Study on the prediction of mineralization degree of groundwater based on grey prediction model

Xue Dong1, Sudan Gu and Tong Li

College of Water Resources and Hydropower, Hebei University of Engineering, Handan, 056000, Hebei, China

1 Email: 1582692462@qq.com

Abstract. The mineralization of groundwater has a direct effect on the safety of water use of residents and irrigation of crops, especially in the areas with low rainfall, large overproduction and high water consumption in Wei County, Handan. In this paper, it takes Wei County as the study area, takes Qiandianpo and Cai Xiaozhuang, two mineralization over the period 2010-2016 for wellhead measurements as dependent variables, and collects well water level, extraction amount and rainfall to be as independent variables, and then establishes the grey prediction model of GM (1,N). The numerical model of groundwater is simulated and forecasted by MATLAB auxiliary software, and the precision of the model is determined. The simulated and forecasted results are compared with the GM (1,N) model and the BP neural network model. It's showed based on the result that the simulation value of GM (1,N) fits best with the actual value, and the precision is the highest. Therefore, the grey GM (1,N) model is feasible and accurate in predicting mineralization of groundwater, and the prediction results of GM (1,N) model are taken as the predicted values for the period 2017-2020 in the area around the two logging heads, such as Qiandianpo, Cai Xiaozhuang. The predicted values show that the degree of mineralization degree of groundwater in Wei County has been on the rise in the past four years, which provides data information for the relevant departments in Wei County to manage, distribute and dispatch groundwater resources.

1. Introduction

Groundwater is an important source of water supply in Handan plain, which has made important contribution to local economic and social development. However, due to the limited amount of surface water resources and the increasing demand, the groundwater has been overexploited in Handan area for many years, especially Wei County, which is used as a research area. Wei County is with arable area of 922000 mu, effective irrigation area of 843800 mu, which is one of the 80 grain production core counties in Hebei Province. Agricultural irrigation water accounts for more than 80% of the total annual water use in the county. Agricultural irrigation accounts for more than 80% of the total annual water use in the county. Irrigation water also depends on groundwater for a long time, As a result, groundwater in the whole county area is overexploited to varying degrees, and the groundwater level continues to decrease, causing land subsidence, water quality deterioration and other environmental and geological problems [1], threatening the people of Wei County in water safety and crop irrigation. The introduction of mineralization degree of groundwater in the study area can effectively have a understanding of the changing trend of groundwater resources in the study area, thus taking appropriate measures to optimize the distribution of groundwater resources.
Based on the measured data of the mineralization degree of groundwater in Wei County from 2010 to 2016, the paper uses the GM (1,N) numerical model [2], and introduces three main influencing factors, such as well water level, mining capacity and rainfall, to simulate and predict the change situation of mineralization degree in Wei County. The simulation results are compared with those of GM (1,1) model and BP neural network model respectively. The feasibility of the model is confirmed, and the final simulation prediction results are determined. The changing trend of mineralization degree of groundwater in Wei County in recent years is determined by means of the predicted results within 4 years.

2. Basic data for the study area

2.1. Physical geography overview of the study area
Wei County is located in the southeast of Handan, at north latitude 36°03′6″~36°26′30″, east longitude 114°43′40″~115°07′24″, next to Guangping in the north, Chengan and Linzhang in the west, and connecting with Danming in the east. It belongs to the Heilong River basin in North China, formed by impact and precipitation of the Yellow River and Zhang River. The topography is slowly inclined from the southwest to the northeast, open and gentle, and the surface soil is mainly light loamy. The climate belongs to the temperate semi-humid semi-arid continental monsoon climate, with the same period of rain and hot. The economic sources of Wei County mainly focuses on agriculture, which is the national grain production capacity county. Because of the need of agricultural water use, there are different degrees of overexploiting in Wei County, which also affect the change of the mineralization degree of groundwater. The geographical location map for Wei County is shown in figure 1.

2.2. Data for study area
This paper selects two well logs with relatively complete data collection: such as Qiandianpo, Cai Xiaozhuang, and collects the value and the water level value of mineralization degree of groundwater of the two logging during 2010-2016 at the frequency of once a month. In order to reflect the change situation of the whole Wei County, this paper chooses the total mining quantity and the total rainfall to take the place of the mining amount of well and the rainfall in the region. In order to improve the accuracy of the data and reflect the changing trend directly, the mean values of the mineralization
degree of groundwater and water level are taken as the simulation data. Collected data are as shown in Table 1.

| Well logging point | Year | Mineralization degree (mg/L) | Well water level (m) | mining amount of well (m) | Rainfall (mm) |
|--------------------|------|------------------------------|---------------------|--------------------------|--------------|
| Qiandianpo         | 2010 | 3585                         | 31.39               | 1.8179                   | 543.9        |
|                    | 2011 | 3460                         | 32.23               | 1.062                    | 521.6        |
|                    | 2012 | 3220                         | 33.32               | 1.9676                   | 469.4        |
|                    | 2013 | 2920                         | 34.08               | 1.914                    | 548.4        |
|                    | 2014 | 3575                         | 34.95               | 1.058                    | 509.1        |
|                    | 2015 | 3615                         | 35.42               | 0.908                    | 459.8        |
|                    | 2016 | 3250                         | 35.51               | 0.87                     | 597.8        |
| Cai Xiaozhuang     | 2010 | 2611                         | 39.94               | 1.8179                   | 543.9        |
|                    | 2011 | 2360                         | 41.62               | 1.062                    | 521.6        |
|                    | 2012 | 2135                         | 40.81               | 1.9676                   | 469.4        |
|                    | 2013 | 2597                         | 41.22               | 1.914                    | 548.4        |
|                    | 2014 | 2825                         | 41.82               | 1.058                    | 509.1        |
|                    | 2015 | 2743                         | 45.81               | 0.908                    | 459.8        |
|                    | 2016 | 2523                         | 45.93               | 0.87                     | 597.8        |

3. Establishment of grey prediction GM (1,N) model

GM model is a Grey Model. The common grey model includes GM (1,N), GM (1,1) model. Grey modeling is called the core of grey system theory and also the bridge of combining grey system theory with practice. Based on the establishment of grey model, the law masked by messy raw data can be founded, and it can provide a new method for forecasting and analyzing the data, and organize the irregular raw data into regular generating sequence. Therefore, the data obtained directly from the grey model is not the prediction data, but the generated data. The messy raw data built by the grey process is called the grey model.

3.1. Simulation steps of GM(1,N) prediction model

To determine the target quantity and the influence factor of GM (1,N) . The objective variable (dependent variable), N associated variables (independent variables) is determined around the content to be studied in the research area, and the trend of variable change is investigated and the simulation and prediction of target variable is realized based on the associated variable. There are three related variables in this paper, which are well water level, rainfall and extraction amount, and the target variable is the mineralization degree of monitoring well. Only by accurately determining the information of 1+N variables can we lay the foundation for the establishment of the model.

\[
X_1^{(0)} = \{x_1^{(0)}(1), x_1^{(0)}(2), \ldots, x_1^{(0)}(7)\} \\
X_2^{(0)} = \{x_2^{(0)}(1), x_2^{(0)}(2), \ldots, x_2^{(0)}(7)\}
\]
where \( X_3^{(0)} = \{ x_1^{(0)}(1), x_2^{(0)}(2), \ldots, x_7^{(0)}(7) \} \) and \( X_4^{(0)} = \{ x_1^{(0)}(1), x_2^{(0)}(2), \ldots, x_7^{(0)}(7) \} \)

Calculation coefficient matrix B and constant matrix Y. According to the GM (1,N) model, the system state equation matrix is established, and the calculation is as follows:

\[
B = \begin{bmatrix}
-\frac{1}{2}(x_1^{(1)}(1) + x_1^{(1)}(2)) & x_2^{(1)}(2) & x_3^{(1)}(2) & x_4^{(1)}(2) \\
-\frac{1}{2}(x_1^{(1)}(2) + x_1^{(1)}(3)) & x_2^{(1)}(3) & x_3^{(1)}(3) & x_4^{(1)}(3) \\
-\frac{1}{2}(x_1^{(1)}(3) + x_1^{(1)}(4)) & x_2^{(1)}(4) & x_3^{(1)}(4) & x_4^{(1)}(4) \\
-\frac{1}{2}(x_1^{(1)}(4) + x_1^{(1)}(5)) & x_2^{(1)}(5) & x_3^{(1)}(5) & x_4^{(1)}(5) \\
-\frac{1}{2}(x_1^{(1)}(5) + x_1^{(1)}(6)) & x_2^{(1)}(6) & x_3^{(1)}(6) & x_4^{(1)}(6)
\end{bmatrix}
\]

and \( Y = \begin{bmatrix} X_1^{(0)}(2) \\ X_1^{(0)}(3) \\ X_1^{(0)}(4) \\ X_1^{(0)}(5) \\ X_1^{(0)}(6) \end{bmatrix} \)

Determination of whitening differential equation and solution of whitening equation for grey state model. Coefficient matrix B and constant matrix Y are used to solve the coefficient sequence of the parameters and the whitening equation is determined.

\[
\frac{dx_1^{(1)}}{dt} + ax_1^{(1)} = b_2x_2^{(1)} + b_3x_3^{(1)} + b_4x_4^{(1)}
\]

\[
\hat{x}_1^{(1)} = e^{-at}[x_1^{(1)}(0) - \hat{\alpha} \sum_{i=2}^{4} \int b_i x_i^{(1)}(t)e^{at} dt]
\]

\[
\hat{\alpha} = (a, b_2, b_3, b_4)
\]

In which \( x_1 \) is to monitor the mineralization degree of well, \( x_2 \) monitor the well water level, \( x_3 \) monitor the exploitation of well water level and \( x_4 \) is regional precipitation.

To solve the simulation and predicted values, and to determine the fitting degree of the model. Using the calculated \( \hat{\alpha} \) value to calculate the simulation and predicted values of \( X_1^{(0)} \), determining the model fitting accuracy of monitoring wells, calculating the relative error and variance ratio between the simulated values and the measured values, and analyzing the result of model simulation. After knowing the modeling process of GM (1,N) model, it’s programmed the running process of GM (1,N) by MATLAB software, and then input the original data such as mineralization degree, groundwater level, groundwater exploitation, regional rainfall and so on from 2010 to 2016 into the compiled program so that the corresponding simulation and prediction values are obtained.

3.2. Application effect of GM (1,N) model in Wei County

The data collected in Table 1 are input into the program compiled by MATLAB, and the simulated and predicted values of the mineralization degree of groundwater in Wei County are obtained after running in which simulation interval is 2010-2016, and the variation of the mineralization degree of groundwater in 2017-2020 is predicted. Finally, the curve is drawn and the trend line of the fitting curve is made. The simulation and prediction values are shown in Table 2, and the fitting curve graph is shown in figure 2 and figure 3.
Table 2. Simulation and predicted values of well logging points.

| Year | Qiandianpo | Cai Xiaozhuang |
|------|------------|----------------|
| 2010 | 3585       | 2611           |
| 2011 | 3462.5     | 2501.1         |
| 2012 | 3038.7     | 2137.6         |
| 2013 | 2967.4     | 2205.7         |
| 2014 | 3671.9     | 2681.8         |
| 2015 | 3599.2     | 2697.8         |
| 2016 | 3242.8     | 2591.6         |
| 2017 |            | 2774.4         |
| 2018 | 3583.5     | 2474.3         |
| 2019 | 3810.6     | 2522.1         |
| 2020 | 3796.3     | 2475.4         |

Figure 2. Qiandianpo fitting curve graph. Figure 3. CaiXiaozhuang fitting curve graph.

The prediction model must pass statistical test to forecast its grade of model accuracy [3]. The methods of test are usually as follows, such as residual test method, posteriori error test method, correlation test method and index value difference test method respectively. In this paper, the residual test method is used to calculate the accuracy value $p_0$. For model accuracy $p_0$, it's generally requestment is $p_0>80\%$, it is best to achieve $p_0>90\%$.

Table 3. Model fitting accuracy $p_0$.

| Monitoring well | Qiandainpo | Cai Xiaoazhuang |
|-----------------|------------|----------------|
| $p_0(\%)$       | 98.47      | 97.53          |
4. Feasibility comparison of GM(1,N) model

This paper intends to use other commonly used models for further simulation and prediction of Wei County to determine a more accurate model. In addition to GM (1,N), there are also BP neural network prediction models, GM (1,1) model. The two models are applied to two observation wells in Wei County and the simulated values are obtained. In order to determine the better feasibility and accuracy of GM (1,N) model, the relative error and root mean square error of the simulated and actual values of the two observation wells under the BP neural network prediction model[4], GM (1,1) model are calculated respectively and the results are compared and analyzed with those of the GM (1,N) model one by one.

4.1. BP neural network prediction model

BP neural network [5-6] is a multi-layered artificial neural network, composed of several layers of neurons, which can be divided into input layer, hidden layer and output layer and the action of neurons varies from layer to layer. The input information of the BP neural network model passes from the input layer to the output layer through the hidden layer (one or more layers). If there is no desired output in the output layer, the input information is turned back and the error signal is returned along the original path. To modify the weights of each layer of neurons by learning, the error signal is minimized, and the state of each layer of neurons will have an effect on the state of the next layer of neurons. On the basis of understanding the principle of the model, the MATLAB application software is used to simplify the calculation of the data. In the end, the simulation value table 5, the mean error and the mean square error comparison table 3 are obtained.

| Monitoring well   | GM(1,N) | BP          |
|-------------------|---------|-------------|
|                   | Mean error | Mean square error | Mean error | Mean square error |
| Qiandianpo        | 1.53     | 80.01       | 4.39       | 364.60           |
| Cai Xiaozhuang    | 2.47     | 81.70       | 8.05       | 317.99           |

The results show that the mean error of Qiandianpo under the GM(1,N) model is 1.53 and mean square deviation is 80.01; and the mean error is 4.39 and mean square deviation is 364.60 respectively in BP neural network model. The mean error of Cai Xiaozhuang under the GM(1,N) model is 2.47 and mean square deviation is 81.70; and the mean error is 8.05 and mean square deviation is 317.99 respectively in BP neural network model. It is apparent that the GM (1,N) model is the closest to zero error, the fitting accuracy is higher and the simulation effect is better. The reason is that the BP neural network model is more suitable for the case of longer series data and more complete data, and is with more superiority for the short series data GM (1,N) model.

4.2. GM(1,1) model prediction model

GM(1,1)model [7] is a commonly used grey model and the basic model of grey prediction in which the first "1" in GM (1,1) is expressed as the first order equation, and the second "1" is expressed as a variable. According to the known information, the original data is added and the newly generated sequence is established to deal with the information by the first order linear differential equation to study the internal laws and characteristics of things. On the basis of understanding the principle of the model, the MATLAB application software is used to simplify the calculation of the data. In the end, the simulation value table 5, the mean error and the mean square error comparison table 4 are obtained.
Table 5. Comparison table of GM (1,N) model and GM (1,1) model simulation.

| Monitoring well | GM(1,N) Mean error | GM(1,N) Mean square error | GM(1,1) Mean error | GM(1,1) Mean square error |
|----------------|-------------------|--------------------------|-------------------|--------------------------|
| Qiandianpo     | 1.53              | 80.01                    | 5.6               | 218.81                   |
| Cai Xiaozhuang | 2.47              | 81.70                    | 8.58              | 234.80                   |

The results show that the mean error of Qiandianpo under the GM(1,N) model is 1.53 and mean square deviation is 80.01; and the mean error is 5.6 and mean square deviation is 218.81 respectively under the GM(1,1) model. The mean error of Cai Xiaozhuang under the GM(1,N) model is 2.47 and mean square deviation is 81.70; and the mean error is 8.58 and mean square deviation is 234.80 respectively under the GM(1,N) model. It is apparent that the mean error of GM (1,N) model is the closest to 0, the mean square deviation is the smallest, the fitting accuracy is higher, and the simulation effect is better. The reason is that the, GM (1,N) model takes into account other external influence factors, and can consider the effects of multiple factors simultaneously, which is more suitable for the practical situation than the GM (1,1) model.

As table 6 shown, it can be seen that the fitting accuracy of the two monitoring wells under the GM (1,N) model is higher than that of the other two methods, which fully indicates the accuracy of simulation of the mineralization degree of groundwater in Wei County by GM (1,N) model. Therefore, the result predicted by GM (1,N) model is used as the final predicted result of Wei County in this paper.

Table 6. Numerical comparison table of model simulation.

| Year | GM(1,N) Simulation value (mg/L) | Relative error (%) | BP Simulation value (mg/L) | Relative error (%) |
|------|---------------------------------|--------------------|----------------------------|--------------------|
| 2010 | 3585                            | 0                  | 3578.5                     | -0.18              |
| 2011 | 3283.4                          | -5.1               | 3449.7                     | -0.3               |
| 2012 | 3305.8                          | 2.66               | 3219.4                     | -0.02              |
| 2013 | 3328.4                          | 13.99              | 2927.1                     | 0.24               |
| 2014 | 3351.1                          | -6.26              | 3567.6                     | -0.21              |
| 2015 | 3374.1                          | -6.66              | 3610.4                     | -0.13              |
| 2016 | 3397.1                          | 4.53               | 4214.5                     | 29.68              |

5. Prediction and evaluation of Wei County achievements

In this paper, GM (1,N) model is constructed in Wei County to predict the mineralization degree of groundwater from 2017 to 2020 to determine the change status of groundwater. The two observation wells, Qiandianpo and Cai Xiaozhuang, are used to collect the four indicators of mineralization degree of groundwater, well water level, pumping capacity and rainfall to simulate and predict and
then to determine the usability of the model. From the fitting curves and fitting accuracy of GM (1,N) model for predicting mineralization degree of two monitoring wells, it can be seen that the fitting effect is good, which indicates that GM (1,N) model can be applied to the study of mineralization degree of groundwater in Wei County; the results of GM (1,N) model are compared with GM (1,1) model and BP neural network model and the final results show that the GM (1,N) model has higher fitting accuracy, which can be better applied in the prediction of mineralization degree of groundwater in Wei County.

From the trend of fitting trend line, it can be seen that the mineralization degree of groundwater in monitoring wells is on the rise, which is higher than that of a few years ago, in which the mineralization degree of groundwater of Qiandianpo monitoring well will reach 3796.3 mg/L in 2020, and that of Cai Xiaozhuang monitoring well will reach 2475.4mg/L in 2020. It's showed based on the result of analysis that, in the next four years, the mineralization degree of groundwater in different areas of Wei County still has rising trend with different degrees. The study provides a preliminary reference for groundwater management, dispatcher and regulation of water and salt in Wei County.

Reference

[1] Xu Tianpeng, Kong Lingjun, Hao Weiwei 2016 The current situation and management opinions of water saving irrigation project in Wei County [J]Science and Technology and Innovation245
[2] Zhang Ziyue, Du Fuhui, Li Xinde et al. 2018 Study on mineralization degree of groundwater in Linzhang County based on GM(1,N) model [J]China Rural Water and Hydropower774-777
[3] Liu Lu 2016 Study on dynamic reliability assessment method of software system based on component [D]Beijing: China University of Petroleum20-34
[4] Xu Youping, Du Hailong, Hu Banghui, et al. 2018 Study on refined prediction method of surface temperature during ice flood period [J]People's Yellow River40(2) 30-35
[5] Li Haikuo, Ceng Xiaojie 2018 Research on pricing scheme based on neural network model " to make money by photography" [J]Science and Technology and Innovation636-38
[6] Dong Qiguang, Zhou Weibo, Hu Banghui et al. 2012 Application of BP neural network in prediction of bury of groundwater in Weibei dryland area [J]Journal of Water Resources and Water Engineering23(4) 112-118
[7] Chen Yu, Pei Jingzu, Meng Dongxia et al. 2006 Prediction and Analysis on Water Quality of the Minxin in Shijiazhuang [J]Research on Forestry in Hebei21(4) 468-470