Prediction of Winter Wheat Yield Based on Multi-Source Data and Machine Learning in China

Jichong Han [ID]**, Zhao Zhang *, Juan Cao, Yuchuan Luo, Liangliang Zhang, Ziyue Li and Jing Zhang**

State Key Laboratory of Earth Surface Processes and Resource Ecology/ MoE Key Laboratory of Environmental Change and Natural Hazards, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China; 201921051199@mail.bnu.edu.cn (J.H.); caojuan@mail.bnu.edu.cn (J.C.); luo_yuchuan@mail.bnu.edu.cn (Y.L.); zhangliangliang@mail.bnu.edu.cn (L.Z.); 201821051202@mail.bnu.edu.cn (Z.L.); jingzhang@mail.bnu.edu.cn (J.Z.)

* Correspondence: zhangzhao@bnu.edu.cn; Tel.: +86-10-58800409

check for
updates

Abstract: Wheat is one of the main crops in China, and crop yield prediction is important for regional trade and national food security. There are increasing concerns with respect to how to integrate multi-source data and employ machine learning techniques to establish a simple, timely, and accurate crop yield prediction model at an administrative unit. Many previous studies were mainly focused on the whole crop growth period through expensive manual surveys, remote sensing, or climate data. However, the effect of selecting different time window on yield prediction was still unknown. Thus, we separated the whole growth period into four time windows and assessed their corresponding predictive ability by taking the major winter wheat production regions of China as an example in the study. Firstly we developed a modeling framework to integrate climate data, remote sensing data and soil data to predict winter wheat yield based on the Google Earth Engine (GEE) platform. The results show that the models can accurately predict yield 1~2 months before the harvesting dates at the county level in China with an $R^2 > 0.75$ and yield error less than 10%. Support vector machine (SVM), Gaussian process regression (GPR), and random forest (RF) represent the top three best methods for predicting yields among the eight typical machine learning models tested in this study. In addition, we also found that different agricultural zones and temporal training settings affect prediction accuracy. The three models perform better as more winter wheat growing season information becomes available. Our findings highlight a potentially powerful tool to predict yield using multiple-source data and machine learning in other regions and for crops.

Keywords: wheat yield prediction; multi-source data; machine learning; Google Earth Engine (GEE); *Triticum aestivum* L.

1. Introduction

Wheat (*Triticum aestivum L.*), as one of the three top grains (wheat, rice, and corn) and one of the most productive cereals in the 21st century [1], provides the most calories and protein for the global food supply [2]. The accurate prediction of crop yields in advance plays an important role in the grain circulation market, famine prevention, and food security [3]. Crop yield forecasting is also valuable for managing field activities, such as fertilization [4]. China is the world's top wheat producer, accounting for 11.26% of the world's total wheat acreage and 17.98% of the world's total production [5]. Especially for winter wheat, China's production assumes an absolutely dominant role, with nearly 85% of total summer grain production [6]. Therefore, accurately and timely estimating winter wheat yield in China is highly required, considering its significant influence on agricultural development and national food security, or even global scale.

In recent decades, many researchers have been increasingly focused on improving crop yield prediction by different methods, including empirical statistical models and process-oriented crop growth models [7,8]. Conventional statistical models predict yields by developing regression equations between weather variables (temperature, precipitation, solar radiation, etc.) and measured yields at different temporal and spatial scales [2,9–13]. Such regression results did show distinctly how climatic factors affected yields, however their relative lower explanation ability was commonly debated, and the dominant factors controlling yields often varied by geographical location, crop variety, and growing season [14]. Thus, the spatial generalization ability of these models is very low, that is to say they were difficult to apply to larger areas.

On the other hand, as an alternative statistical model, process-based crop models are simultaneously applied to simulate and predict yields throughout the world, e.g. decision support system for agrotechnology transfer (DSSAT) [15], agricultural production systems sIMulator (APSIM) [16], model to capture the crop–weather relationship over a large area (MCWLA) [8,17], and world food studies (WOFOST) [18]. Although these models can simulate crop yields with higher accuracy, lots of inputs (e.g., climatic variables, fertilizers, irrigations, soil, and hydrological features) are required to run a model, which are time-consuming, cost-intensive, and difficult to popularize into a larger region or a developing country [14,19,20]. Climate variables are the primary inputs for the above two approaches. Using more vegetation growth status variables, such as normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI), would obtain better results for yield predictions [2,21–23]. The rapid development of remote sensing technology makes these indexes available both on longer temporal and wider spatial resolutions [24].

Machine learning has demonstrated its powerful performance in data mining [25] and agricultural analyses, including crop type classification and yield prediction [2,26]. Crop yield is a function of the interaction between spatial and temporal changes of variables. Considering the strong ability in treating multi-dimensional datasets [14,25]. Accordingly, machine learning techniques could provide powerful supports for improving yield prediction models. Several publications conducted recently in the world [2,27] have substantiated this viewpoint.

However, the variables selected in many studies were based on the entire growing season, which means the final yield wouldn't be estimated until harvesting date [2,28,29]. To our knowledge, few studies have focused on determining the optimal temporal training settings for China's winter wheat yield prediction. Intuitively, determining the best time window, which can capture well the evolution of winter wheat, would improve greatly crop yield prediction models [30]. Hence, we identify the optimal time window for the training settings, and such potential findings will benefit forecast crop yield in advance. Finding the best time to capture the growth characteristics of crops can improve the performance of the yield prediction model. Most studies have used empirical models or crop growth models to predict yield of winter wheat in local areas of China, mainly in the Huang Huai Hai Plain [31–33]. There are few studies on the yield prediction of Chinese winter wheat on a large scale using data mining technology. It is quite important for macro-control economic policy to develop a large-scale yield prediction model or a prediction framework. Furthermore, the current research tends to study the impact of climate change on yield [34–36], and the research on the difference of yield prediction accuracy in different regions is insufficient.

Although previous studies have greatly improved yield prediction accuracy from spatial and temporal domains, they only focused on partial regions due to the complicated data process [20,37]. Crop yield prediction at a larger-area scale generally requires a large amount of data and complex data processing, suggesting high costs for acquiring and processing large data sets. Fortunately, GEE is provided freely as a cloud-based computing platform to store and process huge data sets (at petabyte-scale) for geospatial analysis and visualization of geospatial datasets [38–40].

In view of the shortcomings of current research, referring to previous studies, we integrated 21 indicators derived from remote sensing data, climate data, and soil properties data based on GEE platform, to build machine learning models for predicting winter wheat yield in China. We adopted

eight machine learning algorithms for predicting winter wheat yield at the county scale, including K-nearest neighbor (KNN), neural network (NN), decision tree (DT), support vector machine (SVM), Gaussian process regression (GPR), random forest (RF), boost trees (BST) and bagging trees (BGT). Our main objectivities are: (1) to construct a winter wheat yield prediction framework; (2) to select the better machine learning algorithms for yield prediction; (3) to identify the best time window for the training settings of winter wheat; (4) to explore the regional differences of yield prediction and the relative importance of variables.

2. Materials and Methods

2.1. Study Area

The studied area extends from 100.20 °E to 121.89 °E and from 23.13 °N to 118.2 °N, consisting of 629 counties. Due to differences in climate and planting pattern, agricultural division of China are divided into nine typical types, six of which are planted by winter wheat (Figure 1). The eastern part (Zone III), as the main planting areas for winter wheat, is located in the North China Plain, with an average annual temperature of 16 °C and an annual rainfall of 800 mm. The growing season of winter wheat is nearly eight months, from the end of September to early or mid-June of the following year [29].
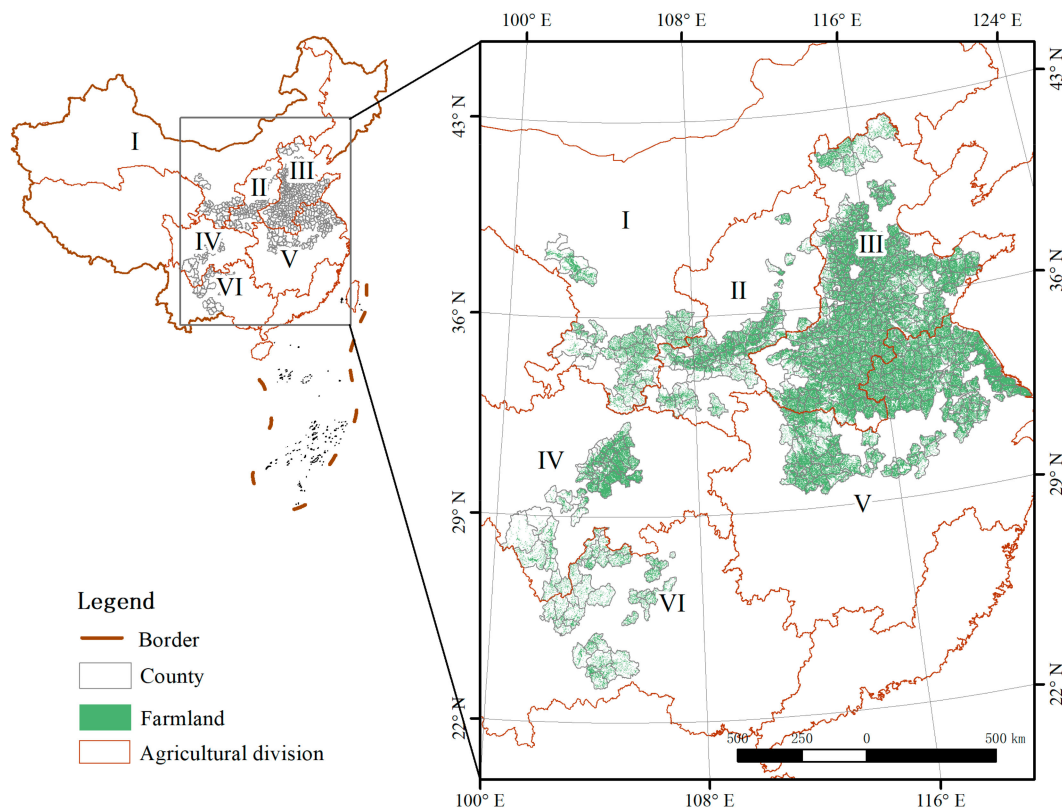

Figure 1. Study area and China's agricultural divisions (Zone I: Northern arid and Semiarid Region; Zone II: Loess Plateau; Zone III: Huang-Huai-Hai Plain; Zone IV: Sichuan Basin and Surrounding Regions; Zone V: Middle-lower Yangtze Plain; Zone VI: Yunnan-Guizhou Plateau and Southern China).

2.2. Data Sources

The data (2001–2014) obtained in the study include remote sensing data, climate data, soil data, yield data and crop map. An overview of the data is presented in Table A1. First, all data were resampled to a common 1×1 km spatial resolution and 1-month temporal resolution. Then all variables were further masked based on the winter wheat planting areas. Finally, we calculated the averages in

each county. Data processing was mainly carried out on GEE platform and ArcGIS. The distribution of the variables used is shown in Figure 2.
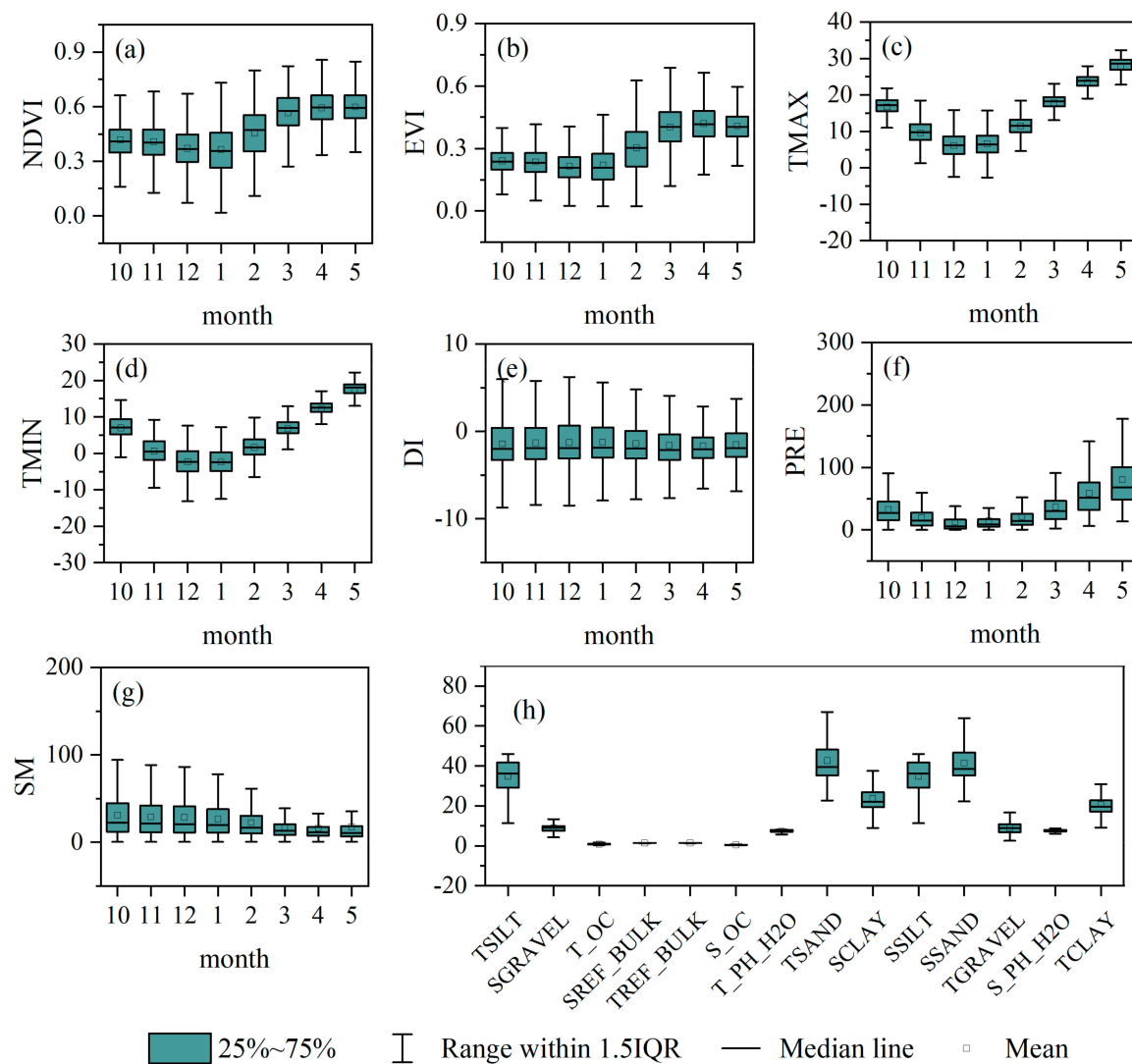

Figure 2. The distribution of NDVI (**a**), EVI (**b**), TMAX (**c**), TMIN (**d**), DI (**e**), PRE (**f**), SM (**g**) and the physical and chemical properties of the soil (**h**) for available data (2001–2014) for the monthly period. The distribution of soil properties for available data for 629 counties in China. 1: January; 2: February; 3: March; 4: April; 5: May; 10: October; 11: November; 12: December. NDVI: normalized vegetation index; EVI: enhanced vegetation index; TMAX: monthly maximum temperature; TMIN: monthly minimum temperature; DI: palmer drought severity index; PRE: monthly precipitation accumulation; SM: soil moisture; SILT: silt content; GRAVEL: volume percentage of crushed stone; OC: organic carbon content; REF_BULK: soil bulk density; PH_H2O: hydrogen ion concentration; SAND: sand content; CLAY: clay content; T and S represent the topsoil layer (0–30 cm) and the subsoil layer (30–100 cm), respectively.

2.2.1. Remote Sensing Data

We collected two types of vegetation index data (VI), the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI). A vegetation index can monitor the dynamic change of vegetation. Currently, many studies have shown that NDVI and EVI have a good correlation with crop yield [41,42]. NDVI is calculated from the red and near-infrared spectral bands, and EVI is obtained from a combination of the red band, near-infrared band, and blue band. Compared with NDVI, EVI does not saturate at high canopy density, it can reduce the canopy background signal and

atmospheric influence and enhance the effects of monitoring vegetation dynamics in high biomass areas. The combination of NDVI and EVI can provide more crop information, which will contribute to crop yield prediction [41]. The two VIs during 2001-2014 in China were derived from the MOD13Q1, which has a spatial resolution of 250 × 250 m (https://ladsweb.modaps.eosdis.nasa.gov/). The recurrence period of this dataset is 16 days. The sequential NDVI and EVI were resampled to monthly scale by the MVC (maximum synthesis method).

2.2.2. Climate Data

Climate variables (e.g., temperature, precipitation) are important drivers for crop production [43–45]. Extreme temperatures and droughts show adverse impacts on crops in the context of global climate change [35,46,47]. We selected monthly maximum temperature (TMAX), monthly minimum temperature (TMIN), monthly drought index (DI) and precipitation (PRE) for predicting winter wheat yield. We used Terra Climate, a dataset of high spatial resolution (1/24°) monthly climate and climatic water balance for global terrestrial surfaces from 2001–2014 [48] (http://doi.org/10.7923/G43J3B0R). The GEE platform was used to process Terra Climate datasets and calculate the climate variables for each county from 2001 to 2014.

2.2.3. Soil Data

Soil is a crucial factor affecting crop yield, including soil water and the physical and chemical properties of soil [49]. For example, soil moisture can increase winter wheat yield in dryland [50,51]. We selected 14 variables (Appendix A Table A1) that describe the physical and chemical properties of the soil. For example organic carbon content, PH for the topsoil layer (0–30 cm) and the subsoil layer (30–100 cm) as the soil variables in this study. Soil moisture (SM) data were obtained from Terra Climate and soil physicochemical properties data from the Harmonized World Soil Database (HSWD) [52]. More details in Appendix A Table A1.

2.2.4. Wheat Yield Data and Planting Area

The winter wheat yield (kg/ha) data were collected from 2001 to 2014 in 629 counties in China, provided by the Agricultural Statistical Yearbook and some unpublished county-level statistics [53]. We calculated the average and standard deviation (SD) of the yield time series in each county, and then defined the observed values as outliers and removed them if they were not within the range of the biophysical attainable yields or ± 3SD [54]. The planting areas were extracted from our previous study on the phenology information of three main crops [55]. Based on the extracted wheat phenology, we defined the pixels including three key phenology stages (green-up, anthesis, and mature stages) for more than 7 years as the winter wheat planting areas. Overall, we selected 629 counties across the whole main wheat planting areas in China.

2.3. Identify the Better Time Window for the Training Settings of Wheat

Different crop growing stages contain different information, and it is essential to determine a suitable time window to retrieve wheat features for yield prediction [56]. Various environmental factors associated with crop yields may vary by the width of time-window during wheat growing season, so we conducted a controlled trial to examine the seasonal sensitivities of winter wheat yields. In China, winter wheat is generally sowed at the end of September and harvested in early or mid-June the next year [29]. To determine the ideal interval for training, we developed machine learning models during different month windows during 2001~2013 and tested them in 2014.

The starting point is triggered by the initial growing period (October and November), and the ending point is determined just before the harvest period (April and May). Accordingly, four time windows are set by combing two different starts and ends as follows: October~May, October~April, November~May, and November~April, respectively.

2.4. Machine-Learning Methods for Estimating Crop Yield

Eight advanced machine learning algorithms were applied here. All variables in the first 13 years (2001–2013) were defined as training samples and those of the last year (2014) as test samples. We standardized all the variables in the datasets by the z-score method before developing each model. All algorithms were implemented in Weka3.8 and matlab2019a.

2.4.1. K-Nearest Neighbor Regression

The K-nearest neighbor (KNN) approach is a type of instance-based learning, which is based on the distance of the predictor variables to the nearest training group known to the model [57]. Aha et al. firstly proposed the new framework and methodology for KNN [58]. KNN can tolerate noise and unrelated properties and has a relatively relaxed concept bias [58].

2.4.2. Neural Network (NN)

Neural networks consist of different elements highly interconnected and have been extensively employed in recent years. BP neural network (BPNN) is one of the most widely used artificial neural networks [59]. BPNN typically includes one input layer, one output layer, and multiple hidden layers. The input layer only inputs data, then the neurons in the hidden layer begin to analyze and process the data, and the results are transmitted to the output layer through the transfer function in the end. When dealing with the issue of nonlinear functions, BPNN would usually be well qualified to detect the complex relationship between independent variables [60].

2.4.3. Decision Tree (DT)

The decision tree (e.g., C4.5) is an effective tool for solving classification and regression problems and has been widely used in remote sensing application [61]. The tree consists of a root node (containing all data), internal nodes, and several leaves. Each node makes a binary decision to separate different categories until the leaf node is reached. The algorithm is non-parametric and can deal with large and complex datasets effectively without complex parameter structure [62]. C4.5 decision tree is a method of approximating discrete value function, which is robust to noisy data [63]. The confidence factor used for pruning is 0.25 and the minimum number of instances per leaf is 2.

2.4.4. Support Vector Machine (SVM)

SVM is a supervised non-parametric algorithm, which is characterized by using the kernels and acting on the margins [64]. During SVM regression, the input is mapped to a high-dimensional feature space using a kernel function, and then a linear regression model is constructed in the new feature space to balance between minimizing errors and overfitting [2,65]. Kernel functions (linear, polynomial, Gaussian, etc.) are one of the important hyper-parameters that need tuning. By comparing different kernel functions, the Gaussian kernel function performed the best in this study.

2.4.5. Gaussian Process Regression (GPR)

GPR is a generalized Gaussian probability distribution for nonlinear regressions [66] and a nonparametric method for a variety of situations, especially for high dimensional space problems. The Gaussian process is a collection of random variables whose properties are any finite number of subsets with a joint Gaussian distribution [67]. However, matrix inversion is a necessary challenge to handle, which increases the computational complexity and causes a very slow run of the model. The GPR used in this paper is based on the kernel function of exponential [67], and the parameters were optimized using Bayesian.

2.4.6. Random Forest (RF)

Random forests are a combination of tree predictors and are more robust with respect to noise [68]. Each tree is built by selecting random variable sets and dataset samples, and all the trees in the forest have the same distribution characteristic. After generating a large number of individual trees, they will vote for the most popular classes. Therefore, RF shows the efficiency to handle high-dimensional datasets and avoids overfitting during the past decade [69,70]. Additionally, RF can quantify the relative importance of measured variables and is a reasonable method for variable selection [68,71].

2.4.7. Ensembles of Learning Machines

Ensemble methods first construct a set of classifiers and then classify new data points by voting on their predictions [72]. This approach can often perform better than any single classifier because it generates classifiers with high precision by combining diverse classifiers with lower precision, which is widely applied to solve practical problems [73]. However, it usually takes more computing time to evaluate the prediction accuracy of an ensemble model. We used two typical ensemble learning algorithms: boost trees (BST) and bagging trees (BGT) [74].

2.5. Model Evaluation

In order to evaluate the eight ML techniques, cross-validation (CV) is a widely used strategy for algorithm selection because of its simplicity, universality, and efficiency in avoiding over-fitting issue [75,76]. It is generally accepted that a model with the smallest estimation error is the best model. We used 5-fold cross-validation to select the model in this study. We adopted the root-mean-square error (RMSE), the coefficient of determination(R^2), and the mean absolute error (MAE) to evaluate the performance of the machine learning model, which can be calculated as follows:

$$R^2 = \frac{\left(\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right)\left(f_i - \overline{f}_i\right)\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right)^2 \sum_{i=1}^{n}\left(f_i - \overline{f}_i\right)^2} \tag{1}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - f_i)^2} \tag{2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - f_i\right| \tag{3}$$

where n ($i = 1, 2, \ldots, n$) is the number of samples used for machine learning model, y_i is the observed winter wheat yield, $\overline{y}_i$ is the corresponding mean value, f_i is the predict winter wheat yield, $\overline{f}_i$ is the corresponding mean value. The closer R^2 is to 1, the higher the prediction performance of the model is. Small RMSE and MAE values indicate less discrepancy within the observed yield and predicted yield.

3. Results

3.1. Comparison of Training Accuracy of Winter Wheat Yield Prediction Models in Different Time Windows

In this study, eight machine learning models were trained with the observed yields and 21 variables of winter wheat from 2001 to 2013 at county level. The evaluated results, based on the five-fold cross-validation, were summarized according to different models and time windows (Figure 3). Comprehensively considering three evaluation indicators (R^2, RMSE and MAE), RF, GPR and SVM models showed the higher accuracy, with higher R^2 (0.79 ~ 0.81) and lower RMSE (<750 kg/ha) and MAE (<531 kg/ha). Although, R^2 of other models are above 0.6, all their RMSEs were >780 kg/ha, even over 1000 kg/ha (KNN), indicating an insignificant relationship between the predicted and the observed yields, and larger errors. Thus, RF, GPR, and SVM are more suitable for winter wheat yield

prediction than other algorithms in China. Moreover, we found that the training accuracy varied by time windows even with the same machine learning algorithm, especially for the RMSE and MAE. However, the time windows showed less impact on R^2 values of RF, GPR and SVM. Finally, three algorithms (SVM, GPR, and RF) were selected to establish prediction models for winter wheat yield at the county level, and all analyses later will only be based on these three models.
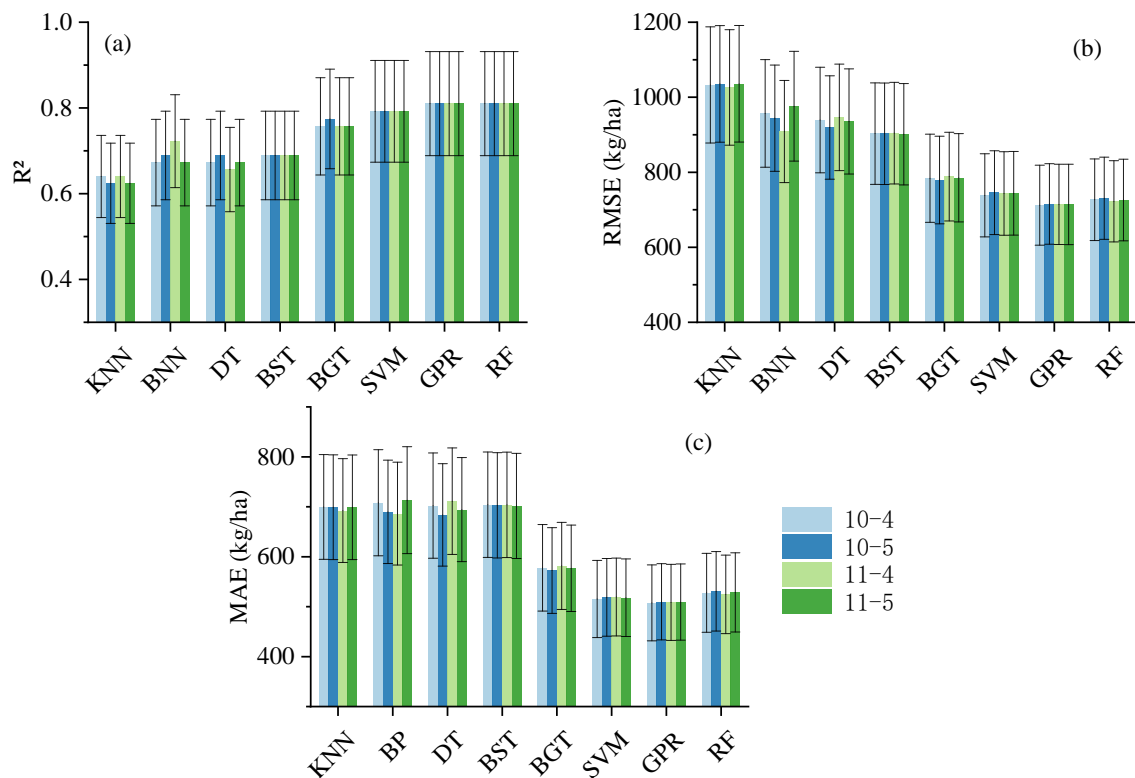

Figure 3. R2 (**a**), RMSE (**b**) and MAE (**c**) skill scores of eight models for winter wheat in different growth periods at the county scale based on five-fold cross-validation results in different time windows. (The unit of RMSE and MAE is kg/ha; 10-4: from October to April; 10-5: from October to May; 11-4: from November to April; 11-5: from November to May.). The error bar represents percentage error, instead of standard error. The top of the line stands for the value plus 15%, and the bottom of the line stands for the value minus 15%.

3.2. Winter Wheat Yield Predictions

Based on the trained models of RF, GPR and SVM in Section 3.1, winter wheat yields of 629 counties in 2014 were predicted. The residuals of the prediction results of these models all passed Kolmogorov-Smirnov test and obey normal distribution, which showed that these regression models were acceptable (Appendix B Figure A1). The scatter diagrams of the predicted and observed yields of the models in different growing periods are shown in Figure 4. We found that the predicted and observed yields showed a good linear fit with R^2 of about 0.80. Such results indicated that the three machine learning models can predict the yield of winter wheat at the county level with higher accuracy, in the order is RF > GPR > SVM. Although all predicted yields were much closer to the 1:1 line, however consistent underestimations were found for all models and all time windows. Moreover, the prediction areas were overestimated for low yields observed with smaller deviations, while underestimated for high observed yields with relatively greater deviations. Nevertheless, all errors of the predicted yields are within 10%, suggesting that RF, GPR, and SVM models perform well for crop prediction at larger scales.
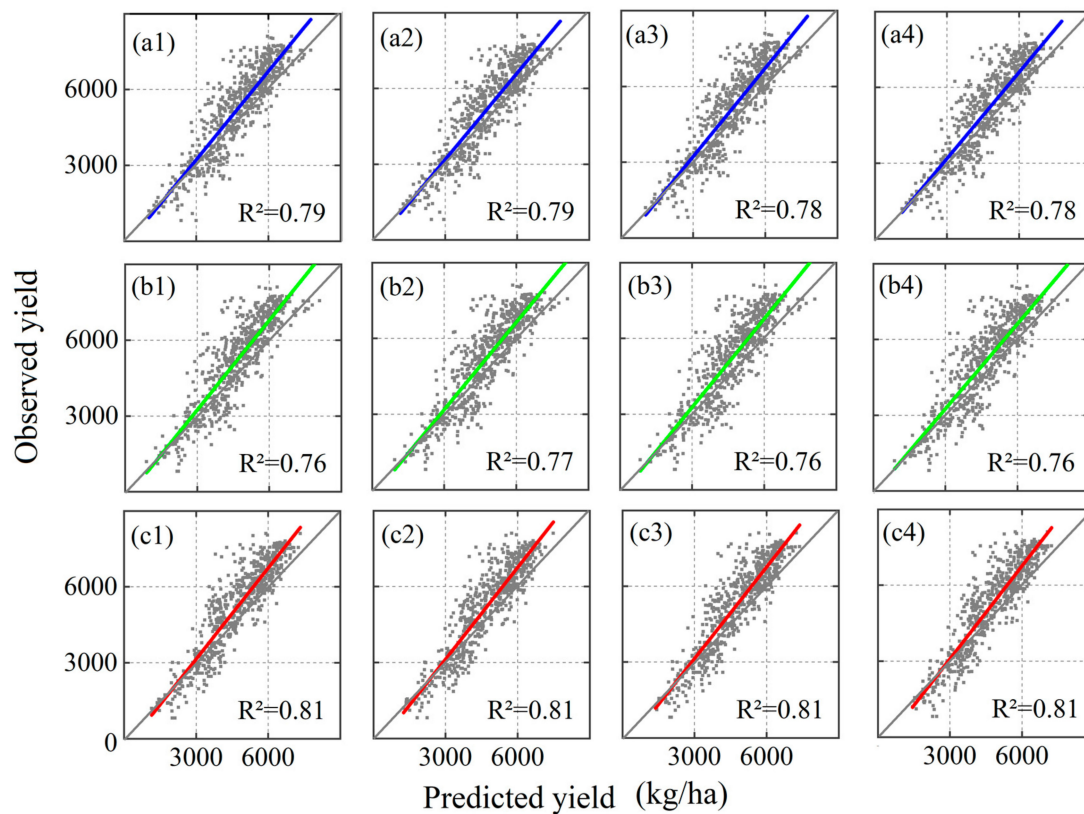
Figure 4. Scatter plots of observed yield and predicted yield of GPR, SVM and RF models for different growth periods of winter wheat at the county scale. One point represents a county (the alphabets (**a**), (**b**) and (**c**) represent GPR, SVM and RF; numbers 1~4 represent October~April, October~May, November~April, and November~May, respectively).

3.3. Impacts of Selecting Time Windows on Prediction Accuracy

We plotted the RMSEs and MAEs of three models by the four growing periods to investigate the impacts of the selected time window on the prediction accuracy (Figure 5). RMSEs (Figure 5a) and MAEs (Figure 5b) were shown in the order of November~April > October~April > Novemeber~May > October~May, regardless of the algorithms or evaluation indicators. The highest accuracy was indicated by the time window of October~May, with the smallest errors (RMSE and MAE). Thus, the closer to the sowing date (the end of September) for time window opening, and the same ending to harvesting date (June), the higher the prediction accuracy will be, which was more strongly substantiated by SVM models. The finding suggests that the prediction accuracy will be significantly improved as more observations during growing seasons are provided by such within-seasons predictor variables. Johnson [77] further pointed out that more dates included will accumulate the entire season information and the developed model would perform better than those from only some dates. Moreover, many studies have also found good relationships between wheat yield and late-season NDVI at the regional scale [78,79], suggesting NDVI prior to harvest can provide good estimates of regional yield. Beyond the above, other attributes such as climate and soil conditions during early sowing have important implications for later crop growth [80,81]. Hence, more information available during crop growing season could improve the accuracy of winter wheat yield prediction.

Moreover, we found that the sensitivities of RMSE and MAE to time window varied by algorithms, with the most significant difference for SVM, followed by GPR and RF (Figure 5). The fewer changes in RF suggest that RF has higher generalization ability than other models [68], with relatively smaller RMSE and MAE values (6.89 and 0.19 kg/ha). It is noted that RF predicted accurately wheat yield 1~2 months in advance before harvest dates ($R^2 > 0.8$).
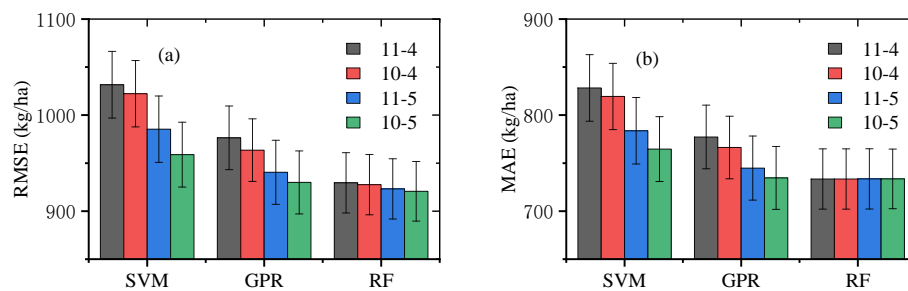
Figure 5. RMSE (**a**) and MAE (**b**) skill scores of the best performing prediction models for winter wheat in different growth periods at the county scale. (SVM: Support vector machine; GPR: Gaussian process regression; RF: Random forest, and 10-4: from October to April; 10-5: from October to May; 11-4: from November to April; 11-5: from November to May). Error bar represent standard error.

3.4. Spatial Patterns of Winter Wheat Yield Predicted

The spatial pattern of the predicted yields by three machine learning models is almost the same as that of the observed yields despite a slight difference (Figure 6). Moreover, the higher yields are located mainly in the eastern areas, while lower yields in the western areas (Figure 6a). However, the areas with higher yields predicted in the east are less than those of observations, especially for SVM (Figure 6b) and GPR (Figure 6c). The eastern areas in the study are the so-called Huang-Huai-Hai plain, the main producing area of winter wheat in China, with flat terrain, fertile soil, suitable climate, and good irrigation conditions. The southwest areas are mainly dominated by rice cultivation, with steep terrain, where winter wheat yields are generally lower and planting areas are fewer. We also found that some extremely high yields (>6500 kg/ha) are difficult to predict, indicating the limitations of the machine learning technique.
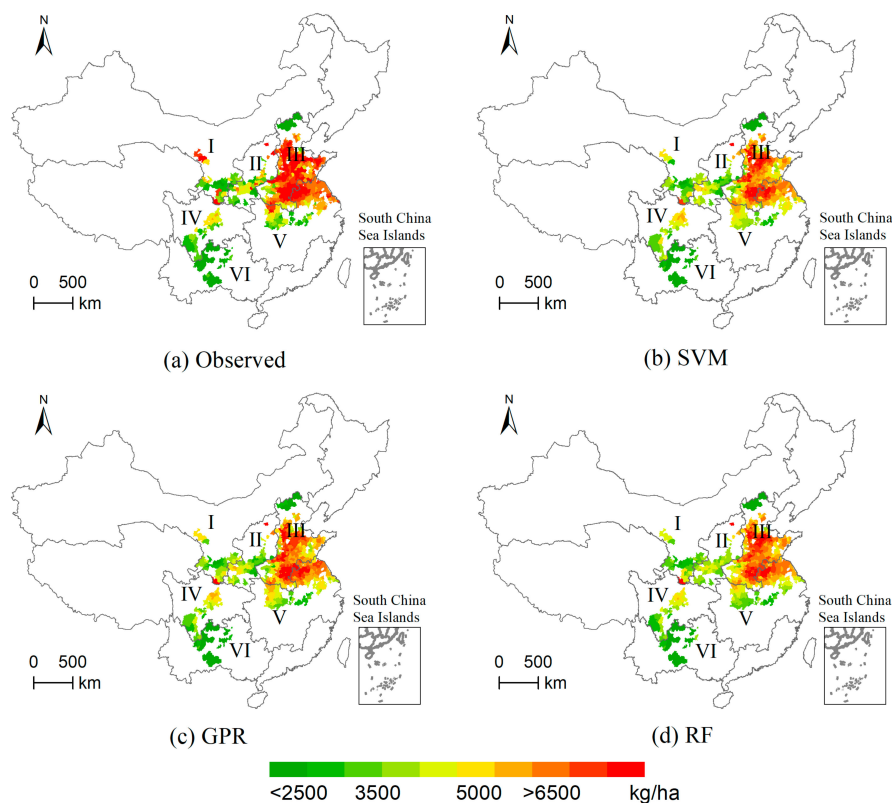

Figure 6. Spatial distribution of observed and predicted yields of winter wheat in 2014. (**a**) yields observed at county scale; yields predicted by SVM (**b**), GPR (**c**), and RF (**d**) at grid-scale. The wheat growth period selected by prediction is from October to next May.

3.5. Comparison of Forecast Errors in Different Areas

Crop yield is driven by the interaction of management, soil, and weather conditions. The yield would vary not only from season to season but also from location to location [14]. To investigate the prediction errors comprehensively, we summarized them according to the areas and the machine learning types (Figure 7). The results showed that the errors of machine learning algorithms do vary by the agricultural zones. With exception of zone III and zone V, the errors of other agricultural zones are all <10%. Moreover, the positive errors suggest overestimations in zone IV and zone VI, and underestimation for the other four areas. Overestimated yields are generally located in the humid and semi-humid areas.
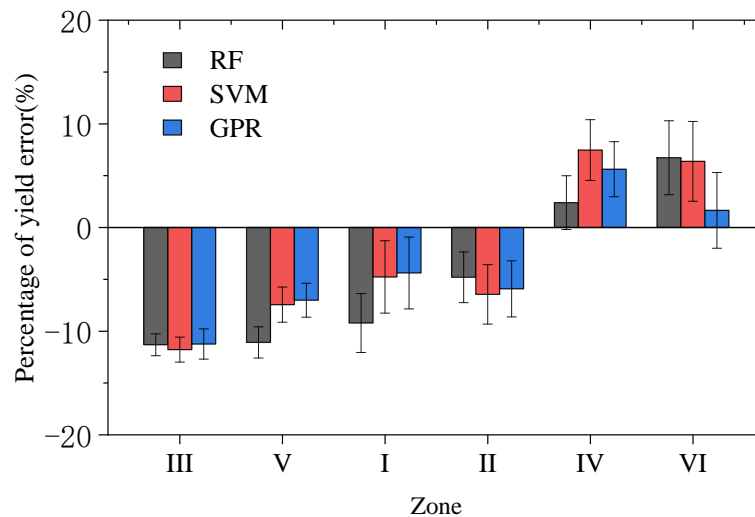

Figure 7. Percentage errors of winter wheat in different agricultural zones. The wheat growth period selected by prediction is from October to next May. Percentage error = (predicted yield - observed yield)/observed yield * 100. Error bar represents one standard error.

We further found that the prediction accuracy of machine learning algorithms does vary by agricultural zones. For example, RF shows the highest accuracy in the zone VI with 2.40% error; SVM and GPR perform the best in zone I (−4.78%, −4.39%). The accuracy which is depended on areas suggests potential other variables should be involved to develop a model. That is to say, a good model developed in one area could reduce its performance in other areas because other key variables exist controlling crop yields, such as crop variety, field managements (e.g., fertilizer application). The crop yield model is very sensitive to the studied area where we are focusing on based on some machine learning algorithms. Therefore, it is essential for us to know the source area as closely resembling the target one as possible when building a crop yield prediction model [14]. In the meantime, optimal spatial area coverage of machine learning methods should be taken into account.

3.6. Prediction Variables and the Order of Relative Importance

To investigate the importance of different variables for the winter wheat yield prediction in China, we calculated the decreased accuracy (mean square error) after removing one variable from the RF model. The mean decrease in accuracy indicates the amount by which the random forest model's prediction accuracy would decrease if one variable is excluded. The larger the decreased accuracy for a prediction variable, the more important the variable is (Figure 8). Mean square error represents the relative importance of the variables. Soil physicochemical properties are generally kept constant during a short period, so we haven't selected soil physicochemical variables. Importance of prediction variables is ordered as: EVI > TMIN > PRE > NDVI > SM > TMAX > DI. EVI is more important for winter wheat yield prediction than NDVI, which is consistent with the study by Bolton et al. [82]. They stated that MODIS-EVI provided relatively better results for predicting crop yields than MODIS-NDVI [82].

As two variables related to temperature, TMIN contributed more significantly than TMAX to the accurate prediction of wheat yield, implying a greater impact of TMIN on wheat yield than TMAX [35].
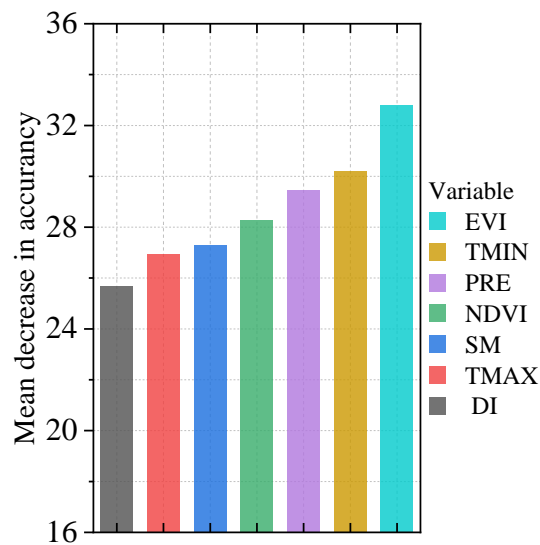

Figure 8. Predictor variable importance based on the RF model, from October to May next year.

4. Discussion

4.1. Model Performance for Estimating Yields in Different Time Windows

We proposed the framework for yield prediction through developing a comprehensive yield prediction model driven by weather, satellite remote sensing, and soil datasets for food security decision-making. The differences in various machine learning algorithms have been discussed above. Among the three best machine learning algorithms (GPR, SVM, and RF), the average computing time of GPR is six times that of RF, but its overall prediction accuracy is slightly lower than that of RF, and the prediction ability of RF is consistently better in different time windows. These results indicate that RF is a promising method for yield prediction [22,83].

Furthermore, we found that various environmental factors related to winter wheat yield showed disparate sensitivities to the growing season [84,85], which is in accord with previous studies demonstrating that the longer time series would better reflect the diversity of growth conditions for estimating yields [84,86]. The time period selected in our study could apply to other wheat planting areas in China, and the combination of months with the optimal time window might vary by the crop phenology in different regions. Generally, the heading period of winter wheat in the study is around in April and the grain filling period in May. Remote sensing vegetation indices (e.g., NDVI and EVI) can reflect the physiological characteristics of crops in different growth stages [87,88]. The close correlations between vegetation indexes and crop yields, especially during the flowering and grain filling stages [78,89,90], have strengthened the important role of later time window for yield prediction. Higher values of vegetation indexes are generally associated with faster growth rates and higher biomass accumulation in the vegetative period, which can prolong the grain filling period by delaying leaf senescence at the maturity stage to increase yield [88,91], and vice versa. Thus, more details of vegetation growing conditions with longer time window will really seize the impacts from environments on the final yields [92].

Moreover, the maximum and minimum temperatures can also characterize the impacts of temperature during the growing season in some degree. For example, Slafer et al. [93] have pointed out that during the grain filling stage, the crop is highly sensitive to temperature, and the increase in temperature causes a contraction of grain filling time [94,95]. However, the dry grain weight (quantified as the product of time and grain filling rate) increases linearly with the extension of the

filling time under a suitable temperature condition [96]. Hence heat stress during the grain filling stage will negatively affect crop yield [97–99] through preventing the transfer of photosynthetic products to the grains [100], or damaging photosynthesis in winter wheat leaves and causing faster senescence and shorter maturity [101,102]. In contrast, the low temperature in winter (frost damage) is a crucial stress affecting the growth of winter wheat, which usually causes damage to, or even the death of wheat seedlings, reducing the number of spikelets or seeds, consequently leading to a decreased yield [93]. Similarly, other climatic variables (e.g., rainfall, soil moisture, and drought index) during the whole growing season of wheat also contribute significantly yield prediction [103].

The weaker associations between EVI/NDVI and yields during the early growth stages, climatic factors could provide more spatial and temporal information for more accurate yield prediction. Combing the appendix role of climatic factors with significant contributions from EVI/NDVI during later growth stages, the yield models have performed robustly for predicting winter wheat yields. Thus, integrating the advantages of remote sensing vegetation indexes and climatic variables, the climate variables in the early time and the late time of the growing period may contribute more to yield forecast. We found the predictors with a longer time windows would better predict the yield because of their better characterization of the diversity of growth conditions. The result is consistent with the previous statement that the effects of different variables on wheat growth and yield are complex and diverse [104,105]. Since machine learning itself is a kind of black box, the conclusions drawn from our analysis are limited in terms of plant physiology. The application of crop process models in the future may help to elaborate the mechanism between variables and yield in more detail.

4.2. Model Performance for Regional Differences

We also found that the accuracy of yield prediction varied by the time window and across different regions. The regional disparate is mainly caused by the regional differences in environmental characters, which have been supported by the previous studies [35,85,86,106]. Ji et al. pointed out that the yield prediction accuracy of the regression models increased with the shrinking of the geographical area [86]. Larger spatial scales include more areas and greater variability in planting conditions. In our study, the variation of winter wheat yield in the Huang-Huai-Hai Plain is mainly affected by irrigation applications and shows a lower sensitivity to rainfall, while a higher sensitivity to rainfall is indicated in other rainfall-fed areas. Moreover, the sample size in diverse regions also will affect yield prediction errors [35]. We established a larger-area-scale winter wheat yield prediction model applicable to average climatic conditions in China and achieved good prediction results. However, the framework of crop yield prediction proposed by us will provide a scientific paradigm for all regions. Attention should be paid to the discrepancy in predictive variables (e.g., winter wheat cultivars) as the spatial region varies.

4.3. Feature Importance in Yield Estimations

The relative impact of a single variable cannot be quantified independently of other variables, and the RF method provides a measure for assessing the relative importance of variables to the prediction results [22]. The results show that vegetation and climate data are crucial for yield prediction. Previous studies have also emphasized the importance of vegetation index, precipitation and temperature in predicting wheat yield [22,23,107,108]. EVI can reflect the phenological characteristics, which is a significant growing state described in many crop growth models (e.g., WOFOST) [109,110]. Due to the phenological changes of plant growth under different climatic conditions, the correlation between crop biomass and remote sensing vegetation index in different phenological periods is different [111–113]. In our study, EVI plays more important role for estimating yield than NDVI. We might attribute the different order to the following: (1) the final crop yield is significantly related to the duration of green biomass [91], especially for the maximum biomass during heading stage [114]; (2) the saturation problem of NDVI under high biomass condition may cause inaccurate yield estimation at the heading stage (April) [88,115]; (3) EVI overcomes the problems of NDVI saturation and soil noise

by decoupling of the canopy background signal and reducing the atmosphere influence, and hence EVI will improve the sensitivity to high biomass condition and canopy structure [88,90,116]. Therefore, EVI is an effective indicator to track crop phenological events and to evaluate and monitor seasonal changes of crops and evergreen vegetation [82,90,117]. Furthermore, in many other crop (corn and rice) yield prediction studies, EVI has proven consistently to be more effective than NDVI [82,88]. The important role of EVI will benefit scientists to estimate winter wheat yield more timely and accurately at a larger spatial scale to ensure regional food security.

Climate data affect crop growth rates, which directly determine the final crop yield because of its strong correlation with spike number, spike length and wheat grouting time [35]. The impact of climate on winter wheat is mainly reflected by the annual variabilities in yields. Additionally, we found that the relative importance of TMIN is higher than that of TMAX (Figure 8), in accord with the previous studies [35]. Tao et al. found that crop yields will increase with the increase of TMIN, which is mainly due to the decrease in frost damage [35]. We also found that the importance of the DI is relatively low, and the potential reasons might be the relatively small difference of DI during the whole growing season (Figure 2e).

4.4. Uncertainties in the Study

There are still some uncertainties, mainly from data sources. For example, the most limitations come from uncertainties in data quality such as the low resolution of the remote sensing, meteorological and soil datasets, which lower the predictive ability of machine learning and make the predicted yields potentially uncertain [118]. The launch of higher-resolution remote sensing satellites, such as sentinel-2 (with the maximum resolution of 10×10 m) [119,120], offers new opportunities to provide more accurate yield prediction. In addition, management measures of human activities such as irrigation, fertilization, and wheat varieties are very important for crop growth, which have been extensively studied in many studies [121]. For example, Ji et al. [86] believe that soil fertility is the key input variable of the artificial neural network model. Because of the limitation of data availability, these variables are only supplied in very limited places, and availability for crop yield prediction at a larger scale remains a challenge. Currently, many studies [122,123] are trying to use more new variables for yield prediction, such as solar-induced chlorophyll fluorescence (SIF) and radar. SIF can reflect crop photosynthesis and microwave remote sensing (either passive or active) can provide canopy biomass and water content. Such additional information will potentially improve yield prediction and should be considered in the future.

In contrast, we should note that no any mechanism processes of crop growth have been included by machine learning models due to their internal black box. Although machine learning is able to mine data efficiently, such unknown processes included will inevitably increase the uncertainty of model performance [85]. Alternatively, crop growth models are developed and validated by many experts during decades of researches. Crop models have characterized the internal growth and development mechanism of crops in some degree and have been widely applied with higher accuracy in many regions. Thus, combining machine learning with the crop growth model is an idea for the future study of yield prediction [123].

5. Conclusions

We predicted winter wheat yield at the county scale based on multi-source data and multiple machine learning models. It was found that RF, GPR, and SVM predicted wheat yields with higher accuracy, and RF demonstrated the best generalization ability among the three methods. RF model can estimate wheat yields accurately in advance (before the harvesting dates) in China. Moreover, we investigated the impact of the time window selected on the prediction accuracy and found the window has a high prediction accuracy throughout the growing period. We also found that the prediction accuracy varied by agricultural zones and algorithms, and the regional difference will affect the yield prediction accuracy. In addition, EVI is the most important predictor used in our study for winter

wheat yield. We are sure the framework to forecast winter wheat yield by multi-source data and the GEE platform is generic and applicable to other crops in the world.

Author Contributions: Conceptualization, Z.Z. and J.H.; Data curation, J.C.; Methodology, J.H.; Supervision, Z.Z.; Validation, Y.L. and L.Z.; Writing—original draft, J.H.; Writing—review & editing, Z.Z. and J.C.; visualization, Z.L. and J.Z. All authors have read and agreed to the published version of the manuscript.

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

Appendix A

Table A1. Data sources used in this study.

Data Type	Variable	Data Description	Period	Resolution	Reference
Yield	-	Yield monitor data(kg/ha)	Year	Regional	[53]
Cultivated land pixel	-	Winter wheat maps	Year	1 km * 1 km	[55]
Remote sensing data	NDVI	MOD13Q1	16-day	250 m * 250 m	NASA LP DAAC
	EVI	MOD13Q1	16-day	250 m * 250 m	
Meteorological data	TMAX	unit: % °C.	Monthly	1/24°, ~4 km	[48]
	TMIN	unit: % °C.	Monthly	1/24°, ~4 km	
	DI	-	Monthly	1/24°, ~4 km	
	PRE	unit: % mm.	Monthly	1/24°, ~4 km	
Soil data	SM	unit: % mm.	Monthly	1/24°, ~4 km	[52]
	T_SILT	unit: % wt.	-	1 km	
	S_SILT	unit: % wt.	-	1 km	
	T_GRAVEL	unit: % vol.	-	1 km	
	S_GRAVEL	unit: % vol.	-	1 km	
	T_OC	unit: % weight.	-	1 km	
	S_OC	unit: % weight.	-	1 km	
	T_REF_BULK	unit: %kg/dm^3.	-	1 km	
	S_REF_BULK	unit: %kg/dm^3.	-	1 km	
	T_PH_H$_2$O	unit: %-log(H$^+$).	-	1 km	
	S_PH_H$_2$O	unit: %-log(H$^+$).	-	1 km	
	T_SAND	unit: % wt.	-	1 km	
	S_SAND	unit: % wt.	-	1 km	
	T_CLAY	unit: % wt.	-	1 km	
	S_CLAY	unit: % wt.	-	1 km	

NDVI: normalized vegetation index; EVI: enhanced vegetation index; TMAX: monthly maximum temperature; TMIN: monthly minimum temperature; DI: palmer drought severity index; PRE: monthly precipitation accumulation; SM: soil moisture, derived using a one-dimensional soil water balance model; SILT: silt content; GRAVEL: volume percentage of crushed stone; OC: organic carbon content; REF_BULK: soil bulk density; PH_H$_2$O: hydrogen ion concentration; SAND: sand content; CLAY: clay content; T and S represent the topsoil layer (0–30 cm) and the subsoil layer (30–100 cm), respectively.
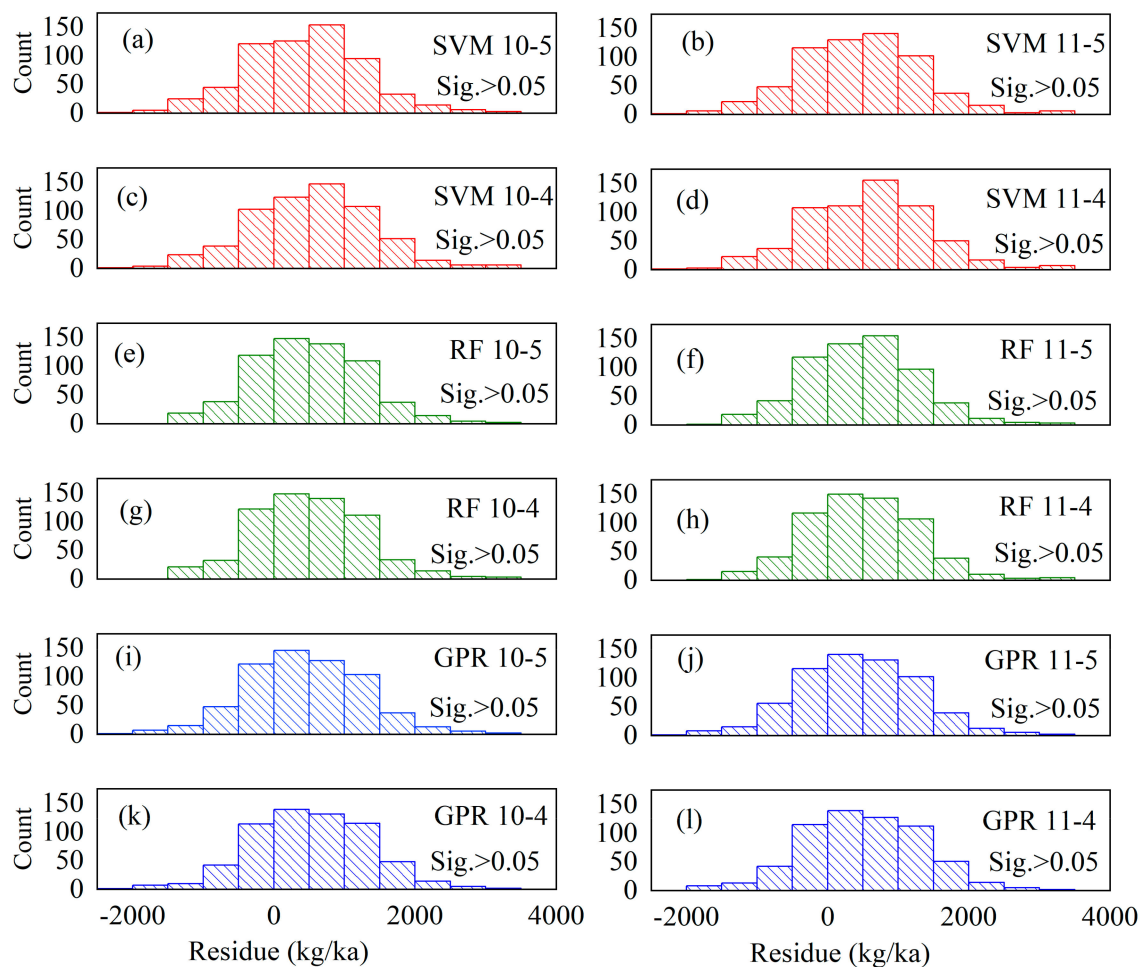
Appendix B

Figure A1. Yield residual distribution predicted by the combination of different models and growth periods. For (**a**), the combination is SVM 10-5, SVM 11-5 (**b**), SVM 10-4 (**c**), SVM 11-4 (**d**), RF 10-5 (**e**), RF 11-5 (**f**), RF 10-4 (**g**), RF 11-4 (**h**), GPR 10-5 (**i**), GPR 11-5 (**j**), GPR 10-4 (**k**), and GPR 11-4 (**l**). The residual distribution of all models had passed Kolmogorov-Smirnov test (Sig. > 0.05), and the distribution of residual conforms to normal distribution. (10-4: from October to April; 10-5: from October to May; 11-4: from November to April; 11-5: from November to May.).

References

1. Curtis, T.; Halford, N.G. Food security: The challenge of increasing wheat yield and the importance of not compromising food safety. *Ann. Appl. Biol.* **2014**, *164*, 354–372. [CrossRef] [PubMed]
2. Cai, Y.; Guan, K.; Lobell, D.; Potgieter, A.B.; Wang, S.; Peng, J.; Xu, T.; Asseng, S.; Zhang, Y.; You, L.; et al. Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agric. For. Meteorol.* **2019**, *274*, 144–159. [CrossRef]
3. Balaghi, R.; Tychon, B.; Eerens, H.; Jlibene, M. Empirical regression models using NDVI, rainfall and temperature data for the early prediction of wheat grain yields in Morocco. *Int. J. Appl. Earth Obs. Geoinf.* **2008**, *10*, 438–452. [CrossRef]
4. Farrell, M.; Macdonald, L.M.; Butler, G.; Chirino-Valle, I.; Condron, L.M. Biochar and fertiliser applications influence phosphorus fractionation and wheat yield. *Biol. Fertil. Soil.* **2014**, *50*, 169–178. [CrossRef]
5. He, Z.; Xia, X.; Zhang, Y. Breeding noodle wheat in China. In *Asian Noodles: Science, Technology, and Processing*; Wiley: Hoboken, NJ, USA, 2010; pp. 1–23.

6. Huang, J.; Tian, L.; Liang, S.; Ma, H.; Becker-Reshef, I.; Huang, Y.; Su, W.; Zhang, X.; Zhu, D.; Wu, W. Improving winter wheat yield estimation by assimilation of the leaf area index from Landsat TM and MODIS data into the WOFOST model. *Agric. For. Meteorol.* **2015**, *204*, 106–121. [CrossRef]

7. Lobell, D.B.; Burke, M.B. On the use of statistical models to predict crop yield responses to climate change. *Agric. For. Meteorol.* **2010**, *150*, 1443–1452. [CrossRef]

8. Tao, F.; Yokozawa, M.; Zhang, Z. Modelling the impacts of weather and climate variability on crop productivity over a large area: A new process-based model development, optimization, and uncertainties analysis. *Agric. For. Meteorol.* **2009**, *149*, 831–850. [CrossRef]

9. Tao, F.; Yokozawa, M.; Liu, J.; Zhang, Z. Climate-crop yield relationships at provincial scales in China and the impacts of recent climate trends. *Clim. Res.* **2008**, *38*, 83–94. [CrossRef]

10. Lobell, D.B.; Schlenker, W.; Costa-Roberts, J. Climate trends and global crop production since 1980. *Science* **2011**, *333*, 616–620. [CrossRef]

11. Shi, W.; Tao, F.; Zhang, Z. A review on statistical models for identifying climate contributions to crop yields. *J. Geogr. Sci.* **2013**, *23*, 567–576. [CrossRef]

12. Tao, F.; Zhang, Z.; Shi, W.; Liu, Y.; Xiao, D.; Zhang, S.; Zhu, Z.; Wang, M.; Liu, F. Single rice growth period was prolonged by cultivars shifts, but yield was damaged by climate change during 1981–2009 in China, and late rice was just opposite. *Glob. Chang. Biol.* **2013**, *19*, 3200–3209. [CrossRef] [PubMed]

13. Zhang, Z.; Song, X.; Tao, F.; Zhang, S.; Shi, W. Climate trends and crop production in China at county scale, 1980 to 2008. *Theor. Appl. Climatol.* **2016**, *123*, 291–302. [CrossRef]

14. Filippi, P.; Jones, E.J.; Wimalathunge, N.S.; Somarathna, P.D.S.N.; Pozza, L.E.; Ugbaje, S.U.; Jephcott, T.G.; Paterson, S.E.; Whelan, B.M.; Bishop, T.F.A. An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. *Precis. Agric.* **2019**, *20*, 1015–1029. [CrossRef]

15. Jones, J.W.; Hoogenboom, G.; Porter, C.H.; Boote, K.J.; Batchelor, W.D.; Hunt, L.A.; Wilkens, P.W.; Singh, U.; Gijsman, A.J.; Ritchie, J.T. The DSSAT cropping system model. *Eur. J. Agron.* **2003**, *18*, 235–265. [CrossRef]

16. Keating, B.A.; Carberry, P.S.; Hammer, G.L.; Probert, M.E.; Robertson, M.J.; Holzworth, D.; Huth, N.I.; Hargreaves, J.N.; Meinke, H.; Hochman, Z.; et al. An overview of APSIM, a model designed for farming systems simulation. *Eur. J. Agron.* **2003**, *18*, 267–288. [CrossRef]

17. Tao, F.; Zhang, Z.; Liu, J.; Yokozawa, M. Modelling the impacts of weather and climate variability on crop productivity over a large area: A new super-ensemble-based probabilistic projection. *Agric. For. Meteorol.* **2009**, *149*, 1266–1278. [CrossRef]

18. Van Diepen, C.V.; Wolf, J.; Van Keulen, H.; Rappoldt, C. WOFOST: A simulation model of crop production. *Soil Use Manag.* **1989**, *5*, 16–24. [CrossRef]

19. Lobell, D.B. The use of satellite data for crop yield gap analysis. *Field Crop. Res.* **2013**, *143*, 56–64. [CrossRef]

20. Aghighi, H.; Azadbakht, M.; Ashourloo, D.; Shahrabi, H.S.; Radiom, S. Machine Learning Regression Techniques for the Silage Maize Yield Prediction Using Time-Series Images of Landsat 8 OLI. *IEEE J. Sel. Top. App. Earth Obs. Remote Sens.* **2018**, *11*, 4563–4577. [CrossRef]

21. Stas, M.; Van Orshoven, J.; Dong, Q.; Heremans, S.; Zhang, B. A comparison of machine learning algorithms for regional wheat yield prediction using NDVI time series of SPOT-VGT. In Proceedings of the 2016 Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Tianjin, China, 18–20 July 2016; pp. 1–5.

22. Saeed, U.; Dempewolf, J.; Becker-Reshef, I.; Khan, A.; Ahmad, A.; Wajid, S.A. Forecasting wheat yield from weather data and MODIS NDVI using Random Forests for Punjab province, Pakistan. *Int. J. Remote Sens.* **2017**, *38*, 4831–4854. [CrossRef]

23. Johnson, M.D.; Hsieh, W.W.; Cannon, A.J.; Davidson, A.; Bédard, F. Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. *Agric. For. Meteorol.* **2016**, *218–219*, 74–84. [CrossRef]

24. Satir, O.; Berberoglu, S. Crop yield prediction under soil salinity using satellite derived vegetation indices. *Field Crop. Res.* **2016**, *192*, 134–143. [CrossRef]

25. Witten, I.H.; Frank, E.; Hall, M.A.; Pal, C.J. *Data Mining: Practical Machine Learning Tools and Techniques*; Morgan Kaufmann: Burlington, MA, USA, 2016.

26. Cai, Y.; Guan, K.; Peng, J.; Wang, S.; Seifert, C.; Wardlow, B.; Li, Z. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. *Remote Sens. Environ.* **2018**, *210*, 35–47. [CrossRef]

27.　Crane-Droesch, A. Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. *Environ. Res. Lett.* **2018**, *13*, 114003. [CrossRef]

28.　Iqbal, M.A.; Shen, Y.; Stricevic, R.; Pei, H.; Sun, H.; Amiri, E.; Penas, A.; Del Rio, S. Evaluation of the FAO AquaCrop model for winter wheat on the North China Plain under deficit irrigation from field experiment to regional yield simulation. *Agric. Water Manag.* **2014**, *135*, 61–72. [CrossRef]

29.　Chen, Y.; Zhang, Z.; Tao, F. Improving regional winter wheat yield estimation through assimilation of phenology and leaf area index from remote sensing data. *Eur. J. Agron.* **2018**, *101*, 163–173. [CrossRef]

30.　Russello, H. Convolutional Neural Networks for Crop Yield Prediction Using Satellite Images. Master's Thesis, University of Amsterdam, Amsterdam, The Netherlands, 2018.

31.　Ren, J.; Chen, Z.; Zhou, Q.; Tang, H. Regional yield estimation for winter wheat with MODIS-NDVI data in Shandong, China. *Int. J. Appl. Earth Obs. Geoinf.* **2008**, *10*, 403–413. [CrossRef]

32.　Mo, X.; Liu, S.; Lin, Z.; Xu, Y.; Xiang, Y.; McVicar, T.R. Prediction of crop yield, water consumption and water use efficiency with a SVAT-crop growth model using remotely sensed data on the North China Plain. *Ecol. Model.* **2005**, *183*, 301–322. [CrossRef]

33.　Lv, Z.; Liu, X.; Cao, W.; Zhu, Y. Climate change impacts on regional winter wheat production in main wheat production regions of China. *Agric. For. Meteorol.* **2013**, *171*, 234–248. [CrossRef]

34.　Xiao, G.; Zhang, Q.; Yao, Y.; Zhao, H.; Wang, R.; Bai, H.; Zhang, F. Impact of recent climatic change on the yield of winter wheat at low and high altitudes in semi-arid northwestern China. *Agric. Ecosyst. Environ.* **2008**, *127*, 37–42. [CrossRef]

35.　Tao, F.; Xiao, D.; Zhang, S.; Zhang, Z.; Rötter, R.P. Wheat yield benefited from increases in minimum temperature in the Huang-Huai-Hai Plain of China in the past three decades. *Agric. For. Meteorol.* **2017**, *239*, 1–14. [CrossRef]

36.　Tao, F.; Zhang, Z.; Zhang, S.; Rötter, R.P. Heat stress impacts on wheat growth and yield were reduced in the Huang-Huai-Hai Plain of China in the past three decades. *Eur. J. Agron.* **2015**, *71*, 44–52. [CrossRef]

37.　Gandhi, N.; Armstrong, L.J.; Petkar, O.; Tripathy, A.K. Rice crop yield prediction in India using support vector machines. In Proceedings of the 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE 2016), Khon Kaen, Thailand, 13–15 July 2016; pp. 1–5.

38.　Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]

39.　Mutanga, O.; Kumar, L. Google Earth Engine Applications. *Remote Sens.* **2019**, *11*, 591. [CrossRef]

40.　Jin, Z.; Azzari, G.; You, C.; Di Tommaso, S.; Aston, S.; Burke, M.; Lobell, D.B. Smallholder maize area and yield mapping at national scales with Google Earth Engine. *Remote Sens. Environ.* **2019**, *228*, 115–128. [CrossRef]

41.　Huang, J.; Wang, H.; Dai, Q.; Han, D. Analysis of NDVI data for crop identification and yield estimation. *IEEE J. Stars.* **2014**, *7*, 4374–4384. [CrossRef]

42.　Mkhabela, M.S.; Bullock, P.; Raj, S.; Wang, S.; Yang, Y. Crop yield forecasting on the Canadian Prairies using MODIS NDVI data. *Agric. For. Meteorol.* **2011**, *151*, 385–393. [CrossRef]

43.　Feng, P.; Wang, B.; Liu, D.L.; Xing, H.; Ji, F.; Macadam, I.; Ruan, H.; Yu, Q. Impacts of rainfall extremes on wheat yield in semi-arid cropping systems in eastern Australia. *Clim. Chang.* **2018**, *147*, 555–569. [CrossRef]

44.　Zhang, Z.; Chen, Y.; Wang, C.; Wang, P.; Tao, F. Future extreme temperature and its impact on rice yield in China. *Int. J. Climatol.* **2017**, *37*, 4814–4827. [CrossRef]

45.　Challinor, A.J.; Watson, J.; Lobell, D.B.; Howden, S.M.; Smith, D.R.; Chhetri, N. A meta-analysis of crop yield under climate change and adaptation. *Nat. Clim. Chang.* **2014**, *4*, 287–291. [CrossRef]

46.　Webber, H.; Ewert, F.; Olesen, J.E.; Müller, C.; Fronzek, S.; Ruane, A.C.; Bourgault, M.; Martre, P.; Ababaei, B.; Bindi, M.; et al. Diverging importance of drought stress for maize and winter wheat in Europe. *Nat. Commun.* **2018**, *9*, 4249. [CrossRef] [PubMed]

47.　Ummenhofer, C.C.; Xu, H.; Twine, T.E.; Girvetz, E.H.; McCarthy, H.R.; Chhetri, N.; Nicholas, K.A. How Climate Change Affects Extremes in Maize and Wheat Yield in Two Cropping Regions. *J. Clim.* **2015**, *28*, 4653–4687. [CrossRef]

48.　Abatzoglou, J.T.; Dobrowski, S.Z.; Parks, S.A.; Hegewisch, K.C. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Sci. Data* **2018**, *5*, 170191. [CrossRef] [PubMed]

49. Khataar, M.; Mohammadi, M.H.; Shabani, F. Soil salinity and matric potential interaction on water use, water use efficiency and yield response factor of bean and wheat. *Sci. Rep.* **2018**, *8*, 2679. [CrossRef] [PubMed]

50. He, G.; Wang, Z.; Li, F.; Dai, J.; Li, Q.; Xue, C.; Cao, H.; Wang, S.; Malhi, S.S. Soil water storage and winter wheat productivity affected by soil surface management and precipitation in dryland of the Loess Plateau, China. *Agric. Water Manag.* **2016**, *171*, 1–9. [CrossRef]

51. Li, H.; Xue, J.; Gao, Z.; Xue, N.; Yang, Z. Response of yield increase for dryland winter wheat to tillage practice during summer fallow and sowing method in the Loess Plateau of China. *J. Integr. Agric.* **2018**, *17*, 817–825. [CrossRef]

52. Fischer, G.; Nachtergaele, F.; Prieler, S.; Van Velthuizen, H.T.; Verelst, L.; Wiberg, D. *Global Agro-Ecological Zones Assessment for Agriculture (GAEZ 2008)*; IIASA: Laxenburg, Austria; FAO: Rome, Italy, 2008.

53. Tao, F.; Zhang, Z.; Zhang, S.; Zhu, Z.; Shi, W. Response of crop yields to climate trends since 1980 in China. *Clim. Res.* **2012**, *54*, 233–247. [CrossRef]

54. Chen, Y.; Zhang, Z.; Tao, F.; Wang, P.; Wei, X. Spatio-temporal patterns of winter wheat yield potential and yield gap during the past three decades in North China. *Field Crop. Res.* **2017**, *206*, 11–20. [CrossRef]

55. Luo, Y.; Zhang, Z.; Chen, Y.; Li, Z.; Tao, F. ChinaCropPhen1km: A high-resolution crop phenological dataset for three staple crops in China during 2000–2015 based on LAI products. *Earth Syst. Sci. Data Discuss.* **2019**, *2019*. [CrossRef]

56. Sakamoto, T.; Yokozawa, M.; Toritani, H.; Shibayama, M.; Ishitsuka, N.; Ohno, H. A crop phenology detection method using time-series MODIS data. *Remote Sens. Environ.* **2005**, *96*, 366–374. [CrossRef]

57. Appelhans, T.; Mwangomo, E.; Hardy, D.R.; Hemp, A.; Nauss, T. Evaluating machine learning approaches for the interpolation of monthly air temperature at Mt. Kilimanjaro, Tanzania. *Spat. Stat.* **2015**, *14*, 91–113. [CrossRef]

58. Aha, D.W.; Kibler, D.; Albert, M.K. Instance-based learning algorithms. *Mach. Learn.* **1991**, *6*, 37–66. [CrossRef]

59. Wang, B.; Gu, X.; Ma, L.; Yan, S. Temperature error correction based on BP neural network in meteorological wireless sensor network. In Proceedings of the International Conference on Cloud Computing and Security, Nanjing, China, 29–31 July 2016; Springer: Cham, Switzerland, 2016; pp. 117–132.

60. Bélisle, E.; Huang, Z.; Le Digabel, S.; Gheribi, A.E. Evaluation of machine learning interpolation techniques for prediction of physical properties. *Comput. Mater. Sci.* **2015**, *98*, 170–177. [CrossRef]

61. Xu, M.; Watanachaturaporn, P.; Varshney, P.; Arora, M. Decision tree regression for soft classification of remote sensing data. *Remote Sens. Environ.* **2005**, *97*, 322–336. [CrossRef]

62. Song, Y.Y.; Lu, Y. Decision tree methods: Applications for classification and prediction. *Shanghai Arch. Psychiatry* **2015**, *27*, 130–135. [PubMed]

63. Polat, K.; Güneş, S. A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach for multi-class classification problems. *Expert Syst. Appl.* **2009**, *36*, 1587–1592. [CrossRef]

64. Gunn, S.R. Support vector machines for classification and regression. *ISIS Tech. Rep.* **1998**, *14*, 5–16.

65. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; Scholkopf, B. Support vector machines. *IEEE Intell. Syst. Appl.* **1998**, *13*, 18–28. [CrossRef]

66. Rasmussen, C.E. Gaussian processes in machine learning. In *Summer School on Machine Learning*; Springer: Cambridge, MA, USA, 2003; pp. 63–71.

67. Williams, C.K.; Rasmussen, C.E. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006.

68. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

69. Rhee, J.; Im, J. Meteorological drought forecasting for ungauged areas based on machine learning: Using long-range climate forecast and remote sensing data. *Agric. For. Meteorol.* **2017**, *237–238*, 105–122. [CrossRef]

70. Vincenzi, S.; Zucchetta, M.; Franzoi, P.; Pellizzato, M.; Pranovi, F.; De Leo, G.A.; Torricelli, P. Application of a Random Forest algorithm to predict spatial distribution of the potential yield of Ruditapes philippinarum in the Venice lagoon, Italy. *Ecol. Model.* **2011**, *222*, 1471–1478. [CrossRef]

71. Strobl, C.; Boulesteix, A.L.; Zeileis, A.; Hothorn, T. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinform.* **2007**, *8*, 25. [CrossRef]

72. Dietterich, T.G. Ensemble Methods in Machine Learning. In Proceedings of the International Workshop on Multiple Classifier Systems, Cagliari, Italy, 21–23 June 2000; Springer: Berlin/Heidelberg, Germany, 2000; pp. 1–15.

73. Voyant, C.; Notton, G.; Kalogirou, S.; Nivet, M.; Paoli, C.; Motte, F.; Fouilloy, A. Machine learning methods for solar radiation forecasting: A review. *Renew. Energy* **2017**, *105*, 569–582. [CrossRef]

74. Quinlan, J.R. Bagging, Boosting, and C4. 5. In Proceedings of the Thirteenth National Conference on Artificial Intelligence, Portland, OR, USA, 4—8 August 1996; Volume 1, pp. 725–730.

75. Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. *Stat. Surv.* **2010**, *4*, 40–79. [CrossRef]

76. Picard, R.R.; Cook, R.D. Cross-validation of regression models. *J. Am. Stat. Assoc.* **1984**, *79*, 575–583. [CrossRef]

77. Johnson, D.M. An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. *Remote Sens. Environ.* **2014**, *141*, 116–128. [CrossRef]

78. Labus, M.P.; Nielsen, G.A.; Lawrence, R.L.; Engel, R.; Long, D.S. Wheat yield estimates using multi-temporal NDVI satellite imagery. *Int. J. Remote Sens.* **2002**, *23*, 4169–4180. [CrossRef]

79. Doraiswamy, P.C.; Cook, P.W. Spring wheat yield assessment using NOAA AVHRR data. *Can. J. Remote Sens.* **1995**, *21*, 43–51. [CrossRef]

80. Tsimba, R.; Edmeades, G.O.; Millner, J.P.; Kemp, P.D. The effect of planting date on maize grain yields and yield components. *Field Crop. Res.* **2013**, *150*, 135–144. [CrossRef]

81. Tsimba, R.; Edmeades, G.O.; Millner, J.P.; Kemp, P.D. The effect of planting date on maize: Phenology, thermal time durations and growth rates in a cool temperate climate. *Field Crop. Res.* **2013**, *150*, 145–155. [CrossRef]

82. Bolton, D.K.; Friedl, M.A. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agric. For. Meteorol.* **2013**, *173*, 74–84. [CrossRef]

83. Jeong, J.H.; Resop, J.P.; Mueller, N.D.; Fleisher, D.H.; Yun, K.; Butler, E.E.; Timlin, D.J.; Shim, K.; Gerber, J.S.; Reddy, V.R.; et al. Random Forests for Global and Regional Crop Yield Predictions. *PLoS ONE* **2016**, *11*, e156571. [CrossRef] [PubMed]

84. Kim, N.; Lee, Y. Machine Learning Approaches to Corn Yield Estimation Using Satellite Images and Climate Data: A Case of Iowa State. *J. Korean Soc. Surv. Geod. Photogramm. Cartogr.* **2016**, *34*, 383–390. [CrossRef]

85. Zhao, Y.; Lobell, D.B. Assessing the heterogeneity and persistence of farmers' maize yield performance across the North China Plain. *Field Crop. Res.* **2017**, *205*, 55–66. [CrossRef]

86. Ji, B.; Sun, Y.; Yang, S.; Wan, J. Artificial neural networks for rice yield prediction in mountainous regions. *J. Agric. Sci.* **2007**, *145*, 249. [CrossRef]

87. Lai, Y.R.; Pringle, M.J.; Kopittke, P.M.; Menzies, N.W.; Orton, T.G.; Dang, Y.P. An empirical model for prediction of wheat yield, using time-integrated Landsat NDVI. *Int. J. Appl. Earth Obs.* **2018**, *72*, 99–108. [CrossRef]

88. Son, N.T.; Chen, C.F.; Chen, C.R.; Minh, V.Q.; Trung, N.H. A comparative analysis of multitemporal MODIS EVI and NDVI data for large-scale rice yield estimation. *Agric. For. Meteorol.* **2014**, *197*, 52–64. [CrossRef]

89. Shanahan, J.F.; Schepers, J.S.; Francis, D.D.; Varvel, G.E.; Wilhelm, W.W.; Tringe, J.M.; Schlemmer, M.R.; Major, D.J. Use of remote-sensing imagery to estimate corn grain yield. *Agron. J.* **2001**, *93*, 583–589. [CrossRef]

90. Fontana, D.C.; Potgieter, A.B.; Apan, A. Assessing the relationship between shire winter crop yield and seasonal variability of the MODIS NDVI and EVI images. *Appl. GIS* **2007**, *3*, 1–16.

91. Hatfield, J.L. Remote sensing estimators of potential and actual crop yield. *Remote Sens. Environ.* **1983**, *13*, 301–311. [CrossRef]

92. Chen, Y.; Zhang, Z.; Tao, F.; Palosuo, T.; Rötter, R.P. Impacts of heat stress on leaf area index and growth duration of winter wheat in the North China Plain. *Field Crop. Res.* **2018**, *222*, 230–237. [CrossRef]

93. Slafer, G.A.; Savin, R. Developmental Base Temperature in Different Phenological Phases of Wheat (Triticum aestivum). *J. Exp. Bot.* **1991**, *42*, 1077–1082. [CrossRef]

94. Garg, D.; Sareen, S.; Dalal, S.; Tiwari, R.; Singh, R. Grain filling duration and temperature pattern influence on the performance of wheat genotypes under late planting. *Cereal Res. Commun.* **2013**, *41*, 500–507. [CrossRef]

95. Lobell, D.B.; Sibley, A.; Ivan Ortiz-Monasterio, J. Extreme heat effects on wheat senescence in India. *Nat. Clim. Chang.* **2012**, *2*, 186–189. [CrossRef]

96. Biscoe, P.V.; Gallagher, J.N. Weather, Dry Matter Production and Yield. In Proceedings of the Environmental Effects on Crop Physiology, a Symposium Held at Long Ashton Research Station, University of Bristol, Bristol, UK, 13–16 April 1975; Academic Press: New York, NY, USA, 1977.

97. Semenov, M.A.; Shewry, P.R. Modelling predicts that heat stress, not drought, will increase vulnerability of wheat in Europe. *Sci. Rep.* **2011**, *1*, 66. [CrossRef]

98. Kern, A.; Barcza, Z.; Marjanović, H.; Árendás, T.; Fodor, N.; Bónis, P.; Bognár, P.; Lichtenberger, J. Statistical modelling of crop yield in Central Europe using climate data and remote sensing vegetation indices. *Agric. For. Meteorol.* **2018**, *260–261*, 300–320. [CrossRef]

99. Webber, H.; Ewert, F.; Kimball, B.A.; Siebert, S.; White, J.W.; Wall, G.W.; Ottman, M.J.; Trawally, D.; Gaiser, T. Simulating canopy temperature for modelling heat stress in cereals. *Environ. Model. Softw.* **2016**, *77*, 143–155. [CrossRef]

100. Siebert, S.; Webber, H.; Rezaei, E.E. Weather impacts on crop yields-searching for simple answers to a complex problem. *Environ. Res. Lett.* **2017**, *12*, 81001. [CrossRef]

101. Porter, J.R.; Gawith, M. Temperatures and the growth and development of wheat: A review. *Eur. J. Agron.* **1999**, *10*, 23–36. [CrossRef]

102. Zhao, H.; Dai, T.; Jing, Q.; Jiang, D.; Cao, W. Leaf senescence and grain filling affected by post-anthesis high temperatures in two different wheat cultivars. *Plant Growth Regul.* **2007**, *51*, 149–158. [CrossRef]

103. Eitzinger, J.; Štastná, M.; Žalud, Z.; Dubrovský, M. A simulation study of the effect of soil water balance and water stress on winter wheat production under different climate change scenarios. *Agric. Water Manag.* **2003**, *61*, 195–217. [CrossRef]

104. Carew, R.; Smith, E.G.; Grant, C. Factors Influencing Wheat Yield and Variability: Evidence from Manitoba, Canada. *J. Agric. Appl. Econ.* **2009**, *41*, 625–639. [CrossRef]

105. Siebert, S.; Ewert, F. Future crop production threatened by extreme heat. *Environ. Res. Lett.* **2014**, *9*, 41001. [CrossRef]

106. Su, Y.; Xu, H.; Yan, L. Support vector machine-based open crop model (SBOCM): Case of rice production in China. *Saudi J. Biol. Sci.* **2017**, *24*, 537–547. [CrossRef]

107. Tack, J.; Barkley, A.; Nalley, L.L. Effect of warming temperatures on US wheat yields. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 6931–6936. [CrossRef] [PubMed]

108. Tao, J.; Wu, W.; Yong, Z.; Yu, W.; Jiang, Y. Mapping winter wheat using phenological feature of peak before winter on the North China Plain based on time-series MODIS data. *J. Integr. Agric.* **2017**, *16*, 348–359. [CrossRef]

109. Zhou, G.; Liu, X.; Liu, M. Assimilating Remote Sensing Phenological Information into the WOFOST Model for Rice Growth Simulation. *Remote Sens.* **2019**, *11*, 268. [CrossRef]

110. Sokoto, M.B.; Abubakar, I.U.; Dikko, A.U. Correlation analysis of some growth, yield, yield components and grain quality of wheat (*Triticum aestivum* L.). *Niger. J. Basic Appl. Sci.* **2012**, *20*, 349–356.

111. Zhang, X.; Friedl, M.A.; Schaaf, C.B.; Strahler, A.H.; Hodges, J.C.; Gao, F.; Reed, B.C.; Huete, A. Monitoring vegetation phenology using MODIS. *Remote Sens. Environ.* **2003**, *84*, 471–475. [CrossRef]

112. Motohka, T.; Nasahara, K.N.; Oguma, H.; Tsuchida, S. Applicability of green-red vegetation index for remote sensing of vegetation phenology. *Remote Sens.* **2010**, *2*, 2369–2387. [CrossRef]

113. Mutanga, O.; Skidmore, A.K. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *Int. J. Remote Sens.* **2004**, *25*, 3999–4014. [CrossRef]

114. Santin-Janin, H.; Garel, M.; Chapuis, J.; Pontier, D. Assessing the performance of NDVI as a proxy for plant biomass using non-linear models: A case study on the Kerguelen archipelago. *Polar Biol.* **2009**, *32*, 861–871. [CrossRef]

115. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]

116. Wang, C.; Lin, W. Winter wheat yield estimation based on MODIS EVI. *Nongye Gongcheng Xuebao/Trans. Chin. Soc. Agric. Eng.* **2005**, *21*, 90–94.

117. Kouadio, L.; Newlands, N.; Davidson, A.; Zhang, Y.; Chipanshi, A. Assessing the Performance of MODIS NDVI and EVI for Seasonal Crop Yield Forecasting at the Ecodistrict Scale. *Remote Sens.* **2014**, *6*, 10193–10214. [CrossRef]

118. Taylor, J.A.; McBratney, A.B.; Whelan, B.M. Establishing Management Classes for Broadacre Agricultural Production. *Agron. J.* **2007**, *99*, 1366–1376. [CrossRef]

119. Battude, M.; Al Bitar, A.; Morin, D.; Cros, J.; Huc, M.; Sicre, C.M.; Le Dantec, V.; Demarez, V. Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 like remote sensing data. *Remote Sens. Environ.* **2016**, *184*, 668–681. [CrossRef]

120. Vergara-Díaz, O.; Zaman-Allah, M.A.; Masuka, B.; Hornero, A.; Zarco-Tejada, P.; Prasanna, B.M.; Cairns, J.E.; Araus, J.L. A novel remote sensing approach for prediction of maize yield under different conditions of nitrogen fertilization. *Front. Plant Sci.* **2016**, *7*, 666. [CrossRef]

121. Jin, Z.; Azzari, G.; Burke, M.; Aston, S.; Lobell, D. Mapping Smallholder Yield Heterogeneity at Multiple Scales in Eastern Africa. *Remote Sens.* **2017**, *9*, 931. [CrossRef]

122. Vereecken, H.; Weihermüller, L.; Jonard, F.; Montzka, C. Characterization of crop canopies and water stress related phenomena using microwave remote sensing methods: A review. *Vadose Zone J.* **2012**, *11*. [CrossRef]

123. Murdoch, W.J.; Singh, C.; Kumbier, K.; Abbasi-Asl, R.; Yu, B. Definitions, methods, and applications in interpretable machine learning. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 22071–22080. [CrossRef] [PubMed]