



Article

Analysis and Prediction of the Impact of Socio-Economic and Meteorological Factors on Rapeseed Yield Based on Machine Learning

Jiaping Liang ^{1,2,3}, Hang Li ^{1,2,3}, Na Li ^{1,2,3,*}, Qiliang Yang ^{1,2,3,*} and Linchao Li ^{4,*}

¹ Faculty of Modern Agricultural Engineering, Kunming University of Science and Technology, Kunming 650500, China

² Yunnan Provincial Field Scientific Observation and Research Station on Water-Soil-Crop System in Seasonal Arid Region, Kunming University of Science and Technology, Kunming 650500, China

³ Yunnan Provincial Key Laboratory of High-Efficiency Water Use and Green Production of Characteristic Crops in Universities, Kunming University of Science and Technology, Kunming 650500, China

⁴ College of Natural Resources and Environment, Northwest A&F University, Yangling 712100, China

* Correspondence: lina_nafu1221@163.com (N.L.); yangqilianglovena@163.com (Q.Y.); llinchao1@163.com (L.L.)

Abstract: Rapeseed is one of China's major oil crops, and accurate yield forecasting is crucial to the growth of the rapeseed industry and the country's food security. In this study, the data on natural and socio-economic factors from 2001 to 2020 and the yield of rapeseed in China were used as the data basis. The Pearson correlation coefficient was used to analyze the relationship between the influencing factors and the yield of rapeseed, and the prediction effect of four machine learning models (linear regression (LR), decision tree (DTR), random forest (RF), and support vector machine (SVM)) on the yield of rapeseed was compared in China's main rapeseed-producing area. The results demonstrate that the yield of rapeseed in China showed an increasing trend, but fluctuated greatly. Rural electricity consumption, gross agricultural production, the net amount of agricultural fertilizer application, effective irrigation area, total power of agricultural machinery, and consumption of agricultural plastic film had a positive effect on the increase in rapeseed yield. However, due to the impact of climate change and disasters, the yield of rapeseed has had significant fluctuations. A Pearson correlation analysis showed that socio-economic factors (rural electricity consumption, gross agricultural production, effective irrigation area, total power of agricultural machinery, consumption of agricultural plastic film, etc.) played a dominant role in rapeseed yield changes. The RF model had a good prediction effect on rapeseed yield, and natural factors and socio-economic factors had different effects on spring rapeseed and winter rapeseed. Winter rapeseed yield was mainly affected by socio-economic factors, accounting for as high as 89% of the importance. Among them, the sown area of rapeseed and the effective irrigation area had the greatest impact. The effects of natural factors and socio-economic factors on spring rapeseed yield were similar, accounting for 47% and 53%, respectively, and the mean annual precipitation, sunshine duration, and sown area of rapeseed were the most influential variables.

Keywords: rapeseed; yield prediction; machine learning; random forest; variable importance



Citation: Liang, J.; Li, H.; Li, N.; Yang, Q.; Li, L. Analysis and Prediction of the Impact of Socio-Economic and Meteorological Factors on Rapeseed Yield Based on Machine Learning. *Agronomy* **2023**, *13*, 1867. <https://doi.org/10.3390/agronomy13071867>

Academic Editor: Roberto Marani

Received: 15 June 2023

Revised: 10 July 2023

Accepted: 12 July 2023

Published: 14 July 2023



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

1. Introduction

Rapeseed is one of the four major oil crops in the world, and is a major source of edible vegetable oil and vegetable protein, occupying an important position in agricultural products. In China, where rapeseed is the oil crop with the largest planting area and the widest distribution, it significantly contributes to national oil supply security [1]. China is one of the largest rapeseed producers in the world, with an average rapeseed production of 13 million tons and 690,000 hectares of planting area in the past two decades. However, due to strong consumer demand in China, domestic rapeseed products are in short supply,

leading to a high reliance on imports, and the yield has been lower than consumption for more than ten consecutive years [2]. Under the background of global climate change, the production conditions and yield of rapeseed are easily affected by different climate variables, and the temperature, sunshine, and precipitation during the growth period of rapeseed have the greatest impact on yield [3–5]. In addition, socioeconomic factors are also important factors affecting crop yield [6,7]. The impact of these factors poses a threat to rape production and industrial development, and the stability and sustainability of the rape industry is at risk.

The development of crop yield prediction models has made significant contributions to agricultural advancement. Crop yield predictions typically use a process-based crop model and statistical analysis methods [8,9]. The crop models have a strong physical mechanism and fully consider various factors affecting crop growth, but they have high requirements for field observed data to calibrate models. Shahhosseini et al. (2021) coupled crop models and machine learning to predict maize yield in the US corn belt, reducing the root mean square error (*RMSE*) of maize yield predictions from 20% to 7% [10]. The CSM-CROPGRO Canola model was used to predict the impact of climate change on rapeseed production. It was found that rapeseed production decreased by nearly 30% on average under the predicted climate change conditions [3]. However, using a crop model to achieve crop yield prediction requires large amounts of data, such as inputs on meteorology, soil, management measures, and crop growth to achieve crop yield prediction, and it requires a lot of professional knowledge to calibrate and validate the model to reduce uncertainty [11–13]. In addition, running crop models, especially at larger scales, requires substantial computational resources and involves significant workloads [8].

In this study, linear regression (*LR*), decision tree regression (*DTR*), random forest (*RF*), and support vector machine regression (*SVM*) were used for modeling. *LR* is a widely used method with an intuitive mathematical expression equation, which can be used to quantitatively evaluate the linear relationship between independent variables and dependent variables, but it cannot consider the nonlinear relationship between variables [14,15]. Decision tree is a kind of tree structure, which is easy to understand and explain. It is easy to over-fit for processing complex data. However, decision trees have many successful applications on small datasets in agriculture. Tanaka et al. (2015) and Banerjee et al. (2022) used decision tree analysis, determined the key factors affecting crop yield changes, and predicted crop yield [16,17]. Rossi Neto et al. (2017) also used the decision tree algorithm to evaluate crop yield by combining soil and climate data [18]. *RF* is an ensemble algorithm composed of decision trees. Because of the ensemble algorithm, random forest can deal with nonlinear problems with high accuracy [19]. Therefore, random forests have been successfully used to solve many important agricultural problems. Pang et al. (2022) used random forest combined with wheat yield data, meteorological variables, and satellite images to establish a yield prediction model to accurately predict wheat yield in south-eastern Australia [20]. Jeong et al. (2016) used a random forest model to compare it with a multiple linear regression model to predict crop yields of wheat, corn, and potatoes at different regional scales [21]. *SVM* is a supervised non-parametric algorithm; it can be used to process multi-dimensional data and has good prediction performance on small sample data. *SVM* is used to participate in the establishment of agricultural prediction systems and accurately predict the yield of various crops in combination with factors such as atmosphere and soil [22,23]. The statistical analysis method based on observation data has the advantage that the data are empirical. It is mainly used in the absence of field management data, and it is difficult to use crop models for parameter adjustment and calibration. It has the characteristics of strong operability, a long research cycle, a wide range, and a large amount of information. In recent years, with the vigorous development of artificial intelligence technology, machine learning algorithms have gradually demonstrated their advantages in constructing prediction models and analyzing the relative importance of feature variables. Researchers have carried out a large number of studies by integrating machine learning methods, such as random forest (*RF*) and support vector machines (*SVM*),

with multi-source environmental data to predict crop yield [24]. Niedbała et al. (2019) used an artificial neural network with a multi-layer perceptron topology to construct a prediction model, and accurately predicted the winter rape yield in southern Poland using meteorological and mineral fertilization data, with the error value MAPE less than 10% [25]. Lischeid et al. (2022) used RF and SVM to identify the meteorological and soil factors that affected the spatial-temporal variation in the yield of four crops (silage major, winter barre, winter rapeseed, and winter wheat yield) in Germany from 1978 to 2017. They found that crop yield was more related to summer temperature and precipitation, but not to soil moisture [26]. Jhajharia Kavita et al. (2023) used machine learning algorithms such as RF and SVM to predict rapeseed yield combined with meteorological and soil data [27]. However, current research on predicting rapeseed yield lacks a comprehensive consideration of the effects of natural and socio-economic factors, especially in the overall region of China.

Therefore, this study investigates and analyzes the main climatic and socio-economic factors affecting rapeseed yield in China, and uses multiple linear regression (LR), decision tree regression (DTR), random forest (RF), and support vector machine regression (SVR) methods combined with climatic and socioeconomic data to develop a rapeseed yield forecasting model to quantify and identify the important variables affecting rapeseed yield, thus providing research ideas for accurate and efficient rapeseed yield forecasting.

2. Materials and Methods

2.1. Study Area

According to the planting time, rapeseed can be divided into spring rapeseed and winter rapeseed. Winter rapeseed is generally sown in autumn and matures in the following summer. Spring rapeseed is generally planted from the end of April to May and harvested in September. The winter-rapeseed-producing areas are mainly distributed in the Yangtze River Basin, including the Yunnan–Guizhou Plateau, Sichuan–Chongqing region, and the middle and lower reaches of the Yangtze River, involving a total of 11 provinces and cities. But, the climatic conditions in various provinces and cities are quite different, as the middle and lower reaches of the Yangtze River have a mild climate and abundant rainfall, and the Yunnan–Guizhou Plateau and the Sichuan Basin are warm and dry in winter. The dry and wet seasons in Yunnan are distinct, and the dry and wet seasons in Guizhou are not obvious. The cold wave in the Sichuan Basin is not easy to invade, but the sunshine is less. Spring rape is mainly planted in five provinces and cities in Inner Mongolia, Shaanxi, Qinghai, Gansu, and Xinjiang. These areas are cold in winter, with less precipitation, and long and strong sunshine. Since 2000, the planting area and yield of winter rape in the middle and lower reaches of the Yangtze River have increased significantly, accounting for more than 85% of the total oilseed rape production in China, due to its mild climate, abundant rainfall, and the support of national policies [28,29]. Spring rapeseed is mainly planted in the north and northwest regions, with only a small amount distributed in the northeast region (Figure 1).

2.2. Data Source and Variable Selection

The rapeseed yield and related socio-economic data affecting rapeseed yield from 2001 to 2020 were obtained from the *China Statistical Yearbook*. The climate data were collected on the China Meteorological Data Sharing Network (<http://data.cma.cn/>, accessed on 1 July 2022) in July 2022, including daily precipitation, minimum temperature, maximum temperature, sunshine hours, average relative humidity, wind speed, and other data from various meteorological stations in China from 1961 to 2020. Temperature is a key factor that determines the rate of crop development, affects the development process, and ultimately affects yield and quality; Water is an essential raw material for photosynthesis and plays an important role in the physiological and ecological processes of crop production. Light is the foundation of plant biological activities [30,31]. Against the backdrop of the rapid development of agricultural technology, production technologies related to

agriculture, such as pesticides, fertilizers, and agricultural machinery, are widely used in crop production [10,32,33]. Moreover, under the influence of national policies and other factors, the planting area of crops is also prone to fluctuations. In addition, the construction of rural infrastructure, such as electricity and water conservancy infrastructure, also affects the development of agricultural production to a certain extent [29,34,35]. In this study, socio-economic data such as gross agricultural production, rural electricity consumption, fertilizer application, and irrigation area, which are closely related to rapeseed yield, as well as natural environmental data such as precipitation, sunshine hours, and temperature were selected as influencing factors [5,32,36–38] (Table 1).

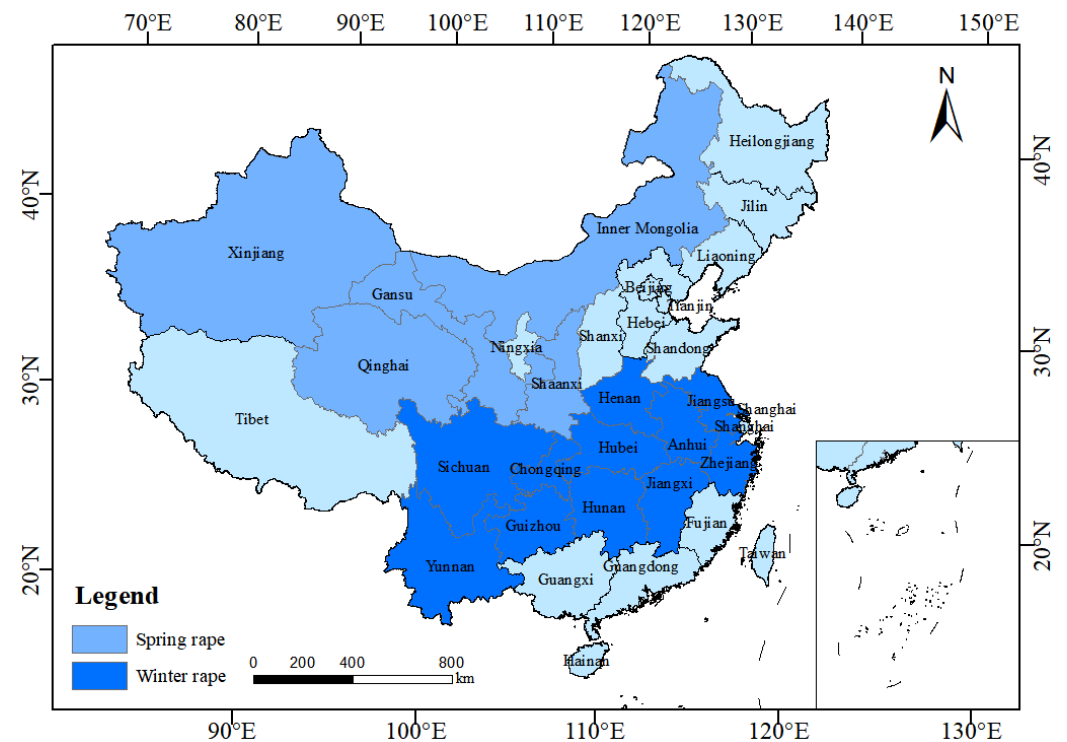


Figure 1. Distribution map of spring rapeseed and winter rapeseed in China.

Table 1. Main natural and socio-economic indicators affecting rapeseed yield.

Type	Definition	Parameters	Unit
Natural	Y	Rapeseed yield	10^4 tons
	X_1	Mean annual precipitation	mm
	X_2	Sunshine duration	hours
	X_3	Mean annual air temperature	$^{\circ}\text{C}$
	X_4	Disaster-affected area	kilo-hectares
Socio-economic	X_5	Sown area of rapeseed	kilo-hectares
	X_6	Rural electricity consumption	kWh
	X_7	Gross agricultural production	10^2 million Yuan (RMB)
	X_8	Net amount of agricultural fertilizer application	kilo-tons
	X_9	Effective irrigation area	kilo-hectares
	X_{10}	Total power of agricultural machinery	kW
	X_{11}	Pesticide usage	kilo-tons
	X_{12}	Consumption of agricultural plastic film	kilo-tons

2.3. Methods

2.3.1. Data Processing

In order to avoid the dimension affecting the accuracy of the model, the data are normalized. In this study, the min–max scaling algorithm is normalized [39,40].

$$y = \frac{x_i - x_{imin}}{x_{imax} - x_{imin}} \quad (1)$$

where y is the processed data of rapeseed yield, climate, and socio-economic data indicators; x_i is the annual value of selected socioeconomic and natural factor indicators in each province. x_{imin} and x_{imax} are the minimum and maximum values in the original variable data.

2.3.2. Correlation Analysis

The Pearson correlation coefficient was used to analyze the relationship between various influencing factors and rapeseed yield, and the significance test was carried out. On this basis, variables with a significant correlation ($p < 0.05$) were selected for comprehensive analysis.

2.3.3. Machine Learning Methods

(1) Linear regression

The linear regression is one of the simplest and most common machine learning algorithms [41]. Linear regression is a linear method of constructing models that describes the relationship between a single output variable (dependent variable) and multiple input explanatory variables (independent variable). The types mainly included are univariate linear regression analysis, multiple linear regression analysis, and nonlinear regression analysis [15]. In this study, the multiple linear regression method was used to predict rapeseed yield, considering natural and socio-economic factors. The model expression is

$$y = b_1x_1 + \dots + b_nx_n + b_0 \quad (2)$$

where, $x_1 \dots x_i$ are the natural and socio-economic indicators, y is the rapeseed yield, $b_1 \dots b_n$ is the regression coefficient, and b_0 is a constant term.

(2) Decision Tree Regression

Decision tree is a basic common machine learning model, and its related algorithms have always been a hot research topic regarding classical machine learning algorithms. Typical algorithms for decision trees include ID3, C4.5, classification and regression tree (CART), etc. [42]. Among them, CART is optimized relative to ID3 and C4.5 [43]. In this study, CART was used to predict the yield of rapeseed. The CART algorithm can generate classification tree and regression tree, which are suitable for the data modeling of multi-feature variables. The algorithm has simple modeling, high accuracy, and strong interpretability, but the algorithm is prone to over-fitting and poor generalization performance. For a given training set, the DTR model, assuming the input space, is divided into M units $R_1, R_2, \dots, R_m$, and there is a fixed output value c_m on each unit; R_m , the regression tree model, can be expressed as

$$f(x) = \sum_{m=1}^M c_m I(x \in R_m) \quad (3)$$

where I is 1 when $x \in R_m$, otherwise I is 0.

(3) Random forest regression

RF is an ensemble-based learning algorithm with a strong anti-noise ability and it does not easily fall into overfitting. It is mainly used for classification prediction and regression prediction. The random forest algorithm is based on the sampling with a replacement (bootstrap) to extract k samples from the original samples, and each sample size is consistent

with the initial training set size. Then, each sample is modeled using the regression decision tree, and k modeling results are obtained. Finally, the final prediction results are obtained by averaging the modeling results of k regression decision trees [19]. In the growth process of each regression tree, the CART method is used to randomly select the variables of the total variables [44]; so, the number of features in the random forest model is $M_{\text{rty}} = 4$. This study uses Python language for random forest regression modeling to obtain the relative importance of each independent variable to the dependent variable, so as to evaluate the importance of each characteristic variable. The sum of the importance values of all feature variables is 1, and the higher the feature importance value is, the more important the feature variable is.

(4) Support Vector Machine Regression

Based on statistical learning theory, support vector machine regression mainly constructs a nonlinear mapping from input space to output space, which can successfully deal with regression and pattern recognition. It maps the input data to a higher dimensional feature space through the kernel function to achieve the best performance of the model [45]. The SVM function can be expressed as

$$f(x) = \sum_{m=1}^M (\alpha_i - \hat{\alpha}_i) k(X_i, X) + b \quad (4)$$

where $\hat{\alpha}_i$ and α_i are Lagrange multipliers, $k(X_i, X)$ is the kernel function, and b is offset. In this study, the radial basis kernel function is selected. The radial basis function is a function that uses vectors as independent variables. It is the most widely used kernel function. Both large and small samples have a better performance.

2.3.4. Model Assessment

In this study, the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) were used to evaluate the models. Among them, R^2 was used to evaluate the accuracy of the prediction model, which is between 0 and 1. The larger the value, the better the model fitting effect. RMSE, MAE, and MAPE were used to evaluate the dispersion between the predicted value and the actual value. The smaller the value, the higher the model accuracy [13,46].

$$R^2 = \left[\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right] \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (6)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (7)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{(\hat{y}_i - y_i)}{y_i} \right| \quad (8)$$

In the formula, n ($i = 1, 2, \dots, n$) is the number of samples used for the machine learning mode, y_i is the actual value of the rapeseed, $\hat{y}_i$ is the predicted value of the rapeseed.

2.3.5. Variable Dependence Analysis

The partial dependence plot (PDP) can be used to explain the machine learning model and show the relationship between the feature variables and the prediction results in the model [47]. The greater the PDP change, the greater the feature importance. On the contrary, the smaller the PDP change, the smaller the feature importance.

3. Result

3.1. The Relationship between Predictors and Rapeseed Yield

The Pearson correlation coefficients between the total yield of spring and winter rapeseed yields and natural and socio-economic indicators are shown in Table 2. There are some differences in the factors affecting the yield of spring rapeseed and winter rapeseed. The yield of spring rapeseed was significantly positively correlated with the sown area of rapeseed, rural electricity consumption, gross agricultural production, net amount of agricultural fertilizer application, effective irrigation area, total power of agricultural machinery, pesticide usage, and consumption of agricultural plastic film ($p < 0.05$). There was a significant negative correlation with the disaster-affected area, for which the correlation coefficient (r) was -0.597 . There was a significant positive correlation between winter rapeseed yield and rural electricity consumption, gross agricultural production, the net amount of agricultural fertilizer application, effective irrigation area, total power of agricultural machinery, and consumption of agricultural plastic film. R is between 0.638 and 0.774. There was a significant negative correlation between winter rapeseed yield and the disaster-affected area, where r is -0.634 .

Table 2. Pearson correlation coefficients between spring and winter rape yields and natural and socioeconomic indicators.

Variables	Spring Rape		Winter Rape	
	r	p -Value	r	p -Value
Mean annual precipitation (X_1)	0.437	0.054	0.079	0.740
Sunshine duration (X_2)	-0.427	0.061	-0.108	0.649
Mean annual air temperature (X_3)	0.197	0.406	0.034	0.888
Disaster-affected area (X_4)	-0.597	0.005	-0.634	0.003
Sown area of rapeseed (X_5)	0.736	<0.001	0.119	0.618
Rural electricity consumption (X_6)	0.837	<0.001	0.754	<0.001
Gross agricultural production (X_7)	0.798	<0.001	0.716	<0.001
Net amount of agricultural fertilizer application (X_8)	0.941	<0.001	0.638	0.002
Effective irrigation area (X_9)	0.878	<0.001	0.679	<0.001
Total power of agricultural machinery (X_{10})	0.892	<0.001	0.774	<0.001
Pesticide usage (X_{11})	0.887	<0.001	0.314	0.178
Consumption of agricultural plastic film (X_{12})	0.877	<0.001	0.768	<0.001

The total output of spring rapeseed increased from less than 0.9 million tons in 2001 to more than 1.6 million tons in 2015 and stabilized above 1.3 million tons after 2008. Among the natural and socioeconomic factors that significantly affect spring rapeseed yield, sown area of seed, rural electricity consumption, gross agricultural production, the net amount of agricultural fertilizer application, effective irrigation area, total power of agricultural machinery, pesticide usage, and the consumption of agricultural plastic film showed an increasing trend. They had a positive effect on the increase in spring rapeseed yield (Figure 2b–i). Spring rapeseed production showed an overall growth trend before 2015, and the sown area and yield decreased significantly after China canceled the temporary storage policy in 2015 (Figure 2b). The disaster-affected area mainly showed a decreasing trend, where spring rapeseed yield increases with the reduction in the affected area (Figure 2a). Overall, the use of pesticides had a certain beneficial effect on the yield of rapeseed. However, from 2004 to 2005, due to a series of policy incentives such as the “No.1 Document” issued by the Chinese government, rapeseed yield significantly increased [48], but the beneficial effect of pesticides was not significant (Figure 2h).

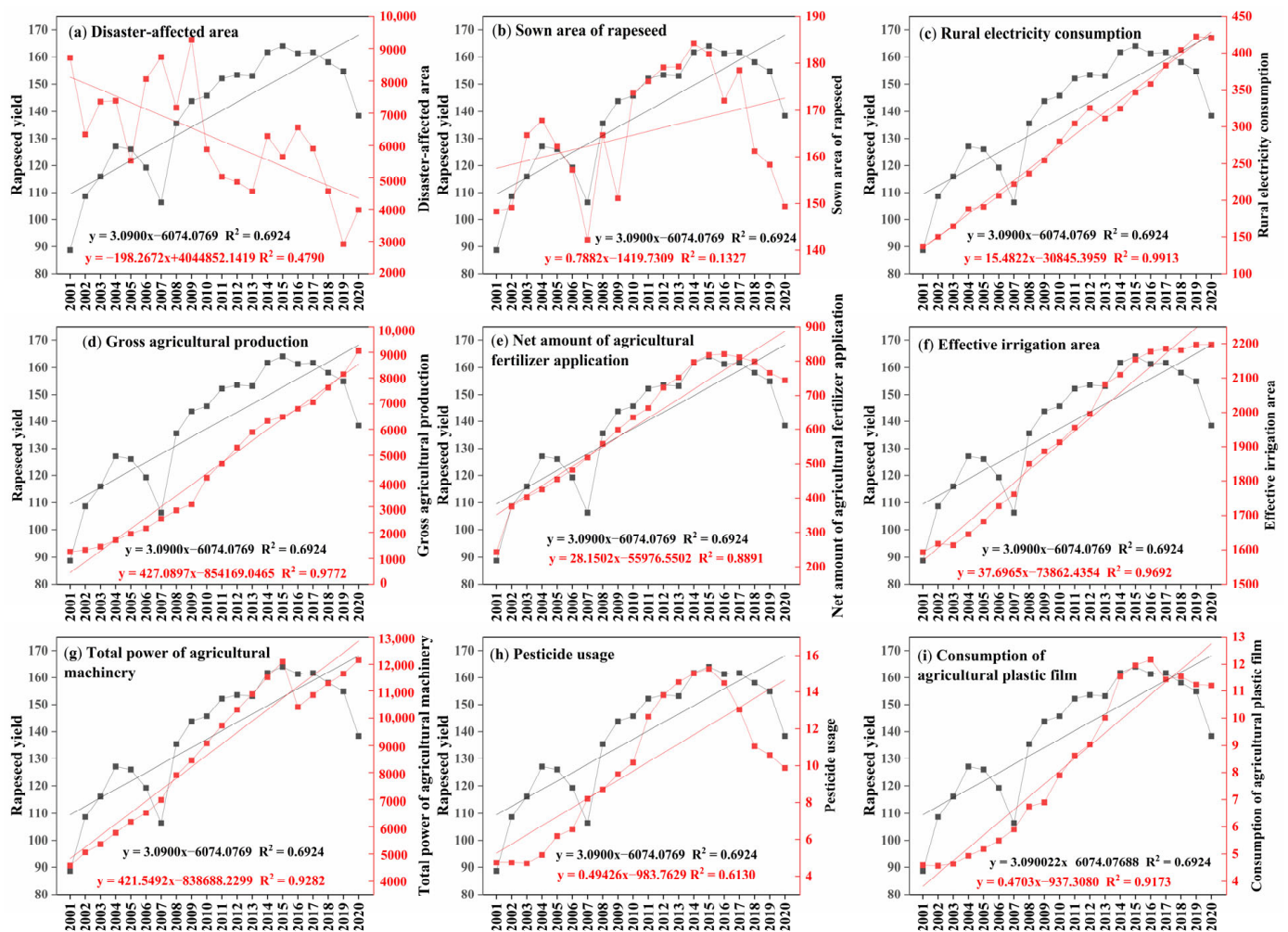


Figure 2. Temporal variations of spring rape yield and significant influence variables ($p < 0.05$).

The yield of winter rapeseed in 2002, 2007, and 2015 showed a large inflection point. The reason was that the continuous rainy weather in the Yangtze River Basin affected the final harvest near the harvest period of rapeseed in 2002, and the yield of rapeseed was reversed from the expected harvest reduction. In 2007, China introduced a series of rape planting support policies, which made the planting area of rape increase significantly and the yield of rapeseed increase significantly. It was the year with the largest change in the past two decades, and the rapeseed yield increased by about one million tons. After the national cancellation of the temporary storage policy of rapeseed in 2015, and under the influence of international competition, the price of rapeseed fell, and the enthusiasm of rapeseed growers was severely affected; as a result, the planting area of rapeseed nationwide significantly decreased, resulting in a decrease in rapeseed production [29]. In addition, the rural electricity consumption, gross agricultural production, net fertilizer application, effective irrigation area, and total power of agricultural machinery have been increasing year after year, which is beneficial for rapeseed cultivation and production (Figure 3b–f). The consumption of agricultural plastic film has been increasing year by year, which is helpful in regard to factors such as water use, temperature regulation, and improvements in nutrient utilization efficiency in rapeseed fields, thereby increasing rapeseed yield (Figure 3g) [49,50].

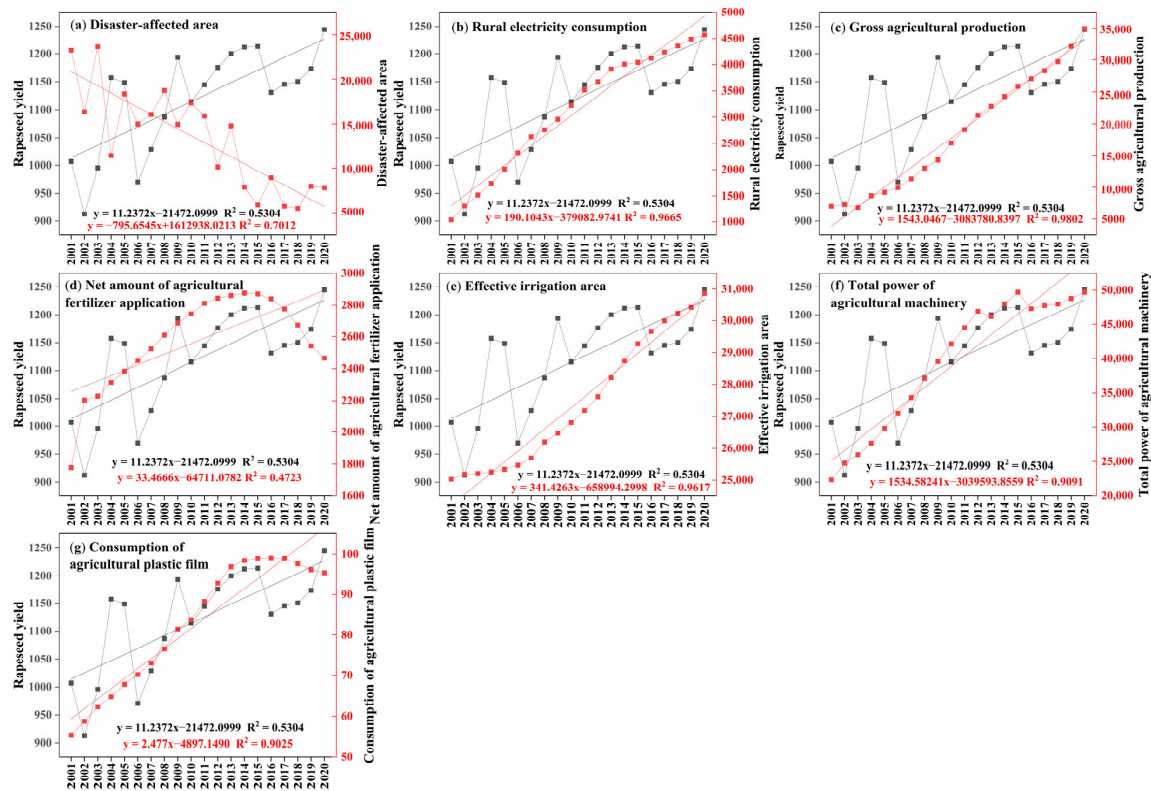


Figure 3. Temporal variations in winter rapeseed yield and significant influencing variables (p -value < 0.05).

3.2. Model Performance

In this study, four machine learning methods (*DTR*, *LR*, *RF* and *SVM*) were used to predict the yield of winter rapeseed and spring rapeseed. The data sequence from 2001 to 2016 was used as the model training set, and the data sequence from 2017 to 2020 was used as the test set. The prediction factors were selected from natural factors and socio-economic factors that affected the yield of winter rapeseed and spring rapeseed. In the yield prediction of winter rapeseed, the random forest model had the highest prediction accuracy, for which R^2 was 0.92 , $RMSE$ was 23.15×10^4 tons, MAE was 13.64×10^4 tons, and $MAPE$ was 12.73% ; the predicted and observed values were close to the 1:1 line, and the prediction performance was better than that of *DTR*, *LR*, and *SVM*. The fitting effects of the decision tree, linear regression model, and support vector machine model are poor, and most of the predicted values are less than the actual observed values. Among the four prediction models, the decision tree model had the lowest prediction accuracy ($R^2 = 0.74$) and a poor fitting effect. *DTR* ($R^2 = 0.75$) and *SVM* ($R^2 = 0.78$) models showed similar performance. In addition, the prediction effect of the models on the high yield of winter rapeseed (>2000 kilo-tons) was general, and the predicted values of the models were generally smaller than the actual observed values (Table 3 and Figure 4).

Table 3. Performance of predicted rape yield based on the machine learning models.

	Model	R^2	RMSE	MAE	MAPE
Spring rape	DTR	0.75	4.65	3.94	14.67
	LR	0.82	3.92	2.96	11.97
	RF	0.87	3.48	2.63	13.20
	SVR	0.78	4.40	3.67	13.02
Winter rape	DTR	0.81	37.20	25.97	25.53
	LR	0.77	40.99	32.63	38.03
	RF	0.92	23.15	13.64	12.73
	SVR	0.79	40.19	28.08	24.62

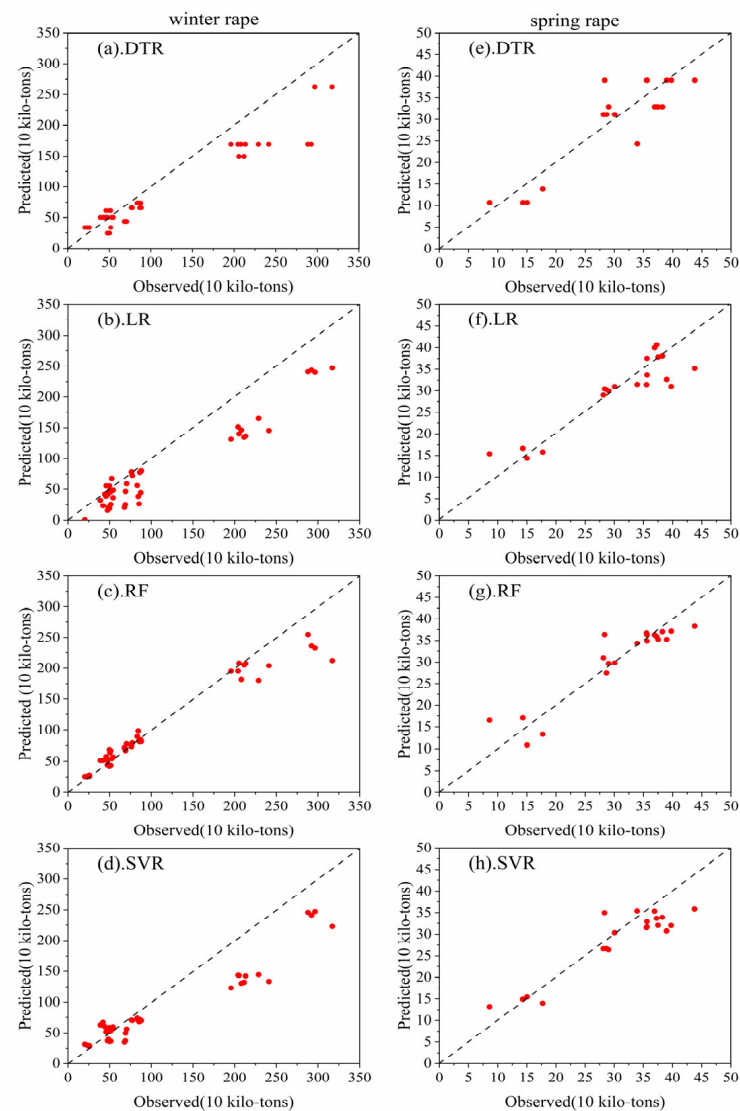


Figure 4. Comparison of observed and predicted winter/spring rapeseed yield in test set. The black dotted lines indicate the 1:1 line.

3.3. Feature Importance and Partial Dependence Analysis

In the prediction model of winter rapeseed yield, the relative importance of socio-economic factors is much greater than that of natural climatic factors, and its importance accounts for nearly 90%. Among them, the sown area of rapeseed, net amount of agricultural fertilizer application, and effective irrigation area are the main variables affecting winter rapeseed yield. This is consistent with the fact that the important environmental factors required for the growth and development of rapeseed, because increasing the planting area of crops can directly provide growth space for crop growth and development, and the effective irrigation area ensure sufficient water supply for the growth and development of rapeseed. Rapeseed production depends largely on the application of exogenous chemical fertilizers, and the rational application of chemical fertilizers can provide effective nutritional conditions for rapeseed growth and development [51–53]. In addition, socio-economic factors such as pesticides, plastic film, and the total power of agricultural machinery also have an important impact on the yield of winter rapeseed. The rational use of agricultural science and technology can greatly promote the scale of rapeseed production, and effectively improve the comprehensive productivity of rapeseed [49,51] (Figures 5a and 6a).

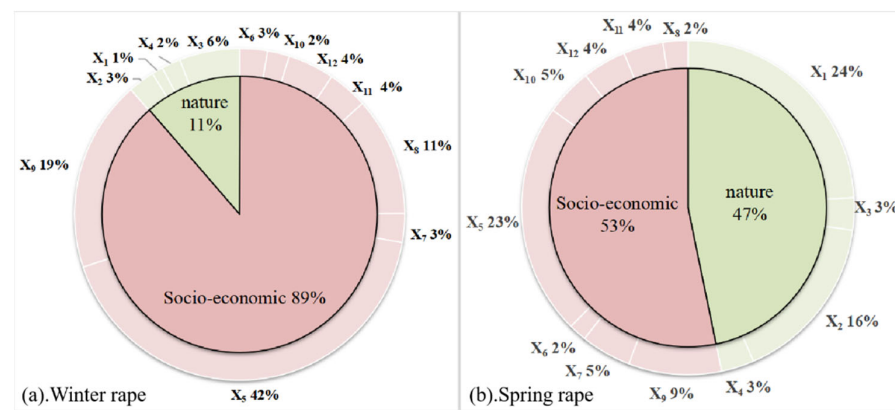


Figure 5. Relative importance ratios of natural and socio-economic indicators based on RF model. The total proportion of variables is 100%.

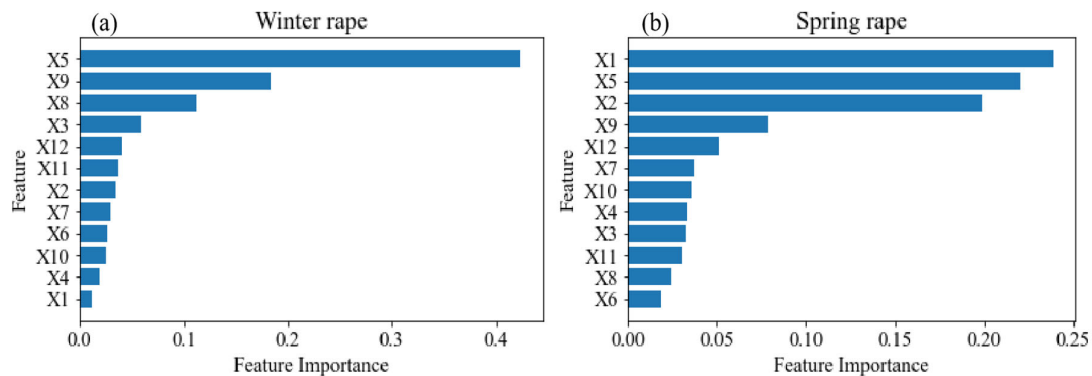


Figure 6. Relative importance of predictor variables from RF models.

Socio-economic and natural factors are similarly important in the prediction model of spring rape yield, with importance values of 53% and 47%, respectively. The three factors that had the greatest impact on spring rapeseed production were mean annual precipitation, sunshine length, and sown area, and their relative weight contributed to 24%, 16%, and 23%, respectively (Figures 5b and 6b). Light and precipitation are important factors affecting crop growth, and they are closely related to the growth and development of crops, photosynthesis, and the absorption and accumulation of nutrients. Moreover, rape is a crop that needs water and light. From flowering to maturity, rape must have more than 300 mm of water and sufficient sunshine time to support its high yield [5,54,55]. The sown area of rapeseed and the effective irrigation area are the two most important socioeconomic factors. The yield of spring and winter rapeseed is mainly affected by socio-economic factors. The sown area of rapeseed and effective irrigation area can reflect the scale of rapeseed production, and the increase in crop planting area and effective irrigation area can increase the yield [56], so it has an important impact on the changes in rapeseed yield.

3.4. Dependence of Rapeseed Yield on Predictors

Based on the random forest model, this study further explored the effects of socio-economic and natural climate factors on rapeseed yield. The association between the yield of winter and spring rapeseed and projected variables such as precipitation, sunshine hours, temperature, rural electricity consumption, and gross agricultural production was revealed by the partial dependent plot (PDP) (Figures 7 and 8). The mean annual precipitation, sunshine duration and disaster-affected area had no significant effect on rapeseed yield. A mean annual air temperature higher than 13 °C harmed winter rapeseed yield. The yield of winter rapeseed increased significantly with the increase in X5, and the increases in X6, X7, X10, X11, and X12 also promoted the increase in winter rapeseed yield. When X8 is

approximately less than 260 kilo-tons, the winter rapeseed yield increases with it. When X9 was less than about 2200 kilo hectares, the yield of winter oilseed rape increased with the increase in X9. When X9 exceeded about 3000 hectares, the yield of winter oilseed rape gradually decreased (Figure 7).

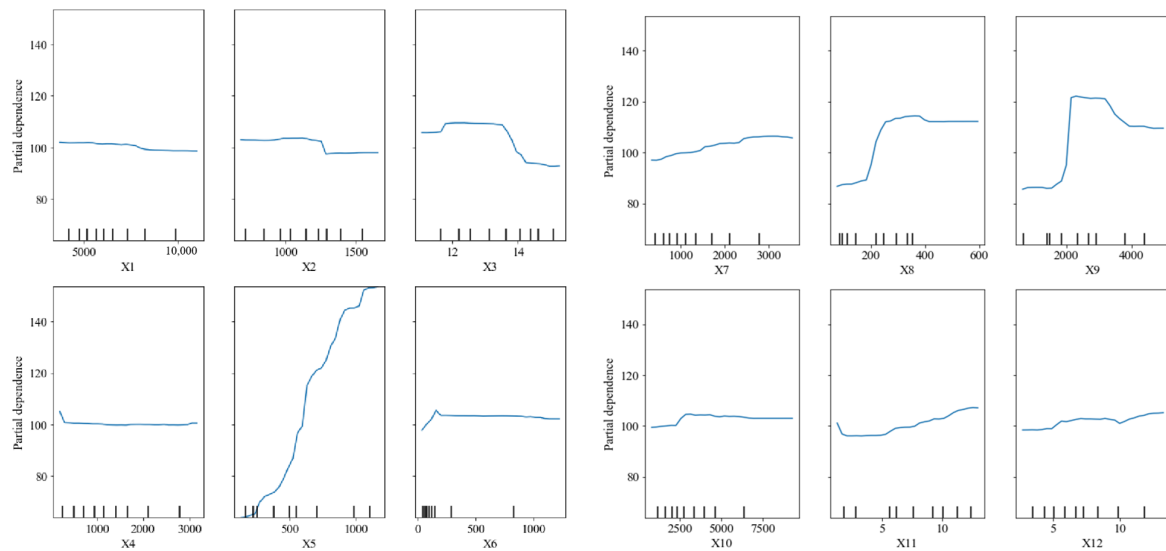


Figure 7. Partial dependence plots of various variables affecting winter rape yield. Note: X1, mean annual precipitation; X2, sunshine duration; X3, mean annual air temperature; X4, disaster-affected area; X5, sown area of rapeseed; X6, rural electricity consumption; X7, gross agricultural production; X8, net amount of agricultural fertilizer application; X9, effective irrigation area; X10, total power of agricultural machinery; X11, pesticide usage; X12, consumption of agricultural plastic film.

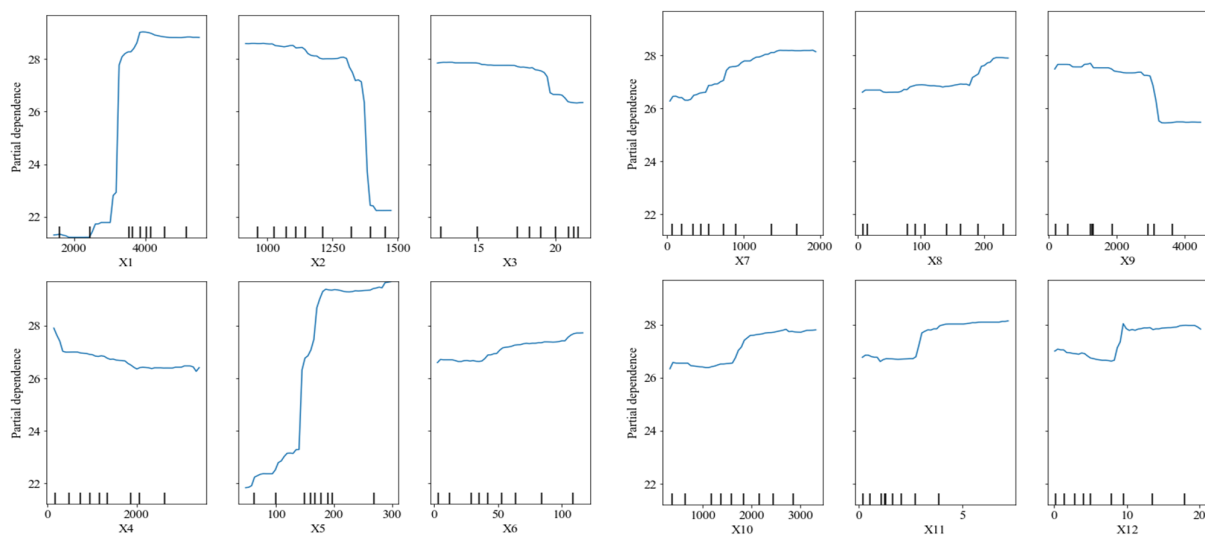


Figure 8. Partial dependence plots of various variables affecting spring rape yield. Note: X1, mean annual precipitation; X2, sunshine duration; X3, mean annual air temperature; X4, disaster-affected area; X5, sown area of rapeseed; X6, rural electricity consumption; X7, gross agricultural production; X8, net amount of agricultural fertilizer application; X9, effective irrigation area; X10, total power of agricultural machinery; X11, pesticide usage; X12, consumption of agricultural plastic film.

The yield of spring rape increased with the increase in X1 and X5. When the mean annual precipitation exceeded about 4000 mm and the sown area of rapeseed exceeded about 190 kilo-hectares, the effect on yield increase gradually slowed down. When X2 was between 1000 and 1500 h and X3 was between 14 and 19 °C, the yield of spring

rapeseed increased with the increase in sunshine duration and mean annual air temperature, but when sunshine duration was greater than 1500 h and X3 was greater than 19 °C, the yield of spring rapeseed decreased with the increases in X2 and X3. The suitable average temperature during the mature period of rape growth is between 15 and 20 °C, which is beneficial to the increase in rapeseed yield. The appropriate sunshine hours and temperature are important factors affecting rapeseed yield [5,57]. The yield of spring rape increased with the increases in X6, X7, and X8. When X10 exceeded about 18 million kW, X11 exceeded about 40,000 tons and X12 exceeded 100,000 tons, and so the yield of spring rape increased rapidly with the increases in X10, X11, and X12, showing that the gain effect was significantly enhanced after exceeding a certain threshold. This may be related to the rapid growth of rapeseed production due to changes in policies, and the scientific and technological development in some provinces. A flat characteristic variable's significance in the PDP plot is lower than that of a PDP input variable with a more dramatic variation in its value range [58] (Figure 8).

4. Discussion

The natural factors of the growth and development time of spring rapeseed and winter rapeseed are quite different. Therefore, this study models the data of spring rapeseed and winter rapeseed separately to reduce the error caused by the large difference in data. Natural factors and socio-economic factors are important factors affecting crop productivity. Under the background of climate change, environmental factors such as precipitation, sunshine and temperature change, and crop growth and development, which in turn affect crop phenology, growth potential, and the planting system, etc., ultimately also affect crop yield [59–61]. The growth and yield of rapeseed are also closely related to temperature, precipitation, and light. The change in the average temperature within the threshold range has a positive effect on the growth and development period of rapeseed, but an extremely high temperature will make rapeseed early in the flowering and filling stage, thus reducing seed yield. Rapeseed is a long-day crop. Sunshine and precipitation can play a positive role in the critical period of rapeseed growth and water demand, promoting biomass accumulation and leaf area increase [5,62,63]. In addition, droughts, floods, and other disasters directly affect crop yield reduction [64,65]. The effective irrigation area, rural electricity consumption, and gross agricultural product, to a certain extent, reflect the speed of the social process of agricultural production. The amount of agricultural fertilizer, pesticide, plastic film, and the total power of agricultural machinery reflect the development of agricultural science and technology [66]. The production of rapeseed is affected by both natural factors and socio-economic factors. The production of rapeseed is mainly affected by socio-economic factors, and the impact of socio-economic factors on rapeseed production is significantly stronger in winter rapeseed production areas than in spring rapeseed production areas (Table 2, Figure 5). On the one hand, in the winter rapeseed production area, the social and economic development conditions in South China, Central China, and East China are better, and the water resources are abundant, and the social and economic factors effectively aid in the production of crops. Spring rapeseed production areas are mainly distributed in northern China, but the social and economic development and agricultural production conditions in the northwest region are relatively poor, water resources are scarce, the climate is dry, and the sunshine time is long [67]. Therefore, spring rapeseed is more susceptible to natural climatic factors than winter rapeseed.

The research shows that the machine learning model has good prediction ability in yield prediction, and the prediction accuracy of crop yields such as for rice, wheat, and corn is higher under different data types [21,68–70]. In this study, four different machine learning algorithms (*LR*, *DTR*, *RF*, and *SVM*) were used to establish a rape yield prediction model that comprehensively considers the effects of natural and socio-economic factors. The results show that the overall performance of the *RF* model is better than that of the *LR*, *DTR*, and *SVM* models, with better estimation capabilities, a higher R^2 ,

and lower error values (*RMSE*, *MAE*, *MAPE*). The RF model has been applied by many researchers to predict the yield of different crops and has shown excellent performance, which is related to its model properties [12,21,69]. The RF model has the advantages of noise insensitivity and unbiased error rate measurement, and can obtain more accurate prediction methods without fitting data [71]. In addition, the RF model can realize the relative importance analysis of variables and the correlation analysis of the partial dependence plot (PDP) [46,72], thereby determining the main variables that affect crop yield and the response of yield to various prediction factors. In addition, the layout of rapeseed production in China is gradually optimized. The production concentration of some provinces is relatively large, and the planting area of rapeseed in the eastern coastal areas is shrinking, which leads to a decline in yield. The spatial layout of rapeseed production generally shows the characteristics of “east reduction, medium stability, west shift and north expansion” [29]. The rapeseed production layout of each province is different, and the natural conditions and socio-economic conditions of each province are different. Under the premise that the overall sample size is small, this paper does not further refine the spring rapeseed and winter rapeseed production areas. Therefore, it is inevitable that there will be excessive differences in rapeseed production in each province. The model trained with fewer samples may lead to greater error values, and the use of provincial data for modeling will inevitably cause errors. If the provincial data are refined to the county level, theoretically a higher prediction accuracy and lower error will be obtained.

5. Conclusions

In this study, the Pearson correlation coefficient and machine learning method were used to analyze the relationship between rapeseed yield and natural factors such as precipitation, temperature, sunshine, and socioeconomic factors such as the sown area of rapeseed, rural electricity consumption, and gross agricultural production in China, and rapeseed yield was predicted based on climate and socioeconomic data. The study found that under the combined influence of natural factors and socio-economic factors, China's rapeseed production fluctuates greatly, but the overall trend is on the rise. The main factors affecting rapeseed yield were rural electricity consumption, gross agricultural production, the net amount of agricultural fertilizer application, effective irrigation area, total power of agricultural machinery, and the consumption of agricultural plastic film; these variables were significantly positively correlated with rapeseed yield ($r > 0$, p -value < 0.001). In terms of predicting rapeseed yield in China, the *LR*, *DTR*, *RF*, and *SVM* models performed well, with the *RF* model outperforming the other three algorithms. The main factors affecting the yield of winter rapeseed are socio-economic factors, which alongside the sown area of rapeseed and effective irrigation area have the greatest influence. The influence of natural factors was small. The effects of natural factors and socio-economic factors on spring rapeseed yield were similar, among which the sunshine duration, sown area of rapeseed, and mean annual precipitation were the most influential variables.

Author Contributions: J.L.: software, data curation, resources, visualization, writing—original draft, writing—review and editing. H.L.: conceptualization, methodology, supervision, formal analysis, writing—review and editing. N.L.: supervision, writing—review and editing, funding acquisition. L.L.: conceptualization, writing—review and editing. Q.Y.: writing revision; financial support. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Yunnan Science and Technology Talent and Platform Program, grant number 202305AM070006, Yunnan Basic Research Program Youth Project, grant number 202301AU070068, Kunming University of Science and Technology “Double First Class” Creation Joint Special Project, grant number 202201BE070001-020.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, Na Li, upon reasonable request.

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

References

1. Fu, D.-H.; Jiang, L.-Y.; Mason, A.S.; Xiao, M.-L.; Zhu, L.-R.; Li, L.-Z.; Zhou, Q.-H.; Shen, C.-J.; Huang, C.-H. Research progress and strategies for multifunctional rapeseed: A case study of China. *J. Integr. Agric.* **2016**, *15*, 1673–1684. [\[CrossRef\]](#)
2. He, W.; Li, J.; Wang, X.; Lin, Q.; Yang, X. Current status of global rapeseed industry and problems, countermeasures of rapeseed industry in China. *China Oils Fats* **2022**, *47*, 1–7.
3. Qian, B.; Jing, Q.; Bélanger, G.; Shang, J.; Huffman, T.; Liu, J.; Hoogenboom, G. Simulated Canola Yield Responses to Climate Change and Adaptation in Canada. *Agron. J.* **2018**, *110*, 133–146. [\[CrossRef\]](#)
4. Dreccer, M.F.; Fainges, J.; Whish, J.; Ogbonnaya, F.C.; Sadras, V.O. Comparison of sensitive stages of wheat, barley, canola, chickpea and field pea to temperature and water stress across Australia. *Agric. For. Meteorol.* **2018**, *248*, 275–294. [\[CrossRef\]](#)
5. Li, X.; Chen, C.; Yang, X.; Xiong, J.; Ma, N. The optimisation of rapeseed yield and growth duration through adaptive crop management in climate change: Evidence from China. *Ital. J. Agron.* **2022**, *17*. [\[CrossRef\]](#)
6. Hampf, A.C.; Carauta, M.; Latynskiy, E.; Libera, A.A.D.; Monteiro, L.; Sentelhas, P.; Troost, C.; Berger, T.; Nendel, C. The biophysical and socio-economic dimension of yield gaps in the southern Amazon—A bio-economic modelling approach. *Agric. Syst.* **2018**, *165*, 1–13. [\[CrossRef\]](#)
7. Rochecoste, J.-F.; Dargusch, P.; Cameron, D.; Smith, C. An analysis of the socio-economic factors influencing the adoption of conservation agriculture as a climate change mitigation activity in Australian dryland grain production. *Agric. Syst.* **2015**, *135*, 20–30. [\[CrossRef\]](#)
8. Cao, J.; Zhang, Z.; Tao, F.; Zhang, L.; Luo, Y.; Zhang, J.; Han, J.; Xie, J. Integrating Multi-Source Data for Rice Yield Prediction across China using Machine Learning and Deep Learning Approaches. *Agric. For. Meteorol.* **2021**, *297*, 108275. [\[CrossRef\]](#)
9. Naghdizadegan Jahromi, M.; Zand-Parsa, S.; Razzaghi, F.; Jamshidi, S.; Didari, S.; Doosthosseini, A.; Pourghasemi, H.R. Developing machine learning models for wheat yield prediction using ground-based data, satellite-based actual evapotranspiration and vegetation indices. *Eur. J. Agron.* **2023**, *146*, 126820. [\[CrossRef\]](#)
10. Shahhosseini, M.; Hu, G.; Huber, I.; Archontoulis, S.V. Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. *Sci. Rep.* **2021**, *11*, 1606. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Shahhosseini, M.; Martinez-Feria, R.A.; Hu, G.; Archontoulis, S.V. Maize yield and nitrate loss prediction with machine learning algorithms. *Environ. Res. Lett.* **2019**, *14*, 124026. [\[CrossRef\]](#)
12. Han, J.; Zhang, Z.; Cao, J.; Luo, Y.; Zhang, L.; Li, Z.; Zhang, J. Prediction of Winter Wheat Yield Based on Multi-Source Data and Machine Learning in China. *Remote Sens.* **2020**, *12*, 236. [\[CrossRef\]](#)
13. Wang, Y.; Zhang, Z.; Feng, L.; Du, Q.; Runge, T. Combining Multi-Source Data and Machine Learning Approaches to Predict Winter Wheat Yield in the Conterminous United States. *Remote Sens.* **2020**, *12*, 1232. [\[CrossRef\]](#)
14. Abdipour, M.; Younessi-Hmazekhanlu, M.; Ramazani, S.H.R.; Omid, A.H. Artificial neural networks and multiple linear regression as potential methods for modeling seed yield of safflower (*Carthamus tinctorius* L.). *Ind. Crops Prod.* **2019**, *127*, 185–194. [\[CrossRef\]](#)
15. Li, B.; Yang, W.; Li, X. Application of combined model with DGM(1,1) and linear regression in grain yield prediction. *Grey Syst. Theory Appl.* **2018**, *8*, 25–34. [\[CrossRef\]](#)
16. Tanaka, A.; Diagne, M.; Saito, K. Causes of yield stagnation in irrigated lowland rice systems in the Senegal River Valley: Application of dichotomous decision tree analysis. *Field Crop Res.* **2015**, *176*, 99–107. [\[CrossRef\]](#)
17. Banerjee, H.; Goswami, R.; Chakraborty, S.; Dutta, S.; Majumdar, K.; Satyanarayana, T.; Jat, M.L.; Zingore, S. Understanding biophysical and socio-economic determinants of maize (*Zea mays* L.) yield variability in eastern India. *Njas-Wagen. J. Life Sc.* **2022**, *70–71*, 79–93. [\[CrossRef\]](#)
18. Rossi Neto, J.; de Souza, Z.M.; de Medeiros Oliveira, S.R.; Kölln, O.T.; Ferreira, D.A.; Carvalho, J.L.N.; Braunbeck, O.A.; Franco, H.C.J. Use of the Decision Tree Technique to Estimate Sugarcane Productivity Under Edaphoclimatic Conditions. *Sugar Tech.* **2017**, *19*, 662–668. [\[CrossRef\]](#)
19. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
20. Pang, A.; Chang, M.W.L.; Chen, Y. Evaluation of Random Forests (RF) for Regional and Local-Scale Wheat Yield Prediction in Southeast Australia. *Sensors* **2022**, *22*, 717. [\[CrossRef\]](#)
21. Jeong, J.H.; Resop, J.P.; Mueller, N.D.; Fleisher, D.H.; Yun, K.; Butler, E.E.; Timlin, D.J.; Shim, K.M.; Gerber, J.S.; Reddy, V.R.; et al. Random Forests for Global and Regional Crop Yield Predictions. *PLoS ONE* **2016**, *11*, e0156571. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Sonal, A.; Sandhya, T. A Hybrid Approach for Crop Yield Prediction Using Machine Learning and Deep Learning Algorithms. *J. Phys. Conf. Ser.* **2021**, *1714*, 012012.
23. Gos, M.; Krzyszczak, J.; Baranowski, P.; Murat, M.; Malinowska, I. Combined TBATS and SVM model of minimum and maximum air temperatures applied to wheat yield prediction at different locations in Europe. *Agric. For. Meteorol.* **2020**, *281*, 107827. [\[CrossRef\]](#)
24. Van Klompenburg, T.; Kassahun, A.; Catal, C. Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric.* **2020**, *177*, 105709. [\[CrossRef\]](#)
25. Niedbala, G. Simple model based on artificial neural network for early prediction and simulation winter rapeseed yield. *J. Integr. Agric.* **2019**, *18*, 54–61. [\[CrossRef\]](#)
26. Lischeid, G.; Webber, H.; Sommer, M.; Nendel, C.; Ewert, F. Machine learning in crop yield modelling: A powerful tool, but no surrogate for science. *Agric. For. Meteorol.* **2022**, *312*, 108698. [\[CrossRef\]](#)

27. Kavita, J.; Pratistha, M.; Sanchit, J.; Sukriti, N. Crop Yield Prediction using Machine Learning and Deep Learning Techniques. *Procedia Comput. Sci.* **2023**, *218*, 406–417.
28. Zhang, S.; Wang, H. Policies and strategies analyses of rapeseed production response to climate change in China. *Chin. J. Oil Crop Sci.* **2012**, *34*, 114–122.
29. Guiping, B.; Xiongze, X.; Jie, X.; Yufeng, Y.; Qianmei, C.; Qingwei, Z. Analysis on the temporal and spatial evolution and influencing factors of oilseed rape production layout in China. *China Oils Fats* **2023**, *48*, 1–6.
30. Song, X.; Zhou, G.; He, Q.; Zhou, H. Stomatal limitations to photosynthesis and their critical Water conditions in different growth stages of maize under water stress. *Agric. Water Manag.* **2020**, *241*, 106330. [[CrossRef](#)]
31. Parthasarathi, T.; Velu, G.; Jeyakumar, P. Impact of Crop Heat Units on Growth and Developmental Physiology of Future Crop Production: A Review. *Res. Rev. A J. Crop Sci.* **2013**, *2*, 1–8.
32. Mousavi-Avval, S.H.; Rafiee, S.; Jafari, A.; Mohammadi, A. Energy flow modeling and sensitivity analysis of inputs for canola production in Iran. *J. Clean. Prod.* **2011**, *19*, 1464–1470. [[CrossRef](#)]
33. Li, W.; Zhang, P. Relationship and integrated development of low-carbon economy, food safety, and agricultural mechanization. *Environ. Sci. Pollut. Res. Int.* **2021**, *28*, 68679–68689. [[CrossRef](#)] [[PubMed](#)]
34. Playán, E.; Mateos, L. Modernization and optimization of irrigation systems to increase water productivity. *Agric. Water Manag.* **2006**, *80*, 100–116. [[CrossRef](#)]
35. Cook, P. Infrastructure, rural electrification and development. *Energy Sustain. Dev.* **2011**, *15*, 304–313. [[CrossRef](#)]
36. Zhang, Z.; Cong, R.-H.; Ren, T.; Li, H.; Zhu, Y.; Lu, J.-W. Optimizing agronomic practices for closing rapeseed yield gaps under intensive cropping systems in China. *J. Integr. Agric.* **2020**, *19*, 1241–1249. [[CrossRef](#)]
37. Fridrihsone, A.; Romagnoli, F.; Cabulis, U. Environmental Life Cycle Assessment of Rapeseed and Rapeseed Oil Produced in Northern Europe: A Latvian Case Study. *Sustainability* **2020**, *12*, 5699. [[CrossRef](#)]
38. Tian, Z.; Ji, Y.; Sun, L.; Xu, X.; Fan, D.; Zhong, H.; Liang, Z.; Gunther, F. Changes in production potentials of rapeseed in the Yangtze River Basin of China under climate change: A multi-model ensemble approach. *J. Geog. Sci.* **2018**, *28*, 1700–1714. [[CrossRef](#)]
39. Maya Gopal, P.S.; Bhargavi, R. A novel approach for efficient crop yield prediction. *Comput. Electron. Agric.* **2019**, *165*, 104968. [[CrossRef](#)]
40. Panda, S.S.; Ames, D.P.; Panigrahi, S. Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques. *Remote Sens.* **2010**, *2*, 673–696. [[CrossRef](#)]
41. Maulud, D.; Abdulazeez, A.M. A Review on Linear Regression Comprehensive in Machine Learning. *J. Appl. Sci. Technol. Trends* **2020**, *1*, 140–147. [[CrossRef](#)]
42. Meng, X.; Zhang, P.; Xu, Y.; Xie, H. Construction of decision tree based on C4.5 algorithm for online voltage stability assessment. *Int. J. Electr. Power Energy Syst.* **2020**, *118*, 105793. [[CrossRef](#)]
43. Huang, Y.; Lan, Y.; Thomson, S.J.; Fang, A.; Hoffmann, W.C.; Lacey, R.E. Development of soft computing and applications in agricultural and biological engineering. *Comput. Electron. Agric.* **2010**, *71*, 107–127. [[CrossRef](#)]
44. Mutanga, O.; Adam, E.; Cho, M.A. High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. *Int. J. Appl. Earth. Obs. Geoinf.* **2012**, *18*, 399–406. [[CrossRef](#)]
45. Wang, H.; Xu, D.; Martinez, A. Parameter selection method for support vector machine based on adaptive fusion of multiple kernel functions and its application in fault diagnosis. *Neural Comput. Appl.* **2018**, *32*, 183–193. [[CrossRef](#)]
46. Hwang, S.-W.; Chung, H.; Lee, T.; Kim, J.; Kim, Y.; Kim, J.-C.; Kwak, H.W.; Choi, I.-G.; Yeo, H. Feature importance measures from random forest regressor using near-infrared spectra for predicting carbonization characteristics of kraft lignin-derived hydrochar. *J. Wood. Sci.* **2023**, *69*, 1. [[CrossRef](#)]
47. Ge, Y.; Zhao, L.; Chen, J.; Li, X.; Li, H.; Wang, Z.; Ren, Y. Study on Soil Erosion Driving Forces by Using (R)USLE Framework and Machine Learning: A Case Study in Southwest China. *Land* **2023**, *12*, 639. [[CrossRef](#)]
48. Xuehua, X.; Yin, Z. Review of domestic rapeseed market in 2004 and outlook for 2005. *Food Sci. Technol. Econ.* **2005**, *30*, 21–23. [[CrossRef](#)]
49. Gu, X.-B.; Li, Y.-N.; Du, Y.-D. Biodegradable film mulching improves soil temperature, moisture and seed yield of winter oilseed rape (*Brassica napus* L.). *Soil Tillage Res.* **2017**, *171*, 42–50. [[CrossRef](#)]
50. Gu, X.; Cai, H.; Zhang, Z.; Fang, H.; Chen, P.; Huang, P.; Li, Y.; Li, Y.; Zhang, L.; Zhou, J.; et al. Ridge-furrow full film mulching: An adaptive management strategy to reduce irrigation of dryland winter rapeseed (*Brassica napus* L.) in northwest China. *Agric. For. Meteorol.* **2019**, *266–267*, 119–128. [[CrossRef](#)]
51. Hu, Q.; Hua, W.; Yin, Y.; Zhang, X.; Liu, L.; Shi, J.; Zhao, Y.; Qin, L.; Chen, C.; Wang, H. Rapeseed research and production in China. *Crop J.* **2017**, *5*, 127–135. [[CrossRef](#)]
52. Feng, J.; Hussain, H.A.; Hussain, S.; Shi, C.; Cholidah, L.; Men, S.; Ke, J.; Wang, L. Optimum Water and Fertilizer Management for Better Growth and Resource Use Efficiency of Rapeseed in Rainy and Drought Seasons. *Sustainability* **2020**, *12*, 703. [[CrossRef](#)]
53. Rong, L.-B.; Gong, K.-Y.; Duan, F.-Y.; Li, S.-K.; Zhao, M.; He, J.; Zhou, W.-B.; Yu, Q. Yield gap and resource utilization efficiency of three major food crops in the world—A review. *J. Integr. Agric.* **2021**, *20*, 349–362. [[CrossRef](#)]
54. Hoffmann, M.P.; Jacobs, A.; Whitbread, A.M. Crop modelling based analysis of site-specific production limitations of winter oilseed rape in northern Germany. *Field Crop Res.* **2015**, *178*, 49–62. [[CrossRef](#)]

55. Huang, J.; Zhou, L.; Zhang, F.; Li, Y. Responses of Yield Fluctuation of Winter Oilseed Rape to Climate Anomalies in South China at Provincial Scale. *Int. J. Plant Prod.* **2020**, *14*, 521–530. [\[CrossRef\]](#)
56. Kang, Y.; Khan, S.; Ma, X. Climate change impacts on crop yield, crop water productivity and food security—A review. *Prog. Nat. Sci.* **2009**, *19*, 1665–1674. [\[CrossRef\]](#)
57. He, Y.; Revell, B.; Leng, B.; Feng, Z. The Effects of Weather on Oilseed Rape (OSR) Yield in China: Future Implications of Climate Change. *Sustainability* **2017**, *9*, 418. [\[CrossRef\]](#)
58. Shahhosseini, M.; Hu, G.; Archontoulis, S.V. Forecasting Corn Yield with Machine Learning Ensembles. *Front. Plant Sci.* **2020**, *11*, 1120. [\[CrossRef\]](#)
59. Zhang, P.; Zhang, J.; Chen, M. Economic impacts of climate change on agriculture: The importance of additional climatic variables other than temperature and precipitation. *J. Environ. Econ. Manag.* **2017**, *83*, 8–31. [\[CrossRef\]](#)
60. Chen, X.; Chen, F.; Chen, Y.; Gao, Q.; Yang, X.; Yuan, L.; Zhang, F.; Mi, G. Modern maize hybrids in Northeast China exhibit increased yield potential and resource use efficiency despite adverse climate change. *Glob. Chang. Biol.* **2013**, *19*, 923–936. [\[CrossRef\]](#)
61. Liu, Z.; Yang, X.; Chen, F.; Wang, E. The effects of past climate change on the northern limits of maize planting in Northeast China. *Clim. Chang.* **2012**, *117*, 891–902. [\[CrossRef\]](#)
62. Zheng, M.; Terzaghi, W.; Wang, H.; Hua, W. Integrated strategies for increasing rapeseed yield. *Trends Plant Sci.* **2022**, *27*, 742–745. [\[CrossRef\]](#) [\[PubMed\]](#)
63. Zhang, J.; Liu, Y. Decoupling of impact factors reveals the response of cash crops phenology to climate change and adaptive management practice. *Agric. For. Meteorol.* **2022**, *322*, 109010. [\[CrossRef\]](#)
64. Lesk, C.; Rowhani, P.; Ramankutty, N. Influence of extreme weather disasters on global crop production. *Nature* **2016**, *529*, 84–87. [\[CrossRef\]](#)
65. Shi, W.; Wang, M.; Liu, Y. Crop yield and production responses to climate disasters in China. *Sci. Total Environ.* **2021**, *750*, 141147. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Liu, D.; Zhu, X.; Wang, Y. China's agricultural green total factor productivity based on carbon emission: An analysis of evolution trend and influencing factors. *J. Clean. Prod.* **2021**, *278*, 123692. [\[CrossRef\]](#)
67. Lu, S.; Bai, X.; Li, W.; Wang, N. Impacts of climate change on water resources and grain production. *Technol. Forecast. Soc. Chang.* **2019**, *143*, 76–84. [\[CrossRef\]](#)
68. Liu, B.; Liu, Y.; Huang, G.; Jiang, X.; Liang, Y.; Yang, C.; Huang, L. Comparison of yield prediction models and estimation of the relative importance of main agronomic traits affecting rice yield formation in saline-sodic paddy fields. *Eur. J. Agron.* **2023**, *148*, 126870. [\[CrossRef\]](#)
69. Obsie, E.Y.; Qu, H.; Drummond, F. Wild blueberry yield prediction using a combination of computer simulation and machine learning algorithms. *Comput. Electron. Agric.* **2020**, *178*, 105778. [\[CrossRef\]](#)
70. Elavarasan, D.; Vincent, D.R.; Sharma, V.; Zomaya, A.Y.; Srinivasan, K. Forecasting yield by integrating agrarian factors and machine learning models: A survey. *Comput. Electron. Agric.* **2018**, *155*, 257–282. [\[CrossRef\]](#)
71. González, C.; Mira-McWilliams, J.; Juárez, I. Important variable assessment and electricity price forecasting based on regression tree models: Classification and regression trees, Bagging and Random Forests. *IET Gener. Transm. Dis.* **2015**, *9*, 1120–1128. [\[CrossRef\]](#)
72. Gyamerah, S.A.; Ngare, P.; Ikpe, D. Probabilistic forecasting of crop yields via quantile random forest and Epanechnikov Kernel function. *Agric. For. Meteorol.* **2020**, *280*, 107808. [\[CrossRef\]](#)

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