction techniques. RF identified the effective variables as blue, green, red, NIR, MIR, SWIR, NDVI, EVI, SAVI, TSAVI, OSAVI, MSAVI, ATSAVI, elevation, slope, precipitation, and temperature, while SVM selected all variables.

Figure 4 displays the model performance of RF and SVM under different hyperparameter combinations at the optimal variable conditions. For the RF, the model achieved high accuracy when the n_estimators was set between 500 and 700, and the best model performance was obtained when the n_estimators and max_features were set to 700 and 4, respectively. As for the SVM, the model achieved high prediction accuracy when gamma was set between 0.02 and 0.05, and the best model performance was achieved when gamma and cost were set to 0.03 and 8, respectively.

Figure 5 compares the prediction accuracy of selected algorithms on the training and testing datasets under effective feature and hyperparameter conditions. The results indicate that nonparametric models outperformed linear models significantly. On the testing dataset, RF achieved the highest prediction accuracy (R^2 = 0.66, RMSE = 39.85, RMSEr = 39.12), followed by SVM (R^2 = 0.60, RMSE = 52.16, RMSEr = 51.21), PLS (R^2 = 0.59, RMSE = 44.47, RMSEr = 43.67), and MLR (R^2 = 0.52, RMSE = 48.85, RMSEr = 47.96). These findings suggest that RF has the best predictive ability for unseen data, which is consistent with the inversion requirements of AGB in Inner Mongolia's desert steppe.

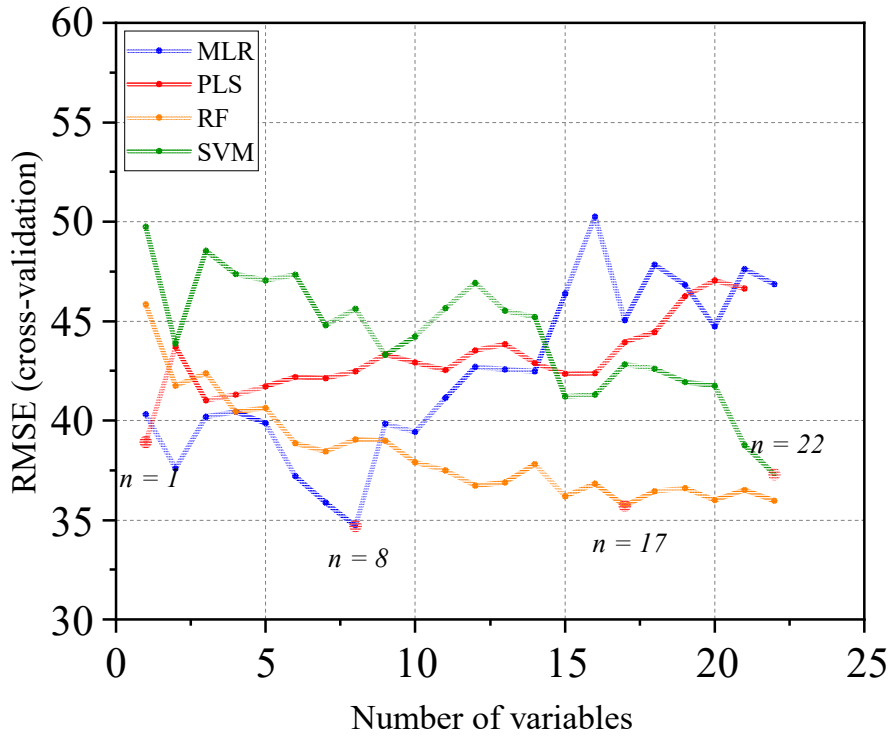

Figure 3. Relationship between the algorithm performances and the number of variables.

Table 4. Selected effective variables for models.

Algorithms	Selected Variables	No. of Variables
MLR	MIR, SWIR, MSAVI, ATSAVI, soil bulk density, clay content, precipitation, temperature	8
PLS	All	22
RF	Blue, green, red, NIR, MIR, SWIR, NDVI, EVI, SAVI, TSAVI, OSAVI, MSAVI, ATSAVI, elevation, slope, precipitation, temperature	17
SVM	All	22

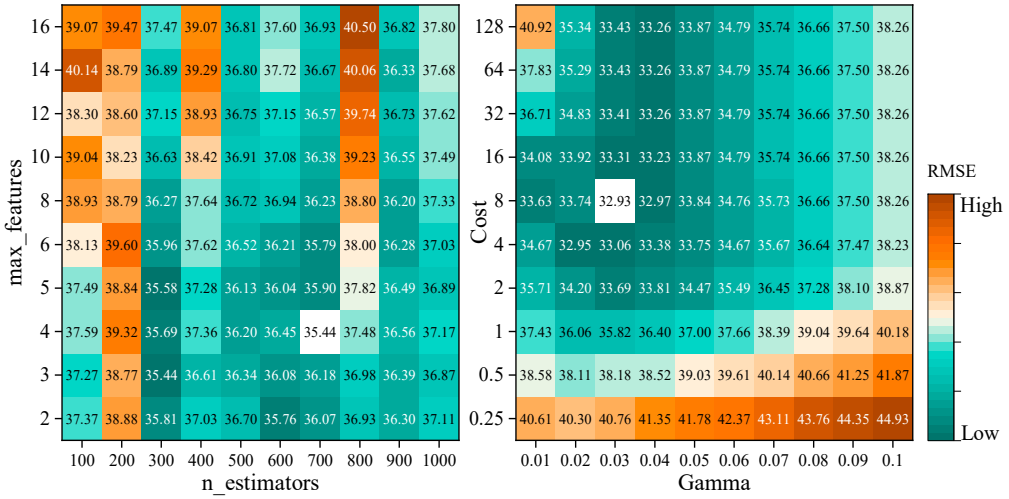

Figure 4. Hyperparameter tuning with grid search for RF and SVM models.

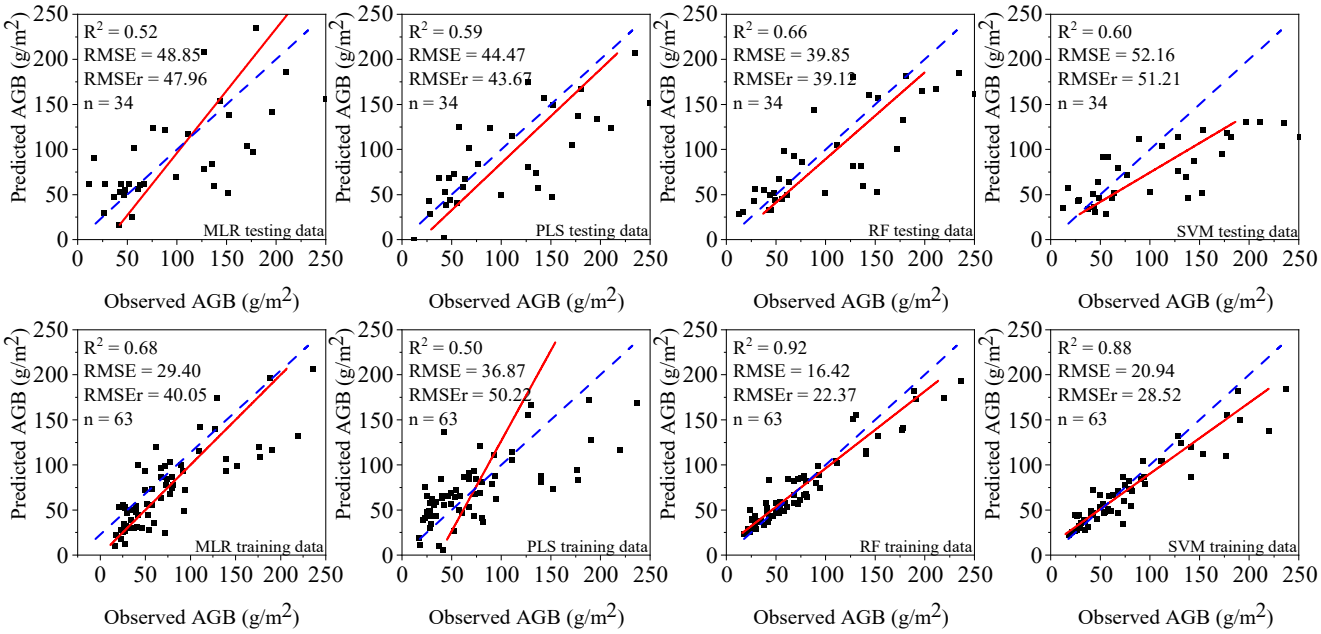

Figure 5. Prediction accuracy of optimal models on training and testing datasets.

3.2. Spatial–Temporal Feature of Desert Steppe AGB in Inner Mongolia during the Past 20 Years

As seen in Figure 6, from 2000 to 2020, the vegetation cover of the desert steppe in the study area exhibited a gradually increasing distribution pattern from north to south, with an average AGB of 58.34 g/m^2. Among them, the SZW had the highest average AGB of 68.91 g/m^2, followed by SZ (60.90 g/m^2), DM (56.52 g/m^2), SY (49.82 g/m^2), and ER (35.32 g/m^2). Within the entire study area, the AGB was highest in the southern parts of SZW and DM, exceeding 120 g/m^2, while the vegetation in most parts of ER, the northern part of SZW, and the western part of SE was sparse, with an AGB ranging from 20–50 g/m^2.

Figure 7a,b demonstrates that while AGB changed in most of the study area, only a small number of regions experienced significant changes ($p < 0.05$). This suggests that over the past 20 years, the AGB changes in the study area have been comparatively stable. Figure 7c shows that the areas with significant changes over the past 20 years were mainly distributed in the southern SZW and central DM, and the years with significant changes mainly occurred in 2010, 2011, and 2012.

Figure 8 displays the spatiotemporal variation of AGB in the study area from 2000 to 2020. Overall, AGB in desert steppe increased with an average annual growth rate of 1.33 g/m^2 during the past two decades (Figure 8a,d). According to the significance testing results (Figure 8b), grassland with a significant increase in AGB accounted for 5.13%, primarily found in the central and southern SZW and central DM, while grassland with a slight increase in AGB accounted for 87.08%, found throughout the study area. The proportion of grassland with a significant decrease in AGB was less than 0.5% and the proportion of grassland with a slight decrease was only 7.76%, with the majority concentrated in the eastern SW, central SE, and northeastern ER (Figure 8).

At the county level, SZW had the most significant restoration, with an annual increase in grassland AGB of 1.84 g/m^2, and the proportion of grassland with a significant increase in AGB was 9.95%, while the proportion of grassland with a slight increase in AGB was 87.17%. DM followed, with an annual increase in grassland AGB of 1.44 g/m^2, with 94.21% of grassland showing an increase in AGB. In SE, the annual increase in grassland AGB was 1.30 g/m^2 and the proportion of grassland with an increase in AGB was 92.38%. In ER, the annual increase in grassland AGB was 0.89 g/m^2, with 90.09% of grassland showing an increase in AGB. In SW, the annual increase in grassland AGB was 0.84 g/m^2 and the proportion of grassland with an increase in AGB was 85.57% (Figure 8c,d).

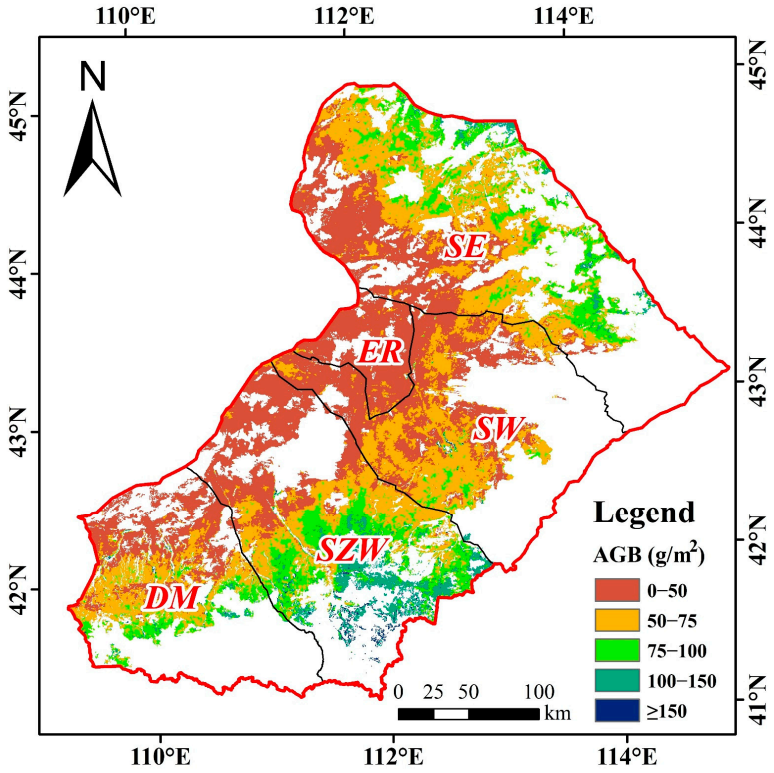

Figure 6. Distribution of AGB in the study area from 2000 to 2020.

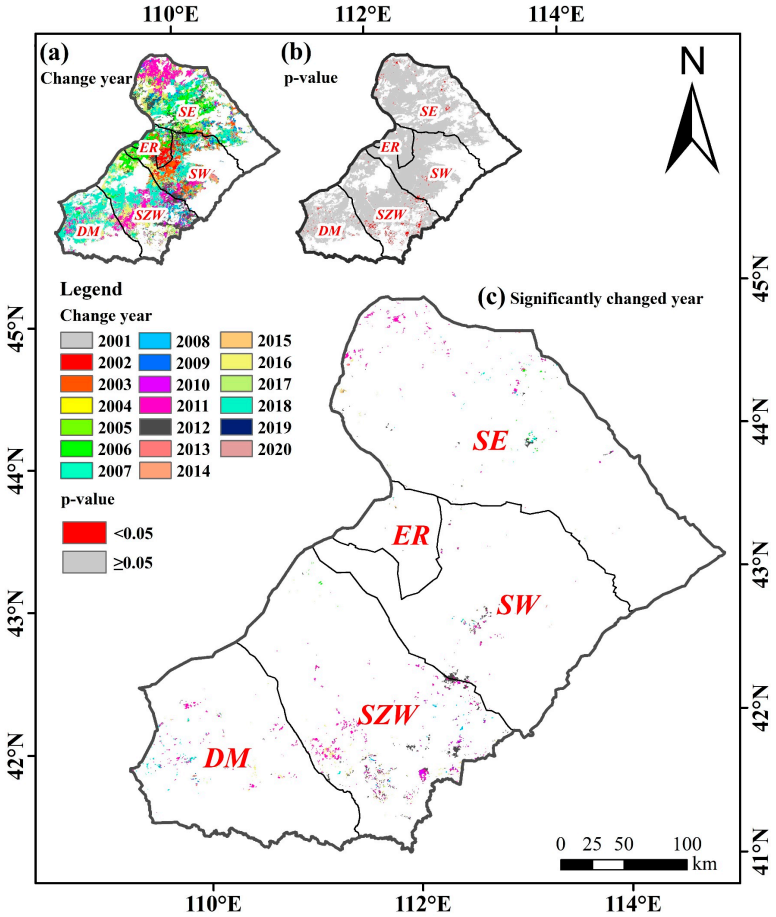

Figure 7. The years with significant changes in AGB in the study area from 2000 to 2020.

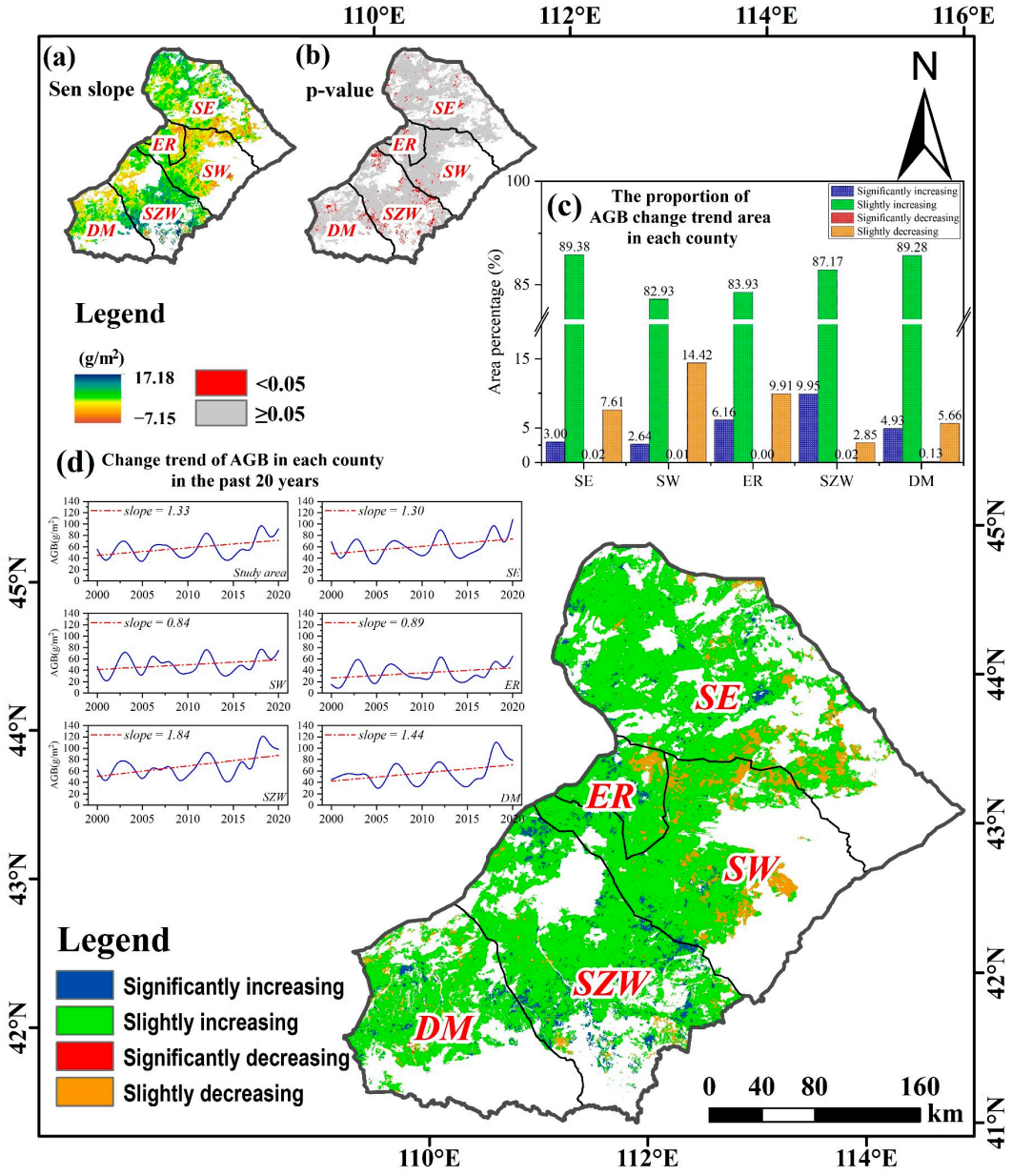

Figure 8. Spatial–temporal variation of desert steppe AGB in the study area from 2000 to 2020.

3.3. Driving Factors for Desert Steppe AGB in Inner Mongolia during the Past 20 Years

Figure 9 displays the drivers behind the desert steppe AGB changes in the study area from 2000 to 2020. Both indicators provide similar rankings. In the MDA ranking, precipitation trends were the most important, followed by evapotranspiration trends > wind trends > temperature trends > livestock growth rate > secondary industry GDP growth rate > population growth rate > tertiary industry GDP growth rate > primary industry GDP growth rate. Similarly, in the MDI, precipitation trends were the most important, followed by temperature trends > wind trends > population growth rate > livestock growth rate > secondary industry GDP growth rate > tertiary industry GDP growth rate > primary industry GDP growth rate. Combining the two indicators, it can be concluded that climate change is the primary factor influencing the changes in AGB in the study area's desert steppe and human activities have a much smaller impact compared to climate change.

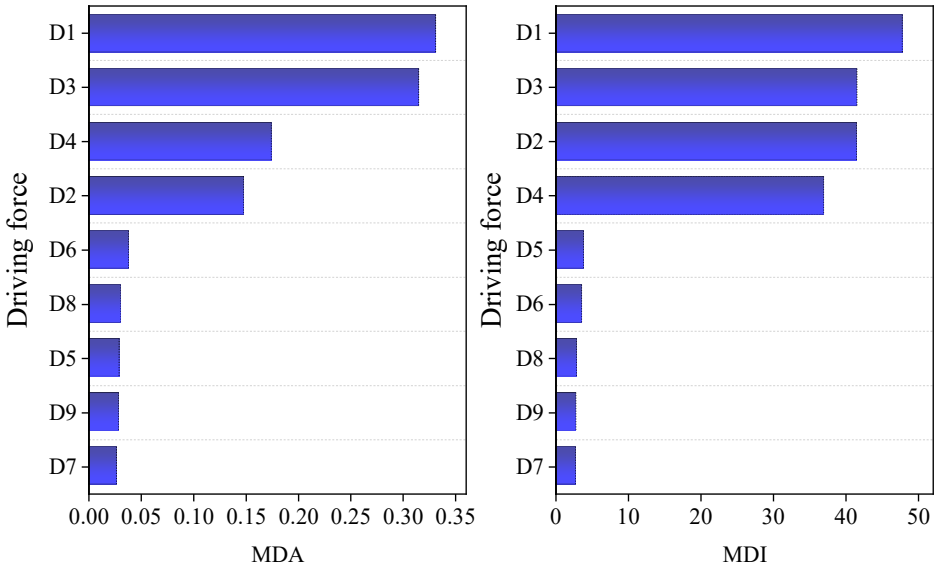

Figure 9. Drivers for the desert steppe AGB changes in the study area from 2000 to 2020.

4. Discussion

4.1. Modeling Algorithm and Feature Selection for Estimating AGB in the Desert Steppe of Inner Mongolia

The desert steppe of Inner Mongolian is characterized by sparse vegetation and simple structure, which poses challenges for reflecting the spatial and temporal characteristics of AGB in the region. In this study, the performance of four popular machine-learning algorithms, MLR, PLS, RF, and SVM, were compared for AGB estimation. From the perspective of prediction accuracy, RF performed the best. Moreover, RF also outperformed other algorithms in terms of noise resistance. This fully demonstrates that RF is suitable for AGB inversion tasks in the Inner Mongolian desert steppe. Similarly, some other studies have observed the same results [24,70]. RF predicts by combining multiple decision-tree models, and due to the ensemble effect of the model, RF has high accuracy and good handling capabilities for high-dimensional and missing data, as well as advantages in interpretability and noise resistance [65]. This makes it one of the commonly used algorithms in machine learning. However, these nonparametric models also have disadvantages, such as low interpretability and reduced model portability due to the influence of spatial heterogeneity [27]. This uncertainty raises the issue of whether the trained RF model is applicable to other desert steppe areas outside the study area, which requires further research for generalization. In comparison, parametric models such as MLR and PLS have advantages in model interpretability, flexibility, and ease of operation. However, based on the results of this study, these models are easily affected by multicollinearity, leading to the exclusion of some key variables from the modeling process. In conclusion, RF is the optimal model for obtaining AGB in the Inner Mongolian desert steppe.

Among the tested variables, VIs contributed the most to AGB inversion in the desert steppe, particularly in nonparametric models, where these indices were all selected as effective variables. However, in the MLR model, only MSAVI and ATSAVI were chosen, indicating the presence of multicollinearity among these VIs. A Spearman correlation test revealed significant correlations among the chosen vegetation indices (Figure A1). This suggests that soil-adjusted VIs do not significantly enhance the spatial heterogeneity representation of AGB in the desert steppe at MODIS spatial resolution scale. Similarly, Baghi and Oldeland found comparable results when comparing soil-adjusted VIs and standard VIs for AGB inversion in Northern Iran's semiarid grasslands using MODIS data [22]. Ren et al. emphasized that the performance of soil-adjusted VIs based on MODIS data was inferior to traditional VIs [21]. This could be attributed to the image resolution, as most of these indices were developed using Landsat imagery, which may affect their

performance on coarser spatial-resolution images. On the other hand, soil–vegetation indices might lose their ability to describe the green vegetation information in areas with low green-vegetation cover [21]. In terms of environmental variables, elevation, slope, precipitation, and temperature all contributed positively to AGB inversion in the desert steppe. This implies that the selected samples exhibit spatial differences in these features. The selection of soil factors should be further determined based on the modeling algorithm.

4.2. Spatial–Temporal Feature and Driving Forces of AGB Changes in the Desert Steppe of Inner Mongolia

Over the past 20 years, the multiyear average AGB of the desert steppe in the study area was 58.34 g/m^2 and it exhibited a gradually decreasing trend from the outer region towards the inner region. This pattern is consistent with the distribution characteristics of annual precipitation in this area. Similarly, some previous studies have also observed the same distribution patterns [24,71]. This indicates that precipitation is the primary factor determining the vegetation distribution characteristics in the region. In terms of the change trend, the AGB exhibited an increasing tendency, with an average annual growth of 1.33 g/m^2, and most regions have not experienced abrupt shifts. That is, the desert steppe in Inner Mongolia has undergone a stable recovery during these 20 years. Human activities and climate change have long been considered the primary factors driving changes in grasslands [44]. However, quantitatively analyzing their impacts on grasslands presents challenges. This study considered that the interaction mechanisms between external factors and grassland vegetation are complex and diverse. Therefore, employing nonparametric algorithms to address these intricate relationships is more consistent with the actual situation. Based on the feature importance ranking from the RF model, climate change emerges as the dominant factor influencing AGB changes in the grasslands of the study area, particularly the contributions of precipitation and evapotranspiration. During the past 20 years, most of the regions in the study area have experienced an upward trend in precipitation, with annual precipitation surpassing annual evapotranspiration. This implies that the water required for vegetation growth has been adequately replenished, promoting vegetation productivity. The desert steppe is primarily dominated by xerophytic, perennial bunchgrasses, but during wet years, annual hydrophilic vegetation can easily form a dominant layer, positively affecting community biomass [46]. Moreover, results show that temperature and wind speed are also major factors causing AGB changes. Nevertheless, the variations in these two factors over the past 20 years have been relatively minor, implying that their associations with AGB are more prominently manifested in spatial distribution patterns. Compared with climatic factors, human activities exert a less pronounced impact on AGB within the desert steppe of the study area. However, it is worth noting that quantifying social statistical indicators such as GDP and livestock numbers for spatial use, particularly on smaller scales, may be challenging and could lead to uncertainties in the results. Therefore, the study of spatial transformation methods for socioeconomic indicators is a key issue for future quantitative analysis of driving forces.

5. Conclusions

This study employed field-survey data and multisource remote-sensing data to compare the performance of four commonly used machine-learning algorithms, including linear and nonparametric methods (MLR, PLS, RF, and SVM), in estimating desert steppe AGB in Inner Mongolia, China. The optimal model was used to analyze the spatiotemporal characteristics of AGB from 2000 to 2020, and the driving forces, including climate change and human activities, on AGB change were quantitatively assessed based on the RF algorithm. The main conclusions are as follows: (1) RF outperformed other algorithms regarding prediction accuracy and model robustness, rendering it suitable for AGB estimation in the desert steppe of Inner Mongolia; (2) among the tested variables, VI contributed the most to the model and no significant difference was found between soil-adjusted VIs and traditional VIs. Elevation, slope, precipitation, and the temperature had positive effects on

model construction; (3) from 2000 to 2020, the multiyear average AGB in the study area was 58.34 g/m^2 and it exhibited a gradually decreasing trend from the outer region towards the inner region; (4) from 2000 to 2020, grassland with a slight increase in AGB accounted for 87.08%, significantly increased accounted for 5.13%, significantly decreased accounted for only 0.05%, and slightly decreased accounted for 7.76% in the study area; (5) over the past two decades, climate change, particularly precipitation, is the dominant driving force behind AGB change in the desert steppe of the study area.

Author Contributions: Conceptualization, N.W. and G.H.; methodology, N.W.; software, N.W.; validation, G.L., D.W. and W.D.; formal analysis, G.L.; writing—original draft preparation, N.W.; writing—review and editing, B.Y. and D.W.; project administration, G.H.; funding acquisition, G.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Innovative Research Team of the Ministry of Education of China (No. IRT_17R59) and the Inner Mongolia Key Project (No. ZDZX2018020).

Data Availability Statement: Not applicable.

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

Appendix A

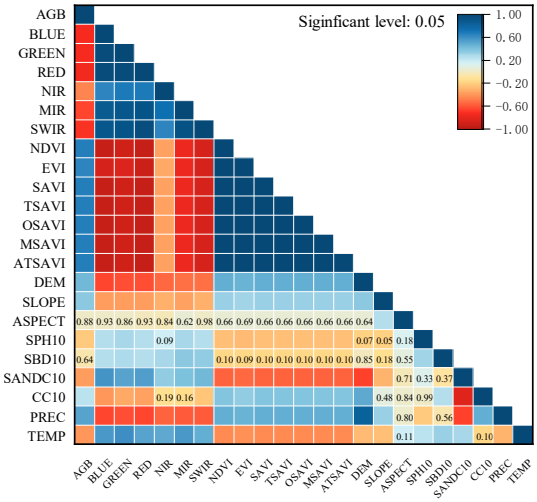

Figure A1. Spearman correlation between variables.

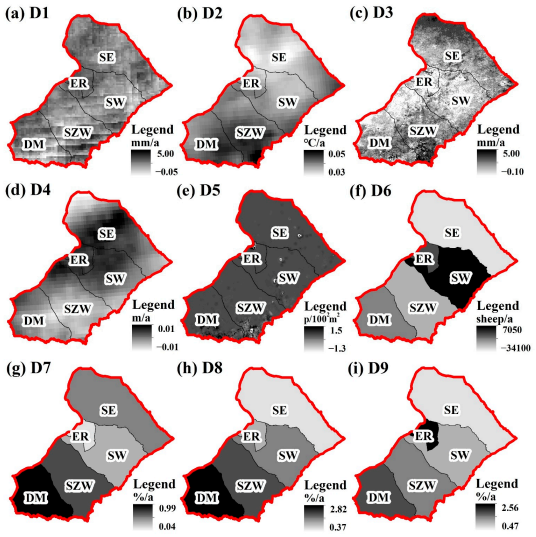

Figure A2. Spatial distribution of driving forces.

References

1. Zhang, R.; Wang, Z.; Han, G.; Schellenberg, M.P.; Wu, Q.; Gu, C. Grazing induced changes in plant diversity is a critical factor controlling grassland productivity in the Desert Steppe, Northern China. *Agric. Ecosyst. Environ.* **2018**, *265*, 73–83. [CrossRef]
2. Yu, D.; Li, Y.; Yin, B.; Wu, N.; Ye, R.; Liu, G. Spatiotemporal variation of net primary productivity and its response to drought in Inner Mongolian desert steppe. *Glob. Ecol. Conserv.* **2022**, *33*, e01991. [CrossRef]
3. Li, L.; Chen, J.; Han, X.-G.; Zhang, W.; Shao, C. Desert Steppe Ecosystem. In *Grassland Ecosystems of China*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 249–283.
4. Akiyama, T.; Kawamura, K. Grassland degradation in China: Methods of monitoring, management and restoration. *Grassl. Sci.* **2007**, *53*, 1–17. [CrossRef]
5. Ali, I.; Greifeneder, F.; Stamenkovic, J.; Neumann, M.; Notarnicola, C. Review of Machine Learning Approaches for Biomass and Soil Moisture Retrievals from Remote Sensing Data. *Remote Sens.* **2015**, *7*, 16398–16421. [CrossRef]
6. Xu, K.; Su, Y.; Liu, J.; Hu, T.; Jin, S.; Ma, Q.; Zhai, Q.; Wang, R.; Zhang, J.; Li, Y.; et al. Estimation of degraded grassland aboveground biomass using machine learning methods from terrestrial laser scanning data. *Ecol. Indic.* **2020**, *108*, 105747. [CrossRef]
7. Wang, Z.; Ma, Y.; Zhang, Y.; Shang, J. Review of Remote Sensing Applications in Grassland Monitoring. *Remote Sens.* **2022**, *14*, 2903. [CrossRef]
8. Kuang, Q.; Yuan, Q.-Z.; Han, J.-C.; Leng, R.; Wang, Y.-S.; Zhu, K.-H.; Lin, S.; Ren, P. A remote sensing monitoring method for alpine grasslands desertification in the eastern Qinghai-Tibetan Plateau. *J. Mt. Sci.* **2020**, *17*, 1423–1437. [CrossRef]
9. Ehlers, D.; Wang, C.; Coulston, J.; Zhang, Y.; Pavelsky, T.; Frankenberg, E.; Woodcock, C.; Song, C. Mapping Forest Aboveground Biomass Using Multisource Remotely Sensed Data. *Remote Sens.* **2022**, *14*, 1115. [CrossRef]
10. Luo, M.; Wang, Y.; Xie, Y.; Zhou, L.; Qiao, J.; Qiu, S.; Sun, Y. Combination of Feature Selection and CatBoost for Prediction: The First Application to the Estimation of Aboveground Biomass. *Forests* **2021**, *12*, 216. [CrossRef]
11. Villacrés, J.; Fuentes, A.; Reszka, P.; Cheein, F.A. Retrieval of Vegetation Indices Related to Leaf Water Content from a Single Index: A Case Study of *Eucalyptus globulus* (Labill.) and *Pinus radiata* (D. Don.). *Plants* **2021**, *10*, 697. [CrossRef]
12. Maimaitijiang, M.; Sagan, V.; Sidike, P.; Maimaitiyiming, M.; Hartling, S.; Peterson, K.T.; Maw, M.J.W.; Shakoor, N.; Mockler, T.; Fritschi, F.B. Vegetation Index Weighted Canopy Volume Model (CVMVI) for soybean biomass estimation from Unmanned Aerial System-based RGB imagery. *ISPRS J. Photogramm. Remote Sens.* **2019**, *151*, 27–41. [CrossRef]
13. Lourenço, P. Biomass Estimation Using Satellite-Based Data. In *Forest Biomass*; Gonçalves, A.C., Sousa, A., Malico, I., Eds.; IntechOpen: London, UK, 2021.
14. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]
15. Huete, A.R.; Liu, H.Q.; Batchily, K.; Leeuwen, W.V. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sens. Environ.* **1997**, *59*, 440–451. [CrossRef]
16. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]
17. Qi, J.G.; Chehbouni, A.R.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A Modified Soil Adjusted Vegetation Index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [CrossRef]
18. Baret, F.; Guyot, G.; Major, D.J. TSAVI: A vegetation index which minimizes soil brightness effects on LAI and APAR estimation. In Proceedings of the 12th Canadian Symposium on Remote Sensing Geoscience and Remote Sensing Symposium, Vancouver, BC, Canada, 10–14 July 1989.
19. Baret, S.F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [CrossRef]
20. Baret, F.; Guyot, G. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens. Environ.* **1991**, *35*, 161–173. [CrossRef]
21. Ren, H.; Feng, G. Are soil-adjusted vegetation indices better than soil-unadjusted vegetation indices for above-ground green biomass estimation in arid and semi-arid grasslands? *Grass Forage Sci.* **2015**, *70*, 611–619. [CrossRef]
22. Baghi, N.G.; Oldeland, J. Do soil-adjusted or standard vegetation indices better predict above ground biomass of semi-arid, saline rangelands in North-East Iran? *Int. J. Remote Sens.* **2019**, *40*, 8223–8235. [CrossRef]
23. Ali, I.; Cawkwell, F.; Dwyer, E.; Green, S. Modeling Managed Grassland Biomass Estimation by Using Multitemporal Remote Sensing Data—A Machine Learning Approach. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3254–3264. [CrossRef]
24. Ge, J.; Hou, M.; Liang, T.; Feng, Q.; Meng, X.; Liu, J.; Bao, X.; Gao, H. Spatiotemporal dynamics of grassland aboveground biomass and its driving factors in North China over the past 20 years. *Sci. Total Environ.* **2022**, *826*, 154226. [CrossRef] [PubMed]
25. Yang, S.; Feng, Q.; Liang, T.; Liu, B.; Zhang, W.; Xie, H. Modeling grassland above-ground biomass based on artificial neural network and remote sensing in the Three-River Headwaters Region. *Remote Sens. Environ.* **2018**, *204*, 448–455. [CrossRef]
26. Wu, C.; Shen, H.; Shen, A.; Deng, J.; Gan, M.; Zhu, J.; Xu, H.; Wang, K. Comparison of machine-learning methods for above-ground biomass estimation based on Landsat imagery. *J. Appl. Remote Sens.* **2016**, *10*, 35010. [CrossRef]
27. Morais, T.G.; Teixeira, R.F.M.; Figueiredo, M.; Domingos, T. The use of machine learning methods to estimate aboveground biomass of grasslands: A review. *Ecol. Indic.* **2021**, *130*, 108081. [CrossRef]
28. Xu, B.; Yang, X.; Tao, W.; Qin, Z.; Liu, H.; Miao, J. Remote sensing monitoring upon the grass production in China. *Acta Ecol. Sin.* **2007**, *27*, 405–413. [CrossRef]

29. Wylie, B.K.; Meyer, D.J.; Tieszen, L.L.; Mannel, S. Satellite mapping of surface biophysical parameters at the biome scale over the North American grasslands: A case study. *Remote Sens. Environ.* **2002**, *79*, 266–278. [CrossRef]
30. Xu, B.; Yang, X.C.; Tao, W.G.; Qin, Z.H.; Liu, H.Q.; Miao, J.M.; Bi, Y.Y. MODIS-based remote sensing monitoring of grass production in China. *Int. J. Remote Sens.* **2010**, *29*, 5313–5327. [CrossRef]
31. Wang, G.; Liu, S.; Liu, T.; Fu, Z.; Yu, J.; Xue, B. Modelling above-ground biomass based on vegetation indexes: A modified approach for biomass estimation in semi-arid grasslands. *Int. J. Remote Sens.* **2018**, *40*, 3835–3854. [CrossRef]
32. Mutanga, O.; Skidmore, A. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *Int. J. Remote Sens.* **2004**, *25*, 3999–4014. [CrossRef]
33. Lu, D. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote Sens.* **2007**, *27*, 1297–1328. [CrossRef]
34. Magiera, A.; Feilhauer, H.; Waldhardt, R.; Wiesmair, M.; Otte, A. Modelling biomass of mountainous grasslands by including a species composition map. *Ecol. Indic.* **2017**, *78*, 8–18. [CrossRef]
35. Liang, T.; Yang, S.; Feng, Q.; Liu, B.; Zhang, R.; Huang, X.; Xie, H. Multi-factor modeling of above-ground biomass in alpine grassland: A case study in the Three-River Headwaters Region, China. *Remote Sens. Environ.* **2016**, *186*, 164–172. [CrossRef]
36. Ohsowski, B.; Dunfield, K.; Klironomos, J.; Hart, M. Improving Plant Biomass Estimation in the Field Using Partial Least Squares Regression and Ridge Regression. *Botany* **2016**, *94*, 501–508. [CrossRef]
37. Otgonbayar, M.; Atzberger, C.; Chambers, J.; Damdinsuren, A. Mapping pasture biomass in Mongolia using Partial Least Squares, Random Forest regression and Landsat 8 imagery. *Int. J. Remote Sens.* **2018**, *40*, 3204–3226. [CrossRef]
38. Zhang, B.; Zhang, L.; Xie, D.; Yin, X.; Liu, C.; Liu, G. Application of Synthetic NDVI Time Series Blended from Landsat and MODIS Data for Grassland Biomass Estimation. *Remote Sens.* **2015**, *8*, 10. [CrossRef]
39. Wang, J.; Xiao, X.; Bajgain, R.; Starks, P.; Steiner, J.; Doughty, R.B.; Chang, Q. Estimating leaf area index and aboveground biomass of grazing pastures using Sentinel-1, Sentinel-2 and Landsat images. *ISPRS J. Photogramm. Remote Sens.* **2019**, *154*, 189–201. [CrossRef]
40. Wu, X.; Li, P.; Jiang, C.; Liu, P.; He, J.; Hou, X. Climate changes during the past 31 years and their contribution to the changes in the productivity of rangeland vegetation in the Inner Mongolian typical steppe. *Rangel. J.* **2014**, *36*, 519–526. [CrossRef]
41. Zhou, W.; Gang, C.; Zhou, L.; Chen, Y.; Li, J.; Ju, W.; Odeh, I. Dynamic of grassland vegetation degradation and its quantitative assessment in the northwest China. *Acta Oecologica* **2014**, *55*, 86–96. [CrossRef]
42. Zhou, W.; Yang, H.; Huang, L.; Chen, C.; Lin, X.; Hu, Z.; Li, J. Grassland degradation remote sensing monitoring and driving factors quantitative assessment in China from 1982 to 2010. *Ecol. Indic.* **2017**, *83*, 303–313. [CrossRef]
43. Yang, Y.; Wang, Z.; Li, J.; Gang, C.; Zhang, Y.; Zhang, Y.; Odeh, I.; Qi, J. Comparative assessment of grassland degradation dynamics in response to climate variation and human activities in China, Mongolia, Pakistan and Uzbekistan from 2000 to 2013. *J. Arid Environ.* **2016**, *135*, 164–172. [CrossRef]
44. Li, D.; Xu, D.; Wang, Z.; You, X.; Zhang, X.; Song, A. The dynamics of sand-stabilization services in Inner Mongolia, China from 1981 to 2010 and its relationship with climate change and human activities. *Ecol. Indic.* **2018**, *88*, 351–360. [CrossRef]
45. Ma, Y.; Fan, S.; Zhou, L.; Dong, Z.; Zhang, K.; Feng, J. The temporal change of driving factors during the course of land desertification in arid region of North China: The case of Minqin County. *Environ. Geol.* **2007**, *51*, 999–1008. [CrossRef]
46. Wu, N.; Liu, A.; Ye, R.; Yu, D.; Du, W.; Chaolumeng, Q.; Liu, G.; Yu, S. Quantitative analysis of relative impacts of climate change and human activities on Xilingol grassland in recent 40 years. *Glob. Ecol. Conserv.* **2021**, *32*, e01884. [CrossRef]
47. Zhang, Y.; Wang, Q.; Wang, Z.; Yang, Y.; Li, J. Impact of human activities and climate change on the grassland dynamics under different regime policies in the Mongolian Plateau. *Sci. Total Environ.* **2020**, *698*, 134304. [CrossRef]
48. Liu, Y.; Wang, Q.; Zhang, Z.; Tong, L.; Wang, Z.; Li, J. Grassland dynamics in responses to climate variation and human activities in China from 2000 to 2013. *Sci. Total Environ.* **2019**, *690*, 27–39. [CrossRef]
49. Xu, D.Y.; Kang, X.W.; Zhuang, D.F.; Pan, J.J. Multi-scale quantitative assessment of the relative roles of climate change and human activities in desertification—A case study of the Ordos Plateau, China. *J. Arid Environ.* **2010**, *74*, 498–507. [CrossRef]
50. Wu, H.; Lin, A.; Xing, X.; Song, D.; Li, Y. Identifying core driving factors of urban land use change from global land cover products and POI data using the random forest method. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *103*, 102475. [CrossRef]
51. Li, Z.; Liu, M.; Hu, Y.; Xue, Z.; Sui, J. The spatiotemporal changes of marshland and the driving forces in the Sanjiang Plain, Northeast China from 1980 to 2016. *Ecol. Process.* **2020**, *9*, 24. [CrossRef]
52. Ye, R.; Liu, G.; Chang, H.; Shan, Y.; Mu, L.; Wen, C.; Te, R.; Wu, N.; Shi, L.; Liu, Y.; et al. Response of plant traits of Stipa breviflora to grazing intensity and fluctuation in annual precipitation in a desert steppe, northern China. *Glob. Ecol. Conserv.* **2020**, *24*, e01237. [CrossRef]
53. Vermote, E.; Wolfe, R. MOD09GA MODIS/Terra Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V006. Available online: https://ladsweb.modaps.eosdis.nasa.gov/api/v1/productPage/product=MOD09GA (accessed on 19 March 2023).
54. Tachikawa, T.; Kaku, M.; Iwasaki, A.; Gesch, D.B.; Oimoen, M.J.; Zhang, Z.; Danielson, J.J.; Krieger, T.; Curtis, B.; Haase, J.; et al. *ASTER Global Digital Elevation Model Version 2—Summary of Validation Results*; NASA: Washington, DC, USA, 2011; p. 27.
55. Funk, C.; Peterson, P.; Landsfeld, M.; Pedreros, D.; Verdin, J.; Shukla, S.; Husak, G.; Rowland, J.; Harrison, L.; Hoell, A.; et al. The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Sci. Data* **2015**, *2*, 150066. [CrossRef] [PubMed]

56. Muñoz-Sabater, J.; Dutra, E.; Agustí-Panareda, A.; Albergel, C.; Arduini, G.; Balsamo, G.; Boussetta, S.; Choulga, M.; Harrigan, S.; Hersbach, H.; et al. ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* **2021**, *13*, 4349–4383. [CrossRef]
57. Mu, Q.; Heinsch, F.A.; Zhao, M.; Running, S.W. Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sens. Environ.* **2007**, *111*, 519–536. [CrossRef]
58. OpenLandMap. 2020. Available online: https://opengeohub.org/about-openlandmap (accessed on 20 June 2022).
59. Hengl, T. OpenLandMap: Using Machine Learning for Global Good. Available online: https://opengeohub.org/article/openlandmap-using-machine-learning-global-good (accessed on 20 June 2022).
60. Gaughan, A.E.; Stevens, F.R.; Linard, C.; Jia, P.; Tatem, A.J. High Resolution Population Distribution Maps for Southeast Asia in 2010 and 2015. *PLoS ONE* **2013**, *8*, e55882. [CrossRef]
61. WorldPop. What Is WorldPop? Available online: https://www.worldpop.org/ (accessed on 22 May 2022).
62. Aiken, L.S.; West, S.G.; Pitts, S.C. Multiple Linear Regression. In *Handbook of Psychology*; Wiley: Hoboken, NJ, USA, 2012; pp. 481–507.
63. Hair, J.F. Multivariate Data Analysis: An Overview. In *International Encyclopedia of Statistical Science*; Lovric, M., Ed.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 904–907.
64. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [CrossRef]
65. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
66. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
67. Guyon, I.; Elisseeff, A. An Introduction to Variable and Feature Selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.
68. Pettitt, A.N. A Non-Parametric Approach to the Change-Point Problem. *J. R. Stat. Society. Ser. C (Appl. Stat.)* **1979**, *28*, 126–135. [CrossRef]
69. Akritas, M.G.; Murphy, S.A.; LaValley, M.P. The Theil-Sen Estimator With Doubly Censored Data and Applications to Astronomy. *J. Am. Stat. Assoc.* **1995**, *90*, 170–177. [CrossRef]
70. Wang, Y.; Wu, G.; Deng, L.; Tang, Z.; Wang, K.; Sun, W.; Shangguan, Z. Prediction of aboveground grassland biomass on the Loess Plateau, China, using a random forest algorithm. *Sci. Rep.* **2017**, *7*, 6940. [CrossRef] [PubMed]
71. Gao, T.; Xu, B.; Yang, X.; Jin, Y.; Ma, H.; Li, J.; Yu, H. Using MODIS time series data to estimate aboveground biomass and its spatio-temporal variation in Inner Mongolia's grassland between 2001 and 2011. *Int. J. Remote Sens.* **2013**, *34*, 7796–7810. [CrossRef]