Original Research Paper | doi https://doi.org/10.46488/NEPT.2021.v20i02.021 | **Open Access Journal**

Prediction Model of Agricultural Non-point Source Water Pollution Based on Grey Correlation Method

R. Wang*(**)†, F. X. Yang* and G. M. Qu***
*Environmental Engineering College, Nanjing Polytechnic Institute, Nanjing-210048, China
**College of Environment, Hohai University, Nanjing-210098, China
***School of Chemical Engineering and Material Science, Nanjing Polytechnic Institute, Nanjing-210048, China
†Corresponding author: R.Wang; wangruijs@163.com

ABSTRACT

Accurate prediction of non-point source water pollution is conducive to the prevention and control of rural water pollution. To improve the prediction accuracy of agricultural non-point source water pollution and achieve timely prevention and control, a prediction model of agricultural non-point source water pollution based on the grey correlation method is designed. According to the historical data of agricultural non-point source water pollution, the influencing factors of water pollution degree are determined by using the grey correlation method, and standardized pretreatment is carried out. The pretreatment results are transformed into function expression forms, and the original sequences of different influencing factors are generated, which are brought into the function table to achieve the results, and the whitening differential equation is constructed to measure the concentration of agricultural non-point source water pollution. The structure design of the prediction model of agricultural non-point source water pollution can realize the prediction of agricultural non-point source water pollution. The results show that the prediction model of agricultural non-point source water pollution based on the grey correlation method has high prediction accuracy and small prediction error.

INTRODUCTION

When the control level of point source pollution reaches a certain degree, non-point source pollution becomes the main reason for water environment pollution. Around the world, 30-50% of surface water has been affected by non-point source pollution. At present, the effective treatment of point source pollution has been realized, but the non-point source pollution, including nitrate and insecticide, is still in urgent need of treatment (Long et al. 2017, Anooob 2017). Nowadays, non-point source water pollution has become a major threat to human drinking water, which seriously restricts the sustainable development of the economy and society. Compared with point source pollution, it is more difficult to study and control non-point source pollution. In the control of the water environment, it is difficult to fundamentally improve the water quality if only point source pollution control is carried out, and non-point source pollution research is not paid attention to. Water resources are important to support and guarantee economic and social development. It is of great significance to strengthen the research of non-point source pollution for the prevention and control of water pollution. It is because non-point source water pollution has such a bad impact on society and the economy, as well as the important

significance of non-point source water pollution control that more and more scholars join in the research of non-point source water pollution, and have made their achievements in related fields.

Non-point source water pollution has attracted the attention of many scholars at home and abroad, who have taken a variety of methods to carry out relevant research on non-point source water pollution. For example, Joy (2018) specifically studied a variety of non-point source models. To analyze these non-point source pollutants, 10 non-point source models were studied, and the advantages and disadvantages of their analysis of SARS source pollutants were tested. According to the results of the study, due to the operational limitations in the test and the difficulty in the calculation, different models are limited in the analysis of non-point source water pollutants. Bai et al.(2018), to analyze the characteristics and laws of non-point source pollution load transfer, selected the Dongjiang River Basin for comprehensive analysis, and the HSPF model was used to simulate the spatiotemporal distribution characteristics of non-point source pollution load such as BOD, TN and TP in the Dongjiang River Basin. The research results show that the HSPF model in the Dongjiang River Basin has a

better simulation application of its SARS source pollution load, which is water for other regions. It provides a way of thinking and reference for the measurement of non-point source pollution load. Zhanget al.(2018) paid close attention to the pollution status of non-point source water. To analyze the pollution load of COD, NH_3-N and TP in Kaijiang River Basin, the three pollution sources in 2015 were estimated by using the emission coefficient method for research, which were rural living source, livestock breeding source and farmland runoff source. According to this method, the total amount of COD, NH_3-N and TP of rural non-point source pollution in Kaijiang River Basin in 2015 is measured, which provides scope guidance for the key treatment of rural non-point source pollution in the Kaijiang River Basin, and provides some guidance for the research in other areas.

The control of non-point source pollution has reached the point of urgency. If the rural areas can effectively predict and control, it will greatly improve the status of non-point source pollution (Dadras et al.2018). It is the most effective way to realize the quantification, impact assessment and pollution control of non-point source pollution to establish a mathematical model to simulate the form and transfer of various non-point source pollution loads in time and space.

This paper focuses on how to improve the prediction accuracy of agricultural non-point source water pollution to achieve timely prevention and control, and designs the prediction model of agricultural non-point source water pollution based on the grey correlation method. Through the structural design of the prediction model of agricultural non-point source water pollution, the accurate prediction of agricultural non-point source water pollution is realized, and the predicted data can be reflected to the relevant departments in time, which can be more effective. The importance of using the model to predict agricultural non-point source water pollution can be seen from the effective formulation of the next development plan.

DESIGN OF AGRICULTURAL NON-POINT SOURCE WATER POLLUTION PREDICTION MODEL

Pretreatment of Agricultural Non-Point Source Water Pollution Historical Data

To determine the influencing factors of water pollution degree, a certain scale of sample size is needed. To obtain a representative and high-quality enough sample set, this paper extensively searches the academic dissertations at home and abroad studying agricultural non-point source water pollution, and obtains the historical sample data (Baranov et al. 2018, Wen et al. 2018) meeting the training requirements of the prediction model through a comprehensive analysis

of research methods, research areas, research conclusions, etc. At the same time, the prediction of agricultural non-point source water pollution at home and abroad mostly focuses on the road, so the prediction model is also established for the road runoff. Table 1 gives the historical data sources of this prediction model.

Based on the analysis of the research results on the influencing factors of agricultural non-point source water pollution at home and abroad, the agricultural non-point source water pollution is mainly affected by rainfall characteristics, air pollution, traffic flow and other urban functional areas (Wang 2017). The influencing factors of the model are summarized as shown in Table 2.

The following three points need to be paid attention to in the selection of water pollution indexes: first, cod is an important index to evaluate the pollution degree and control of rainfall-runoff and has a significant correlation with TN, NH_3-N and TP; second, after determining the input

Table 1: Historical data sources of agricultural non-point source water pollution prediction model.

Number	Author	City	Particular year	Forecast times
1	Wang He Yi	Shanghai	2004	6
2	Wang Ye Lei	Nanchang	2007-2008	9
3	Mao Yan Jing	Chongqing	2009	3
4	Chen Ying	Xi'an	2009	36
5	Hao Li Ling	Chongqing	2010-2011	13
6	Han Jing Chao	Fuzhou	2011-2012	17
7	Chen Zi Yu	Guangzhou	2012	4
8	Jiang Yi Zi	Shenzhen	2012	7
9	Song Qian Feng	Chongqing	2011-2012	7
10	Tian Huan	Shenzhen	2014	4
11	Yuan Yan	Suzhou	2014	12
12	Wang Zhao	Xi'an	2014-2015	13
13	Wang Shan	Wuhan	2017	4

Table 2: Influencing factors.

Number	Influence factor	Unit
X1	Pollution types	-
X2	Water pollution intensity	mm/h
X3	Peak pollution intensity	mm/h
X4	Water pollution duration	min
X5	Dry period before water pollution	h
X6	Atmospheric dust fall	$t/km^2/$ month
X7	PM10	mg/m^3
X8	Vehicle flow	Vehicle/h
X9	Pollution source	-
X10	Functional area	-

Table 3: Sample size of data set.

Water pollution prediction data set	Training sample size	Validation sample size
EMC	45	20
FF30	45	20

(influencing factors) and output (decision variables) of data samples required by the model, it should be directly obtained from the literature, retrieval, research and other methods to supplement the influencing factors of the selected dataset (Liouane et al. 2017). Finally, because the agricultural non-point source water pollution prediction model does not allow the missing value to exist in the prediction of new samples, the missing information cannot be supplemented by the data to be eliminated. Table 3 gives the sample size of the data set after data cleaning.

The influencing factors and decision variables, shown in Table 1, have different dimensions and dimension units, and their numerical values will have different weight effects during model operation, resulting in the reduction of model convergence speed and the distortion of data analysis results. In the support vector machine model, RBF uses the Euclidean distance calculation of sample data, which can avoid the variables with larger values in the influencing factors and control the variables with smaller values. However, the larger numerical difference will still affect the calculation of the water pollution prediction model and reduce the convergence speed of the model (Pierre et al. 2017). Therefore, to eliminate the dimensional influence among variables, it is necessary to standardize the historical data to make the different variables comparable. There are two kinds of commonly used data standardization methods.

The standardization of the maximum-minimum value of the influencing factors, also known as the deviation standardization, is a linear transformation of the historical original data of the influencing factors of water pollution so that the result value is mapped between [0,1]. The conversion function is shown in formula (1):

$$x^* = \frac{x - \min}{\max - \min} \qquad \ldots(1)$$

In the formula, max represents the maximum value of water pollution historical raw data sample, min represents the minimum value of water pollution historical raw data sample.

Z-score standardization, also known as normal standardization, uses the mean value and standard deviation of historical raw data of water pollution influencing factors to standardize the historical raw data of water pollution. The processed historical data conform to the standard normal distribution, i.e. the mean value is 0, the standard deviation is 1, and the conversion function is shown in formula (2):

$$x^* = \frac{x - \mu}{\sigma} \qquad \ldots(2)$$

In the formula, μ represents the average value of water pollution historical data samples, and σ represents the standard deviation of water pollution historical data samples. The above two methods are used to process the historical data of water pollution, and the grey correlation method is used to standardize the historical data of water pollution influencing factors. The processing results are transformed into function expression forms to generate the original sequence of different influencing factors, which are brought into the function expression results to build the whitening differential equation, and measure the concentration of agricultural non-point source water pollution, to carry out the agricultural treatment design of non-point source water pollution prediction model.

Design Agricultural Non-Point Source Water Pollution Prediction Model

Verhulst put forward the water pollution prediction model for the first time when he studied the law of biological reproduction in 1837 (Wang et al. 2019). The basic idea of the model is that the number of biological individuals growing exponentially is limited by the surrounding environment, and their growth rate gradually slows down, and finally they are fixed at a fixed value. This model is mainly used to describe the process with saturation state, i.e. "S-type process". When Verhulst adds a limited development item into the linear model of agricultural water pollution prediction, it will become a Verhulst nonlinear model as follows:

If the original sequence $X^{(0)} = \left\{ X^{(0)}(1), X^{(0)}(2), \cdots, X^{(0)}(r) \right\}$, 1-AGO is $X^{(1)} = \left\{ X^{(1)}(1), X^{(1)}(2), \cdots, X^{(1)}(r) \right\}$ and the sequence $Z^{(1)}\left(X^{(1)}(k) \right) = \frac{1}{2}\left(X^{(1)}(k) + X^{(1)}(k-1) \right)$ is generated next to the mean value, the nonlinear model of water pollution can be obtained:

$$X^{(0)}(k) + aZ^{(1)}(k) = b\left(Z^{(1)}(k) \right)^{\alpha} \qquad \ldots(3)$$

That is the power model of the point source water pollution prediction model.

$$\frac{dX^{(0)}(k)}{dt} + aX^{(1)}(k) = b\left(X^{(1)}(k) \right)^{a} \qquad \ldots(4)$$

It is the whitening differential equation of power model of point source water pollution prediction model.

When $a = 2$, say:

$$X^{(0)}(k) + aZ^{(1)}(k) = b\left(Z^{(1)}(k) \right)^{2} \qquad \ldots(5)$$

It is a grey non-point source water pollution prediction model.

$$\frac{X^{(1)}(k)}{dt} + aX^{(1)}(k) = b\left(X^{(1)}(k)\right)^2 \qquad \ldots(6)$$

It is the whitening differential equation of grey non-point source water pollution prediction model. Namely:

$$\hat{X}^{(1)}(k) = \left[\frac{b}{a} + \left(\frac{1}{X^{(0)}(1)} - \frac{b}{a}\right)e^{ak}\right]^{-1} \qquad \ldots(7)$$

It is the time response formula of grey non-point source water pollution prediction model.

The progressive reduction value is:

$$\hat{X}^{(0)}(k) = \hat{X}^{(1)}(k) - \hat{X}^{(1)}(k-1) \qquad \ldots(8)$$

The least square estimation of grey non-point source water pollution prediction model is as follows:

$$\hat{\Phi} = \left(B^T B\right)^{-1} \cdot B^T Y = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} \qquad \ldots(9)$$

Among them,

$$B = \begin{pmatrix} -Z^{(1)}(2) & \left(Z^{(1)}(2)\right)^2 \\ \vdots & \vdots \\ -Z^{(1)}(n) & \left(Z^{(1)}(n)\right)^2 \end{pmatrix} \qquad \ldots(10)$$

From the solution of Verhulst equation, we can see that if $a > 0$, then $X^{(1)}(k) \to 0 (k \to \infty)$; ; if $a < 0$, then $X^{(1)}(k) \to \frac{b}{a}(k \to \infty)$, that is, there is an infinite k, which makes $X^{(1)}(k)$ and $X^{(1)}(k-1)$ fully close, at this time $X^{(0)}(k+1) = X^{(1)}(k+1) + X^{(1)}(k) \approx 0$. If the original prediction data of water pollution history is "s" or part of it is "s", the original prediction data of water pollution history can be directly taken as $X^{(1)}$, and its 1-AGO is $X^{(0)}$. The agricultural non-point source water pollution prediction model is established to directly predict the original data of water pollution
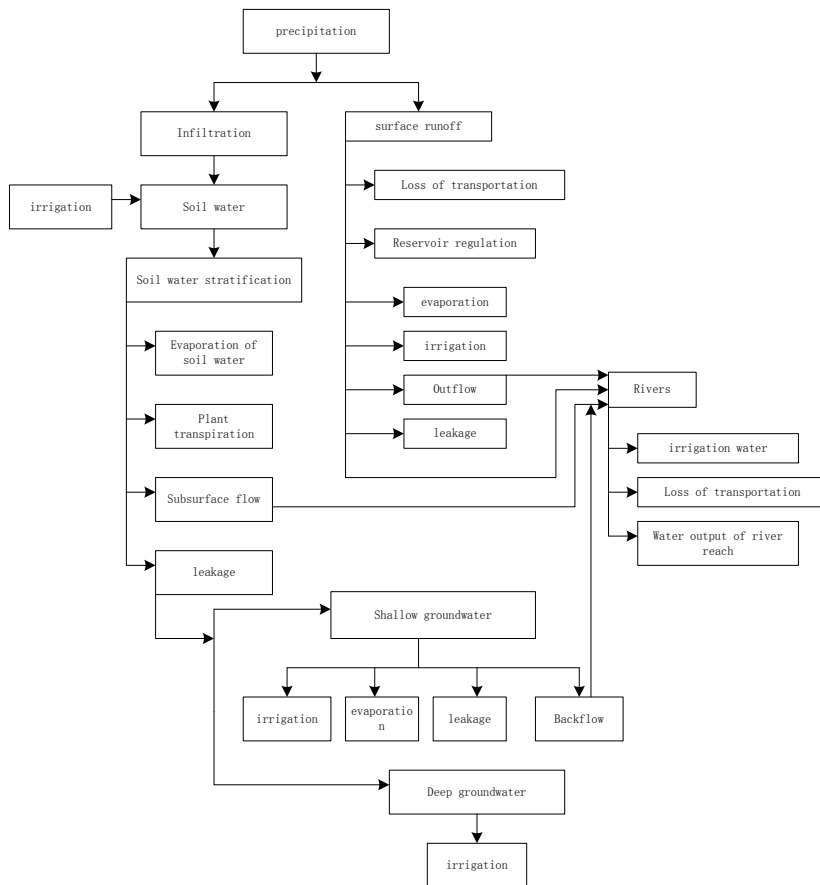

Fig. 1: Structure of agricultural non-point source water pollution prediction model.

history. The structure of agricultural non-point source water pollution prediction model based on grey correlation method is shown in Fig. 1.

To sum up, on the basis of the grey correlation method, the influencing factors of the model are summarized according to the historical data sources of agricultural non-point source water pollution. The data standardization method is used to standardize the historical data of agricultural non-point source water pollution and complete the pretreatment of the historical data of agricultural non-point source water pollution. The power model of the point source water pollution prediction model is used to establish the whitening differential equation, Through the structural design of agricultural non-point source water pollution prediction model, the establishment of agricultural non-point source water pollution prediction model was realized (Alexander et al. 2017).

EXPERIMENTAL ANALYSIS

Experimental Design

In the actual prediction experiment, it is more effective and difficult to determine a specific method than other methods, while the agricultural non-point source water pollution prediction model has the following two advantages over the point source water pollution prediction model.

1. Agricultural non-point source water pollution prediction model has better prediction accuracy and robustness, and can reduce a lot of working time. In the early stage of prediction, generally all the point source water pollution prediction models are applied and then compared, but it is difficult to find the best prediction model through this method.

2. The uncertainty of the point source water pollution prediction model can be dispersed by the combination method, so as to improve the accuracy of the prediction model.

The prediction value $f_i (i = 1, 2, \ldots, k)$ of k models is made for a certain prediction object f, and the error of each water pollution prediction model is as follows:

$$e_i = f_i - f \qquad \ldots(11)$$

The variance is:

$$\sigma_{ii} = \left(f_i - \frac{1}{n} \sum_{t=1}^{n} f_{it} \right)_i^2 \qquad \ldots(12)$$

Where f_{it} represents the t prediction value of the i prediction model.

If the weight of various water pollution prediction models is $W = (\omega_1, \omega_2, \cdots, \omega_k)^T$, it meets the following requirements:

$$\sum_{i=1}^{k} \omega_i = 1 \quad (\omega_i \geq 0) \qquad \ldots(13)$$

Then the basic form of agricultural non-point source water pollution prediction model is:

$$f = \sum_{i=1}^{k} \omega_i f_i \qquad \ldots(14)$$

The weighted geometric average prediction model is expressed as follows:

$$f = \prod_{i=1}^{k} f_i^{w_i}, i = 1, 2, \cdots, k \qquad \ldots(15)$$

Grey Correlation Method to Determine the Weight

The key to establish the agricultural non-point source water pollution prediction model is how to determine the weight of each point source water pollution prediction model. The weight determination of common non-point source water pollution prediction models includes grey correlation method, arithmetic average method, standard deviation method and simple weighting method. These weighting methods are briefly introduced below.

(1) Traditional Weight Determination Method

Arithmetic average method: This method gives the same weight to each point source water pollution prediction model, and can only be used when the importance of each point source water pollution prediction model is not understood. This method is simple in calculation, but the accuracy of prediction is not high because all prediction models are treated equally.

$$\omega_i = \frac{1}{k}, i = 1, 2, \cdots, k \qquad \ldots(16)$$

Reciprocal method of variance: This method is an improvement of arithmetic average method, which requires a certain understanding of the prediction model of source water pollution at each point. The greater the square sum of the prediction errors of the point source water pollution prediction model, the lower the prediction accuracy, so it should be given a greater weight in the non-point source water pollution prediction model.

$$w_i = \frac{D_i^{-1}}{\sum\limits_{i=1}^{k} D_i^{-1}}, i = 1, 2, \cdots, k \qquad \ldots(17)$$

Where D_i^{-1} is the sum of the error squares of the i prediction model.

Simple weighting method: The smaller the sum of square error of point source water pollution prediction model is, the greater the weight it should be given to the combined prediction model.

$$w_i = \frac{i}{\sum\limits_{i=1}^{k} i} = \frac{2i}{k(k+1)}, i = 1, 2, \cdots, k \qquad \ldots(18)$$

(2) Grey Relation Method to Determine Weight

In information theory, grey relation method is a measure of uncertainty. The larger the amount of information, the smaller the uncertainty, the smaller the grey value, and vice versa. Therefore, we can use the grey correlation method to judge the prediction accuracy of a certain prediction model. The higher the prediction accuracy is, the greater the influence of the prediction model on the combined prediction model is, and the smaller the grey value is. The modeling steps of grey relation method are as follows:

Step 1: Assuming that there are n months and m water pollution prediction models, x_{ij} represents the value ($i = 1, 2, \ldots, n; j = 1, 2, \ldots, m$) of the j-th prediction model in the i-th month;

Step 2: Calculate the proportion of the value in the i-th month under the j-th prediction model in the prediction model:

$$p_{ij} = \frac{x_{ij}}{\sum\limits_{i=1}^{n} x_{ij}} (i = 1, 2, \cdots, n; j = 1, 2, \cdots, m) \qquad \ldots(19)$$

Step 3: Calculate the grey value of the j-th prediction model;

$$e_j = -k \sum\limits_{i=1}^{n} p_{ij} \ln\left(p_{ij}\right) \qquad \ldots(20)$$

Among them, $k = \frac{1}{\ln(n)} > 0$ meets $e_j > 0$.

Step 4: Grey redundancy of calculation information;

$$d_j = 1 - e_j \qquad \ldots(21)$$

Step 5: Calculate the weight of each water pollution prediction model;

$$w_j = \frac{d_j}{\sum\limits_{j=1}^{m} d_j} \qquad \ldots(22)$$

Step 6: Calculate the prediction value of non-point source water pollution prediction model in n months;

$$s = \sum\limits_{i=1}^{n} \left(\sum\limits_{j=1}^{m} w_j \cdot p_{ij} \right) \qquad \ldots(23)$$

Because the grey correlation method can reflect the utility value of the point source water pollution prediction model more objectively, and the weight obtained has high credibility, this paper uses the grey correlation method to determine the weight value of the non-point source water pollution prediction model.

Model Comparison Results

First, the grey prediction model, the point source water pollution prediction model and the non-point source water pollution prediction model are predicted and tested; secondly, the three prediction models are used to compare the predicted value and the actual value of the three agricultural pollutants, as given in Tables 4 to 9.

According to the comparison of chloride concentration prediction model results in Table 7, the maximum relative error of grey prediction model is 0.103, and the minimum relative error is 0.002; the maximum relative error of point source water pollution prediction model is 0.099, and the minimum relative error is 0.001; the maximum relative error of non-point source water pollution prediction model based on grey correlation method is 0.072, and the minimum relative error, 0.0015, and the total error of the grey prediction model is 0.408, the total error of the point source water pollution prediction model is 0.518, and the total error of the non-point source water pollution prediction model based on the grey correlation method is 0.309. Therefore, the chloride concentration prediction model of this method has the smallest error and the highest precision.

Table 4: Accuracy comparison of prediction models for chloride concentration.

Model	MSE	MAE	MAPE
Grey prediction model	0.1358	0.0399	2.47%
Prediction model of point source water pollution	0.1465	0.0354	5.18%
Prediction model of non-point source water pollution based on grey correlation method	0.1011	0.0269	3.09%

Table 5: Comparison of prediction model accuracy for total calcium concentration.

Model	MSE	MAE	MAPE
Grey prediction model	0.0902	0.0144	5.12%
Prediction model of point source water pollution	0.0617	0.0354	5.18%
Prediction model of non-point source water pollution based on grey correlation method	0.0407	0.0045	4.96%

Table 6: Accuracy comparison of prediction models for TSS concentration.

Model	MSE	MAE	MAPE
Grey prediction model	0.0442	0.0038	3.28%
Prediction model of point source water pollution	0.0425	0.0030	16.54%
Prediction model of non-point source water pollution based on grey correlation method	0.0342	0.0020	1.97%

Table 7: Comparison of prediction model results for chloride concentration.

Sequence	Actual value	Grey prediction model		Prediction model of point source water pollution		Prediction model of non-point source water pollution based on grey correlation method	
		predicted value	relative error	predicted value	relative error	predicted value	relative error
1.	29285	29985	0.103	29367	0.003	29329	0.0015
2.	30490	30293	-0.065	29469	-0.034	30103	-0.0127
3.	32087	30777	-0.041	29595	-0.078	30497	-0.0496
4.	29709	31268	0.052	29752	0.001	30515	0.0271
5.	31830	31767	-0.002	29948	-0.059	30857	-0.0306
6.	32465	32273	-0.006	30193	-0.070	31153	-0.0404
7.	31962	32788	0.025	30502	-0.046	31884	-0.0025
8.	34278	33312	-0.028	30890	-0.099	31893	-0.0696
9.	32234	33843	0.050	31384	-0.026	32331	0.0030
10.	35660	34383	-0.036	32014	-0.102	33093	0.0720

According to the comparison of the results of the total calcium concentration prediction model in Table 8, the maximum relative error of the grey prediction model is 0.188, and the minimum relative error is 0.008; the maximum relative error of the point source water pollution prediction model is 0.254, and the minimum relative error is 0.015; the maximum relative error of the non-point source water pollution prediction model based on the grey correlation method is 0.1318, and the minimum relative error, 0.0032, and the total error of the grey prediction model is 0.705, the total error of the point source water pollution prediction model is 1.279, and the total error of the non-point source water pollution prediction model based on the grey correlation method is 0.4857. Therefore, the error of the total calcium concentration prediction model of this method is the smallest and the accuracy is the highest.

From the comparison of the results of Table 9 TSS concentration prediction model, it can be seen that the maximum relative error of the grey prediction model is 0.243, and the minimum relative error is 0.007; the maximum relative error of the point source water pollution prediction model is 0.256, and the minimum relative error is 0.061; the maximum relative error of the non-point source water pollution prediction model based on the grey correlation method is 0.1454, and the minimum relative error, 0.0001, and the total error of the grey prediction model is 0.652, the total error of the point source water pollution prediction model is 2.444, and the total error of the non-point source water pollution prediction model based on the grey correlation method is 0.5591. It can be seen that the TSS concentration prediction model of this method has the smallest error and the highest precision. It can be seen that the prediction model of agricultural non-point source water pollution based on the grey correlation degree method is better than that based on the point source water pollution prediction model. From the comparison of MAE, MSE and MAPE results, the prediction accuracy of

Table 8: Comparison of results of prediction models for total calcium concentration.

Sequence	Actual value	Grey prediction model		Prediction model of point source water pollution		Prediction model of non-point source water pollution based on grey correlation method	
		predicted value	relative error	predicted value	relative error	predicted value	relative error
1.	2111	2485	0.188	2463	0.167	2218	0.0508
2.	2297	2506	0.090	2829	0.232	2564	0.0063
3.	2587	2751	0.071	3198	0.236	2844	0.0993
4.	2837	3019	0.064	3558	0.254	3211	0.1318
5.	3345	3313	-0.009	3897	0.165	3557	0.0633
6.	3844	3635	-0.054	4208	0.095	3832	-0.0032
7.	4304	3989	-0.073	4486	0.042	4181	-0.0284
8.	4658	4378	-0.060	4728	0.015	4440	-0.0469
9.	4844	4805	-0.008	4934	0.019	4721	-0.0254
10.	4844	5273	0.088	5106	0.054	4991	0.0303

Table 9: Comparison of the results of various prediction models for TSS concentration.

sequence	actual value	Grey prediction model		Prediction model of point source water pollution		Prediction model of non-point source water pollution based on grey correlation method	
		predicted value	relative error	predicted value	relative error	predicted value	relative error
1.	618	786	0.343	721	0.167	649	0.0509
2.	664	710.8	0.070	834	0.256	681	0.0259
3.	776	797.8	0.028	955	0.231	844	0.0879
4.	889	895.4	0.007	1083	0.218	945	0.0632
5.	968	1004.9	0.038	1215	0.255	1108	0.1454
6.	1163	1127.8	-0.030	1349	0.159	1188	0.0215
7.	1324	1265.8	-0.044	1482	0.119	1324	0.0001
8.	1466	1420.6	-0.031	1611	0.098	1532	0.0452
9.	1635	1594.4	-0.025	1734	0.061	1694	0.0359
10.	1698	1789.4	0.036	1848	0.088	1839	0.0831

the agricultural non-point source water pollution prediction model based on the grey correlation method is also higher than the other two prediction models. Therefore, the prediction model of agricultural non-point source water pollution based on the grey correlation method can be used for the prediction of agricultural non-point source water pollution in the near future.

DISCUSSION AND SUGGESTIONS

In view of the shortcomings of the application of the prediction model of agricultural non-point source water pollution based on the grey correlation method, the following suggestions are put forward:

1. Due to the limited level of relevant theories and lack of knowledge about agricultural non-point source water pollution, only common and main influencing factors are listed when selecting the influencing factors of water pollution, and many indexes are not considered fully, so the index system in this paper does not represent all the influencing factors in the process of shale gas development. In addition, the determination of index system weight will be biased, which will affect the final prediction results. Therefore, in the prediction of such problems, we should try to combine the actual situation, preferably with the participation of professionals, so that the prediction model can be more applied to the actual shale gas production project.

2. In the agricultural non-point source water pollution prediction model, neural network method can also be used to determine the weight, and this paper determines the constant weight. To make the agricultural non-point source water pollution prediction model better applied in practice, the combination model with variable weight

coefficient needs further study. In this paper, the point source water pollution prediction model, grey model and colony grey prediction model are combined to improve the prediction accuracy, and no more other prediction models are involved. The advantages and disadvantages of more single prediction models need further study, to get a better combination prediction model for practical application.

3. Water environment pollution is the main factor affecting agricultural development. In the detection process of agricultural development wastewater and backwater reuse, special standards should be formulated by the government departments to strictly restrict the treatment process of various types of wastewater. At the same time, the government should encourage the wastewater to be reused as fracturing fluid after it reaches the standard completely, to avoid the waste of water resources. Besides, the developers should strive to develop new environmental protection pressure Crack liquid to reduce water pollution.

CONCLUSION

In this paper, a model of agricultural non-point source water pollution prediction based on the grey correlation method is designed. Firstly, the historical raw data of agricultural non-point source water pollution are preprocessed; secondly, the advantages and disadvantages of three models in small sample data prediction are objectively analysed, and the agricultural non-point source water pollution prediction model is established. In the experiment, the weight is determined by the grey correlation method, and the point source water pollution prediction model, the grey prediction model and the agricultural non-point source water pollution prediction model

model based on the grey correlation method are compared. The model is improved, so the prediction ability and accuracy of the model are improved. According to the model, the forecast data can be reported to relevant departments in time, to make the next development plan. Although this paper has completed the design of the agricultural non-point source water pollution prediction model based on the grey correlation degree, the preliminary experimental results are good. But the model has a large amount of calculation and high difficulty of calculation. In the future, we can strengthen the research in this area and design a model with lower difficulty.

ACKNOWLEDGEMENT

REFERENCES

Alexander, K., Pia-Melich, B., Georg, S.K., Thomas, V., Daniel, S., Stuart, M., Julien, M., Joseph, Z., Alessandro, S., Rupert, L. and Siegfried, K. 2017. A new prediction model for evaluating treatment-resistant depression. Journal of Clinical Psychiatry, 78(2): 215.

Anooob, P., Santhoshkumar, A.V. and Roby, C.P. 2017. Impact of particulate pollution on photosynthesis, transpiration and plant water potential of teak (*Tectonagrandis*L.). Current Science, 112(6): 1272-1276.

Bai, X.Y., Wei, S., Shi, X. and Chen X.H. 2018. Using the HSPF model to study the effects of precipitation on nonpoint source pollution in Dongjiang Basin. Journal of Irrigation and Drainage, 37(7): 112-119.

Baranov, A., Salvesen, K.Å., Vikhareva, O. 2018. Validation of prediction model for successful vaginal birth after cesarean delivery based on sonographic assessment of hysterotomy scar. Ultrasound in Obstetrics &Gynecology, 51(2): 189-193.

Dadras, E.Y., Yazdani, A., Nicknam, A. and Eftekhari S.N. 2018. Incorporating source rupture characteristics into the Near-Fault pulse prediction model. Bulletin of the Seismological Society of America, 108(1): 200-209.

Joy, T.A. 2018. Assessing non-point source pollution models: Areview. Polish Journal of Environmental Studies, 27(5): 1913-1922.

Li, Y.P. 2018. Prevention and control of water pollution. China Water-Power Press.

Liouane, Z., Lemlouma, T., Roose, P., Weis, F. and Messaoud, H. 2017. An improved extreme learning machine model for the prediction of human scenarios in smart homes. Applied Intelligence, 48(8): 1-14.

Long, Y.Q., Cui, T.T., Li, W., Wu, C.Y. and Li, Y.G. 2017. Application and sensitivity analysis of geostatistical approach to groundwater pollution source identification. Journal of Hydraulic Engineering, 48(7): 816-824.

Ploton, P., Barbier, N., Couteron, P., Antin, C.M., Ayyappan, N., Balachandran, N., Barathan, N., Bastin, J.F., Chuyong, G., Dauby, G., Droissart, V., Gastellu-Etchegorry, J.P., Kamdem, N.G., Kenfack, D., Libalah, M., Mofack II, G., Momo, S.T., Pargal, S., Petronelli, P., Proisy, C., Rejou-Mechain, M., Sonke, B., Texier, N., Thomas, D., Verley, P., Zebaze Dongmo, D., Berger, U. and Pelissier, R. 2017. Toward a general tropical forest biomass prediction model from very high resolution optical satellite images. Remote Sensing of Environment, 200: 140-153.

Wang, H., Zhang, Feng., Xue Hui-feng. 2019. Construction of Prediction Model and Its Application of Water Pollution Load Intensity Based on Grey Effectiveness. Water Saving Irrigation,(04):72-76.

Wang, L., Chang, Z. and Bao, J. 2017. Prediction model for the thermal conductivity of concrete based on its composite structure. Journal of Hydraulic Engineering, 48(7): 765-772.

Wen, F., Wu, N. and Gong, X. 2020. China's carbon emissions trading and stock returns. Energy Economics, 86: 104627.

Wen, F.H., Zhao, Y.P. and Zhang, M.Z. 2019. Forecasting realized volatility of crude oil futures with equity market uncertainty. Applied Economics, 51(59): 6411-6427.

Wen, S., Xing L.Q. and Zhang H. 2018. Simulation for measurement data pre-processing scheme of 3D target tracking. Computer Simulation, 35(5): 391-396.

Zhang, P.Y., Gou, C.X., Wu, Y., Zhang, D., Li, F. and Mou, Z.S. 2018. Characteristics of rural non-point source pollution in Kaijiang Basin of Sichuan Province. Journal of Agricultural Resources and Environment, 35(5): 398-404.