Article

Rice Yield Simulation and Planting Suitability Environment Pattern Recognition at a Fine Scale

Daichao Li [ID], Jianqin Liang, Xingfeng Wang, Sheng Wu *, Xiaowei Xie and Jiaqi Lu

The Academy of Digital China (Fujian), Fuzhou University, Fuzhou 350116, China; lidc@fzu.edu.cn (D.L.);
205527019@fzu.edu.cn (J.L.); n185527026@fzu.edu.cn (X.W.); N195520021@fzu.edu.cn (X.X.);
n195527024@fzu.edu.cn (J.L.)
* Correspondence: wusheng@fzu.edu.cn

Abstract: Analyzing rice yields and multidimensional environmental factors at a fine scale facilitates the discovery of the planting environment patterns that guide the spatial layout of rice production. This study uses Pucheng County, Fujian Province, a demonstration county of China Good Grains and Oils, as the research area. Using actual rice yield sample data and environment data, a yield simulation model based on random forest regression is constructed to realize a fine-scale simulation of rice yield and its spatial distribution pattern in Pucheng County. On this basis, we construct a method system to identify spatial combination patterns between rice yields and fine-scale multidimensional environmental planting suitability using rice yield data and environmental planting suitability evaluation data. We categorize the areas into four combination model areas to analyze the spatial correlation model of planting suitability, multidimensional environment, and yield: higher-yield and higher-suitability cluster–comprehensive environmental-advantage areas, high-yield and high-suitability cluster–soil condition-limited areas, moderate-yield and moderate-suitability cluster–irrigation and drainage condition-limited areas, and low-yield and low-suitability cluster–site condition-limited areas. The following results are found. (1) The rice yield simulation model, which is based on random forest regression, considers the various complex relationships between yield and natural as well as human factors to realize the refined simulation of rice yields at a county scale. (2) The county rice yield has a strong positive spatial correlation, and the spatial clustering characteristics are obvious; these relationships can provide a basis for effectively implementing intensive rice planting in Pucheng County. (3) We construct a spatial combination pattern recognition method based on rice yield and environmental planting suitability. We can use this method to effectively identify the spatial relationship between yield and planting suitability as well as the shortcomings and advantages of different regions in terms of the climate, soil, irrigation, site, mechanical farming, and similar factors. On this basis, we can provide regional rice planting guidance for Pucheng County. In addition, this method system also provides a new perspective and method for research into spatial combination models and related spatial issues.

Keywords: GIS; rice yield simulation; multidimensional environment; cluster analysis; spatial pattern recognition

1. Introduction

Rice is currently one of the most popular food crops worldwide, and its planting environment significantly affects its quality and yield. Evaluation of the rice-planting environment is important to effectively utilize regional environmental resources, explore arable land, and achieve large yields of high-quality rice [1]. The spatial distribution of the crop-planting environment provides an important decision-making basis for adjusting agricultural planting strategies. Identifying agricultural planting environmental patterns can improve the natural and human–environmental resources and promote sustainable agricultural development [2].

Rice yields are affected by natural environmental factors such as soil, topography, climate, farmland management, agricultural mechanization, and other artificial environmental factors [3–5]. Due to the large workload and slow speed of traditional on-site sampling survey methods [6], sampling point data can be obtained from only some fields. Methods for high-precision simulation of rice yields that combine this rice yield data with environmental factor data have become a research hotspot. Existing research indicates that the relationship between rice yield and environmental factors is complicated [7–9]. Therefore, using a single model to simulate a regional rice yield cannot accurately determine the relationship between rice yield and environmental factors, and the model's simulated rice yield may be unstable when there are many changing environmental factors. Therefore, when comparing the accuracy and fit of different rice yield simulation models, choosing a high-precision rice yield simulation model with a better fit would make it possible to obtain yield data for an entire region using the rice yield data from the sampling points. Such a model would provide researchers with a method for obtaining high-precision prediction data with minimal time and economic cost.

Analyzing the spatial relationship between yield and multidimensional environmental factors and then exploring effective strategies can achieve the rice yield–environment dual goal of rice production spatial pattern. These patterns can support decision making that optimizes the rice production structure and the quality of agricultural development [10]. Most existing research on the spatial relationship between rice yield and environmental factors focuses on natural factors, such as climate and soil, or other factors, such as crop models [11–13]. There is little comprehensive analysis of the impact of natural factors and human factors. Human factors include accessibility to cultivated land, irrigation, and drainage, and the feasibility of mechanization. Most of the available research considers whether environmental factors impact yield [14–16] and the mechanism of effect [17–20], while research on the identification and analysis of the spatial combination pattern between yield and environmental factors is relatively lacking. The scale of most studies has been at the macro level, examining an entire country and its provinces [21–23]; relatively few related studies consider mining a combined model that includes crop yield and the planting environment at the micro level with the county as the regional unit.

This study uses Pucheng County, Fujian Province, as the research area. Firstly, rice yield simulation models are constructed based on a geographically weighted regression model, multilayer perceptron model, and random forest regression model. After comparing the three models, random forest is ultimately selected for this study due to its relatively high simulation accuracy, low mean absolute error, and low root mean square error. The random forest model simulates a refined spatial rice yield distribution on the county's arable land. Secondly, based on the rice yield simulation data and the multidimensional environmental rice planting suitability evaluation results, the local Moran's I index is chosen to analyze the rice yield separately; in addition, k-means clustering and variance analysis are used to identify the spatial combination pattern of rice yield and environmental planting suitability. Finally, on this basis, we find obstacles in the model and propose corresponding countermeasures and suggestions that take advantage of regional environmental resources and optimize the grain-planting structure and industrial layout.

The remainder of this paper is structured as follows. Section 2 introduces the datasets and methods used in this study. Section 3 presents the results and analysis. Section 4 provides a discussion. Finally, a summary of the conclusions is presented in Section 5.

2. Materials and Methods

2.1. Study Area

Pucheng County is under the jurisdiction of Nanping City, Fujian Province and is located in the province's northernmost region. Its geographical position is between $118°11'$ E and $118°49'$ E and between $27°32'$ N and $28°22'$ N (Figure 1). Pucheng County is a China Good Grain and Oil demonstration county, and it is the largest grain-producing area in Fujian Province [24]. It is in a mountainous and hilly area with a significant temperature dif-

ference between day and night. The county's annual average sunshine hours is 1893 h, the annual average temperature ranges from 13.3 to 18.0 °C, the total accumulated temperature is 6391.5 °C, the ≥10 °C accumulated temperature ranges from 4500 to 5300 °C, the annual average precipitation is 1100–2400 mm, and the distribution of cultivated land in the county is generally 200–500 m above sea level. The total area of Pucheng County is 3374.7 km^2, of which the total area of cultivated area accounts for about 11.87%. The soil types in Pucheng County vary with altitude, and most of the river valley basins at an altitude of 168–300 m are paddy soil. Paddy soil is the primary cultivated soil in Pucheng County, accounting for 99.37% of the county's total cultivated area [25]. The natural environmental advantages and resources in Pucheng County have given it a reputation as the "North Fujian Granary".
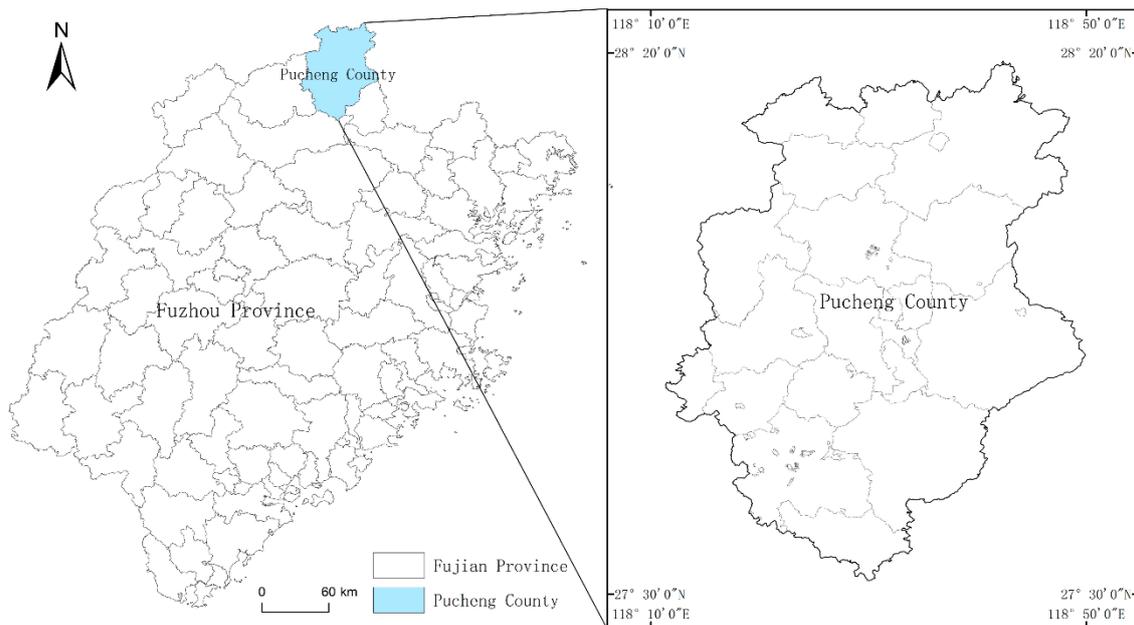

Figure 1. Location of the study area.

2.2. Introduction to Data Sources

The data used in this study include actual rice yield sample data and a multidimensional environmental rice planting suitability evaluation dataset. The specific data sources are as follows:

Actual rice yield sample data: The actual rice yield sample data came from the Department of Agriculture and Natural Resources of Pucheng County, which ensures the reliability of the data. We use an average value of 571 rice yield samples in Pucheng County from 2010 to 2018. We spatialize the actual rice yield sample data to obtain rice yield vector data.

Multidimensional environmental rice planting suitability evaluation dataset: These data are derived from Xingfeng Wang's research results [26]. This study constructed a raster dataset of rice planting environmental suitability evaluation indicators with the spatial resolution of 5 m. The dataset contains environmental indicator data (Table 1) as well as the corresponding planting suitability values, where the planting suitability value ranges from 0 to 100. We use GIS software to assign planting suitability values to the corresponding plots to obtain rice planting suitability evaluation vector data. Figure 2 shows the distribution of rice planting suitability values in Pucheng County.

Table 1. Multidimensional environmental index dataset *.

	Dimension	Index
Multidimensional environmental index	soil conditions	pH soil texture organic matter total nitrogen total potassium available potassium available phosphorus alkali-hydrolyzable nitrogen
	site conditions	altitude slope aspect topsoil thickness
	climate conditions	annual average temperature annual sunshine duration ≥10 °C accumulated temperature
	irrigation and drainage conditions	irrigation conditions drainage conditions
	mechanical farming conditions	field size field regularity accessibility to cultivated land

* Data accuracy: 5 m × 5 m.
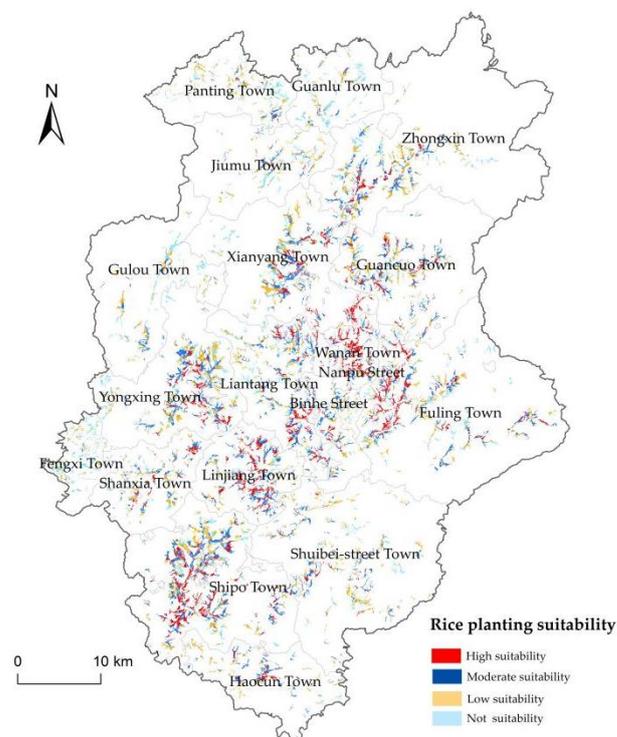
Figure 2. Results of rice planting suitability evaluation in Pucheng County.

This dataset is based on the principles of regionality, dominance, intra-county differences, and relative stability [27], which is combined with expert experience as well as select soil conditions (pH, soil texture, organic matter, total nitrogen, total potassium, available potassium, available phosphorus, alkali-hydrolyzable nitrogen), site conditions (altitude, slope, aspect, and topsoil thickness), irrigation and drainage conditions (irrigation condi-

tions and drainage conditions), climate conditions (annual average temperature, annual sunshine duration and $\geq 10\,^{\circ}C$ accumulated temperature), and mechanical farming conditions (field size, field regularity, and accessibility to cultivated land), totaling 20 indicators in 5 categories. These indicators construct an evaluation index system for rice planting suitability in Pucheng County.

The suitability value described above is based on the calculated weight of each evaluation index, where each index's weight is multiplied by the level experience score of the corresponding index, and the suitability of each sample is accumulated. The formula for the rice planting suitability evaluation model is expressed as:

$$SV = \sum_{i=1}^{n} W_i \times P_i \, , \tag{1}$$

where SV is the rice planting suitability value, W_i represents the weight of environmental indicators, and P_i represents the experience score of multidimensional environmental indicators [26].

In the dataset, the meteorological data originate from the average value of data from 1989 to 2018 of the China Meteorological Data Network (http://data.cma.cn/ (accessed on 2 January 2021)). The sampling data of the cultivated land fertility survey, soil testing, and formula fertilization pilot data are from the average value of the corresponding data from the Pucheng County Agricultural Department from 2010 to 2018. The road vector data and DEM data are from the Pucheng County Natural Resources Department.

The multidimensional environmental indicator data in this dataset provide support for the rice yield simulation, and the planting suitability value provides environmental suitability values that contribute to the combined analysis of rice yield and planting suitability in this study.

2.3. Research Process and Methods

Figure 3 presents the overall framework of our study. First, based on the multidimensional environmental rice planting suitability evaluation raster dataset and actual rice yield sample vector data, a geographically weighted regression (GWR) model, a multilayer perceptron (MLP) model, and a random forest regression (RFR) model were used to construct rice yield simulation models. Through testing, the optimal model was selected by examining the simulation results from the three models to determine the rice yield spatial distribution at a fine scale. Then, Moran's I index was used to analyze the rice yield's spatial clustering characteristics. K-means clustering and variance analysis were used to analyze the spatial combination pattern between the rice yield and environmental planting suitability; then, we searched for obstacles in the planting environment and identified the spatial model that appropriately combines the rice yield and environmental planting suitability. Finally, we put forward corresponding countermeasures and suggestions.

2.3.1. Simulation Model of Rice Yield Spatial Distribution

Based on the GWR, MLP, and RFR models, three rice yield spatial distribution simulation models were constructed in this study. First, we cleaned and normalized the original data. Then, the 571 sets of measured rice output vector data and the environmental factor data were used as the training set, and finally, the optimal model was selected to simulate the spatial distribution of Pucheng County's rice output data.

(1) GWR Model

A GWR [28,29] adds the data's spatial coordinates to the model and uses the observed values from adjacent areas in the sample data to fit the regression parameters of the local area. When the spatial coordinates of the authority domain change, these parameters change accordingly. While the ordinary least squares regression model (OLS) uses global data to estimate parameters [30], the GWR model extends the OLS model.

The structure of the GWR model is as follows:

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{ik} + \varepsilon_i, \tag{2}$$

where y is an $n \times 1$-dimensional rice yield simulation value, $\beta_k(u_i, v_i)$ represents the regression coefficient of multidimensional environmental factor variable k at regression point i, x is an $n \times k$-dimensional matrix of environmental factor variables, (u_i, v_i) represents the spatial coordinates of the i-th sample point, and ε is an $n \times 1$-dimensional vector.
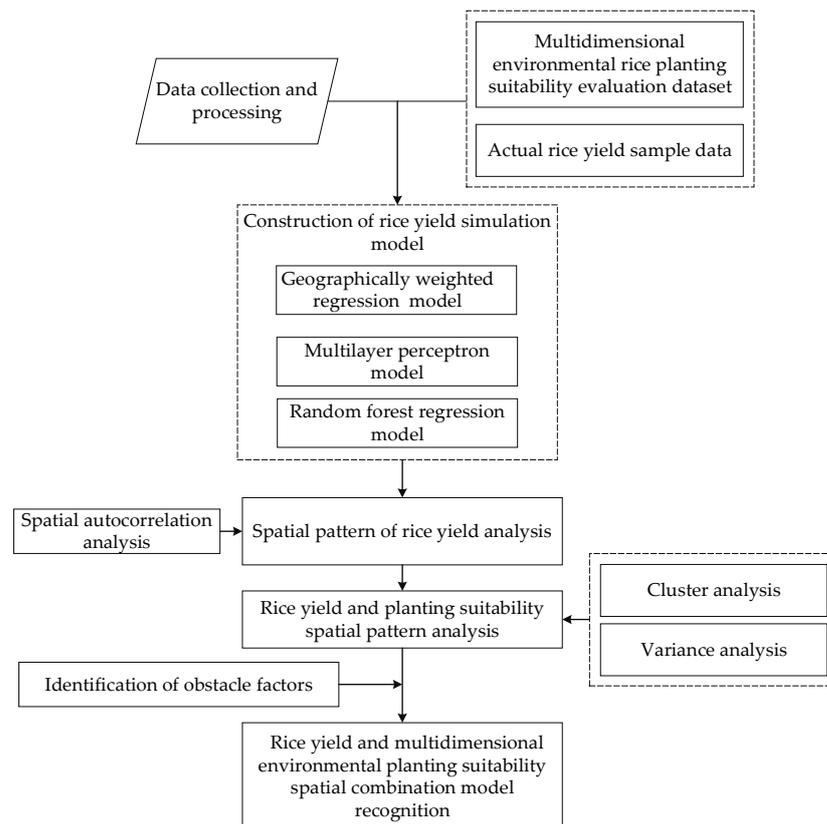

Figure 3. Flow chart of the study.

In this study, GWR4 software was used to perform geographically weighted regression. During the modeling process, the traditional geographic weighted regression model is selected as the model parameter, that is, the Gaussian function is selected as the weight function. The adaptive Gaussian was used as the kernel function for calculating geographic weights, the interval search method is used to generate a series of bandwidths, and corrected Akaike information criterion (AICc) was used as the criterion for judging the pros and cons, and then, the model was solved. After multiple geographically weighted regressions, the best bandwidth, which minimizes the AICc value and demonstrates the best performance, was finally obtained. That is, the best bandwidth (the number of neighbors) was 59, and its AICc value was 9818.182. After adjustment, R^2 was 0.526.

(2)   MLP Model

An MLP [31,32], a common artificial neural network model, is composed of an input layer, a hidden layer, and an output layer. In this paper, the MLP model uses only one hidden layer. The MLP process is supervised, and the model is trained by continuously inputting and outputting complete data. This mode establishes the best-fitting model between the input and output layers through continuously training and learning the sample data; finally, it uses the model that maximizes the fit for the simulation. The specific modeling steps are as follows:

- Take the rice yield of the measured sample point as the dependent variable, use the normalized data of 20 environmental impact factors as covariates, and use the partition variables to allocate the training set and test set data.
- The system structure is automatically selected, the minimum number of units in the hidden layer is set to one, the maximum number of units is set to 50, and Softmax is selected as the activation function.
- The model is trained with the training set data, and a simulation model is established.
- The test set data tests the model's simulation ability.

This paper used SPSS software to build a multilayer perceptron simulation model. In the process of model training, after many adjustments, it was found that 70% of the 571 sets were used as simulation set data. The remaining 30% of the sample data was used to test the model's simulation ability, and the effect of the obtained rice yield simulation model is the best.

(3) RFR Model

RFR obtains a more accurate result than the other models [33,34]. Breiman [35] proposed that a random forest is a combination of tree predictors. The RFR method is a random combination of multiple classification and regression trees that uses the calculated results of each regression tree to comprehensively define the output results. The model-building steps are as follows:

- The data are used as the sample dataset for RFR. Bootstrap is repeatedly used to randomly select a certain number of subsamples from the dataset. After each subsample is randomly selected, it is put back into the total sample.
- When generating a decision tree, an environmental factor feature variable is randomly selected from the multidimensional environmental factor dataset and designated the split feature set. Then, the mean square error is used to select each node in the decision tree.
- The extracted subsample sets are used to build classification regression trees. The decision tree is allowed to grow freely without pruning. Due to the random nature of the RFR model, the classification and regression trees will not appear to fit.
- Calculate the weighted average of the output results of the independent and equally important decision trees as the value of the rice yield simulation result for the RFR model.

This study used the RandomForestRegressor package in Python to simulate the spatial distribution of rice production in Pucheng County. A certain proportion of sample data were randomly selected from the training data set as the training set samples. After testing, the model simulation ability was the best when the training set samples were divided into 469 groups. During the modeling process, all environmental factor data were used as characteristic variable sets, and the rice yield was used as the simulation target variable when constructing the model. After repeated training and testing, and based on historical experience and test results, the number of decision trees in the random forest regression simulation model was set to 100, and the number of multi-dimensional environmental factor features was set to 17; that is, the model used 100 decision trees and 17 feature variables. An additional 102 groups of data were used as test set samples to test the model's simulation accuracy.

2.3.2. Spatial Autocorrelation Analysis

Spatial autocorrelation analysis is based on Tobler's first law of geography [36,37], in which the correlation of a specific variable is analyzed at different locations. It encompasses global and local autocorrelation analysis [38]. In this study, local autocorrelation was used to explore the spatial distribution patterns of the rice yield in Pucheng County. Moran's I is calculated as follows:

$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij}(x_i - \overline{x})(x_j - \overline{x})}{S^2 S_0} \tag{3}$$

where x_i is the rice yield, $\{x_i\}$ is the attribute value in position i, $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ indicates the average value of the rice yield, $S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$ represents the variance of the rice yield, $S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}$ represents the sum of the spatial weights of all rice yields in the region, n is the total number of rice production samples, and W_{ij} is an element in the spatial weight matrix W, which refers to the spatial weight between location i and j. We set $W_{ij} = 1$ if the space unit i and j are adjacent, and otherwise, $W_{ij} = 0$.

The key to calculating Moran's I index is the spatial weight matrix of all variables in the region. The calculated result of the spatial weight matrix varies with the variable's adjacency rule algorithm at different spatial positions [39]. The adjacent relationship refers to assigning a spatial weight value according to whether the spatial units are adjacent to each other. If they are adjacent, $W_{ij} = 1$, and otherwise, $W_{ij} = 0$. The distance relationship refers to the preset distance threshold L. If the space unit is adjacent and the distance is less than or equal to L, $W_{ij} = 1$, and otherwise, $W_{ij} = 0$. The nearest K point relationship refers to setting the number of adjacent units or units closest to the space unit as K (the distance between the space units is not considered); if it belongs to one of the sets of K-adjacent or similar spatial units, $W_{ij} = 1$, and otherwise, $W_{ij} = 0$. Considering the applications of different spatial weight matrices and the data in this paper, the study chose the nearest K-point relationship first-order adjacency rule to construct the spatial weight matrix.

Therefore, this paper used the simulated rice yield vector data from Pucheng County, which is based on grid cells, the simulated rice yield in Pucheng County as the research variable, and Moran's I index to analyze the spatial clustering characteristics of the rice production areas. Among them, the spatial adjacency relationship was defined by the K-nearest neighbor (K = 6) first-order adjacency method, and the spatial weight matrix was calculated using this rule.

2.3.3. Pattern Recognition of Rice Yield–Multidimensional Environmental Planting Suitability

Based on evaluating the results of rice planting suitability, this study used k-means clustering to perform attribute clustering. The rice yield and rice planting suitability values in each grid unit were used as the clustering attributes. The number of clusters was based on the determination coefficient, semi-biased determination coefficient [40], and the regional environmental characteristics. Four types of clusters were obtained, and then, variance analysis was used to analyze the soil condition suitability, site condition suitability, climate condition suitability, and irrigation and drainage conditions in different concentration areas. Differential statistical analysis was carried out on the degree and suitability values of mechanical farming conditions to identify the spatial patterns of the rice yield–multidimensional environmental planting suitability. When the environmental planting suitability values from different spatial patterns were statistically significant, the obstacle factors were determined separately for each rice yield–multidimensional environmental planting suitability spatial pattern.

(1) K-Means Clustering

The k-means clustering algorithm is an analysis and calculation method that continuously updates the center value of the cluster through multiple iterations to obtain the optimal solution of the clustering result. The k-means clustering algorithm is popular due to its simplicity, convenience, and high computational efficiency suitable for continuous data [41,42]. The algorithm's specific steps are as follows:

- The rice yield and planting suitability values are selected as clustering indicators.
- The number of clusters based on the determination coefficient, semi-biased determination coefficient, and regional environmental characteristics is determined.
- The distance between each object and each cluster center is calculated.
- The data are divided into the closest clusters.
- Multiple iterations are performed until certain conditions are met and the clustering and partitioning of rice yield–multidimensional environmental planting suitability values are complete.

(2)   Variance analysis

The basic principle of the variance method is to divide the total dispersion of data indicators into two parts: the dispersion caused by level changes and the dispersion caused by errors. Then, the F statistic is calculated and tested to verify the significance of the factors (i.e., determine whether they are usable) [43,44]. The specific calculation process is as follows:

- Various environmental planting suitability values from different clusters are selected for analysis.
- Calculate and count the degrees of freedom and squares of the suitability values from various dimensions in different clusters.
- Calculate the F value and make judgments based on the *p*-value corresponding to the F value.

3. Results

3.1. Rice Yield Simulation Model Comparison

Mean absolute error (MAE) and root mean square error (RMSE) are the two most commonly used indicators to measure the accuracy of the variables. MAE [45] represents the average absolute values of the error between the simulated and actual rice yield values. RMSE [46,47] measures the deviation between the simulated rice yield and the actual rice yield values. The smaller the MAE and RMSE values, the smaller the deviation between the simulated values and the fitted values, the better the simulation performance, and vice versa. R^2 represents the model's fit: that is, the degree to which the variance of the independent variables can be explained by the model results. Its value ranges from 0 to 1. The larger the value of R^2, the better the fit of the model [48,49].

Table 2 shows the MAE, RMSE, and R^2 of the simulated and true rice yield values based on the three models: GWR, MLP, and RFR. It can be seen that the R^2 of RFR is bigger than that of the GWR and MLP models. Therefore, we initially concluded that the RFR model has the best overall fit. Then, we compared the MAE values of different models and found that the RMSE value of RFR is smaller than those of the GWR and MLP models. At the same time, we found that the RMSE value of RFR is also lower than the other models. Then, we comprehensively considered the fitting effect reflected by the three indicators and concluded that the RFR model performs well in terms of fit, model stability, and simulation error, and thus, this study used it to simulate rice yield in Pucheng County.

Table 2. Error comparison of simulation results of different models.

Model	MAE	RMSE	R^2
GWR	836.035	1056.360	0.526
MLP	1056.971	1348.318	0.386
RFR	492.368	765.827	0.762

3.2. Spatial Pattern of Rice Yield

Using the RFR model to simulate the spatial distribution of the rice output in Pucheng County reveals obvious regional differences (Figure 4). The cultivated land in higher-yield areas lies primarily in Xianyang Town, Guancuo Town, Nanpu Town, Wanan Town, Binhe Town, Liantang Town, Yongxing Town, Linjiang Town, and Shipo Town, which are located in the central and southwestern regions of Pucheng County. The moderate-yield and high-yield areas lie primarily in Zhongxin Town, Fuling Town, Shuibeijie Town, Panting Town, Jiumu Town, Guanlu Town, and other peripheral areas of Pucheng County such as the eastern, northern, and southern regions. The low-yield areas lie primarily in the southern and western regions of Pucheng County, such as Gulou Town, Fengxi Town, Shanxia Town, and Haocun Town. According to the statistical data for Pucheng County, Nanping City, and Fujian Province [50], the rice yield in Pucheng County is higher in the

middle, east, and north, and it is lower in the west and south. It can be seen that the spatial distribution of the simulated yield is approximately the same as the actual yield.
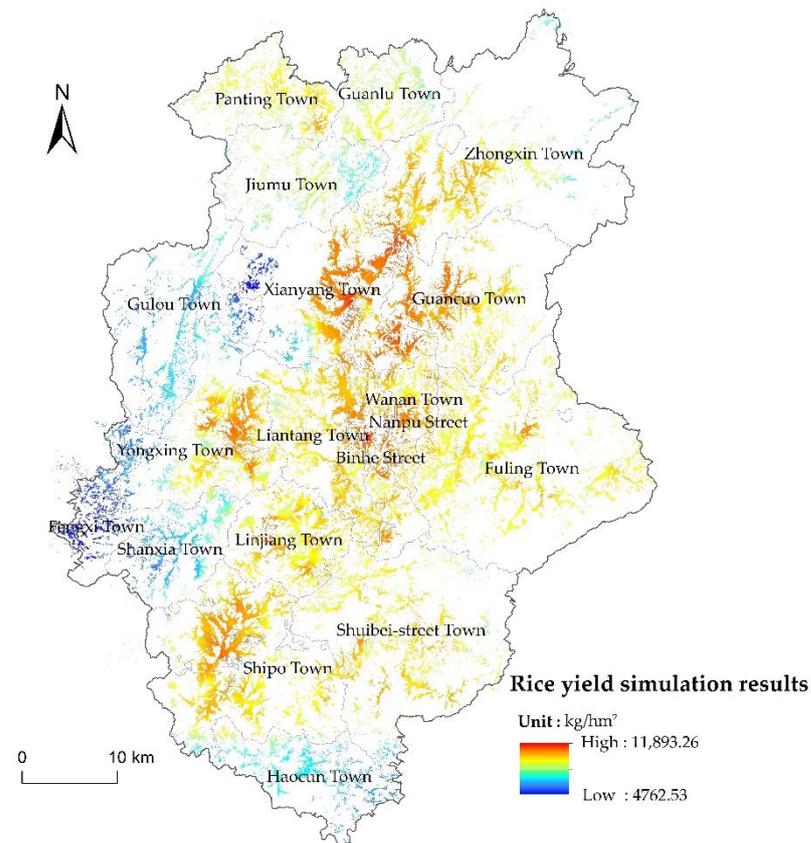

Figure 4. Simulated rice yield results on cultivated land in Pucheng County.

Using ArcGIS and GeoDa, a local Moran's I analysis was carried out on farmland rice yield. The Moran's I value of the farmland rice yield in the entire county was 0.999, which shows a high degree of spatial aggregation, indicating that Pucheng County's rice yield has a very strong positive spatial correlation. This result was further categorized according to the local indicator of spatial analysis (LISA), which clustered the cultivated land by type (Figure 5). The high yield–high yield areas, which have a positive correlation for cultivated rice yield, accounted for 47.65% of the total cultivated area and primarily include Nanpu Town, Wan'an Town, Binhe Town, Liantang Town, Yongxing Town, Linjiang Town, and other areas. These high yield–high yield areas should be developed intensively to improve industrial efficiency while reducing production costs and increasing farmers' income. The number of units in low yield–low yield positively correlated clustering areas accounts for 38.54% of the cultivated land, and these are primarily located in Zaigulou Town, Fengxi Town, Shanxia Town, and Haocun Town. The positive-correlation type accounts for 86.19% of the total cultivated area, indicating that the rice yield from cultivated land in Pucheng County has a strong, positive spatial correlation. The spatial clustering characteristics are relatively obvious. Such areas should adopt methods such as farming reform, fertilization, and soil improvement, and they should improve environmental defects in the area. Low yield–high yield and high yield–low yield clustering areas account for 1.95% and 0.53% of the total cultivated area, respectively. The number of negatively correlated types of units only accounted for 2.48% of the total cultivated area. Rice yields alternate between high and low values in the regions. The factors that affect the productivity of the two types of regions are different; the primary defects should be addressed, and we should prescribe the appropriate solutions to improve regional production capacity.
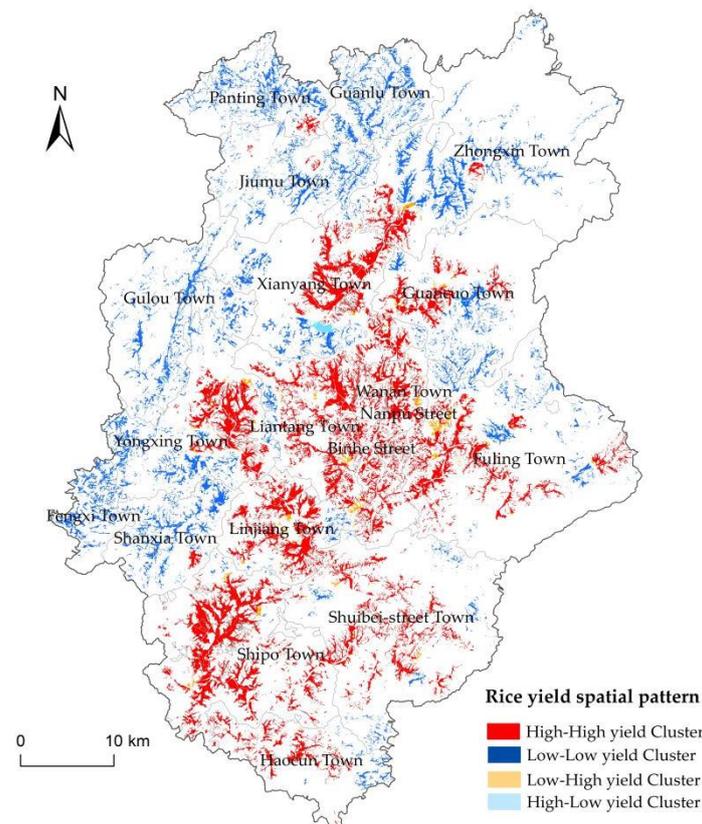
Figure 5. Spatial pattern of rice yield in Pucheng County.

3.3. Spatial Pattern of Rice Yield and Environmental Planting Suitability

To identify the spatial combination patterns of rice yield and rice planting environmental suitability, and to conduct a more in-depth analysis of the spatial characteristics and influencing factors of these different patterns, this study is based on rice yield simulation data and a rice environmental suitability evaluation dataset. Using the planting suitability values and rice yield simulation data as clustering indicators as well as the GeoDa platform, K-means clustering was performed. The number of clusters was determined based on the determination coefficient, semi-biased determination coefficient [40], and regional environmental characteristics, and finally, four types of cluster partitions are obtained. On this basis, the ArcGIS platform was used to perform statistical analysis on the four types of clusters (Table 3).

Table 3. Group analysis results.

Mean	Planting Suitability Value *	Yield (kg/hm^2)
Class I cluster	79.016	10,989.60
Class II cluster	77.737	10,240.50
Class III cluster	72.696	9618.83
Class IV cluster	68.483	8422.50

* $0 \leq$ Planting Suitability Value ≤ 100.

The four types of cluster areas, which include higher yield and higher suitability, high yield and high suitability, moderate yield and moderate suitability, and low yield and low suitability, are hereafter referred to as class I, II, III, and IV clusters (Figure 6).
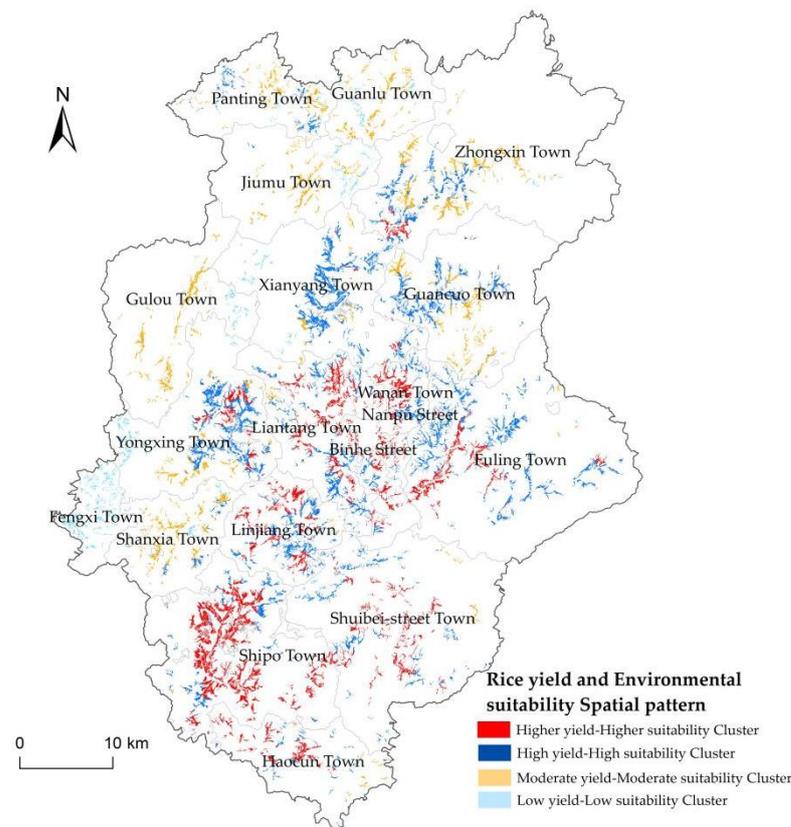
Figure 6. Spatial pattern of rice yield and environmental planting suitability in Pucheng County.

Based on the clustering results, variance analysis was used for various statistical analyses of the soil, site, climate, irrigation and drainage, and mechanical farming condition suitability values in different clustering areas. The results of the variance analysis demonstrated that the environmental planting suitability values from different clustering areas showed highly significant differences ($p < 0.01$) (Table 4).

Table 4. Variance analysis results.

Planting Suitability Value ***	Cluster Area (Mean $\pm$ Standard Deviation)				F	p
	I	II	III	IV		
Irrigation and drainage suitability value	72.59 $\pm$ 20.99	75.07 $\pm$ 18.28	71.50 $\pm$ 19.24	73.49 $\pm$ 14.40	684.281	0.000 **
Mechanical farming suitability value	83.55 $\pm$ 2.67	82.96 $\pm$ 2.99	80.65 $\pm$ 4.16	81.31 $\pm$ 3.76	13,818.739	0.000 **
Site suitability value	80.63 $\pm$ 11.73	80.72 $\pm$ 10.74	76.43 $\pm$ 10.29	67.83 $\pm$ 10.73	14,020.066	0.000 **
Climate suitability value	84.15 $\pm$ 3.15	81.16 $\pm$ 4.46	71.79 $\pm$ 4.94	65.84 $\pm$ 3.14	91,877.973	0.000 **
Soil suitability value	70.10 $\pm$ 8.81	68.95 $\pm$ 9.06	68.48 $\pm$ 9.04	67.38 $\pm$ 7.68	924.390	0.000 **

** $p < 0.01$ *** $0 \leq$ Suitability Value ≤ 100.

3.4. Rice Yield and Multidimensional Environmental Planting Suitability Spatial Model

Rice planting suitability and rice yield are affected by many factors, such as the soil, site, climate, irrigation and drainage, and mechanical farming. This study considered the actual Pucheng County soil, site, climate, irrigation and drainage, and mechanical farming conditions, which were selected from 20 indicators in five dimensions to analyze the rice yield and environmental planting suitability factors (Table 1). The multidimensional environmental characteristics were analyzed based on the grouping results and the spatial distribution characteristics of the environmental indicators (Figure 7). The influencing factors on various clusters of cultivated land in Pucheng County were determined based on

these results. Finally, we obtained the rice yield and the environmental planting suitability spatial model for Pucheng County (Figure 8).
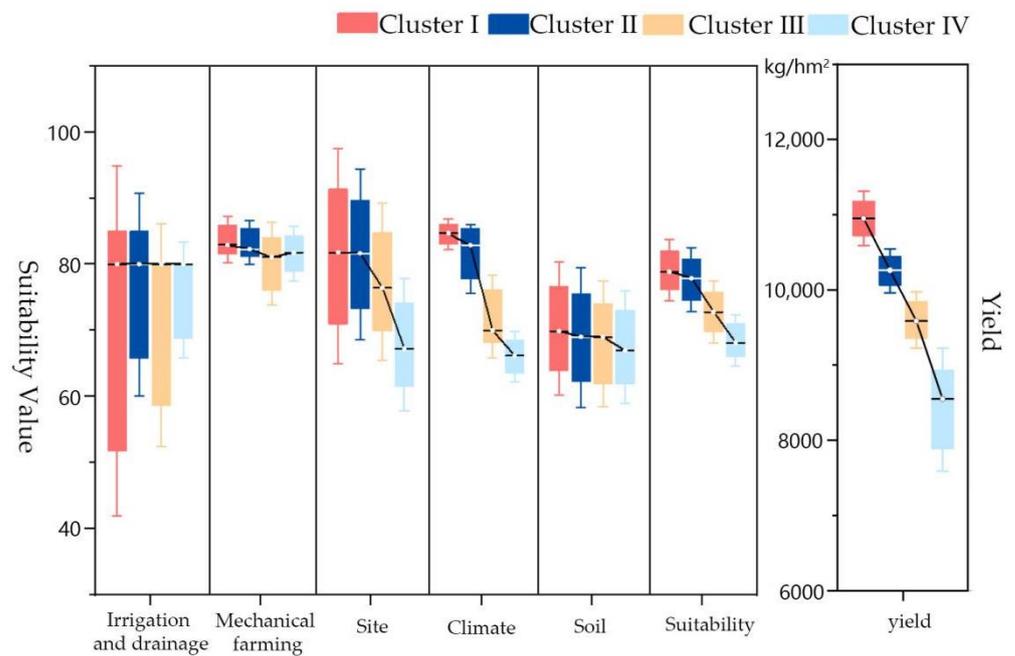

Figure 7. Multidimensional environmental suitability characteristics of the four clusters.
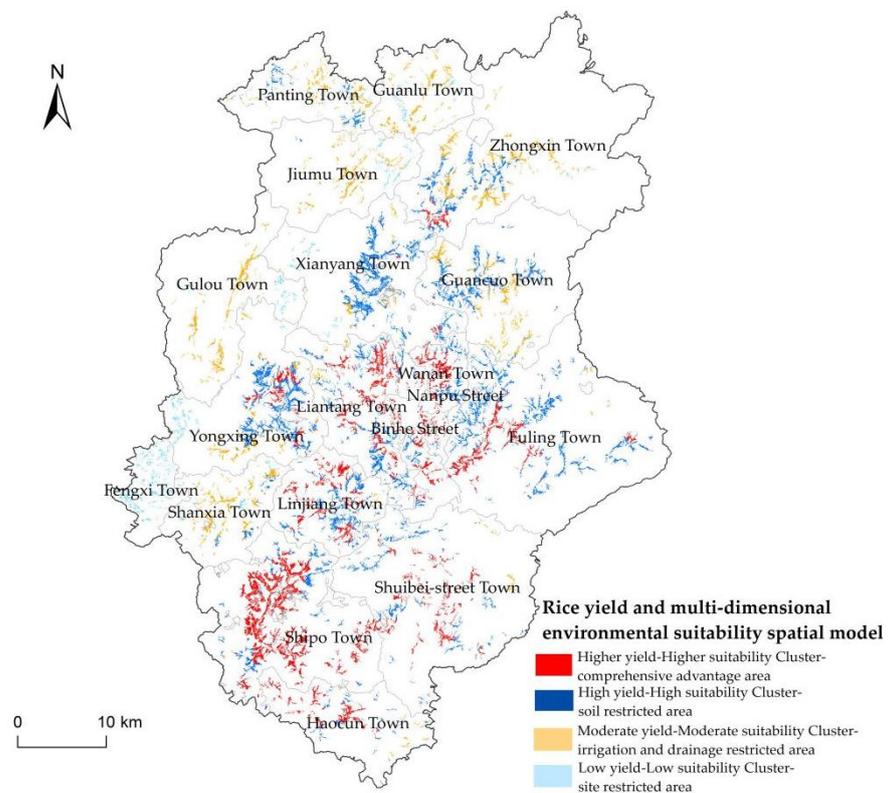

Figure 8. Rice yield and multidimensional environmental planting suitability spatial model.

3.5. Identification of Obstacle Factors and Planting Guidance

3.5.1. Higher Yield, Higher Suitability—Comprehensive Environmental-Advantage Areas

This type of arable land accounts for 32.46% of the county's total arable land and lies primarily in Shipi Town, Haocun Town, central Beishuijie Town, Liantang Town, Nanpu Street, Binhe Street, and other areas. These areas are characterized by flat terrain and excellent indicators, among which the climate and mechanical farming conditions have obvious advantages (Figure 7). The environmental characteristics of the clustering area are as follows: the ≥10 °C accumulated temperature >5000 °C, the annual average temperature >14.0 °C, the annual average sunshine hours >1500 h, and there is abundant light and heat provided by a beneficial climate; in addition, the soil nutrients are relatively high. The basic infrastructure and various other conditions are relatively good. Generally speaking, these areas meet the environmental conditions required for the growth and development of rice. For higher-yield and higher-suitability areas, intensive development should be adopted to make full use of these conditions to ensure high yields, reduce production costs, improve farmers' income, and effectively protect the fertility of the existing soil.

3.5.2. High Yield, High Suitability—Soil Condition-Limited Areas

This type of arable land accounts for 42.92% of the county's total arable land and is located primarily in Linjiang Town, Fuling Town, Xianyang Town, west of Guancuo Town, Yongxing Town, and other areas. The irrigation and drainage, site, and mechanical farming conditions in these areas are good, However, the soil condition suitability index is lower than that of higher yield and higher suitability areas (Figure 7). Combined with the local arable land fertility survey sampling data and soil testing as well as formula fertilization pilot data, we comprehensively analyzed the spatial distribution characteristics of various indicators, such as organic matter, available potassium, total nitrogen, and pH. The results revealed that available phosphorus is relatively scarce, and most of the arable land is sandy loam and gravel soil with poor soil texture; therefore, the biggest obstacle in this area is the soil quality. All regions should fertilize reasonably; increase soil nitrogen, phosphorus, and potassium content; focus on applying organic fertilizers; control soil pollution; moderately adjust soil pH; strengthen soil improvement measures; popularize high-quality rice varieties with good planting adaptability; and increase the rice planting quality to guarantee rice production.

3.5.3. Moderate Yield, Moderate Suitability—Irrigation and Drainage Condition-Limited Areas

This type of arable land accounts for 19.03% of the county's total arable land and is located primarily in Shanxia Town, Gulou Town, west of Jiumu Town, in the middle of Guanxi Town, in the middle of Zhongxin Town, and south of Guancuo Town. The irrigation and drainage suitability in this area is low, as most of the arable land has "normal" or "poor" irrigation conditions, and the drainage capacity is "average". Therefore, irrigation conditions and drainage capacity are the primary obstacles affecting the planting suitability in the clustering area (Figure 7). For this type of arable land, farmland improvement should be carried out first, water conservancy facilities should be built, field irrigation channels should be optimized, detailed farmland management should be strengthened, and engineering funds should be invested appropriately to ensure water conservancy for the arable land.

3.5.4. Low Yield, Low Suitability—Site Condition-Limited Areas

This type of arable land accounts for 5.58% of the county's total arable land and is located primarily in Fengxi Town, east of Jiumu Town, and west of Xianyang Town. Such areas lie primarily in high-altitude areas with steep slopes. A comprehensive analysis of the spatial distribution characteristics of various indicators revealed that the biggest obstacle is site condition (Figure 7). The high altitude results in a lower average temperature. Simultaneously, this land type is located in a mountainous area, which often results in insufficient sunshine. Furthermore, the mountain-dominated topography causes the soil

in this area to be dominated by sand, and the soil elements such as organic matter are also low. Various factors have led to lower production in this area. Therefore, for such areas, scientific planting should be strengthened, and appropriate varieties of rice should be selected and combined with the changes in regional temperature to achieve "cold tail and warm head, grab clear sowing". The area can also adopt indirect irrigation to ensure soil moisture to increase the regional production capacity as much as possible.

4. Discussion

4.1. Rice Yield Simulation Model

The method of rice yield simulation can solve the time-consuming and labor-intensive problem of traditional on-site sampling survey methods, which will improve the efficiency of rice yield acquisition. Jianghua [51] used the improved ORYZA model to simulate rice yield with different soil organic carbon content, and Roushani et al. [52] used the AquaCrop model to simulate and predict rice yield based on water management conditions. However, due to the complexity of the relationship between rice yield and the environmental factors, the use of a single simulation model may not produce adequate results. Therefore, it is important to construct different rice yield simulation models and compare their accuracy and fitting to choose a high-precision rice yield simulation model.

In this paper, we constructed rice yield simulation models based on GWR, MLP neural networks, and RFR. Among these, the GWR model found the mathematical expression that best represents the relationship between output and influencing factors by studying the relationship between multiple variables; however, it is not applicable when there are many influencing factors or nonlinear factors [53]. Although the MLP model can adequately solve nonlinear problems [32,54], analyzing many influencing factors with a small amount of training data causes the performance to be unstable, which may cause the model to produce significant errors between the simulated yield results and actual rice yields [55]. The RFR model can solve both linear and nonlinear complex problems, accounting for multiple factors in crop yield simulation [56], and its algorithm performance is also better than the MLP model. In addition, the model construction process is relatively straightforward [57,58]. The results show that using the RFR model to simulate rice yield produces the RAE and RMSE that are smaller than those of the other two models, and it has the best fit.

4.2. Rice Yield Has a Strong Spatial Clustering

Studying the spatial rice yield distribution pattern is of great significance to understanding the environmental suitability of different planting areas. Pucheng County, as a typical example, is located in a mountainous and hilly area with a complex topography, which is of great significance because many farmland regions exhibit this topography. We found that although Pucheng County is located in a mountainous region, the rice yield of cultivated land in the county has strong spatial clustering characteristics. The cultivated land rice output has a strong spatial positive correlation: the proportion of positive correlation types in the spatial clustering area of cultivated land rice yield is as high as 86.19%. Its spatial clustering characteristics are more obvious. Therefore, it is necessary to effectively implement intensive rice planting in this area to realize precision agriculture.

4.3. Rice Yield–Multidimensional Environmental Planting Suitability Spatial Pattern Recognition

Mastering the rice yield–multidimensional environmental planting suitability spatial combination model, and analyzing the environmental factors using different models, can provide scientific guidance for taking advantage of regional environmental resources and optimizing the grain planting structure and industrial layout. In previous research, Lobell et al. [59] studied the spatial distribution of global crop production under the influence of the climate, and Lasini et al. [60] used statistical and machine learning techniques to analyze the spatial relationship between rice yield and climate variables in a major region in SriLanka. Ostrowski et al. [61] analyzed the effects of climate and environmental

factors on European grass species and the influence of the spatial distribution pattern of the relationship with wheat planting. Gazolla-Neto et al. [62] used precision farming techniques to assess the spatial dependence between soil chemistry and yield components in soybeans. Resop et al. [63] constructed a new crop model based on natural factors, such as weather and soil, and analyzed the spatial distribution of potato production in Maine using that model. Parry et al. [64] analyzed the impact of global human and climate scenarios on the spatial distribution of cereal production based on the Special Report on Emissions Scenarios (SRES). These studies analyzed only some environmental factors and rice yields, and they did not systematically identify and analyze the combined patterns of yield and multidimensional environmental spatial distribution. However, our research found that in the actual agricultural production process, yield is often the result of a combination of many factors. By effectively identifying the combination model of yield and planting suitability, the shortcomings and advantages of different regions in terms of climate, soil, irrigation, site, and mechanical farming can be obtained. Therefore, in the process of rice production, it is necessary to adjust planting strategies according to different rice yield–multidimensional environmental suitability combination model areas, so as to improve the yield and quality of rice.

5. Conclusions

We have constructed a method system for the identification of rice yield and multidimensional environmental suitability spatial combination pattern at a fine scale. First, we constructed a rice yield simulation model that simulates the refined spatial distribution of rice yield at the county level, analyzed the rice yield spatial pattern, and finally identified the spatial combination pattern of rice yield and multidimensional environmental suitability. Our research results show that Pucheng County's rice output has a strong positive spatial correlation with obvious spatial clustering characteristics. The spatial combination pattern recognition of rice yield and multidimensional environmental suitability provides Pucheng County with regional rice planting guidance, and at the same time, it provides a new perspective for exploring the spatial pattern of the rice planting environment. However, there several limitations of this research. First, the impact of human activity is not comprehensive, and due to the lack of data on pests and diseases, this study does not take these into consideration. Second, the feasibility of mechanization of agricultural operations is an important indicator to measure the suitability of cultivated land. However, due to limited data, our measurement of the feasibility of mechanization is not perfect. Future work will improve the relevant indicator system of the model. Third, when simulating rice yield, the sensitivity analysis of numerical models can be used to realize the output variations of numerical models are studied in function of the input variations. Therefore, the sensitivity analysis of the numerical model can be considered to optimize the rice yield simulation model in future research.

Author Contributions: Conceptualization, Daichao Li, Jianqin Liang and Xingfeng Wang; methodology, Daichao Li, Jianqin Liang and Xingfeng Wang; validation, Daichao Li and Jianqin Liang; formal analysis, Xingfeng Wang, Xiaowei Xie and Jiaqi Lu; data curation, Xiaowei Xie and Jiaqi Lu; writing—original draft preparation, Daichao Li and Xingfeng Wang; writing—review and editing, Daichao Li and Jianqin Liang; supervision, Daichao Li and Sheng Wu; project administration, Daichao Li and Sheng Wu. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data of this work can be shared to the readers depending on the request.

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

References

1. Ayoade, M.A. Suitability assessment and mapping of Oyo State, Nigeria, for rice cultivation using GIS. *Theor. Appl. Climatol.* **2017**, *129*, 1–14. [CrossRef]
2. Tang, J.; Wu, W.; Yang, P.; Zhou, Q.; Chen, Z. Research progress of remote sensing monitoring of crop spatial pattern. *Chin. Agric. Sci.* **2010**, *43*, 2879–2888.
3. Du, Y.; Zhang, W.; Wu, X.; Li, G.; Wang, S.; Liu, Z.; Tang, S.; Gong, Y. Temporal and Spatial Variation Characteristics of Rice Yield in Jiangsu Province. *J. Nanjing Agric. Univ.* **2014**, *37*, 7–12.
4. Yi, Z.; Tian, P.; Deng, M.; Fei, Y.; Wang, Z.; Wang, X.; Hu, Z.; Yang, D.; Zhao, P. Correlation analysis between annual rice yield fluctuations and meteorological factors in the growing season. *Jiangsu Agric.* **2021**, *49*, 51–63.
5. Sun, Y.; Cao, Y.; Fu, Q.; Gu, C. Soil Temperature Changes and Rice Yield Effect under Water-saving Irrigation Conditions in the Area of Well Irrigated Rice in Cold Region. *J. Irrig. Drain.* **2008**, *27*, 67–70.
6. Wang, M. Research on Yield Estimation Method Based on the Simulation of Rice Yield Elements. Master's Thesis, Nanjing University of Information Technology, Nanjing, China, 2016.
7. Tang, H. Simulation Study on Rice Yield under Different Irrigation System. *Water Conserv. Tech. Superv.* **2018**, *3*, 74–76, 108.
8. Xiong, W.; Tao, F.; Xu, Y.; Lin, D. Simulation of my country's rice yield changes under climate change scenarios. *Chin. Agric. Meteorol.* **2001**, *22*, 2–6.
9. Feng, H. Research on Dynamic Prediction Technology of Rice Yield at County Level Based on Crop Model and GIS. Master's Thesis, Anhui Agricultural University, Anhui, China, 2010.
10. Shen, X. Research on the Economic and Environmental Effects of Rice Planting Patterns and Their Spatial Layout Optimization Strategies. Ph.D. Thesis, Huazhong Agricultural University, Hubei, China, 2020.
11. Du, C. Research on Spatio-temporal Simulation of Chinese Rice Yield Based on DSSAT Model. *Mod. Agric. Technol.* **2015**, *10*, 211–213, 217.
12. Liu, H.; Wu, W.; Shen, G.; Huang, Q. Spatiotemporal evolution of traditional soybean planting structure in Songnen Plain from 1996 to 2016. *J. Appl. Ecol.* **2014**, *69*, 353–364.
13. Gómez, D.; Salvador, P.; Sanz, J.; Casanova, J.L. Modelling wheat yield with antecedent information, satellite and climate data using machine learning methods in Mexico. *Agric. For. Meteorol.* **2021**, *300*, 108317. [CrossRef]
14. Liu, X.; Zhang, C.; Shuai, J.; Wang, P.; Shi, W.; Chen, Y.; Tao, F. The effect of chilling injury on rice yield in Heilongjiang Province. *Acta Geogr. Sin.* **2012**, *67*, 1223–1232.
15. Li, Q.; Kang, G.; Li, X.; Li, B. Factors Influencing Grain Production of Henan Province Based on Gray Correlation. *Asian Agric. Res.* **2009**, *1*, 23–27.
16. Qi, J.-Y.; Zhang, A.L.; Su, Y.-Q. Analysis of Factors Influencing on Rice Yield in Main Rice Production Areas in Yongsheng County, Yunnan Province. *J. Yunnan Agric. Univ.* **2010**, *25*, 868–874.
17. Wang, L.; Xu, X.; Li, Q.; Su, H.; Li, M. Effects of high temperature stress in different periods on the growth period and yield of rice in Jiangsu Province. *Crop. J.* **2015**, *2*, 95–100.
18. An, Y.; Tan, X.; Tan, J.; She, H.; Wang, Z.; Li, W. Evolution and influencing factors of crop planting structure in Hunan Province. *Econ. Geogr.* **2021**, *41*, 156–166.
19. Zeng, Z. A review of research on the impact of climate change on rice production. *Northeast. Agric. Sci.* **2021**, *46*, 89–93.
20. Whetton, R.L.; Zhao, Y.; Nawar, S.; Mouazen, A.M. Modelling the Influence of Soil Properties on Crop Yields Using a Non-Linear NFIR Model and Laboratory Data. *Soil Syst.* **2021**, *5*, 12. [CrossRef]
21. Luh, Y.-H.; Chang, Y.-C. Effect of Climate Change on Staple Food Production: Empirical Evidence from a Structural Ricardian Analysis. *Agronomy* **2021**, *11*, 369. [CrossRef]
22. Xiao, D. Reasonable distribution and high-yield cultivation of winter green manure in Jiangxi. *Jiangxi Agric. Sci. Technol.* **1981**, *9*, 7–9.
23. Tan, J.; Li, Z.; Yang, P.; Liu, Z.; Li, Z.; Zhang, L.; Wu, W.; You, L.; Tang, H. Temporal and spatial distribution characteristics of spring corn in the three provinces of northeast china based on crop spatial distribution model. *Acta Geogr. Sin.* **2014**, *69*, 353–364.
24. Huang, Y. Research on Development Strategy of Characteristic Agriculture Industrialization in Pucheng County. Master's Thesis, Fujian Agriculture and Forestry University, Fujian, China, 2014.
25. Jiang, S. Impacts of Climate and Soil Temporal and Spatial Changes in Pucheng County on Cultivated Land's Potential for Food Crops. Master's Thesis, Fujian Agriculture and Forestry University, Fujian, China, 2010.
26. Wang, X.F.; Li, D.C.; Wu, S.; Lu, J. Research on comprehensive suitability evaluation method of rice planting environment. *J. Geo-Inf. Sci.* **2021**, *23*, 1484–1496.
27. Li, W.X.; Zhang, Y.P.; Mao, W.; Chen, X.; Wang, C.; Zhang, B. Study on cultivated land suitability evaluation for rice and index determination. *J. Agric.* **2013**, *3*, 19–24.
28. Xue, J.; Han, J.; Liu, Y.; Zhang, F. Changes of grain production pattern and its influencing factors in Henan province. *Reg. Res. Dev.* **2013**, *32*, 150–155.
29. Dong, X.; Yang, Y.; Zhao, X. Research on resource misallocation and China's industrial total factor productivity and regional differences under spatial effects. *Study Pract.* **2021**, *4*, 45–56.

30. Sassi, M. OLS and GWR Approaches to Agricultural Convergence in the EU-15. *Int. Adv. Econ. Res.* **2010**, *16*, 96–108. [CrossRef]
31. Phitakwinai, S.; Auephanwiriyakul, S.; Theera-Umpon, N. Multilayer perceptron with Cuckoo search in water level prediction for flood forecasting. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016.
32. Chen, W.; Qi, W.; Yuan, F.; Li, Z. Comparative analysis of rice yield forecast in Jilin Province based on time series and cross-sectional data. *China Agric. Inf.* **2018**, *30*, 91–101.
33. Yang, B.; Chen, S.; Yu, H.; An, T. Remote sensing estimation of rice yield based on random forest regression method. *J. China Agric. Univ.* **2020**, *25*, 26–34.
34. Liu, J.; Dong, Z.; Xia, J.; Wang, H.; Meng, T.; Zhang, R.; Han, J.; Wang, N.; Xie, J. Estimation of soil organic matter content based on CARS algorithm coupled with random forest. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2021**, *258*, 119823. [CrossRef]
35. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
36. Tiefelsdorf, M. Some practical applications of Moran's I's exact conditional distribution. *Pap. Reg. Sci.* **1998**, *77*, 101–129. [CrossRef]
37. Tobler, W. A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* **1970**, *46*, 234–240. [CrossRef]
38. Ge, Y.; Ren, Z.; Fu, Y. Understanding the Relationship between Dominant Geo-Environmental Factors and Rural Poverty in Guizhou, China. *ISPRS Int. J. Geo.-Inf.* **2021**, *10*, 270. [CrossRef]
39. Liu, Y.; Liu, C.; He, Z.; Zhou, X. Basic farmland delineation based on local spatial autocorrelation of farmland quality at pixel scale. *J. Agric. Mach.* **2019**, *50*, 260–268.
40. Bian, W.; Qiu, X.; Liu, L.; Jia, X. Evaluation method of flexible network opening parameters based on orthogonal simulation experiment. *Comput. Meas. Control.* **2021**, *29*, 238–243, 255.
41. Xia, X.; Gai, J. Urban rail transit station classification and passenger flow characteristic analysis based on K-Means clustering algorithm. *Mod. Urban Rail Transit* **2021**, *4*, 112–118.
42. Hemade, B.A.; Ibrahim, H.A.; Talaat, H. Steady-state Security Assessment Based on K-Means Clustering Algorithm and Phasor Measurement Units. *Recent Adv. Electr. Electron. Eng. (Former. Recent Pat. Electr. Electron. Eng.)* **2020**, *13*, 559–570. [CrossRef]
43. Shen, F.; Han, Y.; Lv, G.-L.; Ma, X.; Tang, Z. Variance Analysis on Environmental Effect of Yield and Yield Components of Rice Varieties newly bred in Liaoning. *North Rice* **2007**, *3*, 72–75.
44. Asadi, A.A.; Valizadeh, M.; Mohammadi, S.A.; Manochehr, K. Genetic Analysis of Response to Water Deficit Stress in Wheat Yield Traits with Generation Means and Variance Analysis. *J. Crop Breed.* **2019**, *11*, 88–99.
45. Chen, X. A Condition of Nonparametric Regression Estimator Exponential Convergence Rate of Mean Absolute Error. *Acta Math. Sin.* **1987**, *4*, 433–443.
46. Sun, D.; Wang, Y.; Huang, Q.; Zhang, L.; Xiao, Y. Selection of typical years for reservoir groups based on the minimum root mean square error criterion. *J. Xi'an Univ. Technol.* **2011**, *3*, 275–279.
47. Wang, W.; Lu, Y. Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *324*, 012049. [CrossRef]
48. Guo, L.; Zhang, H.; Cheng, J.; Li, J.; Qin, C. Comparison of Soil Attribute Spatial Prediction Based on Co-Kriging Interpolation and Geographically Weighted Regression Model. *Acta Pedol. Sin.* **2012**, *49*, 1037–1042.
49. Zhang, J. Research on Spatial Differentiation of Urban Residential Land Price Based on GWR Model. Master's Thesis, Zhejiang University, Zhejiang, China, 2012.
50. *2020 China Statistical Yearbook. Statistical Theory and Practice*; China Statistics Press: Beijing, China, 2021; Volume 2.
51. Jiao, J. Rice Yield Simulation and Model Improvement Dominated by Meteorological Factors under Different Soil Organic Carbon Content. Master's Thesis, Chinese Academy of Agricultural Sciences, Beijing, China, 2020.
52. Roushani, S.; Mohassel, M.H.; Haghighi, R.S.; Amiri, E. The Assessment of AquaCrop Model in Predicting Rice Genotypes Grain and Biological Yield under Water Management Conditions. *Pol. J. Environ. Stud.* **2021**, *30*, 2283–2291. [CrossRef]
53. Xiao, M. *Research on the Correlation between Tourism POI and Regional Economy Based on Geographic Weighted Regression*; Southwest Jiaotong University: Xi'an, China, 2019.
54. Liu, M.; Jiang, Y. Research on port throughput forecast based on improved multilayer perceptron model. *Softw. Eng.* **2021**, *24*, 39–42.
55. Yocu, U.; Bas, E.; Egrioglu, E.; Aladag, C.H. A new multilayer feedforward network based on trimmed mean neuron model. *Neural Netw. World* **2015**, *25*, 587–602.
56. Zhang, X.; Chen, Y.; Xiao, G. A review of research on crop yield prediction based on machine learning. *Anhui Agric. Sci. Bull.* **2021**, *27*, 117–119.
57. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput. Electron. Agric.* **2018**, *151*, 61–69. [CrossRef]
58. Elavarasan, D.; Vincent, D.R.; Sharma, V.; Zomaya, A.Y.; Srinivasan, K. Forecasting yield by integrating agrarian factors and machine learning models: A survey. *Comput. Electron. Agric.* **2018**, *155*, 257–282. [CrossRef]
59. Lobell, D.B.; Schlenker, W.; Costa-Roberts, J. Climate Trends and Global Crop Production since 1980. *Science* **2011**, *333*, 616–620. [CrossRef]
60. Wickramasinghe, L.; Weliwatta, R.; Ekanayake, P.; Jayasinghe, J.; Ghatee, M. Modeling the Relationship between Rice Yield and Climate Variables Using Statistical and Machine Learning Techniques. *J. Math.* **2021**, *2021*, 6646126. [CrossRef]

61. Ostrowski, M.; Prosperi, J.; David, J. Potential Implications of Climate Change on Aegilops Species Distribution: Sympatry of These Crop Wild Relatives with the Major European Crop *Triticum aestivum* and Conservation Issues. *PLoS ONE* **2017**, *11*, e0153974. [CrossRef] [PubMed]
62. Gazolla-Neto, A.; Fernandes, M.C.; Vergara, R.; Gadotti, G.; Villela, F. Spatial distribution of the chemical properties of the soil and of soybean yield in the field. *Rev. Cienc. Agron.* **2016**, *47*, 325–333. [CrossRef]
63. Resop, J.P.; Fleisher, D.H.; Wang, Q.; Timlin, D.J.; Reddy, V.R. Combining explanatory crop models with geospatial data for regional analyses of crop yield using field-scale modeling units. *Comput. Electron. Agric.* **2012**, *89*, 51–61. [CrossRef]
64. Parry, M.L.; Rosenzweig, C.B.; IgIesias, A.C.; Livermore, M.D.; Fischer, E.G. Effects of climate change on global food production under SRES emissions and socio-economic scenarios. *Glob. Environ. Chang.* **2004**, *14*, 53–67. [CrossRef]