Research Article

Prediction of Grain Yield in Henan Province Based on Grey BP Neural Network Model

Bingjun Li , Yifan Zhang , Shuhua Zhang , and Wenyan Li

College of Information and Management Science, Henan Agricultural University, Zhengzhou 450046, China

Correspondence should be addressed to Yifan Zhang; zyf185423@163.com

Received 3 April 2021; Revised 14 August 2021; Accepted 19 August 2021; Published 28 August 2021

Academic Editor: Filippo Cacace

Copyright © 2021 Bingjun Li et al. 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.

BP neural network (BPNN) is widely used due to its good generalization and robustness, but the model has the defect that it cannot automatically optimize the input variables. In response to this problem, this study uses the grey relational analysis method to rank the importance of input variables, obtains the key variables and the best BPNN model structure through multiple training and learning for the BPNN models, and proposes a variable optimization selection algorithm combining grey relational analysis and BP neural network. The predicted values from the metabolic GM (1, 1) model for key variables was used as input to the best BPNN model for prediction modeling, and a grey BP neural network model prediction model (GR-BPNN) was proposed. The long short-term memory neural network (LSTM), convolutional neural network (CNN), traditional BP neural network (BP), GM (1, N) model, and stepwise regression (SR) are also implemented as benchmark models to prove the superiority and applicability of the new model. Finally, the GR-BPNN forecasting model was applied to the grain yield forecast of the whole province and subregions for Henan Province. The forecasting results found that the growth rate of grain production in Henan Province slowed down and the center of gravity for grain production shifted northwards.

1. Introduction

Grain is not only an important strategic material related to the national economy and people's livelihood but also the most basic means of livelihood for the people. Scientific analysis and prediction of grain yield are of great significance to the harmonious and stable development of society and the maintenance of national food security. Henan Province is the main grain export province in China and is the first wheat export province, with the production of all major agricultural products ranking steadily among the top in China. As one of the important grain production areas in China, Henan Province has made great contributions to national food security. Therefore, scientific statistics of Henan Province grain yield data and reasonable prediction of its development trend is helpful to stabilize grain production and guarantee food security.

There are various methods for forecasting grain yield. Traditional grain yield prediction methods mainly include the empirical method, exponential smoothing method,

linear regression method, and time series analysis method. These methods are simple and easy to realize, but they are only applicable to short-term grain yield prediction and are still insufficient in mining complex data information, which has great limitations [1]. In addition, grain production is often influenced by a variety of complex factors such as meteorology, land, human use activities, and institutions making accurate forecasting of grain yield very difficult. In recent years, neural network models have become one of the research hotspots of scholars at home and abroad. For example, LSTM, CNN, and BPNN have shown high accuracy and high timeliness when dealing with multivariate, multitemporal heterogeneous data and mining of nonlinear data [2, 3]. The neural network models have proven their power in data mining and agricultural analysis, including crop type classification and yield prediction [4].

Scholars have carried out a great deal of in-depth research into grain production systems using a variety of different methods, among which the BPNN is one of the most widely used important methods [5–9]. In contrast,

BPNN has received increasing attention due to its fast convergence, high accuracy, and strong nonlinear mapping capability. Gu et al. [10] integrated particle swarm optimization algorithm and BPNN to construct NCPSSO-BP prediction model algorithm to solve the complex problem of grain yield prediction. Li et al. [11] and Zhang and Pan [12] studied the simulation ability and new data prediction ability of multiple linear regression model and BPNN model, and the results showed that the BPNN model was better than the linear regression model in accuracy, stability, generalization degree, and theoretical basis. Some scholars have tried to combine methods such as rough sets, genetic algorithms, and linear regression to improve the prediction speed and accuracy [13, 14]. Guo [15] and Zheng [16] both established a combination prediction model combining principal component analysis and BPNN, which improved the learning convergence speed and prediction accuracy of the neural network model. Wang and Zhu [17] combined the BPNN model with other forecasting models to construct a multiscale combined forecasting model. But because these scholars did not consider the applicability of the combination method and application objects, overfitting problems occur occasionally. Some scholars have also used different algorithms to optimally filter the initial weights and thresholds of BPNN, thus building prediction models suitable for small samples [18]. Hu et al. [19] established the IPSO-BP grain yield prediction model by introducing reproduction and mutation mechanisms and using improved particle swarm optimization (IPSO) to optimize the connection weights and thresholds of the BPNN model. Rong et al. [20] carried out multiple screening and comparisons on the internal nodes and thresholds of the BPNN layer and obtained the optimal network structure. The application results showed that the prediction accuracy of grain output had been significantly improved. But they did not identify the importance of the input variables; the improvement of prediction accuracy has certain limitations. On the whole, scholars have improved and optimized the prediction accuracy of BPNN models from different perspectives, and the results of the relevant research results are remarkable and have a good reference significance.

However, these optimization measures have improved the accuracy of the model to a certain extent. But due to the screening and judgment that the priority of the indicator variables is ignored, it is impossible to fundamentally change the defect that the BPNN model cannot automatically optimize multiple variables, which leads to the learning speed converging slowly and easily falling into the situation of local minimums. The correct choice of input variables determines the validity of the prediction results, while few scholars have considered improving the prediction accuracy of BPNN from the perspective of variable optimization. In addition, grain production is a process of multiple factors acting together, and it is a complex nonlinear dynamic system that includes grayness, randomness, and uncertainty. Existing studies have insufficiently considered the grayness and randomness of data information. The LSTM and CNN, which are similar to BPNN as neural network models, are also lacking at the

level of input variable identification, resulting in these network models still suffering from computationally time-consuming and gradient vanishing defects. The improvement in prediction accuracy is often limited by the omission of valuable time-series information.

One of the key objectives of this study is to enhance the adaptability of BPNN and make it more effective in modeling multivariable complex systems by optimizing the number of input variables. Grey relational analysis is an important content of grey system theory [21], which can provide a basis for advantage analysis, factor discrimination, etc. This method can solve the problem of optimal selection of input variables of the BP neural network algorithm on the basis of taking into account the grey nature of forecast information in just the right way. Furthermore, as a tool for dealing with uncertain information, the method is able to reflect more accurately the actual state of its external information in dealing with a wide range of uncertainties, such as randomness, ambiguity, and greyness, and has a good variable screening capability.

In recognition of the strong nonlinear approximation ability of BP neural networks and the variable screening ability of grey correlation, this study combines grey system theory with BPNN and uses grey relational analysis to improve the BPNN for its shortcomings of not being able to identify the priority and importance of input variables and constructs the variable optimization selection algorithm. In addition, models such as long and short-term memory network LSTM and convolutional neural network CNN are used as benchmark models to test the prediction performance of the proposed GR-BPNN model [22]. Finally, the GR-BPNN was developed by integrating the variable optimization selection algorithm with the metabolic GM (1, 1) model. Through the application of GR-BPNN in grain yield forecasting in Henan Province, another key objective of this study is to improve the accuracy of grain yield forecasting data in Henan Province, providing a reliable basis for the formulation of national grain policies, and also providing a new way to quantify and intellectualize grain yield forecasting.

The rest of this paper is organized as follows. Section 2 presents the main steps of the GR-BPNN prediction model. Section 3 verifies the prediction performance of the GR-BPNN model by comparing it with several commonly used benchmark models. Section 4 gives the prediction result and analysis of grain yield in Henan Province. Section 5 draws conclusions and puts forward relevant countermeasures and suggestions for increasing grain yield in Henan Province.

2. Construction of GR-BPNN Model

2.1. Variable Optimization Selection Algorithm. In order to filter out the key input variables and overcome the defect that BPNN cannot automatically optimize multiple variables, this study combined the grey relational analysis and BPNN to build a variable optimization improved algorithm to improve the BPNN model's recognition ability of important variables. The specific steps are as follows.

2.1.1. Determine Input Variable Priority. In this study, grey relational analysis was used to rank the importance of the input variables and prioritize them. The modeling steps are as follows.

Set $Y'_1, Y'_2, \dots, Y'_j, \dots, Y'_l$ as the sequence of system characteristic behaviors, $X'_1, X'_2, \dots, X'_i, \dots, X'_m$ is the correlation factor sequence, and X'_i and Y'_j have the same length, where

$$\begin{aligned} Y'_j &= (y'_j(1), y'_j(2), \dots, y'_j(k), \dots, y'_j(n)), \quad j = 1, 2, 3, \dots, l, \\ X'_i &= (x'_i(1), x'_i(2), \dots, x'_i(k), \dots, x'_i(n)), \quad i = 1, 2, 3, \dots, m. \end{aligned} \quad (1)$$

Obtain the initial image of each sequence. Make

$$X_i = \frac{X'_i}{x'_i(1)} = (x_i(1), x_i(2), \dots, x_i(n)), \quad i = 1, 2, 3, \dots, m, \quad (2)$$

$$Y_j = \frac{Y'_j}{y'_j(1)} = (y_j(1), y_j(2), \dots, y_j(n)), \quad j = 1, 2, 3, \dots, l. \quad (3)$$

Strive for the degree of incidence.

$$\gamma(x_i(k), y_j(k)) = \frac{\min_i \min_k |y_j(k) - x_i(k)| + \xi \max_i \max_k |y_j(k) - x_i(k)|}{|y_j(k) - x_i(k)| + \xi \max_i \max_k |y_j(k) - x_i(k)|}, \quad (4)$$

is Deng's grey incidence coefficient. $\rho \in (0, 1)$ is the distinguishing coefficient and generally takes $\rho = 0.5$. Well call

$$\gamma_{ij} = \gamma(X_i, Y_j) = \frac{1}{n} \sum_{k=1}^n \gamma(y_j(k), x_i(k)), \quad (5)$$

is Deng's degree of incidence for sequence Y_j and X_i .

Sort relational order.

In general, the factors are always in order as long as they can form a relationship and calculate the degree of correlation. The grey correlation between the sequence of characteristic system behaviors Y and the sequence of related factors X'_i can be noted as $\gamma_1, \gamma_2, \dots, \gamma_m$. The corresponding correlation sequence is obtained by arranging the elements in correlation degree $\gamma_1, \gamma_2, \dots, \gamma_m$ according to their values from largest to smallest.

2.1.2. Establish BP Neural Network Model. Traditional BPNN divides the learning process into two stages: forward propagation of the signal and backward propagation of the error, and adjusts the "connection weights" and "thresholds" between neurons by training on all input variables. The parameters include the number of neurons in each layer, the activation function, the connection weights, and thresholds. In this study, the first p variables $X'_i(k)$ ($i = 1, 2, \dots, p, k = 1, 2, \dots, n$) corresponding to the correlation series obtained from the grey relational analysis were used as input factors (the number of input nodes is p). The characteristic behavior sequences $Y'_j(k)$ ($j = 1, 2, \dots, l, k = 1, 2, \dots, n$) were used as output factors (the number of output nodes is l).

We call $(x_i(k), y_j(k))$ is the k training pair, x_i is the input, y_j is the actual output, and its target output is set as d_j . The input node, intermediate node and output node are, respectively, represented by the subscript i, h, j . The weight from the input layer to the middle layer node h is represented by W_{ih} . The weight of node j from the middle layer to the output layer is expressed by W_{hj} . The thresholds of functional neurons in the middle layer and output layer are represented by a_h and b_j , respectively. $f(\cdot)$ stands for the activation function. The topological structure of the BPNN is shown in Figure 1:

When the training pair k is inputted, the input weighted sum $S_h(k)$ and output $Y_j(k)$ of the middle layer node h are, respectively,

$$S_h(k) = \sum_{i=1}^p x_i(k) W_{ih}, \quad (6)$$

$$y_h(k) = f[S_h(k) - a_h]. \quad (7)$$

The input weighted sum $S_j(k)$ and output $Y_j(k)$ of the output layer node j are, respectively,

$$S_j(k) = \sum_{h=1}^H y_h(k) W_{hj}, \quad (8)$$

$$y_j(k) = f[S_j(k) - b_j]. \quad (9)$$

The error of node j of the output layer is as follows:

$$e_j = d_j(k) - y_j(k). \quad (10)$$

If the sum of error squares of all output nodes for n inputs is used as the total network error, then there is a loss function:

$$E = \frac{1}{2} \sum_{j=1}^l [e_j(k)]^2. \quad (11)$$

The gradient descent method is adopted, and the derivative of E with respect to weight and threshold value is taken one by one, and the gradient that makes E decrease can be obtained and used as the direction of adjusting weight W_{hj} , W_{ih} and threshold value a_h, b_j . The adjusted weight is denoted as W'_{hj} , W'_{ih} , and the adjusted threshold value is denoted as a'_h, b'_j .

The adjustment formula from the middle layer to the output layer weight W_{hj} is as follows:

$$W'_{hj} = W_{hj} + \Delta W_{hj}, \quad (12)$$

$$\Delta W_{hj} = -\eta \frac{\partial E}{\partial W_{hj}} = -\eta \sum_k \delta_j(k) y_h(k), \quad (13)$$

$$\frac{\partial E}{\partial W_{hj}} = -\sum_k [d_j(k) - y_j(k)] f'[S_j(k)] y_h(k). \quad (14)$$

The adjustment formula from the input layer to the middle layer weight W_{ih} is as follows:

FIGURE 1: Schematic diagram of three-layer network.

$$W'_{ih} = W_{ih} + \Delta W_{ih}, \quad (15)$$

$$\Delta W_{ih} = -\eta \frac{\partial E}{\partial W_{ih}} = -\eta \sum_k \delta_h(k) y_i(k), \quad (16)$$

$$\frac{\partial E}{\partial W_{ih}} = -\sum_k [d_j(k) - y_j(k)] f'[S_j(k)] W_{hj} f'[S_h(k)] x_i(k). \quad (17)$$

The adjustment formula of intermediate layer threshold a_h is as follows:

$$a'_h = a_h + \Delta a_h, \quad (18)$$

$$\Delta a_h = -\eta \frac{\partial E}{\partial a_h} = -\eta y_h(k) [1 - y_h(k)] \sum_{j=1}^l w_{hj} e_j. \quad (19)$$

The adjustment formula of output layer threshold b_j is as follows:

$$b'_j = b_j + \Delta b_j, \quad (20)$$

$$\Delta b_j = -\eta \frac{\partial E}{\partial b_j} = \eta e_j. \quad (21)$$

The connection weights and thresholds of each unit layer are dynamically adjusted according to the error signal (loss function). Through the cyclic forward propagation and reverse regulation, the weights between neurons and the thresholds of each functional neuron are continuously revised. When the output error signal meets the accuracy requirements, stop learning and get the neural network model corresponding to the first p variable. By repeating (6)–(21), m variable corresponding to m BP neural network model can be obtained.

2.1.3. Accuracy Evaluation. We take the mean absolute percentage error (MAPE), mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination (R^2) to evaluate the accuracy of BPNN models [23–25]. Denote the number of input variables

corresponding to the model with the highest accuracy as p' . The specific formulas are as follows:

$$\text{MAPE} = 100 \times \frac{\sum_{i=1}^n |y_j - d_j| / y_j}{n}, \quad (22)$$

$$\text{MAE} = \frac{\sum_{i=1}^n |y_j - d_j|}{n}, \quad (23)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_j - d_j)^2}{n}}, \quad (24)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_j - d_j)^2}{\sum_{i=1}^n (y_j - \bar{y}_j)^2}, \quad (25)$$

where n refers to the number of samples, y_j is the observation value, d_j is the predictive value, and $\bar{y}_j$ is the mean of the observed values. The smaller the values of the three evaluation indexes, the higher the accuracy of model fitting.

2.1.4. Flow of Variable Optimization Selection Algorithm. The variable optimization selection algorithm is based on the correlation sequence of input variables, and the optimal BPNN model is obtained by stepwise input of variables and intermodel accuracy evaluation screening. At the same time, the number of corresponding key variables under the model is determined. Then, the predictive modeling is based on the key variables obtained through the screening and the optimal BPNN model. The specific algorithm flow is shown in Figure 2:

2.2. GR-BPNN Prediction Model

2.2.1. Metabolic GM (1, 1) Model. In this study, a metabolic GM (1, 1) model is used to predict p' key variables. The metabolic GM (1, 1) model is based on the GM (1, 1) model by iterating over old and new information, and the resulting modeling sequence is continuously adjusted to reflect the current operating characteristics of the system as it evolves [26].

(1) GM (1, 1) model modeling steps: Let the observed value of a characteristic behavior sequence of the system be $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$. One accumulative generated sequence is $X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))$, where $x^{(1)}(k) = \sum_{\tau=1}^k x(\tau)$, $k = 1, 2, 3, \dots, n$. The mean adjacent to the sequence is $Z^{(1)} = (z^{(1)}(1), z^{(1)}(2), \dots, z^{(1)}(n))$, where $z^{(1)}(k) = (1/2)(x(k) + txn(k-1))$.

The establishment of first-order linear differential equation model for $X^{(1)}$:

$$\frac{dx^1(t)}{dt} + ax^1(t) = b, \quad (26)$$

where a is the development coefficient; b is the grey action. The value of the parameter vectors $\hat{a} = [a, b]^T$ is estimated by the least square method.

FIGURE 2: Flow chart of the variable optimization selection algorithm.

$$\hat{a} = (B^T B)^{-1} B^T Y. \quad (27)$$

Y, B of them are as follows:

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad (28)$$

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}.$$

The solution of the differential equation (22) is as follows:

$$\hat{x}^{(1)}(t) = \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-a(t-1)} + \frac{b}{a}. \quad (29)$$

To make B - b reduction:

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1). \quad (30)$$

The grey prediction model of the original sequence is as follows:

$$\hat{x}^{(0)}(k) = (1 - e^a) \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-a(t-1)}. \quad (31)$$

(2) Metabolic GM (1, 1) model predicts principle: First, the GM (1, 1) model was established from the original sequence $X^{(0)}$ to predict the new data $\hat{x}^{(0)}(n+1)$, which is added to the original sequence and remove $X^{(0)}(1)$ at the same time. Then, the GM (1, 1) model is built again to predict the next data. The earliest data is removed and the new predicted data is added to the prediction sequence, so the metabolism is completed until the prediction purpose is completed.

2.2.2. Concrete Steps of GR-BPNN Prediction Model. The priority of variables was determined by grey relational analysis, and the system variables were ranked in order of importance. The optimal BPNN model structure and the corresponding number of input variable nodes under the model were determined by establishing a BPNN prediction model based on different variable combinations. The sorting processing of a large number of input variables by the grey relational model makes the selected variables representative, which realizes the optimization of the node number of input variables of BPNN without subjective screening. It enhances the modeling ability of the BPNN algorithm for multivariable complex systems and the adaptability of the network. The specific steps of the GR-BPNN prediction model are as follows:

- (i) Step 1. Based on equations (2)–(5), Deng's degree of incidence of the sequence of characteristic behaviors l and the sequence of relevant factors m is obtained.
- (ii) Step 2. Rank correlation according to correlation degree.
- (iii) Step 3. Establish the BPNN models based on the first variable p of the correlation order by formulas (6)–(21).
- (iv) Step 4. According to equations (22)–(24), the accuracy of m models was tested to obtain the optimal model structure, and the number of corresponding key variables p' under the model was determined.
- (v) Step 5. Calculate formulas (26)–(31) to obtain the predicted value of p' key impact factors.
- (vi) Step 6. Substitute the predicted value of p' key impact factors into the best trained BPNN model in step 4 and predict the characteristic behavior sequence.

3. Prediction Performance Evaluation of GR-BPNN Model

3.1. Selection and Treatment of Predictors. In addition to factors that cannot be determined, three aspects of grain production capacity, grain production guarantee, and economic scale are considered to build a system of indicators

related to grain yield forecasting in Henan Province. Relevant indicators are selected according to the basic principles of feasibility, purpose, comprehensiveness, comparability, and a combination of quantitative and qualitative indicators. The grain yield at the target layer is taken as the output of the BPNN prediction model. By consulting relevant literature [27] and combining with expert experience, 14 impact factors at the indicator layer in Table 1 are preliminarily selected as the input of the BPNN prediction model. Data for the relevant indicators are taken from the Henan Statistical Yearbook 2001–2020 [28].

3.1.1. Strives for the Degree of Incidence. Using grey modeling software to analyze the relevant index data, the degree of incidence between the output factor Y and each input factor X_i ($i = 1, 2, 3, \dots, 14$) is as follows:

$$\gamma_{0i} = \begin{pmatrix} 0.9829 & 0.9851 & 0.9917 & 0.9939 & 0.9897 & 0.9808 & 0.9821 \\ 0.9913 & 0.9593 & 0.6710 & 0.8335 & 0.8882 & 0.8953 & 0.9491 \end{pmatrix}. \quad (32)$$

3.1.2. Arrange the Order of Correlation. The relational order is obtained by the magnitude of the degree of incidence, which is as follows:

$$\gamma_{04} > \gamma_{03} > \gamma_{08} > \gamma_{05} > \gamma_{01} > \gamma_{06} > \gamma_{07} > \gamma_{06} > \gamma_{09} > \gamma_{014} > \gamma_{013} > \gamma_{012} > \gamma_{011} > \gamma_{010}. \quad (33)$$

3.2. Determination of Key Variables and the Best BPNN Model.

The typical three-layer feedforward BPNN structure was selected. According to Figure 2, the variable optimization selection algorithm is written using MATLAB software. The training samples are 15 sample data from 2000 to 2014, and the detection samples are 5 independent samples from 2015 to 2019. The Sigmoid function is used for both the hidden and output layer activation functions. The network connection initialization weights and thresholds are set to random numbers on $[0, 1]$ according to the random generator procedure; The number of nodes in the input layer corresponds to the first p variables after the m variables have been sorted. The number of nodes in the output layer is 1. Determine the number of hidden layer nodes according to the empirical method of $2n + 1$ [29]. According to the grey relational analysis results, 14 BPNN models are trained, and the model accuracy comparison is shown in Figure 3.

According to Figure 3, when the first 10 predictors were input, the BP network model reached the highest prediction accuracy. Therefore, the optimal BPNN model topology is 10-10-1 (Nodes of input layer-Nodes of hidden layer-Nodes of output layer). The key factors affecting grain yield are as follows: $X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{14}$. Show the observed and predicted values under the best BPNN obtained by screening in the line chart and get the fitting curve graph in the training sample under the best model. Figure 4 shows that the GR-BPNN model only had a poor fitting

effect in 2002, but it is quickly adjusted in 2003, indicating that the model has good adaptability.

3.3. Comparison between GR-BPNN Model and Benchmark Model

3.3.1. Benchmark Models and Classification of Accuracy Classes. Considering that GR-BPNN is an improvement on the traditional BPNN, LSTM and CNN are suitable for modeling the time series data in this study, and the traditional econometric models SR, GM (1, N) are equally capable of capturing the feature facts and aggregating the essential elements of interest in this study. Therefore, this study chooses LSTM, CNN, BPNN, SR, GM (1, N) as the benchmark model to compare the accuracy of the model proposed in this research and accurately evaluate the predictive ability of the GR-BPNN model.

The benchmark model is correspondingly introduced as follows: LSTM solves the long memory problem that recurrent neural network does not have by introducing a gate mechanism and can show better performance in nonlinear time series forecasting. The network model parameters mainly include the number of neurons in each network layer and the learning rate; CNN uses local connectivity and weight sharing to transform and abstract the original data matrix in a high-dimensional way and can build different dimensional structure models based on the characteristics of the data set. Its network model parameters mainly include learning rate, number of nodes in each layer of the network, activation function, and step size. Traditional BPNN is described in detail in Section 2.1.2. The basic idea of SR is to reduce the degree of multicollinearity by eliminating variables that are less important and highly correlated with other variables. The optimal set of variables is obtained and predictive modeling is carried out by introducing variables one by one and iteratively testing the variables; GM (1, N) model is a development of the one-dimensional series grey prediction model GM (1, 1). The magnitude and sign of the weight coefficients of the factor variables in this model are used to understand the degree of influence of each factor on the behavioral variables and to model the predictions in the form of differential equations.

In order to verify the robustness and sensitivity of the model estimation results, refer to the formula (24) and use the coefficient of determination R^2 to judge the overall fitting accuracy of the model. The small error probability and the posterior difference ratio are used to divide and judge the accuracy levels of different models. Lu [30] and Cao [31] graded the accuracy of the grain yield prediction model based on the posterior difference ratio and the probability of small errors and qualitatively evaluated different grain yield prediction models. The details are shown in Table 2.

3.3.2. Analysis of Comparative Results. The predicted values of the key variables for the metabolic GM (1, 1) model are used as the input of the best BPNN model. The 2015–2019 grain production data in Henan Province is selected as the test sample. For LSTM, a three-layer LSTM network

TABLE 1: Impact factors of grain yield in Henan Province.

Target layer	Rule layer	Index layer
Grain yield (Y , 103 hectares)	Grain yield capacity	Grain sown area (X_1 , 103 hectares)
		Area of cultivated land (X_2 , 103 hectares)
		Consumption of chemical pesticides (X_3 , 104 tons)
		Diesel oil use for agriculture (X_4 , 104 tons)
		Plastic film use for agriculture (X_5 , 104 tons)
		Total power of agricultural machinery (X_6 , 104 kw)
		Irrigated area (X_7 , 103 hectares)
	Grain yield guarantee	Fertilizer application in discounted amounts (X_8 , 104 tons)
		Electricity consumption in rural areas (X_9 , 108 kwh)
		Agriculture, forestry, and water expenditure (X_{10} , 108 yuan)
	Size of the economy	Gross domestic product (X_{11} , 108 yuan)
		Rural per capita consumption expenditure (X_{12} , yuan)
		Rural per capita disposable income (X_{13} , yuan)
		Share of the contributions of primary industry (X_{14} , %)

FIGURE 3: Comparison of model precision under different input variables.

FIGURE 4: GR-BPNN Model training sample fitting curve.

structure is used, and the network model topology is 14-10-1. The loss function refers to formula (11), using the Adam optimizer, the learning rate is initially defined as 0.01 and decreases with iteration. For CNN, a three-layer LSTM neural network is also used, the network model topology is 14-7-1, and the learning rate is initially defined as 0.01. In addition, comprehensively considering three evaluation indicators

(MAPE, MAE, and RMSE), the accuracy of six models is evaluated according to formulae (22)–(24). The evaluation results are shown in Figure 5.

It can be seen from Figure 5 that the prediction performance of the six models tends to decrease in order. Among them, the ST prediction model and the GM (1, 1) model have the worst accuracy. The MAPE reached 20.5%

TABLE 2: Classification standard of precision grade of grain yield prediction model.

Model accuracy level	Probability of small error P	Posterior difference ratio C
Good	>0.95	<0.35
Qualified	>0.80	<0.50
Barely qualified	>0.70	<0.65
Unqualified	≤ 0.60	≥ 0.65

FIGURE 5: Accuracy comparison between different models. SR: stepwise regression; GM (1, N): grey GM (1, N) model; BPNN: backpropagation neural network; LSTM: long short-term memory; CNN: convolutional neural network; GR-BPNN: grey back propagation neural network.

and 18.22%, respectively. The MAPE of BP, LSTM, and CNN were all above 5%, and the MAPE, MAE, while the MAPE, MAE, and RMSE of GR-BPNN are 2.97%, 1.64 tons, 2.83 tons, respectively, which is the least error among all models. The prediction performance of the GR-BPNN algorithm model has been greatly improved due to its own factor screening capability and its ability to describe both linear and nonlinear relationships between variables. Therefore, the GR-BPNN algorithm model is more suitable for the prediction of grain yield in Henan Province.

Reference [30] calculate the P value and C value of each prediction model and R^2 was calculated with reference to formula (25). The calculation results were shown in Table 3. As can be seen from Table 3, excluding the SR and GM (1, N), as compared to BPNN, LSTM, and CNN, the prediction model exhibited values that were, respectively, improved by 16.25%, 8.14%, and 5.68%. Furthermore, SR model has the worst accuracy grade, which may be due to the setting bias caused by the elimination of important related variables by the stepwise regression method. Both GM (1, N) and BP are barely qualified, but there is still a large error probability. LSTM, CNN, and GR-BPNN algorithms all reached the accuracy level of qualified or above. Among them, the GR-BPNN model showed the best performance, and the model level is “good,” which proves that the food output prediction algorithm of this research has high prediction accuracy. In the establishment of the GR-BPNN model, 15 prediction variables were selected and 10 key variables were selected to participate in the construction of the prediction model, which effectively simplified the complexity of the model and had

high sparsity. Therefore, this model is superior to other models in computational efficiency and has a high running speed. In addition, compared with [30, 31] in the literature, the model built in this study achieves higher accuracy, indicating that the GR-BPNN algorithm model has certain advantages. Based on the above data analysis, it shows that the GR-BPNN algorithm has fast calculation speed, high accuracy, and strong reliability in the grain output prediction.

4. Prediction of Grain Yield in Henan Province Based on GR-BPNN Model

4.1. Predicted Results. Based on the original sequence of 10 key impact factors, namely the corresponding index data from 2010 to 2019, the metabolic GM (1, 1) model is established, respectively, to predict the specific measurement values of key impact factors affecting grain yield in Henan Province during 2020–2025. The prediction results are shown in Table 4—prediction of grain yield in Henan Province based on GR-BPNN model.

The predicted values of 10 key impact factors for grain yield (2019–2023) are used as the input of the trained optimal BPNN model, and the predicted values of grain yield in Henan Province from 2019 to 2023 are obtained accordingly. The prediction results are shown in Table 5.

According to the predicted results in Table 5 and the actual value of grain yield in 2000–2019, the trend chart of grain yield in Henan Province is plotted (Figure 6). As can be seen from Figure 6, the total grain yield of Henan Province will show a steady and increasing trend in the next few years, which is a good development situation for the grain supply of Henan Province and even the whole country. In 2003 (about 35.6 947 million tons), there was a serious reduction in grain production, which was 15.21% lower than in 2002 (about 42.099 8 million tons); the reason is that the grain yield increased year by year during this period, but the phenomena of “sell grain” and “increasing production without increasing income” appeared immediately, which seriously dampened the enthusiasm of farmers to grow grain. In addition, there are some serious natural disasters, such as continuous rain, low temperature, lack of light and heat, hail, collapse, and flood in 2003, which seriously affected the grain yield of Henan Province.

4.2. Results Analysis

4.2.1. Temporal Change Trend of Grain Production in Henan Province. Based on the predicted results in Table 4, the trend graph of the volatility for total grain yield shown in Figure 7 was drawn. As shown in Figure 7, the grain yield increased

TABLE 3: Evaluation index of the precision grade of grain yield prediction model.

Predicted models	Coefficient of determination R^2	Probability of small error P	Posterior difference ratio C	Level
SR	0.43	0.24	2.15	Unqualified
GM (1, N)	0.65	0.60	0.32	Barely qualified
BPNN	0.80	0.87	0.23	Barely qualified
LSTM	0.86	0.90	0.17	Qualified
CNN	0.88	0.91	0.08	Qualified
GR-BPNN	0.93	1.00	0.05	Good

SR: stepwise regression; GM (1, N): grey GM (1, N) model; BP: backpropagation neural network; LSTM: long short-term memory; CNN: convolutional neural network; GR-BPNN: grey relational-back propagation neural network.

TABLE 4: Prediction of key impact factors for grain yield in Henan Province based on metabolism GM (1, 1) model.

Year	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{14}
2020	14 629.4	8 185.3	10.4	93.4	14.6	9 554.8	5 460.8	677.5	341.5	4.0
2021	14 571.6	8 205.2	9.9	90.7	14.1	9 432.7	5 485.9	668.5	347.1	3.7
2022	14 514.0	8 225.2	9.4	87.4	13.8	9 312.1	5 511.0	659.6	352.9	3.4
2023	14 456.7	8 245.2	8.9	84.3	13.5	9 193.1	5 536.3	650.8	358.7	3.1
2024	14 399.6	8 265.3	8.6	81.4	13.1	9 075.5	5 561.7	642.1	364.7	2.8
2025	14 342.7	8 285.4	8.1	78.5	12.8	8 959.5	5 587.3	633.6	370.7	2.5

The specific meaning of each key impact factor in the table is as follows. X_1 : Grain sown area (10^3 hectares); X_2 : Area of cultivated land (10^3 hectares); X_3 : Consumption of chemical pesticides (10^4 tons); X_4 : Diesel oil use for agriculture (10^4 tons); X_5 : Plastic Film Use for Agriculture (10^4 tons); X_6 : Total Power of Agricultural Machinery (10^4 kw); X_7 : Irrigated Area (10^3 hectares); X_8 : Fertilizer application in discounted amounts (10^4 tons); X_9 : Electricity Consumption in Rural Areas (10^2 million kwh); X_{14} : Share of the contributions of primary industry (%).

TABLE 5: Prediction of grain yield in Henan province (unit: 10^4 tons).

Year	2020	2021	2022	2023	2024	2025
Production	6 735.40	6 793.77	6 852.648	6 912.04	7 002.22	7 066.75

FIGURE 6: Chart of grain yield trend of Henan province.

FIGURE 7: The fluctuation rate of total grain yield in Henan province.

significantly after 2003, especially in 2004, the grain yield increased by 19.35% compared with the previous year. During this period, the state began to adopt a series of measures to stimulate grain yield, and Henan Province fully implemented a series of national policies for the development of grain production: continuously accelerating economic system reforms, implementing the household contract responsibility system with joint output to adjust rural production relations, put forward to exempt the whole province agricultural tax program and so on. Driven by this series of policies supporting agriculture and benefiting agriculture, farmers' enthusiasm for growing grain has been significantly improved. Therefore, Henan's grain production will increase steadily in the coming years, provided that there are no major changes in the environment and policy. The total grain yield of Henan Province continued to grow from 2004 to 2019, making an important contribution to national grain security. However, it is worth noting that the growth rate has been decreasing year by year. The forecast results show that from 2020 to 2025, Henan Province's grain yield will increase by within 1%, they are 0.72%, 0.82%, 0.92%, 0.93%, and 0.92%. It can be seen that the overall growth rate of grain yield in Henan Province is slowing down, while the ability to stabilize yield is gradually stronger. Annual fluctuation of total grain yield $R = (Y_{t+1} - Y_t)/Y_t \times 100\%$, where Y is the total grain yield and t is the year.

4.2.2. Spatial Evolution Trend of Grain Yield in Henan Province. In order to further explore the spatial differentiation law of grain yield for Henan Province and its dynamic change trend, according to Table 1, the grain production and related indicators of 18 cities in Henan Province in 2000–2019 were selected (data from Henan Statistical Yearbook 2001–2020). Using the GR-BPNN prediction model, the predicted value of grain yield in 2020–2025 for 18 cities in Henan Province is obtained in the same way. The equal interval grading method in ArcGIS was used to classify the grain yield in 2020–2025 of the 18 cities into 5 classes. The spatial distribution map of grain yield shown in Figure 8 was obtained by selecting 2015, 2020, and 2025, respectively. As can be seen from Figure 8, the main areas of grain yield in Henan Province in 2015 were Shangqiu, Zhumadian, Nanyang, and Xinyang; it is estimated that the main grain production areas in Henan Province in 2020 will be Shangqiu, Zhoukou, Zhumadian, and Nanyang; it is estimated that the main grain production areas in Henan Province in 2025 will be Xinxiang, Shangqiu, Zhoukou, Zhumadian, and Nanyang. The main grain-producing areas in Henan Province are mainly located in the eastern plains, where strict measures have been implemented to protect arable land in a balanced manner, develop reserve arable land resources, and rehabilitate the land, and its excellent natural resource conditions and policy support have largely increased the grain production capacity of the region. The areas with low grain yield in three years were all distributed in the northwest of Henan, which has a relatively fragile ecological environment and has seen a dramatic increase in the area returned to the forest since the national policy of

returning farmland to forests was implemented in 1998. Among them, Sanmenxia has a severe lack of water resources and is topographically located in the mountainous, hilly region of western Henan, resulting in extremely low grain production. In addition, Zhengzhou and its surrounding cities are areas of high industrialization and urbanization, and the reduction in the land area caused by the demands of various industrial constructions has put enormous pressure on arable land conservation and food production. On the whole, the main grain production areas in Henan Province showed a trend of moving to the north, reflecting the spatial imbalance of grain supply and demand, and the grain security problem was gradually highlighted in some areas.

The model of elemental transfer center of gravity [32] is an analytical tool to study the spatiotemporal variation law of elements in the process of regional development. The fluctuation of grain yield in a region with time will cause the spatial center of gravity for grain yield in the region to move, which is of great significance to the rational utilization of regional cultivated land and the guarantee of grain security. Based on the measured value (2015–2019) and predicted value (2020–2025) of grain yield in 18 regional units of Henan Province, the following model of grain yield elemental transfer center of gravity was established:

$$\begin{aligned} \text{LON}_j &= \frac{\sum_{i=1}^n (M_{ij} \times \text{lon}_i)}{\sum_{i=1}^n M_{ij}}, \\ \text{DIM}_j &= \frac{\sum_{i=1}^n (M_{ij} \times \text{dim}_i)}{\sum_{i=1}^n M_{ij}}, \end{aligned} \quad (34)$$

where n is the 18 regional units in Henan Province; M_{ij} is the grain yield of zone unit i in Year j ; $(\text{lon}_i, \text{dim}_i)$ is the geospatial barycenter coordinates of the region unit i ; $(\text{LON}_j, \text{DIM}_j)$ is the spatial barycentric coordinate of grain yield in Henan Province in the year j . Using ArcGIS software, the grain production center coordinates in Henan Province from 2015 to 2025 were obtained, and the grain production center trajectory in Henan Province was obtained (Figure 9).

According to the elemental transfer center of the gravity model, the center of gravity should be located in the geometric center of the region if the grain production of each region in Henan Province is in equilibrium. Otherwise, it will lead to a shift of the center of gravity. Based on the calculations in Figure 8, it can be seen that the center of grains production in Henan Province deviates from the regional geometric center. From 2015 to 2025, the geographical coordinates of grain yield gravity center in Henan Province were between $114^\circ 7' 20'' \text{E} \sim 114^\circ 8' 20'' \text{E}$, $33^\circ 59' 0'' \text{N} \sim 34^\circ 0' 20'' \text{N}$, and the grain yield gravity center of Henan Province did not change strongly. Comparing the trajectory of the grain production center for gravity in Henan Province with the spatial distribution of regional grain production, the changes in the two are corresponding. The changing trajectory of the center of gravity reveals that the center of gravity for grain production in Henan Province

FIGURE 8: Spatial distribution of regional grain yield in Henan province. (a) Regional grain yield in 2015. (b) Regional grain yield in 2020. (c) Regional grain yield in 2025.

FIGURE 9: Changes of grain yield gravity center track in Henan province from 2015 to 2025.

is gradually shifting to the north. It also shows that the pressure on regional food security is constantly moving northward, and the northern region has to bear greater production pressure.

5. Conclusions

This study established the GR-BPNN prediction model by combining grey system theory and BPNN, making full use of the inclusiveness of grey relational analysis for sample data, the randomness of grey forecasting model weak data, the

regularity of cumulative data, and the high nonlinearity of neural network. Combining the advantages of two predictive models, it can effectively deal with the analysis and modeling of multivariable complex realistic systems.

Applying the model developed in this paper to grain yield forecasting in Henan Province, it is found that the total grain yield in Henan Province increases with each year; the total grain yield growth slows down and tends to be stable, and the ability to stabilize grain production is gradually enhanced; grain production in Henan Province is strong regionally, and there are great differences among regions.

The center of gravity of grain production keeps moving from the South to the North during 2015–2025. Changes in total grain yield are closely related to grain production capacity and security, economic development level, and input of production factors. The increase in the use of pesticides, plastic films, and chemical fertilizers, the construction and renovation of irrigation and water conservancy facilities, and the increase in labor force and technology input have played a huge role in improving the total grain yield. The following aspects should be considered in future efforts to increase grain yield in Henan Province: (1) Strengthening agricultural infrastructure. (2) Deepen policies to strengthen agriculture and benefit farmers, maintain policy strength, and play a synergistic effect. (3) Increase the role of scientific and technological innovation in grain production. (4) More policy preferences should be provided to major grain-producing regions to boost the development of the grain industry. (5) Strengthen interregional communication and exchange to achieve balanced and coordinated development of regional grain production.

For the forecast of grain yield in Henan Province, only three factors have been selected in this paper, including grain production capacity, food production security, and economic scale. Although there are many indicators, they may be missing, and other aspects such as the natural environment and meteorological conditions have not been considered. Changes in grain yield are the result of the interaction for natural resource endowments, socioeconomic level, agricultural technology progress, degree of marketization, and policy factors. The relevant indicators of Henan Province's grain yield can be established in a comprehensive and multilayered approach to further improve the prediction accuracy of grain yield. This will also be the next step in the research.

Data Availability

The relevant index data of grain production in Henan Province from 2000 to 2020 from the “Henan Statistical Yearbook” presented in this manuscript are open and available. Link for data: <http://www.ha.stats.gov.cn/>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Key Project of Soft Science Research in Henan Province (202400410051).

References

- [1] F. Y. Yang, L. M. Liu, and C. C. Yuan, “Influencing factors of grain production fluctuation in hunan province: analysis and prediction,” *Chinese Agricultural Science Bulletin*, vol. 36, no. 29, pp. 153–160, 2020.
- [2] L. Li, B. Du, Y. Wang, L. Qin, and H. Tan, “Estimation of missing values in heterogeneous traffic data: application of multimodal deep learning model,” *Knowledge-Based Systems*, vol. 194, no. 194, Article ID 105592, 2020.
- [3] L. Li, X. Sheng, B. Du, Y. Wang, and B. Ran, “A deep fusion model based on restricted Boltzmann machines for traffic accident duration prediction,” *Engineering Applications of Artificial Intelligence*, vol. 93, no. 93, Article ID 103686, 2020.
- [4] Y. Cai, K. Guan, J. Peng, C. Seifert, B. Wardlow, and Z. Li, “A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach,” *Remote Sensing of Environment*, vol. 210, pp. 35–47, 2018.
- [5] A. Kross, E. Znoj, D. Callegari et al., “Using artificial neural networks and remotely sensed data to evaluate the relative importance of variables for prediction of Within-Field corn and soybean yields,” *Remote Sensing*, vol. 12, no. 14, p. 2230, 2020.
- [6] I. A. Saleh, W. A. Alawsi, O. I. Alsaif, and K. Alsaif, “A prediction of grain yield based on hybrid intelligent algorithm,” *Journal of Physics: Conference Series*, vol. 1591, no. 1, Article ID 012027, 2020.
- [7] F. Shu, “Prediction of wheat yield in Henan Province based on NAR neural network,” *Journal of Xuchang University*, vol. 38, no. 5, pp. 118–122, 2019.
- [8] H. Zhang, G. W. Wang, C. Yuan, and H. Hongyan, “Research on forecast of grain production based on AIGA-BP neural network,” *Journal of Chinese Agricultural Mechanization*, vol. 33, no. 6, pp. 205–209, 2016.
- [9] H. P. Luo, N. Zou, L. X. Pan, and Q. Q. Zhu, “Research on temporal and spatial attribute of food production security state in Chinese major grain regions under the horizon of ecological footprint: 2007–2025,” *Jiangsu Journal of Agricultural Sciences*, vol. 35, no. 6, pp. 1468–1475, 2019.
- [10] L. L. Gu, Y. Liu, and J. Q. Zhen, “Prediction of beet yield based on BP neural network optimized by improved particle swarm algorithm,” *Journal of Xinjiang University (Natural Science Edition)*, vol. 38, no. 2, pp. 191–196, 2021.
- [11] P. B. Li, X. R. Yan, and D. R. Xu, “Comparison of grain yield spatial distribution forecast between the models of BP neural network and multiple linear regression,” *Journal of Arid Land Resources & Environment*, vol. 28, no. 9, pp. 74–79, 2014.
- [12] J. Y. Zhang and G. Y. Pan, “Comparison and application of multiple regression and BP neural network prediction model,” *Journal of Kunming University of Science and Technology*, vol. 38, no. 6, pp. 61–67, 2013.
- [13] Y. S. Yang, L. J. Dong, and L. Guan, “The Prediction of grain production based on rough set and BP neural network,” *Journal of Agricultural Mechanization Research*, vol. 36, no. 11, pp. 34–37, 2014.
- [14] H. Ren, X. R. Ma, and H. B. Liu, “Improvement of input evaluation for giant projects based on GA-BP neural network,” *Systems Engineering-Theory & Practice*, vol. 35, no. 6, pp. 1474–1481, 2015, in Chinese.
- [15] Y. F. Guo, “Prediction of grain yield prediction using neural network based on principal component analysis and particle swarm optimization,” *Jiangsu Agricultural Sciences*, vol. 47, no. 19, pp. 241–245, 2019.
- [16] J. A. Zheng, “Application of PCA and BP neural networks in grain production prediction,” *Computer Systems and Applications*, vol. 25, no. 11, pp. 274–278, 2016.
- [17] S. P. Wang and Y. Y. Zhu, “Forecasting of wheat price Based on multiscale analysis,” *Chinese Journal of Management Science*, vol. 25, no. 4, pp. 85–91, 2016.
- [18] C. S. Zong, H. X. Zheng, and L. S. Wang, “Grain yield prediction based on BP neural network optimized by improved

- particle swarm optimization,” *Computer Systems and Applications*, vol. 27, no. 12, pp. 204–209, 2018.
- [19] C. L. Hu, Y. H. Liu, and J. L. Gao, “Research on prediction method of grain yield based on IPSO-BP model,” *Journal of Chinese Agricultural Mechanization*, vol. 42, no. 3, pp. 136–141, 2021.
- [20] L. Q. Rong, F. Chen, and O. Y. Hao, “Predicting grain yield of Guangxi province based on GRA&BPNN,” *Chinese Journal of Agricultural Resources and Regional Planning*, vol. 38, no. 2, pp. 105–111, 2017.
- [21] M. Y. Tong, Z. Yan, and L. Chao, “Research on a grey prediction model of population growth based on a logistic approach,” *Discrete Dynamics in Nature and Society*, vol. 2020, Article ID 2416840, 14 pages, 2020.
- [22] L. Li, H. Zhou, H. Liu, C. Zhang, and J. Liu, “A hybrid method coupling empirical mode decomposition and a long short-term memory network to predict missing measured signal data of SHM systems,” *Structural Health Monitoring*, vol. 20, no. 4, pp. 1778–1793, 2021.
- [23] G.-F. Fan, S. Qing, H. Wang, W.-C. Hong, and H.-J. Li, “Support vector regression model based on empirical mode decomposition and auto regression for electric load forecasting,” *Energies*, vol. 6, no. 4, pp. 1887–1901, 2013.
- [24] Y. H. Chen, W. C. Hong, W. Shen, and N. N. Huang, “Electric load forecasting based on a least squares support vector machine with fuzzy time series and global harmony search algorithm,” *Energies*, vol. 9, no. 2, 2016.
- [25] M.-W. Li, Y.-T. Wang, J. Geng, and W.-C. Hong, “Chaos cloud quantum bat hybrid optimization algorithm,” *Non-linear Dynamics*, vol. 103, no. 1, pp. 1167–1193, 2021.
- [26] W. Zhou and D. Zhang, “An improved metabolism grey model for predicting small samples with a singular datum and its application to sulfur dioxide emissions in China,” *Discrete Dynamics in Nature and Society*, vol. 2016, Article ID 1045057, 11 pages, 2016.
- [27] G. Q. Meng, F. Wang, and L. S. Ren, “Correlation analysis of influencing factors of grain yield forecast in Hebei province,” *Journal of Hebei Agricultural University*, vol. 42, no. 6, pp. 122–127, 2019.
- [28] *Henan Province Bureau of Statistics*, <http://www.ha.stats.gov.cn/>.
- [29] C. C. Zhang and S. D. Chen, “Application of BP neural network in predicting grain yield of Henan Province,” *Hubei Agricultural Sciences*, vol. 53, no. 8, pp. 1969–1993, 1971.
- [30] L. P. Lu, J. Zhang, P. H. Jiang, and S. W. Zhang, “Driven analysis and trend forecasting changes in grain production in Gansu Province,” *Economic Geography*, vol. 33, no. 4, pp. 125–131, 2013.
- [31] F. F. Cao, “Application of gray system theory in grain output forecasting,” *Mathematics in Practice and Theory*, vol. 47, no. 13, pp. 310–312, 2017.
- [32] Z. H. Yang, J. K. Li, C. X. Han, and L. H. Guang, “The evolution path of China’s grain production base and the influencing factors,” *Research of Agricultural Modernization*, vol. 40, no. 1, pp. 36–43, 2019.