Research on Forecasting Model of Daily Discharge in Karst Area Based on Mea Grey Neural Network

Xia Li¹,², Xin Jin¹,², * and Pan Guo³

¹Guangxi Key Laboratory of Theory and Technology for Environmental Pollution Control, the University of Technology in Guilin, Guilin, Guangxi 541004, China
²Collaborative Innovation Center for Water Pollution Control and Water Safety Guarantee, the University of Technology in Guilin, Guilin, Guangxi 541004, China
³Guilin Farmland Irrigation Experimental Center Station, Guilin, Guangxi 541105, China

*Corresponding author e-mail: jinxin1390@163.com

Abstract. The water-bearing system in the karst area is complex and changeable. Water in the water-bearing medium has the characteristics of coexistence of fissure flow and pipeline flow, coexistence of laminar and turbulent flow, coexistence of linear and nonlinear flow, coexistence of continuous flow and isolated water body. In areas where economic development is relatively backward, most areas lack data or no data. Based on the characteristics of the karst area in the southwest, this paper proposes a thinking evolution algorithm to optimize the gray neural network model. This method improves the optimization ability of runoff prediction models and effectively overcomes human neural network learning. The speed is slow and there are inherent shortcomings of local minima. After studying the daily flow forecast of Zhenlong Station, it is shown that the relative error of the prediction is small and can be effectively used for short-term runoff prediction.

1. Introduction

Due to the uneven spatial and hydrological conditions, the hydrological process is complex and changeable in time and space, and the monitoring data in the karst area is relatively scarce. There are practical difficulties in establishing a hydrological model of the watershed in the karst area. Currently, when studying the characteristics of rainfall runoff in the karst area, It usually generalizes and combines existing hydrological forecasting models. The application of hydrological models has certain applicability conditions, and the karst area hydrological model is more challenging than other different regions. It is composed of multiple media such as karst gaps, karst caves, and pipelines. The formed karst aquifer system is an evolving complex dynamic system, and its hydrological cycle with atmospheric precipitation, surface water, and soil water is complex. In addition, the water flow in the karst aquifer medium has both linear and non-linear flows, laminar flow with characteristics of coexistence with turbulence, continuous flow and isolated water body, coexistence of fissure flow and pipeline flow, there is no systematic and complete research model that can accurately describe the flow law of water flow in karst aquifer medium[1]. Limitations of the regional hydrological cycle process, combining mathematics such as gray theory, fractal method and artificial neural networkThe
mathematical statistical model established by the method has been applied and developed. Zhang Baoxiong et al. Established a Jinan urban karst groundwater forecasting model based on neural network and genetic algorithm[2], Li Guiming et al. Established a spring flow prediction model coupled with reconstructed phase space and neural network[3], Chen Hongfeng, etc. established a lake using bp artificial neural networkRunoff prediction model of surface karst spring in Ganhe pig farm in South Lota area[4]; Shu Longcang et al. Combined threshold autoregressive model and wavelet bp neural network to predict the outlet flow of Houzhai river basin[5] and other models The predictions have achieved good results. This article attempts to use the evolutionary thinking algorithm to optimize the gray neural network model to forecast the daily flow in the karst area and enrich the research methods and methods of the karst area hydrological forecast model.

2. Mind Evolutionary Algorithm (MEA)
Evolutionary algorithms are a combination of computer science and biological evolution, and gradually develop a type of heuristic random search algorithm. The most famous evolutionary algorithms are: genetic algorithms, evolutionary strategies, and evolutionary planning. The advantages of these evolutionary algorithms are group search, problems and The shortcomings are precocity, slow convergence, etc. In response to the problems of traditional evolutionary algorithms, Sun Chengyi[6] and others proposed a thinking evolutionary algorithm in 1998.

2.1. Overview of Mind Evolutionary Algorithms
For the basic concepts of thinking evolution, algorithm structure, etc., there are detailed introductions in [6] and [7]. Figure 1 shows the structure of the thinking evolution algorithm system.

![Diagram](image)

**Figure 1.** The structure of the thinking evolution algorithm system.

2.2. Basic Thinking of Thinking Evolutionary Algorithm
The basic idea of mea is:

1. Randomly generate individuals of a certain size in the solution space, and search for several winning individuals and temporary individuals with the highest scores according to the score.
2. Take these superior individuals and temporary individuals as the center, and generate some new individuals around each individual, so as to obtain several superior subgroups and temporary subgroups.
3. Perform convergence operations within each subgroup until the subgroup is mature, and use the optimal individual score in the subgroup as the score of the subgroup.
4. After the subgroup is mature, post the scores of the subgroups on the global bulletin board, and perform alienation operations between the subgroups to complete the replacement, abandonment, and
release of individuals in the subgroup between the winning group and the temporary subgroup, and calculate Global optimal individual score.

3. Optimize gray neural network based on ema

3.1. Grey Theory
The gray system theory was first proposed by the Chinese scholar Professor Deng Julong in 1982 [8]. The gray system theory is a method for studying poor information, little data, and uncertainty. The theory does not consider the effect of factors affecting future sequence changes. In view of the poor variation of time series, it is a good method to deal with the difficulties in establishing an accurate model. The gray theoretical data generation methods usually include accumulation, subtraction, and mean generation. In this paper, the model data is calculated by using the accumulation generation (ago). The specific steps are outlined below:

With time series $x^{(0)}$:

$$x^{(0)} = \{x^{(0)}_t | t = 1, 2, \ldots, n\} = (x^{(0)}_1, x^{(0)}_2, \ldots, x^{(0)}_n)$$

(1)

Superimpose a new data sequence $x^{(j)}$ on $x^{(0)}$ once, and the $t$-th term of the new data series $x^{(j)}$ is the sum of the previous $t$-terms of the original data series $x^{(0)}$, that is:

$$x^{(j)} = \{x^{(j)}_t | t = 1, 2, \ldots, n\} = (x^{(j)}_1, \sum_{t=1}^{j} x^{(0)}_t, \sum_{t=2}^{j} x^{(0)}_t, \ldots, \sum_{t=n}^{j} x^{(0)}_t)$$

(2)

Establish the whitening equation based on the new sequence $x^{(j)}$, that is:

$$\frac{dx^{(j)}}{dt} = ax^{(j)} = u$$

(3)

The solution to this equation is:

$$x^{(j)}_t = \left(x^{(0)}_t - \frac{u}{a}\right)e^{-at} + \frac{u}{a}$$

(4)

Do one accumulation of (4) to get the predicted value:

$$x^{(0)}_t = x^{(j)}_t - x^{(j)}_{t-1} | t = 2, 3, \ldots$$

(5)

3.2. Gray neural network
For the system, the output of the neural network can actually be understood as the gray number of the gray system. Therefore, the neural network itself contains gray content and there is a certain similarity in the performance of the two information, which can be fused [9].

The advantage of gray prediction is that short-term prediction requires less sample information. The disadvantage is that the larger the data discrete program, the worse the prediction accuracy [10]. Secondly, the fitting sequence of the model is a non-homogeneous exponential sequence. The prediction result of the exponential law is accurate enough, so it is not suitable for prediction with a long back-up period [11]. The neural network model has the characteristics of local approximation and strong non-linear mapping ability, which can well simulate non-linear problems. Its disadvantages are slow training speed, easy to fall into local minimums, and sensitive selection of learning rate. Grey neural network prediction model combines the advantages of gray model and neural network in univariate prediction and non-linear processing. The determination of the coefficient needs further study.

The differential equation expression of the gray neural network model with $n$ parameters is:
\[
\frac{dy_1}{dt} + ay_1 = b_1y_2 + b_2y_3 + \ldots + b_{n-1}y_n
\] (6)

Where: \(y_2, y_3, \ldots, y_n\) are system input parameters, \(y_1\) are system output parameters; \(a, b_1, b_2, \ldots, b_{n-1}\) is the coefficient of the differential equation.

The time response of equation (6) is:

\[
z(t) = (y_1(0) - \frac{b_1}{a}y_2(t) - \frac{b_2}{a}y_3(t) - \ldots - \frac{b_{n-1}}{a}y_{n}(t)) + \frac{b_1}{a}y_2(t) + \frac{b_2}{a}y_3(t) + \ldots + \frac{b_{n-1}}{a}y_{n}(t)
\] (7)

Make:

\[d = \frac{b_1}{a}y_2(t) + \frac{b_2}{a}y_3(t) + \ldots + \frac{b_{n-1}}{a}y_{n}(t)\]

Transform (7) as follows:

\[
z(t) = ((y_1(0) - d) \times \frac{e^aw}{1 + e^aw}) + d \times \frac{1}{1 + e^{-aw}} \times (1 + e^aw) =
\]

\[
((y_1(0) - d)\times\frac{1}{1 + e^aw}) + d \times \frac{1}{1 + e^{-aw}} \times (1 + e^aw) =
\]

\[
((y_1(0) - d)\times\frac{1}{1 + e^aw}) + 2d \times \frac{1}{1 + e^aw} \times (1 + e^aw)
\] (8)

Map the transformed (8) formula to an extended BP neural network to get a gray neural network with \(n\) input parameters and 1 output parameter. The network topology is shown in Figure 2.

3.3. Mining Evolutionary Algorithms Optimizing Grey Neural Networks

The thinking evolution algorithm has a strong global optimization ability to make up for the shortcomings of the gray neural network, such as slow convergence and weak global search ability. This article uses the thinking evolution algorithm to optimize the initial weights and thresholds of the gray neural network. The basic idea of optimization is: According to the gray neural network topology, the solution space is mapped to the coding space, and each code corresponds to a solution of the problem, which is an individual. The reciprocal of the mean square error of the training set is selected as the score function of each individual and the population, and thinking evolution is used. The algorithm obtains a good global solution, iterates continuously, outputs the optimal individual, and uses it as the initial weight and threshold to train the gray neural network. The design steps are:
(1) Generate training and test sets
The demand sample is large enough and well representative to ensure that the model established has good generalization ability.

(2) Initial population generation
Use the initial population generation function initpop generate and the subpopulation generation function subpop generate

(3) Subpopulation convergence operation
After the winning sub-population is generated, the sub-populations first perform the convergence operation, and the population maturity discriminant function ismature is used to determine whether each sub-population has completed the convergence operation.

(4) Subpopulation alienation operation
After the winning sub-populations and temporary sub-populations complete the convergence operation, they perform alienation operations and supplement new sub-populations based on the alienation results.

(5) Analysis of the optimal individual
When the iterative stop condition is met, the thinking evolution algorithm ends the optimization process. According to the coding rules, it is necessary to find the optimal individual to obtain the corresponding gray neural network weight and threshold.

(6) Training gray neural network
The optimized weights and thresholds are given to the gray God's network's initial weights and thresholds, which are trained and learned using training set samples.

(7) Simulation prediction and result analysis
After training, input the test set samples, perform simulation prediction, and analyze the results.

4. Example applications

4.1. Data source and analysis
The study area is a part of the Zhenljong River and is a tributary of the left bank of the Yujiang River. It belongs to the subregion of the Guangdong-Guizhou karst plain-Fenglin Plain. The karst structure is developed and the underground underground rivers are widely distributed. The daily flow, daily water level, and daily precipitation data in 2012 were used for model training and prediction. Figure 3 shows the daily flow sequence process line of Zhenlong Station from 2010 to 2012.

![Figure 3. Daily flow process line of Zhenlong Station from 2010 to 2012.](image)

It can be seen from Figure 3:
(1) The extreme ratio of discharge in the study area is very large, reflecting that the seasonal distribution of rainfall in the study area is extremely uneven, and there are obvious periods of high water and low water. The annual proportion is large, and the discharge during the dry season is small and stable, reflecting that the precipitation replenishment is mainly in the study area and the groundwater replenishment is secondary.

(2) It can be seen from the flow process line that the flow in the study area has the characteristics of steep increase and slow decline. Due to the existence of large groundwater storage reservoirs in the karst area, the capacity of this space is related to the degree of karst development, and also to the part of this space. When the spatial position of this part is lower than the local average groundwater level, the precipitation during the flood season will first supplement the free capacity of the reservoir (referring to the storage below the average groundwater level) without forming a significant runoff. After the free storage capacity is satisfied, the subsequent precipitation can produce floods. After that, due to the slow discharge of groundwater reservoirs, the late runoff is obviously rich, forming a long-term post-flood outflow process. This kind of low groundwater level at the beginning of the flood has a small runoff coefficient. After a period of time, the groundwater level has increased dramatically, the runoff system is large, floods are frequent, and the characteristics of slow water withdrawal after the flood are obvious. All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper.

4.2. Model establishment
When the model uses a mind evolution algorithm to optimize the initial weights and thresholds of the gray neural network, the population size is set to 200, the number of winning and temporary subpopulations is 5, and the number of iterations is 10. The iteration is satisfied by the mind evolution algorithm. After the conditions are found, the optimal individual output is found and decoded according to the coding rules to generate the initial weights and thresholds of the gray neural network. Figure 4 shows the subpopulation convergence process of the thinking evolution algorithm.

Since the input data of this model is 2 dimensional and the output is 1 dimensional, the gray neural network structure is 1-3-1. A total of 1096 samples were used from 2010 to 2012, of which the first 730 samples were used to train the training data. The last 366 samples are used to evaluate the prediction performance of the network. The statistical percentage of the prediction sample (ape) and
the average absolute percentage of the model error (mape) are used as the criteria for the evaluation model. Figure 5 is the model prediction effect chart, Table 1 is the model prediction error distribution table.

\[
APE = \frac{|v_t - v_p|}{v_t} \times 100%
\]

\[
MAPE = \frac{1}{n} \sum_{s=1}^{n} \frac{|v_t - v_p|}{v_t} \times 100%
\]  

(9)

In the formula: \(v_t\) and \(v_p\) are the true and predicted values of the sample, respectively; \(n\) is the number of samples.

According to Table 1, the average relative error of the model is 9.8%, and the ratio of the error is less than 20% is 88.80%, which indicates that the overall prediction effect of the model is better. As shown in Figure5: (1) The model is suitable for non-abundant water. The forecast accuracy of this period is higher than that of the high-water period, mainly because the data series of the high-water period has a large degree of dispersion, which causes an increase in errors. By normalizing the data, the accuracy can be effectively improved. (2) The model prediction has a large error in the case of a steep daily flow increase, and the error decreases in the withdrawal phase. The main reason is that the model training sample is limited, and the sample is not representative in the steep increase phase. It can be seen that there are a few maximum values in the prediction result, which affects the calculation accuracy of the model.

5. Conclusion
There are many factors that affect the dynamic change of karst flow, which causes the complexity of the daily flow dynamic system. This paper uses the method of ema algorithm optimization and neural network to make full use of the advantages of these two methods. The model is in Zhenlong The application results of the station show that the prediction accuracy of the ema optimization modified by the neural network is higher, and the neural network training convergence is faster after the
optimization. The disadvantage of this method is that it requires more raw data, which is used for gray system prediction. It also needs to be used for the training of neural networks. It should not be used in the case of less data; at the same time, the daily flow data has a large range of ratios. As well as corrections in equal-dimensional gray numbers.

The simulation prediction results of surface runoff in this paper are ideal, especially the prediction accuracy in dry season is high, which is consistent with the characteristics of gray neural network that requires small data level ratio, but there is a large error in the prediction results in high season. In particular, it is difficult to accurately predict the maximum point. The changes of groundwater regimes are relatively stable, so you can consider applying this model to the prediction of groundwater in karst areas.

Acknowledgments
This work was supported by the National Natural Science Foundation of China (Grants No. 51369009), the Natural Science Foundation of Guangxi Province, China (Grant No. 2016GXNSFAA380116), Guangxi mining and metallurgy and Environmental Science Experimental Center and the Project of High Level Innovation Team and Outstanding Scholar in Guangxi Colleges and Universities (Grant No.002401013001).

References
[1] Zheng Zheng, Caihong Hu, Hao Yonghong. Research status and progress of karst spring water [J]. Journal of water resources and water engineering, 2005, 16 (1): 56-59.
[2] Baoxiang Zhang. Research on karst groundwater prediction model based on neural network and genetic algorithm in Jinan urban area [J]. Journal of Shandong Agricultural University (Natural Science Edition), 2004, 35 (3).
[3] Guiming Li. Prediction model of karst spring discharge based on phase space reconstruction and neural network coupling [J]. Engineering Survey, 2006, 4.
[4] Hongfeng Chen. Analysis of BP artificial neural network on the surface karst spring in Lota Ganhe pig farm, Hunan province [J]. China Karst, 2005, 12.
[5] Longcang Shu, Guiming Dong, Yufei Tao. Analysis and prediction of water level changes in skylights of underground rivers [J]. Journal of Hydraulic Engineering, 2009, 40 (5): 529-534.
[6] Wenming Wang, Wenke Wang, Dong Du. Application of grey prediction model gm (1,1) in hydrological prediction [J]. Groundwater, 2007, 29 (3): 10-12.
[7] Chengyi Sun, Mind-Evolution-Based Machine Learning: Framework and TheImplementation of Optimization, Proc. of IEEE Int. Conf. On Intelligent EngineeringSystems (INES’98), Edts. P. Kopacek, IEEE Inc, Sept. 17-19, 1998, Vienna,Austria,pp.355-359.
[8] Julong Deng. Basic method of gray system [M]. Wuhan, Hubei: Huazhong Institute of Technology Press, 1987: 96-108.
[9] Yu Feng, Fengshan Ma, et al. Application of grey system and neural network combined model in groundwater level prediction [J]. Chinese Journal of Geological Hazard and Prevention, 2011, 22 (3): 119-124.
[10] Tanghong Liu, Jun Zhou, Qingchuan Zhu. Application of improved grey prediction model in prediction of land subsidence [J]. Geological Hazards and Environmental Protection, 2007, 18 (3): 104-107.
[11] Pengyu Chen. Grey discrete GM (1, l) model for simulation calculation of land subsidence [J]. Geological Hazards and Environmental Protection, 2009, 20 (4): 101-105.