

Article

Not peer-reviewed version

Leveraging Remotely Sensed and Climatic Data for Improved Crop Yield Prediction in the Chi Basin, Thailand

[Akkarapon Chaiyana](#) , Ratchawatch Hanchoo Wong , [Neti Srihanu](#) , [Haris Prasanchum](#) , Anongrit Kangrang , [Rattana Hormwichian](#) , [Siwa Kaewplang](#) * , [Werapong Koedsin](#) , [Alfredo Huete](#)

Posted Date: 2 February 2024

doi: 10.20944/preprints202402.0166.v1

Keywords: Decision support tools; machine learning; remote sensing, climatic data, predictive model; province scale



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

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.

Article

Leveraging Remotely Sensed and Climatic Data for Improved Crop Yield Prediction in the Chi Basin, Thailand

Akkarapon Chaiyana ¹, Ratchawatch Hanchoo Wong ², Neti Srihanu ³, Haris Prasanchum ⁴, Anongrit Kangrang ¹, Rattana Hormwichian ¹, Siwa Kaewplang ^{1,*}, Werapong Koedsin ⁵, and Alfredo Huete ⁶

¹ Department of Civil Engineering, Faculty of Engineering, Maha Sarakham University, Kantharawichai District, Maha Sarakham 44150, Thailand.

² Department of Civil Engineering, School of Engineering and Industrial Technology, Mahanakorn University of Technology, Bangkok 10530, Thailand.

³ Faculty of Engineering, Northeastern University, Muang District, Khon Kaen 40000, Thailand.

⁴ Faculty of Engineering, Rajamangala University of Technology Isan, Khon Kaen Campus, Khon Kaen, Thailand 40000, Thailand.

⁵ Faculty of Technology and Environment, Phuket Campus, Prince of Songkla University, Phuket 83120, Thailand.

⁶ School of Life Sciences, University of Technology Sydney, Sydney, NSW 2007, Australia.

* Correspondence: siwa.kae@msu.ac.th;

Abstract: Predictions of crop production in the Chi basin are of major importance for decision support tools in countries such as Thailand, which aim to increase domestic income and global food security by implementing the appropriate policies. This research aims to establish a predictive model for predicting crop production for an internal crop growth season prior of harvest at the province scale for fourteen provinces in Thailand's Chi basin between 2011 and 2019. We provide approaches for reducing redundant variables and multicollinearity in remotely sensed (RS) and meteorological data to avoid overfitting models using correlation analysis (CA) and variance inflation factor (VIF). Temperature condition index (TCI), normalized difference vegetation index (NDVI), land surface temperature (LST_{night}), and mean temperature (T_{mean}) were the resulting variables in the prediction model with a p-value < 0.05 and a VIF < 5. The baseline data (2011–2017: June to November) were used to train four regression models, which revealed that eXtreme Gradient Boosting (XGBoost), random forest (RF), and XGBoost achieved R² values of 0.95, 0.94, and 0.93, respectively. In addition, the testing dataset (2018–2019) displayed a minimum root mean square error (RMSE) of 0.18 ton/ha for the optimal solution by integrating variables and applying the XGBoost model. Accordingly, it is estimated that between 2020 and 2022, the total crop production in the Chi basin region would be 7.88, 7.64, and 7.72 million tons, respectively. The results demonstrated that the proposed model is proficient at greatly improving crop yield prediction accuracy when compared to a conventional regression method and that it may be deployed in different regions to assist farmers and policymakers in making more informed decisions about agricultural practices and resource allocation.

Keywords: decision support tools; machine learning; remote sensing, climatic data, predictive model; province scale

1. Introduction

Rice is a major agricultural commodity in Thailand and an important contributor to the country's economy. According to data from the United Nations Food and Agriculture Organization (FAO), Thailand was the world's second-largest exporter of rice in 2018, with exports valued at around \$7.4 billion [1]. Thailand is the world's biggest producer of milled rice, producing approximately 20.3 million metric tons in 2018, equivalent to approximately 17.6 million metric tons of paddy rice [1]. The main rice-producing regions in Thailand are the central, northeastern, and northern regions, with the central region accounting for the largest share of production [2]. The rice sector is a significant

contributor to Thailand's GDP, accounting for approximately 3.3% in 2018 [3]. Rice farming is also a vital source of business and income for many smallholder farmers in Thailand, with the sector employing around 10 million people, or approximately 20% of the country's total workforce [2] and can also account for a significant portion of income of smallholder farmers, with some estimates suggesting that rice can make up to 60% of their income [3]. As a result, rice yield has become an important variable for maximizing the efficiency of rice production and fulfilling the increasing demand for rice, especially as the world's population grows. Yet many factors can affect rice productivity, including environmental factors, physical factors, and farmer quantities. Thailand has faced with the previously stated factors for many decades. Currently, land use change and climate change are the major concern to every sector, particularly developing countries. Land use change and climate change are major drivers of crop yield variations, with both expected to have significant impacts on agricultural productivity [4]. Climate change, through warming temperatures, extreme weather events, and altered precipitation patterns, can lead to yield reductions. Land use transformation, including the conversion of agricultural to urban or industrial use, can also influence crop yields by altering the availability of land and resources for agriculture [5].

Moreover, natural disasters such as droughts and floods can significantly affect rice yield and production by causing damage to crops, disrupting the growing season, and reducing overall yield. These hazards may result in complete crop failures or have a more limited impact, depending on the severity of the incident and the vulnerability of the disturbed region. Rice production is concentrated in some parts of the world, such as Thailand, which may be more vulnerable to natural disasters due to its position and environment. For example, seven typhoons in 2021 caused flooding in Thailand, which may wreak havoc on rice crop production in 0.85 million hectares of agricultural areas and result in farmers losing around 220 million U.S. dollars or 30% of productivity [6].

Additionally, drought is a common occurrence in Thailand, which has a tropical climate and is prone to dry spells and water shortages. According to the report, Thailand suffered from long-term drought conditions that affected to approximately 3.8 million hectares of the whole country in 2021, and it is expected to increase and become more severe every year [7]. Nevertheless, every factor that impacts rice production can affect directly to rice growth phases, for example, reduced leaf area index (LAI), deformation of the leaf, little growth, green to pale-colored leaves, dwarf, and lesions on the leaves.

Crop yield measurement in massive agricultural areas is difficult to verify under current circumstances such as time, budgetary, and surveyor constraints. Recently, a data-driven remote sensing approach has become efficient to measure crop conditions and predict crop yield production from a distance without being physically present in the study area. This can be done by using various sensors and platforms, including satellites, which can collect data on various aspects of the surface of the Earth, including land usage, vegetation, and weather patterns. There are many studies that utilized remote sensing data to forecast agricultural crop production [8–12]. The weather factors have long been used to explain crop yield fluctuations. For instance, the study [13] applied machine learning (ML) with land surface temperature (LST), enhance vegetation index (EVI), and normalized difference vegetation index (NDVI) from MODIS satellite and weather variables to improve soybean yield forecasts with a mean absolute error of around 0.24 to 0.42 Mg/ha. This study [14] employed LST and air temperature to foresee corn outcomes across the US with an r-square of 0.56 to 0.65. In addition, [15] indicated that the eXtreme Gradient Boosting (XGBoost) Machine Learning (ML) method exhibited the best metrics, which can reduce the predicting errors of cereal yield by combining remote sensing data and weather data in Morocco. Besides, several studies used drought and health indices that were obtained from computed indices of remotely sensed data, such as the vegetation health index [16], the temperature condition index (TCI) [17], and the vegetation condition index (VCI) [18,19]. It outperforms the use of health and drought indicators to predict crop production when combined with machine learning technologies [15].

Accurate and up-to-date prediction of crop yields is essential for sustainable food security and agriculture because it helps farmers by providing decision support systems about planting and harvesting and enables policymakers to plan for and address potential food shortages. The

conventional regression approaches have been overcome by ML and deep learning to provide precise and accurate statistical predictions [20,21]. Several studies have recently observed the statistical metrics of ML algorithms, for instance, support vector regression (SVR) [22], random forest (RF) regression [23], and XGBoost regression [24], to predict crop production at local (i.e., province) scales. This study [8] investigated eight different ML classifiers and regressors to forecast the outcome of wheat in the winter season for China. The result indicated that SVR, RF, and gaussian process regression (GPR) denote the top three of the greatest methods for prediction, amongst others with r -square > 0.75 . ML approaches are popular and outperform results when applied to crop yield prediction in many aspects, but there is evidence that the multivariate ordinary least squares approach can provide a lower error rate of soybean yield prediction than RF and long-short-term memory (LSTM) [13]. Then, linear regression and ML regression have been compared [25]. Moreover, hyperparameter tuning of ML models is complicated to adjust, so Gridsearch Cross Validation (CV) has been developed [26] to apply to crop yield prediction. However, a number of studies have attempted to forecast agricultural yield at the regional level using remote sensing data without taking meteorological information into account. These are the primary elements that have a significant impact on crop yield. For instance, this study [27] found that the root mean square error (RMSE), which is based on remote sensing data, ranged from 14% to 49%. This study also [28] illustrated how remote sensing data could be used to predict wheat yields in Australia. According to the findings, the RMSE varies depending on the research locations and is between 0.07 and 0.25 t/ha⁻¹. It is unclear if using solely remote sensing data or combining it with climatic data can produce accurate results especially tropical areas. As a result, the goal of this work is to demonstrate and offer not just input datasets but also model methods that can minimize crop production forecast error.

The objective of this study was to test the capability of MLR models and machine learning (RF, XGBoost, and SVR) to predict crop yields. The model uses several variables, including various indices derived from satellite images and climate variables. Before performing the models, the variable selection process will be carried out. The associated variables with crop yield in the Chi Basin area will be reported. The differences in the combination of predictor variables were performed. All models will be used to predict crop yields at the provincial scale. The R² and RMSE of all models were analyzed. The trends and comparison with the testing data will be performed. The selected model will be used to predict crop yields in future years.

2. Materials and Methods

2.1. Study area

The Chi Basin is a region situated in central Thailand, positioned between 15°13' and 17°40' N latitude and 101°14' and 104°46' E longitude, and ranging in altitude from 104-1060 meters above mean sea level. The study area covers approximately 4.91 million hectares, with approximately 3.22 million hectares of cropland (<https://esa-worldcover.org/en/data-access>). The climate of the Chi Basin is characterized by humid and hot conditions, with average temperatures ranging from 27-32°C. The region experiences two monsoon seasons: the Southwest Monsoon, which brings wet and rainy conditions from May to October, and the Northeast Monsoon, which brings dry and cool conditions from November to April. The rainy season in the Chi Basin typically lasts from May to October, with an average annual rainfall of 1380 mm. Crop cultivation in the region typically occurs from June to November, with harvest occurring in December [29].

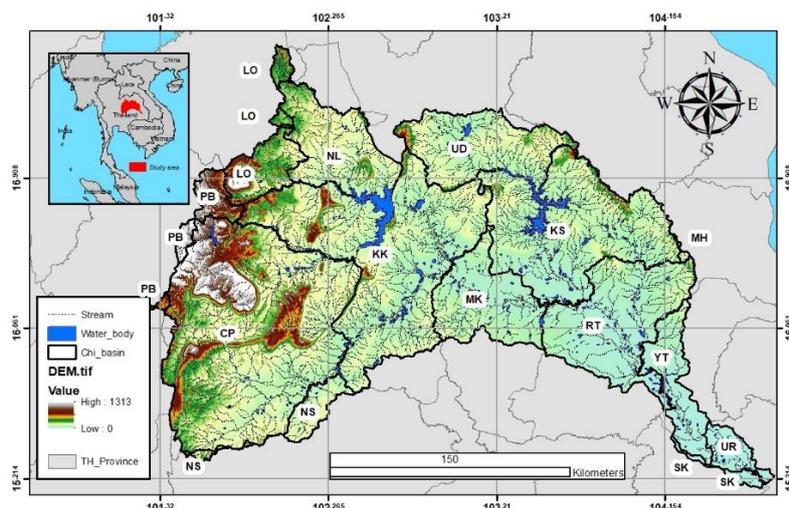


Figure 1. Study area.

2.2. Crop yield data and their phenology

In this study, historical crop yield at the administrative provincial scale is derived between 2011 and 2019 from the Office of Agricultural Economics (OAE) for fourteen provinces (<https://www.oae.go.th/>), which are described in **Error! Reference source not found..** The annual crop production in this study was calculated as the ratio of total crop production divided by harvested area. The annual crop yield production ranges from 1.97 to 4.4 tons/ha, depending on area.

Table 1. List name of fourteen study areas of Chi basin (ton/ha).

No	Province	Acronym	2011	2012	2013	2014	2015	2016	2017	2018	2019
1	NAKHON RATCHASIMA	NS	2.59	2.25	2.31	2.24	2.26	2.22	2.26	2.22	2.27
2	SISAKET	SK	2.51	2.30	2.45	2.28	2.26	2.27	2.29	2.28	2.17
3	UBON RATCHATHANI	UR	2.16	2.15	2.15	2.06	2.06	2.09	2.18	2.27	2.25
4	YASOTHON	YT	2.54	2.27	2.28	2.31	2.21	2.23	2.22	2.27	2.25
5	CHAIYAPHUM	CP	2.46	2.36	2.39	2.21	2.19	2.25	2.32	2.29	2.32
6	NONG BUA LAMPHU	NL	2.41	2.33	1.97	2.01	1.98	2.11	2.16	2.07	2.08
7	KHON KAEN	KK	2.16	2.09	2.11	2.12	2.11	2.15	2.14	2.02	1.98
8	UDON THANI	UD	2.47	2.32	2.24	2.32	2.34	2.37	2.40	2.28	2.23
9	LOEI	LO	2.41	2.42	2.46	2.34	2.31	2.43	2.46	2.33	2.11
10	MAHA SARAKHAM	MK	2.37	2.32	2.33	2.30	2.28	2.30	2.23	2.18	2.25
11	ROI ET	RT	2.37	2.32	2.33	2.34	2.38	2.39	2.37	2.21	2.15
12	KALASIN	KS	2.32	2.26	2.26	2.29	2.30	2.32	2.30	2.31	2.33
13	MUKDAHAN	MH	2.40	2.24	2.26	2.40	2.40	2.40	2.38	2.47	2.19
14	PHETCHABUN	PB	3.54	3.54	3.61	4.37	4.36	3.46	4.33	3.53	4.40

According to the crop calendar period [30], crop transplanted in this region is usually planted around June and July, and flower around late October to November and, harvest around December. In Thailand, rice grows through several stages, starting with the planting of seedlings and ending with the harvest of mature grains. It can be summarized as four stages: nursery stage, vegetative growth stage, reproductive growth stage, and maturity stage. These all stages take about 3–4 months [31], depending on the environmental condition and variety of crop types (**Figure 2**). As the rice plant

progresses through different stages of growth, its reflectance is impacted at various wavelengths. Studies have revealed that indices used for vegetation, for example, the Normalized Difference Vegetation Index (NDVI), can be used to accurately track rice growth. During the early vegetative stage, the NDVI is typically low due to the low percentage of vegetation cover. As the plant continues to grow and the chlorophyll content increases, the absorbance of light in the red and blue regions also increases [32]. The reflectance in the Near-Infrared (NIR) region increases with the development of foliage and tillers. As the plant reaches maturity, the NDVI begins to decrease due to a reduction in biomass, a decrease in chlorophyll content, and an increase in grain filling [33].

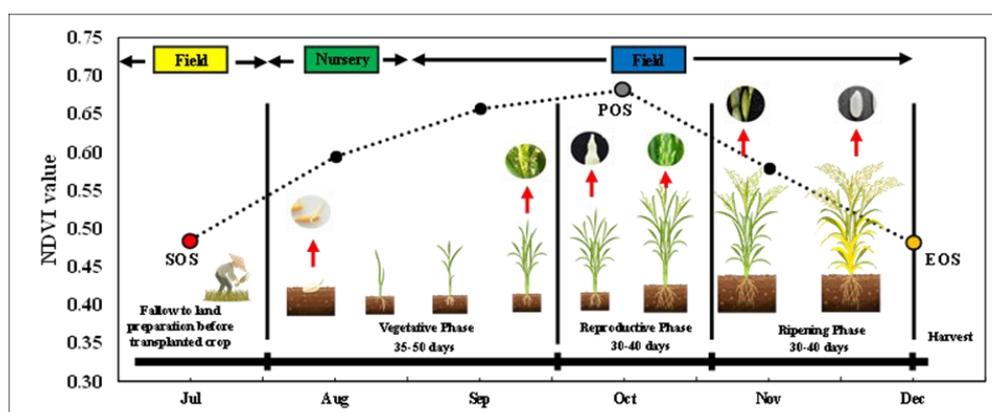


Figure 2. Mean temporal crop NDVI profile over crop growth stages in Chi basin region between 2011 and 2019; SOS, POS and EOS refer to starting, peak and end of season, respectively.

2.3. Remotely sensed data and climate data

The Remote Sensing (RS) data for the proposed study originated from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor. The remotely sensed MODIS data are applied to delimit the spatial extent of the crop area in Chi Basin, Thailand. The crop land use masked in this study was derived from the land use data of the Land Development Department (LDD) of Thailand in 2020. As the vegetation indices are usually used for vegetation tracking and monitoring. The Enhanced vegetation index (EVI), NDVI, and Land Surface Temperature (LST) daytime and nighttime products of MODIS data were used in this study. As mentioned above, the disaster (drought) and climate were the factors that affected crop productivity. Therefore, drought and healthy conditions, Temperature condition index (TCI), Vegetation Condition Index (VCI), and Vegetation Health Index (VHI) were also applied in this study [34–36], which can be calculated from NDVI and temperature [37]. In addition, all remote sensing datasets will be aggregated to monthly mean data. The major climatic factors used in this study were the monthly mean values of rainfall, minimum temperature (Tmin), mean temperature (Tmean), and maximum temperature (Tmax) throughout the crop growth period. The variables (both predictors and response) used in this study can be summarized as shown in **Table 2**.

Table 2. The predictors to be applied for modeling and predicting crop production.

Data type	Product	Variable	Spatial resolution	Temporal resolution	Acquisition date	Data source
Yield recorded		Crop yield	Provincial level	Annual	2011-2019	https://www.oae.go.th/
RS data	MOD 13Q1	NDVI EVI	250 m	16-day interval	2011-2022 2011-2022	https://lpdaac.usgs.gov/products/mod13q1v006/

		LST				
		dayti			2011-2022	
	MOD	me	1 km	8-day		https://lpdaac.usgs.gov/products/mod11a1v006/
	11A2	LST		interval		
		nightt			2011-2022	
		ime				
		Rainf			2011-2022	
		all				https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5
Climatic data	ERA5	Tmean	27.83 km	Monthly	2011-2022	
		Tmin			2011-2022	
		Tmax			2011-2022	

¹ All RS and climatic data were averaged into the crop growth season from June to November.

2.4. Features Selection: Correlation Analysis (CA) and Variance Inflation Factor

The main problem with multiple linear regression is the multicollinear problem, in which some variables are highly correlated together. For the purpose of variable selection, correlation analysis (CA) was used to analyze the correlation between variables. It helps to determine if there is a correlation, or association, between the two variables, along with the intensity and direction of the relationship. Several studies have applied the CA to reduce redundant variables by removing highly co-related variables [38–40]. The variance inflation factor (VIF) is a tool used in multiple regression analysis to assess the degree of multicollinearity between independent factors. When two or more predictor variables are extremely related, multicollinearity occurs, which can lead to unstable and unreliable regression coefficient estimates [39-40]. There are studies that attempt to integrate VIF as indicator to reduce multicollinearity (VIF < 5-10) [41-42]. The VIF score threshold applied in this study was 5, for considerable to moderate correlation. This study applied both statistical methods (i.e, CA and VIF) to analyze the influent factors in crop yield prediction by determining a p-value < 0.05 and a VIF < 5 [45] (Figure 3). The selected variables in this step will be used as the predictor variables for the next step.

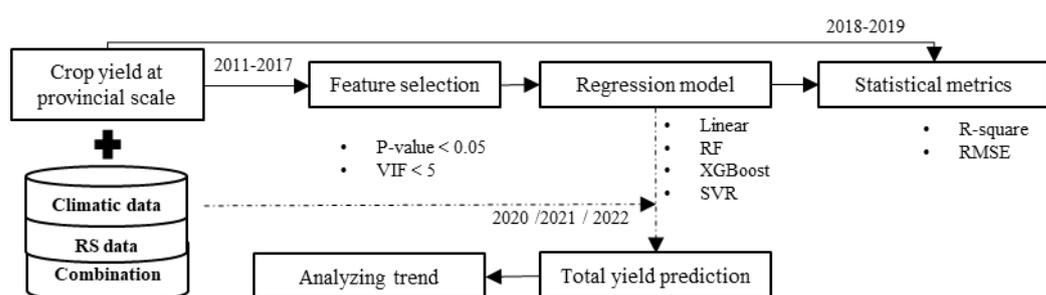


Figure 3. Research framework for predicting crop production.

2.5. Regression Model

The selected variables in previous step were used as the input variable in the regression model. This study utilized four regression models: multiple linear regression (MLR), random forest (RF) regression, XGBoost regression, and the Support Vector Regression (SVR) model. The crop yield dataset (126 samples) was separated into a training dataset (98 samples) (2011–2017) and a testing dataset (28 samples) (2018–2019) (Figure 3). Since machine learning approaches require optimization methods to deal with hyperparameters, Grid Search Cross-Validation (GridsearchCV) was utilized to choose the appropriate hyperparameters for each ML model. Since machine learning approaches require optimization methods to deal with hyperparameters, grid search cross-validation

(GridsearchCV) from the scikit-learn library in Python was utilized to optimize the hyperparameters of a machine learning model [46]. GridSearchCV exhaustively searches through a specified hyperparameter grid to find the optimal combination of hyperparameters for a given model by training and evaluating the model with different combinations and selecting the combination that performs best according to a chosen evaluation metric [47]. The root mean square error (RMSE) and r squared (R^2) were used to evaluate the performance of a regression model. The regression model was performed in a Python environment using the Scikit-Learn library. The reliable predictive model will be applied for predicting crop yield at the provincial scale in fourteen provinces and visualized as a map prediction. On the other hand, analyzing trends for future periods is required for farmers and policymakers to make more informed decisions regarding agricultural practices and resource allocation.

3. Results

3.1. Variables selection

Eliminating redundant variables (keeping the significant variables) is required to reduce misleading and avoid overfitting models. This study applied CA and VIF as an initial step to remove redundant variables and utilized the remaining crucial variables in the model for predicting crop yields in the Chi basin region. The result indicated that RS data showed six significant variables consisting of TCI, NDVI, LST_night, VCI, VHI, and EVI, all with p-values less than 0.05 (Table 3). While climatic data remained only a single variable, that is, Tmean. After applying the VIF, only four variables remain (Table 3). Therefore, the variables selected for the training and testing model that provided a VIF lower than 5 are TCI, NDVI, LST_night for RS data, and Tmean for climatic data, which range from 1.31 to 2.17 (Table 3).

Table 3. The statistical metrics of correlation analysis (CA) and variance inflation factor (VIF).

Data type	Variable	P-Value	VIF
RS data	TCI	0.001***	1.31
	NDVI	0.023*	1.22
	LST_night	0.001***	2.17
	VCI	0.001***	15.49
	VHI	0.001***	65.2
	EVI	0.035*	20.67
Climate data	LST_day	0.37	11.19
	Tmean	0.001***	2.05
	Rainfall	0.213	1.76
	Tmax	0.24	13.44
	Tmin	0.051	5

Note: * and ** refers to the confidence interval for p-value < 0.05 and 0.01, respectively.

3.2. Regression model predictions for province-level crop yield prediction in the Chi basin

In this study, a total of 126 samples were used to examine crop yield production at the provincial scale. These samples were divided into two time periods as training and testing data: 2011–2017 (98 samples) and 2018–2019 (28 samples). Four regression models (MLR and machine learning techniques) were applied to the three categories of data (Table 3): remote sensing (RS), climatic, and a combination of both. The MLR model using RS data provided the lowest R^2 value of 0.42 in the training dataset, while the XGBoost model using fusion data possessed the highest R^2 value of 0.95 (Table 4 **Error! Reference source not found.**). This study is congruent with the report of [15], which stated that a fusion of remote sensing-based drought indicators and climatic and weather indicators can provide high statistical measurement when used with the XGBoost model for cereal yield

forecasting. In terms of validation (RMSE), the XGBoost model with combination data provided the lowest root mean squared error (RMSE) of 0.18 ton/ha, while the support vector regression (SVR) model using climatic data had the second lowest RMSE of 0.18 to 0.3 ton/ha. This error threshold is generally accepted in European agro-statistics [48]. Overall, the XGBoost model was the most reliable for predicting crop yield production (highest R^2 and lowest RMSE) (Table 4).

Table 4. Training and testing of each model and data type for predicting crop yield production.

Category	R-square (Training: 2011-2017)			
	MLR	RF	XGBoost	SVR
RS data	0.42	0.74	0.89	0.64
Climatic data	0.55	0.94	0.93	0.88
Combination	0.63	0.92	0.95	0.81
RMSE (Testing: 2018-2019) (ton/ha)				
RS data	0.36	0.42	0.45	0.4
Climatic data	0.3	0.23	0.21	0.18
Combination	0.26	0.19	0.18	0.29

3.3. Temporal trend of crop production measurement and Changes of crop production validation

To further elaborate on the findings presented in Figure 4, it is crucial that the observed crop yield data and predicted crop yield data are evaluated using four different approaches: three non-parametric approaches (RF, XGBoost, and SVR) and one parametric approach (MLR). These approaches were then used to predict crop yield for a period of one month leading up to the harvest. The results showed that while there were fluctuations in yield among the variables and regression models, these fluctuations were not well reflected in the predicted crop yields. In fact, the peak yields actually observed in 2011 and 2017 (Figure 4) resulted in a reduced yield as observed in 2018.

To further assess the accuracy of the prediction models, the changes in crop yield were calculated for the validation periods (testing datasets) of 2018 and 2019 (Table 5). The outcomes showed that the MLR model executed very well for almost all predictor variables, with a difference of 0.03, 0.01, and 0.01 ton/ha for combination, RS, and climatic data, respectively. In 2019, the XGBoost and RF regression models showed insignificant changes in observed and predicted data, with a difference of around -0.01 ton/ha. Overall, these findings suggest that the non-parametric and parametric approaches used in this study can effectively predict crop yield for the period leading up to harvest, with the XGBoost and MLR models performing particularly well. However, it is notable that the linear regression model can perform well with the testing dataset, but it is not fully agreeable to apply for crop yield prediction if we consider the training statistical result, which has a low r-square when compared to other models.

Table 5. Changes of the crop yield validation relative to historical values of each model.

Model	Year	Mean Actual yield (ton/ha)	Variable			Change		
			Combination	RS	Climate	Δ Combination	Δ RS	Δ Climate
			Mean predicted yield (ton/ha)	Mean predicted yield (ton/ha)	Mean predicted yield (ton/ha)			
Linear	2018	2.34	2.37	2.35	2.34	0.03	0.01	0.01
	2019	2.36	2.45	2.51	2.45	0.10	0.15	0.09
RF	2018	2.34	2.28	2.32	2.26	-0.05	-	-0.07
	2019	2.36	2.35	2.45	2.35	0.00	0.10	-0.01

XGBoost	2018	2.34	2.28	2.36	2.27	-0.06	0.02	-0.07
	2019	2.36	2.35	2.50	2.35	-0.01	0.14	-0.01
SVR	2018	2.34	2.31	2.30	2.31	-0.02	-	-0.02
	2019	2.36	2.41	2.45	2.36	0.05	0.10	0.00

MLR and XGBoost regression are two different techniques that can be utilized to get predictions by using the input data. MLR is a parametric approach which assumes a linear correlation between the input factors and the output variable. This means that the output variable changes in a directly proportional manner with respect to the input variables. In contrast, XGBoost is a non-parametric technique that uses decision trees as weak learners and unites them through boosting to make predictions. Boosting is an ensemble learning approach that trains weak models sequentially, with each model attempting to correct the errors made by the previous model. While MLR is generally easier to understand and implement, XGBoost is more flexible and can model non-linear relationships. However, it can be more complex to implement and may require more computational resources. Therefore, in this study, XGBoost was selected as the main algorithm used for crop yield prediction at the provincial scale due to its ability to handle the complexity of the observed data and predictor variables and produce accurate and reliable results.

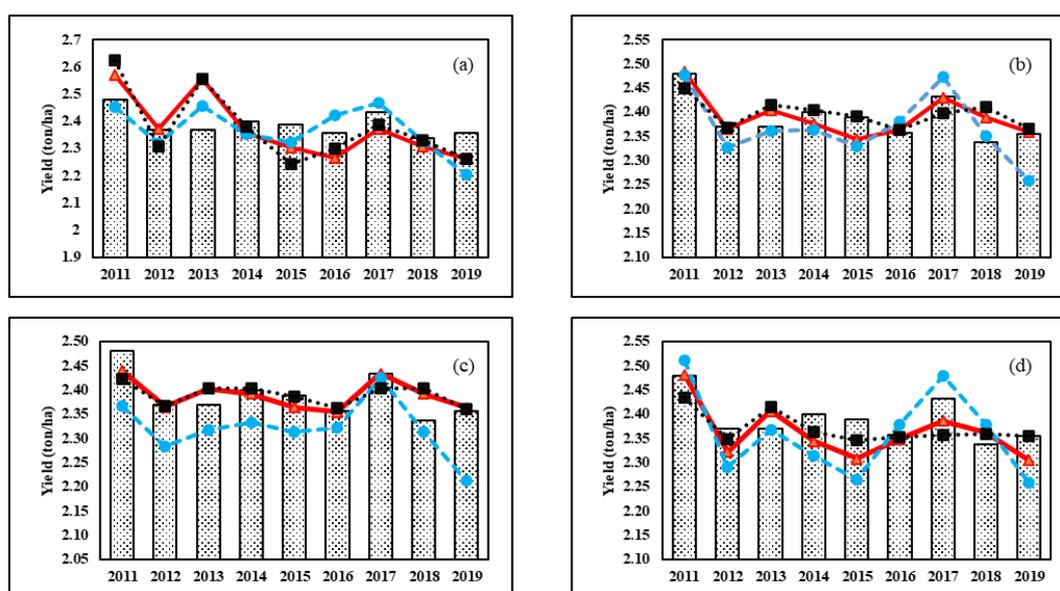


Figure 4. Comparison of annual averaging crop yield prediction (ton/ha) and averaged historical data for a) the MLR model, b) the RF model, c) the XGBoost model, and d) the SVR model; the bar chart is the averaged historical crop yield production; the red, blue, and black dot lines are the combination, RS and climatic data, respectively.

3.4. Crop yield prediction between 2018 and 2022

XGBoost is a selected machine learning algorithm that can be utilized to forecast yield production of crops, as was already mentioned. The crop yield ratio (tons/ha) was calculated over 14 provinces from 2018 to 2022 using the XGBoost model. The results showed that in 2018 and 2019, the highest crop yield ratio was observed in the PB province at 3.77 tons/ha, while the lowest value was observed in the NL province (2.23 tons/ha). In 2020, the PB province still possessed the highest crop yield ratio at 3.60 tons/ha, which is a decrease of 4.5% and 2.9% from 2018 and 2019, respectively. It is worth noting that the CP and KK provinces had the largest areas suitable for crop production, with 0.699 million hectares and 0.683 million hectares, respectively. Finally, in 2022, the crop yield ratio in the PB province decreased by 11.9% from 2021 and 2018. These findings suggest that the XGBoost model can effectively forecast the ratios of crop outcomes at the provincial stage and highlight the

importance of considering both yield and production area when making predictions. On the other hand, crop yield prediction was reproduced in CP province with the following number of areas: 1.61, 1.58, 1.90, 1.59, and 1.74 million tons a year between 2018 and 2022. Additionally, KK province was the second-largest region and produced crop yields ranging from 1.55 to 1.62 million tons a year. The total crop yield production that can reproduce in the Chi basin region ranges from 7.33 to 7.88 million tons a year starting from 2018 to 2022. The total of predicted crop yield production maps at the provincial scale for 2020–2022, shown in Figure 5. Therefore, this prediction may help gauge the overall economic performance of a country and is considered a key indicator of the standard of living of Thailand's citizens.

Table 6. Estimation of crop yield prediction over crop area through the Chi basin between 2018 and 2022.

Area	Crop yield area (ha)	Crop yield ratio (ton/ha)					Total crop yield (Mton)					
		Validation period		Predicting period			Validation period		Predicting period			
		2018	2019	2020	2021	2022	2018	2019	2020	2021	2022	
NS	81,076	2.29	2.26	2.9	2.31	2.36	0.19	0.18	0.23	0.19	0.19	
SK	16,562	2.29	2.29	2.36	2.26	2.45	0.04	0.04	0.04	0.04	0.04	
UR	44,155	2.27	2.27	2.36	2.26	2.42	0.1	0.1	0.1	0.1	0.11	
YT	125,803	2.29	2.26	2.36	2.38	2.36	0.29	0.28	0.3	0.3	0.3	
CP	699,264	2.3	2.26	2.72	2.27	2.49	1.61	1.58	1.9	1.59	1.74	
NL	216,043	2.23	2.23	2.27	2.66	2.2	0.48	0.48	0.49	0.57	0.48	
KK	683,868	2.28	2.26	2.34	2.36	2.36	1.56	1.55	1.6	1.61	1.62	
UD	256,024	2.29	2.26	2.3	2.36	2.32	0.59	0.58	0.59	0.6	0.59	
LO	61,150	2.31	2.24	2.7	2.6	2.6	0.14	0.14	0.16	0.16	0.16	
MK	247,999	2.28	2.28	2.32	2.32	2.45	0.56	0.56	0.58	0.57	0.61	
RT	366,514	2.29	2.28	2.36	2.38	2.45	0.84	0.83	0.86	0.87	0.9	
KS	418,757	2.29	2.28	2.32	2.36	2.26	0.96	0.95	0.97	0.99	0.95	
MH	1,254	2.32	2.26	2.36	2.66	2.33	0.01	0.01	0.01	0.01	0.01	
PB	11,195	3.77	3.66	3.6	3.77	3.32	0.04	0.04	0.04	0.04	0.04	
							Sum	7.40	7.34	7.89	7.65	7.73

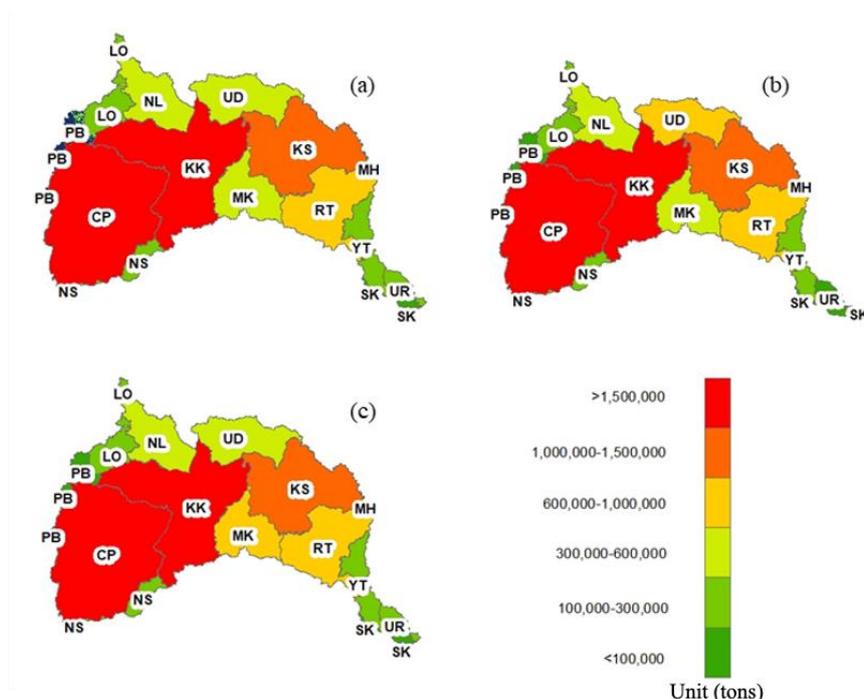


Figure 5. Total of predicted crop yield production at the provincial scale for a) 2020, b) 2021, and c) 2022.

4. Discussion

Monitoring, mapping, and predicting crop production in large regions can help farmers and policymakers make the best decisions for sustainable management, particularly in the Chi basin region, which is a major producer of crops in Thailand. This is especially important at present, as natural hazards often impact tropical monsoon areas. Additionally, climate change is one of the most important problems for the agricultural sector in the global region. Crop yield is crucial for global food security, so it is important to monitor and provide information about threats to crop production. Exact and well-timed early estimation of crop production has potential for trade and proper food management. There are various approaches to estimating the crop yield [49–51]. Predictive models for crop yield have been developed using remote sensing data and ML methods [52,53]. However, these approaches may not always provide accurate results. The study [10] applied the NDVI to forecast crop production in the Canadian Prairies, with results indicating R^2 values ranging from 0.8 to 0.9. While the study [51] used MODIS EVI and LAI data to examine the prediction of rice crop production in Vietnam's Mekong Delta and found that the maximum correlation coefficients at the growing stage of crops were 0.70 and 0.74, respectively.

Agricultural production relies on environmental conditions, such as climatic data (rainfall, temperature, humidity, and solar radiation) [54], so climatic and remote sensing data have been integrated for the prediction of crop yield [55], which is consistent with the findings in this study. This study compared and evaluated various approaches and predictor variables for predicting crop yield at the provincial scale in the Chi basin, Thailand, prior to the one- to two-month harvest period. This study found that combining satellite imaging data with climatic data improved the accuracy of predicting crop yield in the Chi basin. The results showed that the LST_{night}, NDVI, TCI, and Tmean data perform well when used with the XGBoost algorithm and can provide an R^2 value of up to 0.95. The results showed that the LST_{night}, NDVI, TCI, and Tmean data perform well when used with the XGBoost algorithm and can provide an R^2 value of up to 0.95. This combination of data can also improve the RMSE to 0.18 ton/ha. The XGBoost algorithm, which is a non-parametric technique that uses decision trees and joins them through boosting to make predictions, is an excellent method, similar to what was found by [15], which reported that the fusion of remote sensing-based drought, climatic, and weather indicators improved accuracy when used with the XGBoost model for cereal

yield forecasting. The temporal trend of crop yield prediction using XGBoost was rather close to the actual crop yield data; however, in 2018, the crop yield ratio differed by about 0.05 tons/ha due to natural hazards.

In 2018, there were 66 provinces or 420 districts affected by floods [56] that destroyed several agricultural areas, especially the rice crop area, which is located in a lowland area. According to [15], rainfed rice production is expected to decrease by around 5% from 2021 to 2029, which is inconsistent with our study, which predicts that yield will decrease by around 0.078 million tons per year starting in 2020 to 2022. In addition, drought impacts are expected to affect crop yield predictions in Thailand by about 5% mean absolute percentage error (MAPE) [57], this can be tele-connected from El Niño southern oscillation [58]. According to the results of total crop yield predictions for the period from 2020 to 2022 (Table 6), crop yield predictions have fluctuated and are likely to continue to incline in the coming future due to climate conditions. However, climate change has a considerable influence on the agriculture sector, and it could lead to an increase in temperatures by 1.4 to 5.8 degrees Celsius in 2100 [4]. This will increase crop water requirements due to increased evapotranspiration, which will mainly affect crop production [59]. This study shows acceptable accuracy for crop yield prediction that can be used by policymakers for management at the country and province scales. Since the methodology proposed in this study can accurately forecast the crop yield, it is anticipated that this methodology can be used as a guideline for crop yield prediction in other study areas as well as for policymaking to drive the economy at the provincial or country scale. As rice is the main staple crop in Thailand and is an important source of export income for the country.

The rice crop yield in Thailand is important to the overall trade and industry performance of the whole region and contributes to the overall GDP. This can be attributed to a number of aspects, including the adoption of modern agricultural technologies such as hybrid seeds and precision agriculture, as well as improvements in irrigation and fertilization practices. In addition, Thailand has a well-developed infrastructure for agriculture, including a network of roads, ports, and storage facilities that facilitate the transportation and distribution of crops. However, despite these improvements, crop yield production in Thailand can still be affected by various factors, such as drought and extreme weather events, which can lead to fluctuations in yield from year to year. In addition, market demand and prices for crops can also impact production trends, as farmers may choose to plant crops that are more in demand or more profitable. Finally, a decrease in crop yield may lead to greater usage of pesticides, fertilizers, and other chemical inputs, which can negatively impact the environment, including pollution and degradation of natural resources. Therefore, it is important to apply the proposed approach to early crop yield prediction and take steps to maintain high crop yields and sustainable development policies in order to minimize these negative consequences.

5. Conclusions

Crop yield prediction is crucial information for enabling farmers to quickly decide to increase production by enhancing management techniques for the period of developing season before one to two months prior to the harvested period. In this study, we demonstrated the approaches to predict crop production of RS and climatic variables. The study aims of this paper was to provide a predictive model before harvest for approximating crop outcomes in Thailand's Chi basin at the province scale between 2011 and 2019. To do this, we use a variety of remotely sensed and meteorological data and applied correlation analysis and variance inflation factor to identify the most relevant variables. Then the selected variables were used to train four regression models (MLR, RF, XGBoost and SVR), with the XGBoost model performing the best, with a minimum root mean square error of 0.18 ton/ha. To predict total crop production, the XGBoost model was applied in the Chi basin for the years 2020-2022, with the result that total crop production is expected to be approximately 7.88, 7.64, and 7.72 million tons. This research found that using satellite-based drought indicators, vegetation index, and meteorological data with assistant of machine learning algorithms is an effective method for predicting agricultural yields in the study area. This method also provided timely data that can be used for decision making during the crop growth season. The discoveries of the proposed study may

also be exploited to plot crop yields and its gaps at the provincial level in Thailand and neighboring countries, helping farmers and policymakers make informed decisions.

Author Contributions: Conceptualization, A.C., A.H. and S.K.; methodology, A.C., A.H. and S.K.; software, A.C., A.H. and S.K.; validation, R.H., N.S. and S.P.; formal analysis, A.C.; investigation, A.K.; resources, W.K.; data curation, R.H.; writing—original draft preparation, A.C.; writing—review and editing, A.H. and S.K.; visualization, A.C. All authors have read and agreed to the published version of the manuscript.

Funding: “This research was financially supported by Mahasarakham University.”.

Data Availability Statement: The data used in this study are available from the corresponding author upon reasonable request.

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

References

1. Food and Agriculture Organization of the United Nations 2020 Asia Pacific Regional Overview of Food Security and Nutrition: Maternal and Child Diets at the Heart of Improving Nutrition.
2. Department of Agricultural Extension Rice Production in Thailand Available online: https://www.agriculture.gov.au/sites/default/files/documents/annual-report-2019-20-awe-oct-2020_0.pdf (accessed on 30 January 2024).
3. World Bank Thai Economic Monitor Productivity for Prosperity Available online: <https://documents1.worldbank.org/curated/en/394501579357102381/pdf/Thailand-Economic-Monitor-Productivity-for-Prosperity.pdf> (accessed on 30 January 2024).
4. Intergovernmental Panel on Climate Change Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, Switzerland; Geneva, 2014;
5. Majumder, A.; Kingra, P.K.; Setia, R.; Singh, S.P.; Pateriya, B. Influence of Land Use/Land Cover Changes on Surface Temperature and Its Effect on Crop Yield in Different Agro-Climatic Regions of Indian Punjab. *Geocarto Int* 2020, 35.
6. Office of Agricultural Economics, M. of A. and C. in T. Agricultural Statistics of Thailand 2016 Available online: <https://www.oae.go.th/view/1/Home/EN-US> (accessed on 14 December 2022).
7. Land Development Department Annual Report Available online: https://webapp.ldd.go.th/lpd/pdfs/web/viewer.html?File=../../node_modules/file/Report/Annual%20Report%202021.pdf (accessed on 30 January 2024).
8. 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 (Basel)* 2020, 12, doi:10.3390/rs12020236.
9. Gao, Y.; Wang, S.; Guan, K.; Wolanin, A.; You, L.; Ju, W.; Zhang, Y. The Ability of Sun-Induced Chlorophyll Fluorescence from OCO-2 and MODIS-EVI to Monitor Spatial Variations of Soybean and Maize Yields in the Midwestern USA. *Remote Sens (Basel)* 2020, 12, doi:10.3390/rs12071111.
10. 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, doi:10.1016/j.agrformet.2010.11.012.
11. Stepanov, A.; Dubrovin, K.; Sorokin, A.; Aseeva, T. Predicting Soybean Yield at the Regional Scale Using Remote Sensing and Climatic Data. *Remote Sens (Basel)* 2020, 12, doi:10.3390/rs12121936.
12. Mongkolnithithada, W.; Nontapun, J.; Kaewplang, S. Rice Yield Estimation Based on Machine Learning Approaches Using MODIS 250 m Data. *Engineering Access* 2023, 9, 75–79.
13. Schwalbert, R.A.; Amado, T.; Corassa, G.; Pott, L.P.; Prasad, P.V.V.; Ciampitti, I.A. Satellite-Based Soybean Yield Forecast: Integrating Machine Learning and Weather Data for Improving Crop Yield Prediction in Southern Brazil. *Agric For Meteorol* 2020, 284, doi:10.1016/j.agrformet.2019.107886.
14. Pede, T.; Mountrakis, G.; Shaw, S.B. Improving Corn Yield Prediction across the US Corn Belt by Replacing Air Temperature with Daily MODIS Land Surface Temperature. *Agric For Meteorol* 2019, 276–277, doi:10.1016/j.agrformet.2019.107615.
15. Bouras, E.H.; Jarlan, L.; Er-Raki, S.; Balaghi, R.; Amazirh, A.; Richard, B.; Khabba, S. Cereal Yield Forecasting with Satellite Drought-Based Indices, Weather Data and Regional Climate Indices Using Machine Learning in Morocco. *Remote Sens (Basel)* 2021, 13, doi:10.3390/rs13163101.
16. Pei, F.; Wu, C.; Liu, X.; Li, X.; Yang, K.; Zhou, Y.; Wang, K.; Xu, L.; Xia, G. Monitoring the Vegetation Activity in China Using Vegetation Health Indices. *Agric For Meteorol* 2018, 248, doi:10.1016/j.agrformet.2017.10.001.
17. Bouras, E.H.; Jarlan, L.; Er-Raki, S.; Albergel, C.; Richard, B.; Balaghi, R.; Khabba, S. Linkages between Rainfed Cereal Production and Agricultural Drought through Remote Sensing Indices and a Land Data Assimilation System: A Case Study in Morocco. *Remote Sens (Basel)* 2020, 12, doi:10.3390/rs12244018.

18. Baniya, B.; Tang, Q.; Xu, X.; Haile, G.G.; Chhipi-Shrestha, G. Spatial and Temporal Variation of Drought Based on Satellite Derived Vegetation Condition Index in Nepal from 1982–2015. *Sensors (Switzerland)* **2019**, *19*, doi:10.3390/s19020430.
19. Dutta, D.; Kundu, A.; Patel, N.R.; Saha, S.K.; Siddiqui, A.R. Assessment of Agricultural Drought in Rajasthan (India) Using Remote Sensing Derived Vegetation Condition Index (VCI) and Standardized Precipitation Index (SPI). *Egyptian Journal of Remote Sensing and Space Science* **2015**, *18*, doi:10.1016/j.ejrs.2015.03.006.
20. Herrero-Huerta, M.; Rodriguez-Gonzalvez, P.; Rainey, K.M. Yield Prediction by Machine Learning from UAS-Based Multisensor Data Fusion in Soybean. *Plant Methods* **2020**, *16*, doi:10.1186/s13007-020-00620-6.
21. 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*, doi:10.1038/s41598-020-80820-1.
22. Khosla, E.; Dharavath, R.; Priya, R. Crop Yield Prediction Using Aggregated Rainfall-Based Modular Artificial Neural Networks and Support Vector Regression. *Environ Dev Sustain* **2020**, *22*, doi:10.1007/s10668-019-00445-x.
23. Zhang, W.; Wu, C.; Li, Y.; Wang, L.; Samui, P. Assessment of Pile Drivability Using Random Forest Regression and Multivariate Adaptive Regression Splines. *Georisk* **2021**, *15*, doi:10.1080/17499518.2019.1674340.
24. Zhang, Y.; Xia, C.; Zhang, X.; Cheng, X.; Feng, G.; Wang, Y.; Gao, Q. Estimating the Maize Biomass by Crop Height and Narrowband Vegetation Indices Derived from UAV-Based Hyperspectral Images. *Ecol Indic* **2021**, *129*, doi:10.1016/j.ecolind.2021.107985.
25. Kang, Y.; Ozdogan, M.; Zhu, X.; Ye, Z.; Hain, C.; Anderson, M. Comparative Assessment of Environmental Variables and Machine Learning Algorithms for Maize Yield Prediction in the US Midwest. *Environmental Research Letters* **2020**, *15*, doi:10.1088/1748-9326/ab7df9.
26. Memon, N.; Patel, S.B.; Patel, D.P. Comparative Analysis of Artificial Neural Network and XGBoost Algorithm for PolSAR Image Classification. In Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); 2019; Vol. 11941 LNCS.
27. 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*, doi:10.1371/journal.pone.0156571.
28. 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*, doi:10.3390/s22030717.
29. Boonwichai, S.; Shrestha, S.; Babel, M.S.; Weesakul, S.; Datta, A. Climate Change Impacts on Irrigation Water Requirement, Crop Water Productivity and Rice Yield in the Songkhram River Basin, Thailand. *J Clean Prod* **2018**, *198*, doi:10.1016/j.jclepro.2018.07.146.
30. Sujariya, S.; Jongrunklang, N.; Jongdee, B.; Inthavong, T.; Budhaboon, C.; Fukai, S. Rainfall Variability and Its Effects on Growing Period and Grain Yield for Rainfed Lowland Rice under Transplanting System in Northeast Thailand. *Plant Prod Sci* **2020**, *23*, doi:10.1080/1343943X.2019.1698970.
31. Ramadhani, F.; Pullanagari, R.; Kereszturi, G.; Procter, J. Mapping a Cloud-Free Rice Growth Stages Using the Integration of Proba-v and Sentinel-1 and Its Temporal Correlation with Sub-District Statistics. *Remote Sens (Basel)* **2021**, *13*, doi:10.3390/rs13081498.
32. Peñuelas, J.; Filella, L. Technical Focus: Visible and near-Infrared Reflectance Techniques for Diagnosing Plant Physiological Status. *Trends Plant Sci* **1998**, *3*, doi:10.1016/S1360-1385(98)01213-8.
33. Mosleh, M.K.; Hassan, Q.K.; Chowdhury, E.H. Application of Remote Sensors in Mapping Rice Area and Forecasting Its Production: A Review. *Sensors (Switzerland)* **2015**, *15*.
34. Alahacoon, N.; Edirisinghe, M.; Ranagalage, M. Satellite-Based Meteorological and Agricultural Drought Monitoring for Agricultural Sustainability in Sri Lanka. *Sustainability (Switzerland)* **2021**, *13*, doi:10.3390/su13063427.
35. Zhang, L.; Jiao, W.; Zhang, H.; Huang, C.; Tong, Q. Studying Drought Phenomena in the Continental United States in 2011 and 2012 Using Various Drought Indices. *Remote Sens Environ* **2017**, *190*, doi:10.1016/j.rse.2016.12.010.
36. Yu, H.; Li, L.; Liu, Y.; Li, J. Construction of Comprehensive Drought Monitoring Model in Jing-Jin-Ji Region Based on Multisource Remote Sensing Data. *Water (Switzerland)* **2019**, *11*, doi:10.3390/w11051077.
37. Tuvdendorj, B.; Wu, B.; Zeng, H.; Batdelger, G.; Nanzad, L. Determination of Appropriate Remote Sensing Indices for Spring Wheat Yield Estimation in Mongolia. *Remote Sens (Basel)* **2019**, *11*, doi:10.3390/rs11212568.
38. Liao, K.; Xu, S.; Wu, J.; Zhu, Q. Spatial Estimation of Surface Soil Texture Using Remote Sensing Data. *Soil Sci Plant Nutr* **2013**, *59*, doi:10.1080/00380768.2013.802643.
39. Boori, M.S.; Choudhary, K.; Paringer, R.; Kupriyanov, A. Spatiotemporal Ecological Vulnerability Analysis with Statistical Correlation Based on Satellite Remote Sensing in Samara, Russia. *J Environ Manage* **2021**, *285*, doi:10.1016/j.jenvman.2021.112138.

40. Guechi, I.; Gherraz, H.; Alkama, D. Correlation Analysis between Biophysical Indices and Land Surface Temperature Using Remote Sensing and GIS in Guelma City (Algeria). *Bulletin de la Societe Royale des Sciences de Liege* **2021**, *90*, doi:10.25518/0037-9565.10457.
41. Kang, J.; Jin, R.; Li, X.; Zhang, Y.; Zhu, Z. Spatial Upscaling of Sparse Soil Moisture Observations Based on Ridge Regression. *Remote Sens (Basel)* **2018**, *10*, doi:10.3390/rs10020192.
42. Hamzehpour, N.; Shafizadeh-Moghadam, H.; Valavi, R. Exploring the Driving Forces and Digital Mapping of Soil Organic Carbon Using Remote Sensing and Soil Texture. *Catena (Amst)* **2019**, *182*, doi:10.1016/j.catena.2019.104141.
43. Browning, M.H.E.M.; Kuo, M.; Sachdeva, S.; Lee, K.; Westphal, L. Greenness and School-Wide Test Scores Are Not Always Positively Associated – A Replication of “linking Student Performance in Massachusetts Elementary Schools with the ‘Greenness’ of School Surroundings Using Remote Sensing.” *Landsc Urban Plan* **2018**, *178*, doi:10.1016/j.landurbplan.2018.05.007.
44. Alsharif, A.A.A.; Pradhan, B. Urban Sprawl Analysis of Tripoli Metropolitan City (Libya) Using Remote Sensing Data and Multivariate Logistic Regression Model. *Journal of the Indian Society of Remote Sensing* **2014**, *42*, doi:10.1007/s12524-013-0299-7.
45. Maya Gopal, P.S.; Bhargavi, R. Selection of Important Features for Optimizing Crop Yield Prediction. *International Journal of Agricultural and Environmental Information Systems* **2019**, *10*, doi:10.4018/IJAEIS.2019070104.
46. Rtayli, N.; Enneya, N. Enhanced Credit Card Fraud Detection Based on SVM-Recursive Feature Elimination and Hyper-Parameters Optimization. *Journal of Information Security and Applications* **2020**, *55*, doi:10.1016/j.jisa.2020.102596.
47. Dong, W.; Huang, Y.; Lehane, B.; Ma, G. XGBoost Algorithm-Based Prediction of Concrete Electrical Resistivity for Structural Health Monitoring. *Autom Constr* **2020**, *114*, doi:10.1016/j.autcon.2020.103155.
48. Genovese, C.R.; Roeder, K.; Wasserman, L. False Discovery Control with P-Value Weighting. *Biometrika* **2006**, *93*, doi:10.1093/biomet/93.3.509.
49. Sakamoto, T.; Gitelson, A.A.; Arkebauer, T.J. Near Real-Time Prediction of U.S. Corn Yields Based on Time-Series MODIS Data. *Remote Sens Environ* **2014**, *147*, doi:10.1016/j.rse.2014.03.008.
50. Zhuo, W.; Fang, S.; Gao, X.; Wang, L.; Wu, D.; Fu, S.; Wu, Q.; Huang, J. Crop Yield Prediction Using MODIS LAI, TIGGE Weather Forecasts and WOFOST Model: A Case Study for Winter Wheat in Hebei, China during 2009–2013. *International Journal of Applied Earth Observation and Geoinformation* **2022**, *106*, doi:10.1016/j.jag.2021.102668.
51. Ji, Z.; Pan, Y.; Zhu, X.; Wang, J.; Li, Q. Prediction of Crop Yield Using Phenological Information Extracted from Remote Sensing Vegetation Index. *Sensors (Switzerland)* **2021**, *21*, doi:10.3390/s21041406.
52. Ban, H.Y.; Ahn, J.B.; Lee, B.W. Assimilating MODIS Data-Derived Minimum Input Data Set and Water Stress Factors into CERES-Maize Model Improves Regional Corn Yield Predictions. *PLoS One* **2019**, *14*, doi:10.1371/journal.pone.0211874.
53. Sun, J.; Di, L.; Sun, Z.; Shen, Y.; Lai, Z. County-Level Soybean Yield Prediction Using Deep CNN-LSTM Model. *Sensors (Switzerland)* **2019**, *19*, doi:10.3390/s19204363.
54. Lin, B.B. Agroforestry Management as an Adaptive Strategy against Potential Microclimate Extremes in Coffee Agriculture. *Agric For Meteorol* **2007**, *144*, doi:10.1016/j.agrformet.2006.12.009.
55. Ju, S.; Lim, H.; Ma, J.W.; Kim, S.; Lee, K.; Zhao, S.; Heo, J. Optimal County-Level Crop Yield Prediction Using MODIS-Based Variables and Weather Data: A Comparative Study on Machine Learning Models. *Agric For Meteorol* **2021**, *307*, doi:10.1016/j.agrformet.2021.108530.
56. Water Analysis and Assessment Division Water Situation Report Available online: <http://mekhala.dwr.go.th/en/situation.php?numpage=value&Page=1875> (accessed on 30 January 2024).
57. Raksapatcharawong, M.; Veerakachen, W.; Homma, K.; Maki, M.; Oki, K. Satellite-Based Drought Impact Assessment on Rice Yield in Thailand with SIMRIW-RS. *Remote Sens (Basel)* **2020**, *12*, doi:10.3390/rs12132099.
58. Anderson, W.; Seager, R.; Baethgen, W.; Cane, M. Crop Production Variability in North and South America Forced by Life-Cycles of the El Niño Southern Oscillation. *Agric For Meteorol* **2017**, *239*, doi:10.1016/j.agrformet.2017.03.008.
59. Astuti, I.S.; Wiwoho, B.S.; Purwanto, P.; Wagistina, S.; Deffinika, I.; Suchahyo, H.R.; Herlambang, G.A.; Alfarizi, I.A.G. An Application of Improved MODIS-Based Potential Evapotranspiration Estimates in a Humid Tropic Brantas Watershed—Implications for Agricultural Water Management. *ISPRS Int J Geoinf* **2022**, *11*, doi:10.3390/ijgi11030182.

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.