

GA-BP Neural Network Regression Model for Predicting the Soil Moisture of Ecological Slope Protection

Dunwen Liu

Central South University

Chao Liu

Central South University

Yu Tang (✉ tangyu9433@163.com)

Central South University

Chun Gong

Central South University

Research Article

Keywords: GA-BP Neural Network Regression Model, Soil Moisture , Ecological Slope Protection, genetic algorithm(GA), root mean square error (RMSE), mean absolute percentage error (MAPE)

Posted Date: July 8th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-681849/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

GA-BP Neural Network Regression Model for Predicting the Soil Moisture of Ecological Slope Protection

Dunwen Liu¹, Chao Liu¹, Yu Tang^{1,*}, Chun Gong¹

¹school of resources and safety engineering, central south university, Changsha, 410000, China.

*Correspondence should be addressed to Y.T.(email: tangyu9433@163.com)

ABSTRACT

The neural network optimized by genetic algorithm(GA) is an efficient and accurate prediction method, which can quickly find the optimal solution through high-speed computing capability and self-learning function. The neural network model optimized by GA is applied to the prediction of soil moisture of ecological slope protection, which provides reference for practical application of slope vegetation screening. In this paper, nine meteorological factors and soil moisture data were obtained by field monitoring instruments and related meteorological data. Considering the lag of meteorological factors, the neural network optimized by GA is used to predict the soil moisture of 8 meteorological data. The results show that the root mean square error (RMSE) and the mean absolute percentage error (MAPE) of the prediction model are 0.22726 and 0.41234%, respectively, indicating that the prediction model runs well. Through V-fold cross-validation, it is found that the prediction results of the model is accurate and stable. The algorithm combining artificial neural network and GA can well predict the soil moisture of ecological slope protection, with high prediction accuracy, and has a good application prospect in other fields.

Introduction

Soil moisture is one of the main factors in agricultural production and hydrological cycle, which not only plays an important role in maintaining plant growth, but also is a key link in the water cycle of soil-plant-atmosphere continuous system^{1,2}. Studies have shown that suitable soil moisture is one of the most important factors for the healthy growth of plants³. Other studies have shown that soil moisture can improve the natural nutrient absorption of crops and plants. Less water will affect the growth of cells, resulting in a decrease in cell growth. Secondly, excessive water will affect the growth of some stages of plant life^{4,5}. The soil moisture before planting will affect plant growth. Vegetation performs better when the soil is wet or dry⁶. Its accurate prediction is of great significance to the rational utilization and management of water resources. The change of soil moisture directly affects water consumption and crop growth, and the influence of soil moisture can improve the prediction of flowering time and the assessment of abiotic stress risk⁷. It is an important index of drought resistance, flood control⁸ and precise irrigation decision^{9; 10} in agricultural production. Climate and vegetation change are the two main factors affecting the trend of soil moisture. Climate change dominates the change trend of soil moisture, while vegetation change plays a regulatory role in soil dryness and wetness under changing climate conditions¹¹. Soil moisture is closely related to the optimization of slope vegetation, and meteorological factors such as environmental humidity, daily radiation, temperature and rainfall have great influence on soil moisture. Therefore, it is of great significance to use machine learning method to predict the soil

moisture of the slope, and then use meteorological factors and soil moisture to optimize the slope protection plants.

At present, the mainstream soil moisture prediction methods mainly use empirical formula, linear regression and neural network to build prediction models¹². Hummel¹³ uses near infrared reflectance sensors to predict the soil moisture content of surface and underground soil, and uses stepwise multiple linear regression analysis data, and the prediction standard deviation is 5.31%. Joo¹⁴ and others studied the comparison between soil moisture and hydrometeorological data, and found that the change of soil moisture showed seasonal characteristics. Kashif Gill¹⁵ put forward a method to predict soil moisture by using support vector machine, which uses structural risk rather than empirical risk to ensure global optimization, The results show that SVM model is better than artificial neural network model in predicting soil moisture. Prakash et al,¹⁶ used machine learning techniques such as multiple linear regression, support vector regression and recurrent neural network to predict soil moisture for 1 day, 2 days and 7 days, and the results showed that multiple linear regression had better prediction results. Qian Zhu¹⁷ combined support vector machine (SVM) with double set Kalman filter (EnKF), Compared with support vector machine, it can be found that it can effectively improve the efficiency of soil moisture prediction. However, the parameters and kernel functions of support vector machine need to be selected artificially, and the calculation is complicated, which will affect the prediction accuracy to a certain extent. In complex practical applications, the traditional methods of soil moisture prediction often do not perform well, but the neural network model has better performance in soil moisture prediction^{18,19}. Ji Ronghua et al,²⁰ used multi-level neural network to predict soil moisture, According to correlation analysis, rainfall, temperature and wind speed have significant effects on soil moisture. Principal component analysis is used to eliminate the correlation between the initial input parameters, and the first three principal components are specified to reconstruct the samples into lower dimensions, so as to reduce the network scale and improve the network performance. An Xiaoyu et al,²¹ put forward an improved neural network prediction method for ascidian algorithm optimization. The results show that the optimal prediction model has small prediction error. It overcomes the shortcomings of neural network prediction model, such as low convergence accuracy and easy to fall into local optimum.

However, soil moisture involves complex structural effects and meteorological factors, and it is difficult to establish an ideal mathematical model for predicting soil moisture. Although artificial neural network can solve complex nonlinear mapping problems, it also has some shortcomings, such as over-fitting, falling into local optimum easily and slow convergence speed²². The neural network optimized based on GAs can be optimized from two aspects of learning efficiency and global optimization, which greatly improves the structure of the original network. Compared with the traditional neural network, it shows strong computational ability. The neural network has been successfully applied in engineering cost prediction²³, stock price prediction^{24,25}, traffic flow²⁶, image and text recognition^{27,28}, medical treatment^{29,30} and other fields.

In this study, considering the correlation between meteorological factors and soil moisture and the lag of meteorological factors, neural network is used to predict the nonlinear relationship between soil moisture and meteorological factors on ecological slope protection s. On this basis, GA is used for global optimization, and the optimal soil moisture prediction model is determined, which overcomes the problem of low prediction accuracy and achieves high-precision soil moisture prediction.

Methods

GA and neural network

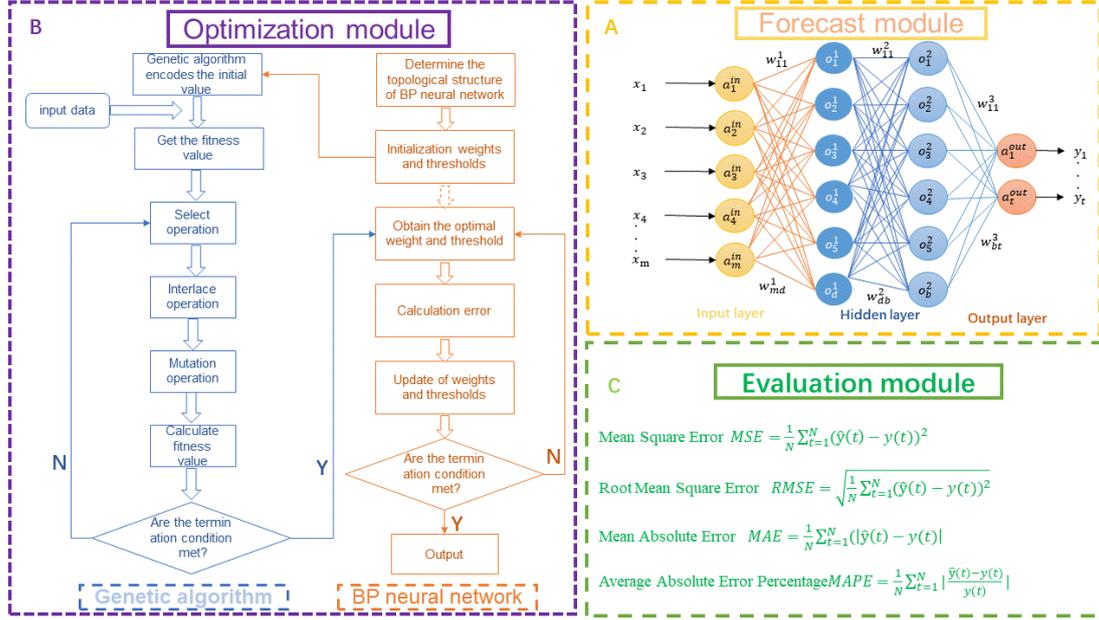


Figure 1. Overall forecasting model; A is the structure diagram of neural network, B is the step diagram of neural network prediction model optimized by genetic algorithm, and C is the error estimation diagram

• BP neural network

BP neural network can be analogized as multi-layer perceptron, in which the hidden layer function mostly adopts S-type transfer function, while the output layer neuron mostly adopts linear transfer function. The fig.1.A shows the four-layer BP network topology with $m \times d \times b \times t$, and the process of calculating the output of pre-BP neural network can be summarized as three steps. Firstly, the data of eight meteorological factors obtained by climate monitoring instruments are used as input layers. Secondly, by comparing the output data with the measurement data, we use the loss function (also called cost function) to calculate the error. Third, we propagate the error back, find its derivative with respect to each weight in the network, and update the model. For BP artificial neural network, each layer transmits data to the next layer without cycle. After we repeat these three steps for multiple cycles and learn the weights, we apply the model to new data to predict multiple target values simultaneously. It can be assumed that the existing sample pairs (\hat{x}, \hat{y}) are $x = [x_1, x_2, \dots, x_m]$, $y = [y_1, y_2, \dots, y_t]$, and the hidden layer neurons are $o = [o_1, o_2, \dots, o_m]$. The network weight matrix between neurons in input layer and hidden layer is w^1 , and the network weight between neurons in input layer and output layer are w^3 respectively.

$$w^1 = \begin{bmatrix} w_{11}^1 & w_{12}^1 & \dots & w_{1d}^1 \\ w_{21}^1 & w_{22}^1 & \dots & w_{2d}^1 \\ \vdots & \vdots & \dots & \vdots \\ w_{m1}^1 & w_{m2}^1 & \dots & w_{md}^1 \end{bmatrix} \quad w^3 = \begin{bmatrix} w_{11}^3 & w_{12}^3 & \dots & w_{1t}^3 \\ w_{21}^3 & w_{22}^3 & \dots & w_{2t}^3 \\ \vdots & \vdots & \dots & \vdots \\ w_{b1}^3 & w_{b2}^3 & \dots & w_{bt}^3 \end{bmatrix} \quad (1)$$

x_i represents the i th input value in a neural network, y is the vector of predicted values. a_i^j represents the i th activation of neurons in the i th layer, w_{lm}^i represents the connection between the m th unit in the i th layer and the first unit in the $(i + 1)$ th layer.

The thresholds of hidden layer neurons and output layer neurons are θ^1 and θ^2 , respectively

$$\theta^1 = [\theta_{11}^1, \theta_{21}^1, \dots, \theta_{l1}^1], \theta^2 = [\theta_{11}^2, \theta_{21}^2, \dots, \theta_{l1}^2] \quad (2)$$

Then the output of hidden layer neurons is

$$o_j = \int \left(\sum_{i=1}^m w_{ji}^1 x_i - \theta_j^1 \right) = f(\text{net}_j), j = 1, 2, \dots, l$$

$$\text{net}_j = \sum_{i=1}^m w_{ji}^1 - \theta_j^1, j = 1, 2, \dots, l; f(\cdot) \quad (3)$$

The output of neurons in the output layer is

$$z_k = \int \left(\sum_{j=1}^l w_{kj}^1 o_j - \theta_k^1 \right) = g(\text{net}_k), k = 1, 2, \dots, n$$

$$\text{net}_k = \sum_{j=1}^l w_{kj}^1 - \theta_k^1, k = 1, 2, \dots, n; g(\cdot) \quad (4)$$

$$\text{The error between network output and expected output is } E = \frac{1}{2} \sum_{k=1}^n (y_k - z_k)^2 \quad (5)$$

•BP neural network optimized by GA

GA is a search algorithm for solving optimization problems, which is the most widely used evolutionary algorithm at present³¹. GA is an adaptive heuristic search algorithm that imitates the biological evolution mechanism of "natural selection, survival of the fittest" in nature, and is combined with artificial neural network, which is widely used to solve optimization problems³². GAs are efficient, practical and robust. GA encodes the data into chromosomes, and exchanges the information in chromosomes through selection, crossover and mutation in iteration, and finally finds out the chromosomes that meet the optimization objectives. By selecting, crossing and changing the population composed of a certain number of problem parameters, GA screens out the individuals with higher fitness and excludes the individuals with lower fitness, and continuously evolves and iterates to find the most satisfied individuals. Using GA to optimize the weights and thresholds of neural network networks can avoid falling into local optimum and effectively improve the training speed. The algorithm steps (see Fig.1.B) are as follows:

Step 1: Determine the neural network topology; In this BP neural network model, the input layer is 8 kinds of weather and meteorological factors of each group of data, and the output layer is daily average soil moisture, which constitutes 8 nodes input layer and 1 node output layer.

Step 2: Obtaining an initial population; By preprocessing the input data, the population number is set, in which each individual contains the ownership value of BP neural network and the real number string of threshold value.

Step 3: Setup of fitness function F ; The formula is $F = \sum_{i=1}^n |f_i - f(x_i)|$ (6)
where f_i is the expected output value and $f(x_i)$ is the actual output value.

Step 4: Select the operation; Selection operation refers to selecting excellent individuals from the original population through a certain probability and producing the next generation of individuals through reproduction. The probability of individual selection is correlated with fitness. The higher the fitness, the higher the probability of individual choice. Roulette is used in this paper. The probability of choice is

$$p^i = \frac{F_i}{\sum_{j=1}^N F_j} \quad (7)$$

where F_i is the population fitness value of individual I , and N is the population number.

Step 5: Cross-operation: Cross-operation refers to randomly selecting two individuals from the population, and transferring the excellent characteristics of their parents to their offspring to produce new individuals with high adaptability through the exchange and combination of their chromosomes. That is to say, there are two individuals a_k 、 a_l who cross-operate at j as follows

$$a_{kj} = a_{kj}(1 - b) + a_{lj} \quad (8)$$

$$a_{lj} = a_{lj}(1 - b) + a_{kj} \quad (9)$$

where b is a random number between $[0,1]$.

Step 6:Mutation operation; The diversity of population can be maintained through population variation. Randomly select an individual from the population, and select a part of the individual to perform mutation operation, so as to produce a better individual. Individual a_i mutates at j , forming a new gene a_{ij} as follows.

$$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{\max})r_2(1 - g/G_{\max}), & r \geq 0.5 \\ a_{ij} + (a_{\min} - a_{ij})r_2(1 - g/G_{\max}), & r < 0.5 \end{cases} \quad (10)$$

where a_{\max} 、 a_{\min} are the upper and lower bounds of a_{ij} , r_2 is a random number, g is the current iteration number, G_{\max} is the maximum evolution number.

Step 7: The fitness function of the individual in the newly generated population is calculated. If the fitness function meets the requirements or the evolution times reach the maximum, the evolution is completed, otherwise, the step 4 is returned.

Step 8: Through GA optimization, the BP network is given initial weights and thresholds, and the prediction results are obtained after meeting the termination conditions.

•Error estimation

In order to quantitatively evaluate the effectiveness and accuracy of neural network prediction model optimized by GA, MSE (mean square error), RMSE (mean square error root), MAE (mean absolute error), and MAPE (mean absolute percentage error) are considered respectively. (see Fig.1.C)

$$MSE = \frac{1}{N} \sum_{t=1}^N (\hat{y}(t) - y(t))^2 \quad (11)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{y}(t) - y(t))^2} \quad (12)$$

$$MAE = \frac{1}{N} \sum_{t=1}^N (|\hat{y}(t) - y(t)|) \quad (13)$$

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{\hat{y}(t) - y(t)}{y(t)} \right| \quad (14)$$

where N is the total number of training or test set; $\hat{y}(t)$ and $y(t)$ is the forecast value and the actual value, respectively.

GA-BP soil moisture prediction model

•Meteorological data monitoring scheme

In order to solve the problems of complicated operation and high monitoring cost of long-

term monitoring instruments for slope environment, a monitoring device and prediction method for slope vegetation growth environment in typhoon area are adopted. By setting temperature and humidity sensors, wind speed and direction sensors, solar radiation sensors, rainfall sensors, soil temperature, humidity and salt comprehensive sensors, and adopting wireless network and cloud platform data transmission means, the long-term remote monitoring of slope growth environment is realized. Finally, based on the collected long-term data and the data provided by China Meteorological Network, the time series method and linear regression model can be used to make short-term prediction of the meteorological parameters on the slope, thus obtaining the time variability of the meteorological parameters on the slope.

The monitoring system for meteorological parameters of ecological slope protection is composed of environmental temperature and humidity sensors, wind speed and direction sensors, solar radiation sensors, rainfall sensors, comprehensive sensors of soil temperature, humidity and salinity, data acquisition instruments, wireless communication modules and solar power supply system powered by the whole test system. The partial structure is shown in Fig. 2.

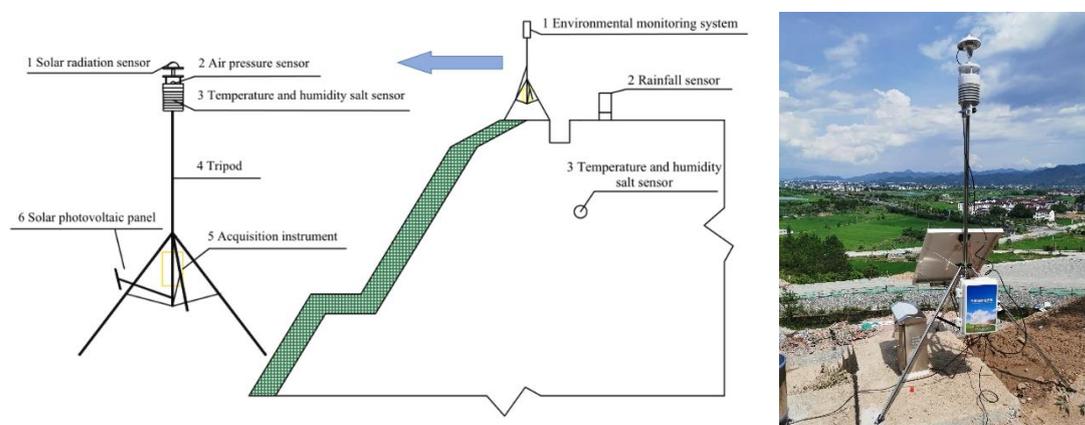


Figure 2. Schematic diagram of environmental monitoring system

The environmental temperature and humidity sensors, wind speed and wind direction sensors and solar radiation sensors in the device are installed on the top of the slope without vegetation shielding, and are erected 2 m away from the top surface of the slope with a triangular bracket. The top of the triangular bracket is provided with a horizontal vacuole indicator to ensure the level of the triangular bracket, and then the three corners of the bracket are fixed with expansion screws. The rain sensors is installed at the top of slope without shielding and fixed by expansion screws. Install the soil temperature, humidity and salt integrated sensor on the top of the slope, dig a hole with a diameter of 50 mm and a depth of 2m in the top of the slope, and put the sensor into the hole; Make the sensor in close contact with the soil, make sure that the soil is full of the sensor, and fill the upper part with soil to prevent water from entering. The wireless communication module and the data acquisition instrument are uniformly placed in the acquisition box and fixed on the triangle bracket. The solar power supply system consists of a solar photovoltaic panel, a storage battery and a debugger. The solar photovoltaic panel inclines to the south and is placed on a reserved bracket; The storage battery is placed in a stainless steel box and fixed on the ground with expansion screws. The environmental temperature and humidity sensor, air pressure sensor, solar radiation sensor, rainfall sensor and soil temperature, humidity and salt comprehensive sensor are connected with the collector through data lines, and the collected data are collected into the data collector through data lines. The data collector uploads the

collected data to the cloud platform server in real time through the wireless communication module, and the client computer can view and download the test data by accessing the cloud, thus realizing long-term monitoring of the vegetation growth environment on the slope. The field test is shown in fig. 3.

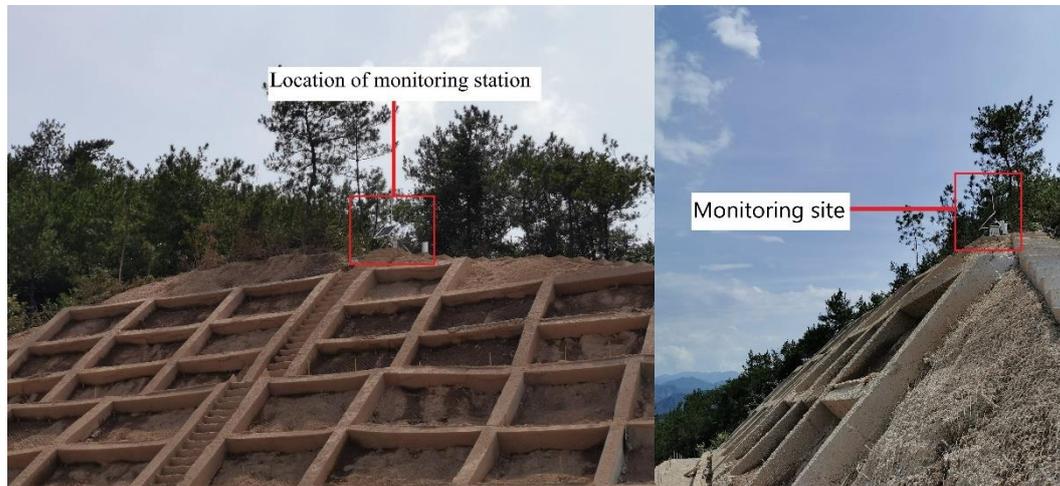


Figure 3. Field experiment diagram

•*Influencing Factors of Soil Moisture*

The soil moisture of ecological slope protection s is closely related to local meteorological factors, and the change of soil moisture is affected not only by its own hydrological factors but also by surrounding meteorological factors. The motion of atmospheric systems is a highly unified process of certainty and randomness. From a macroscopic point of view, the uncertainty of soil moisture change is largely determined by the influence of weather conditions in the early stage. Climate is the main factor that determines the distribution of soil moisture on the earth, and the uncertainty of soil moisture mainly comes from the randomness of atmospheric factors. Therefore, meteorological factors can be used to predict soil humidity. In this paper, nine meteorological elements are obtained, namely, daily maximum environmental temperature (X_1), daily minimum environmental temperature (X_2), daily average temperature (X_3), daily average environmental humidity (X_4), daily average dew point temperature (X_5), daily average wind speed (X_6), daily average rainfall (X_7), daily total radiation (X_8) and daily average soil temperature (X_9). Pearson correlation coefficient and Spearman correlation coefficient are as shown in Fig. 4.

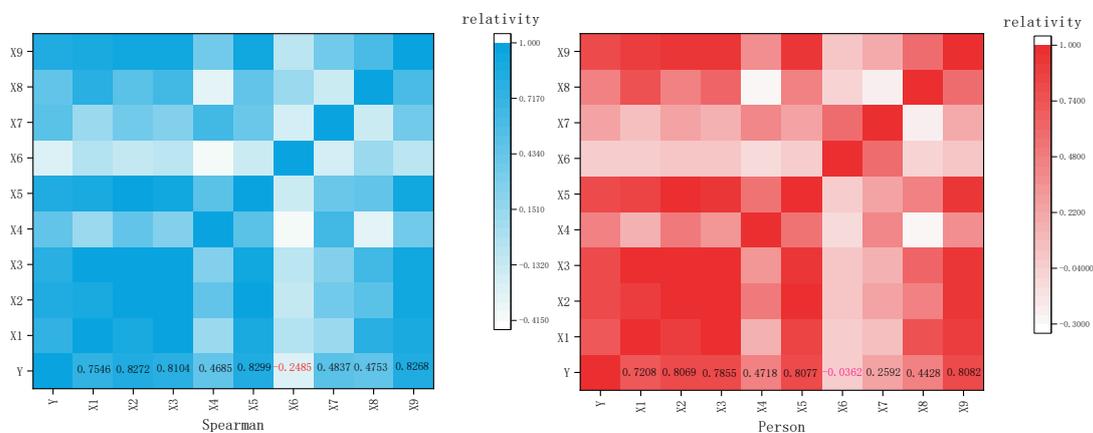


Figure 4. Correlation analysis between soil moisture and meteorological factors

According to Pearson correlation coefficient and Spearman correlation coefficient of

meteorological factor and soil moisture, it can be found that (Y) has strong correlation with (X₁), (X₂), (X₃), (X₅) and (X₉), which is higher than 0.8. (Y) has a weak correlation with (X₄) and (X₈), showing a moderate correlation, while it has a low correlation with (X₆), so this factor is excluded (see Fig.4). Considering that the influence of environmental humidity and rainfall on soil moisture has a certain lag, the increase of daily precipitation and environmental humidity leads to the increase of moisture in the air, and it takes a certain time for moisture to enter the soil, so the lag is taken into account to predict. After considering its lag, the correlation coefficient has been significantly improved, and the correlation degree has increased to more than 0.5, as shown in Table 1.

Table 1. Comparison Table of Hysteresis of X₄, X₈ and Y

	Pearson		Spearman	
	original	Consider lag	original	Consider lag
X ₄	0.4718	0.52151	0.46849	0.54337
X ₈	0.44284	0.471	0.47529	0.54591

Considering the discreteness of field monitoring results due to few samples and many factors, the neural network algorithm optimized by GA is adopted. In the monitoring data, the meteorological data of the first 90 days are selected as training samples, and the data of the last 10 days are predicted samples. The sample data are shown in Table 2, and the GA-BP regression prediction model is established by programming with Matlab language.

Table 2. Sample data

NO	X ₁	X ₂	X ₃	X ₄	X ₆	X ₇	X ₈	X ₉	Y
1	35.2	23	30.3	76.4	24.7	0	16.507	32.8	89.8
2	35.3	23.5	29.0	87.5	24.2	25.2	25.182	32.6	97.3
2	35.5	23.8	29.5	77.4	23.8	0	23.64	32.9	88.4
3	35.2	2.8	29.5	77.4	23.8	0	23.608	33.2	82.7
...
87	13.5	0.7	6.4	57.1	-1.7	0	11.72	16.1	43.4
88	10.3	2.6	6.7	60.5	0.3	0	4.347	16.3	43
89	8.9	2.5	7.5	64.3	2.8	0	1.587	11.281	42.8
90	13.5	-0.8	5.3	62.8	-2.7	0	11.063	14.8	41.7

Results

Comparison of model prediction

The GA-BP regression prediction model was established, and the soil moisture of ecological slope protection s was simulated and predicted by using neural network, The monitoring data of the first 90 days were taken as the training set, and the data of 10 days were taken as the test set. Matlab software is used for prediction simulation experiment, and the prediction results are compared, as shown in Fig. 5 .

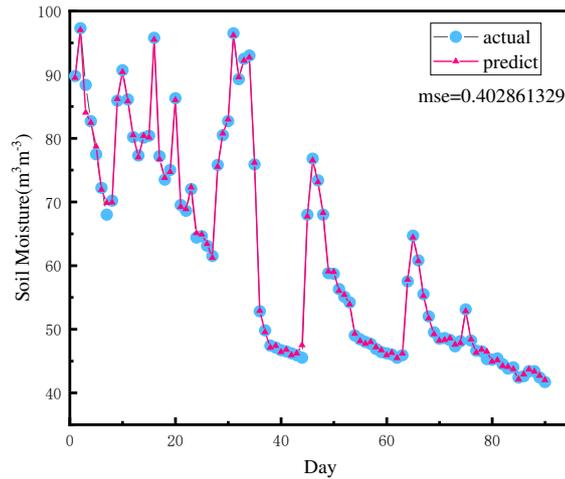


Figure 5. Training set prediction considering lag

After running the program, MAE, MSE, RMSE and MAPE can be obtained by the above calculation formula, The BP neural network prediction model is compared with the neural network prediction model optimized by GA, and the prediction result of the original data is compared with the prediction result considering lag, The results are shown in Fig. 6, 7 and Table 3.

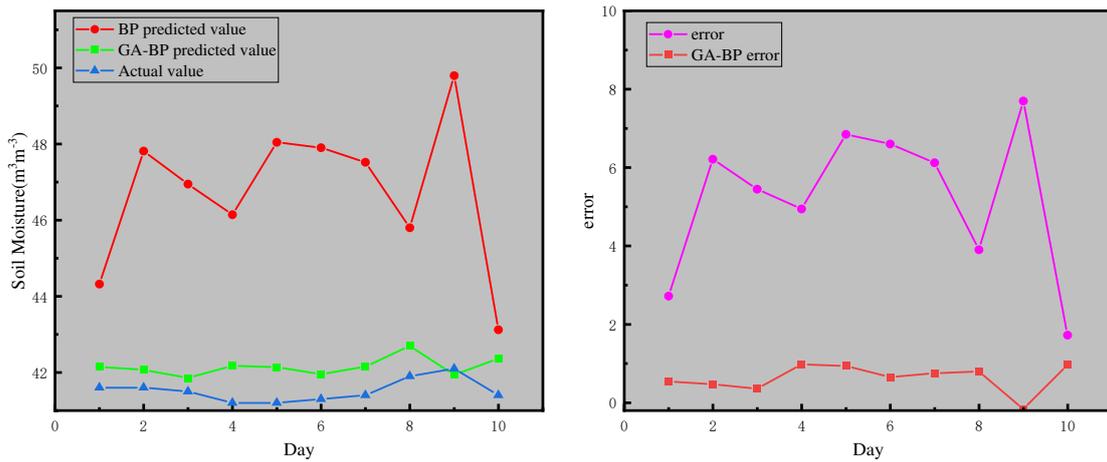


Figure 6. BP and GA-BP prediction and error analysis of original data

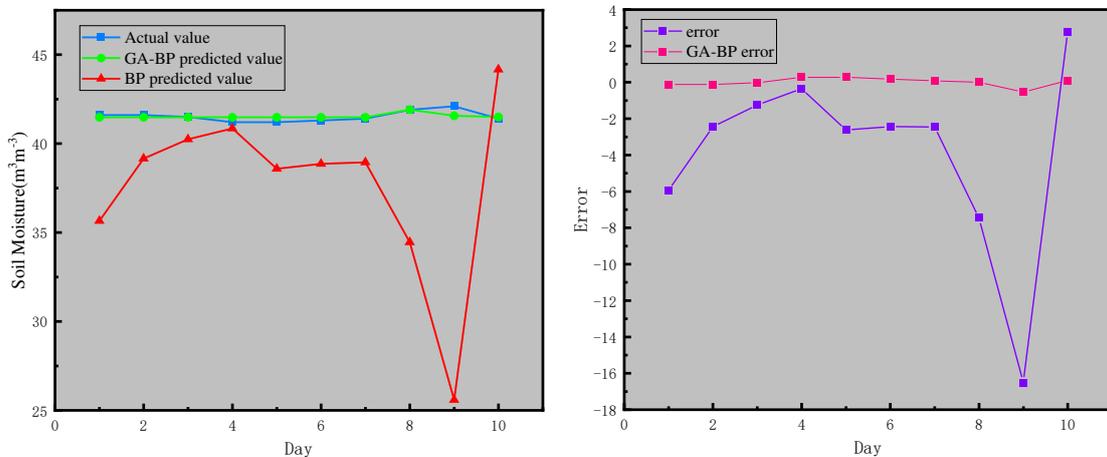


Figure 7. Prediction error analysis of BP and GA-BP considering lag

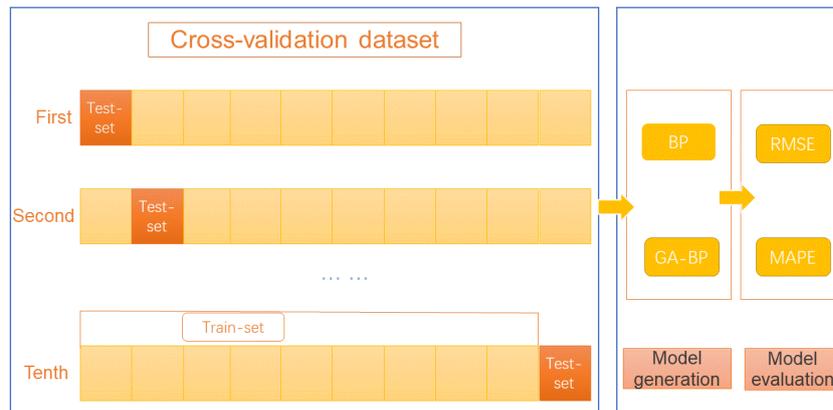
Table 3. Error analysis of BP and GA-BP

		MAEE	MSE	RMSE	MAPE
original	BP	5.2214	30.5491	5.5271	12.5741 %
	GA-BP	0.66081	0.50499	0.71063	1.5945 %
Consider lag	BP	4.4237	39.7918	6.3081	10.592 %
	GA-BP	0.17141	0.051645	0.22726	0.41234 %

The MSE predicted by the training set considering lag is 0.0288 . GA-BP prediction model is obviously superior to BP prediction model. Comparing the original data with the prediction model considering lag, it can be seen that the MSE of GA-BP prediction model of the original data is 0.50499 and the MAPE is 1.5945%. The MSE and MAPE of GA-BP prediction model considering lag are 0.051645 and 0.41234%. The maximum prediction error of BP neural network is 15%, and the average absolute percentage error MAPE is 10.592%, so it is impossible to predict soil moisture accurately. The GA-BP prediction model considering hysteresis has good prediction results, and the error control is close to 0.01, which can accurately predict soil moisture.

Prediction model evaluation

The neural network optimized by GA can obtain the optimal weights and thresholds through GA. Through selection, crossover and mutation in GAs, the population is continuously optimized and the best individual is selected. The optimization of GA mainly involves the initial population size, maximum evolutionary algebra, crossover probability and mutation probability. In the prediction model, the crossover probability is 0.8, the mutation probability is 0.05, the initial population size is 100, and the maximum evolution algebra is 100.

**Figure 8 .** Cross-validation evaluation of forecasting model

In this paper, the cross-validation method is used to evaluate the soil moisture prediction model. Allen³³ and others introduced cross-validation as a method to evaluate the parameters of prediction model, in order to improve the prediction accuracy. Cross-validation is a method to eliminate the deviation caused by random sampling in the process of sample training, and it is a method to estimate the generalization error directly without any preconditions. Cross-validation refers to dividing the original data into training set and testing set according to the prediction setting of neural network. The model is trained with training set, and the trained model is verified

with test samples to evaluate the prediction accuracy of the model. There are many cross-validation methods, among which the V-fold cross-validation method proposed by Geisser³⁴ has the characteristics of high computational efficiency and simple operation. Therefore, the V-fold cross-validation method is used to evaluate the prediction model. The V-fold cross-validation method refers to dividing the data into V parts, selecting one part as the test set, V-1 part as the test set, repeating v times, and finally analyzing the average error of v times as the performance index of model evaluation. In this paper, 100 groups of data are divided into 10 parts, and the GA-BP prediction model considering lag is evaluated by V-fold cross validation, as shown in the fig.8.

Table 4. Cross-validation evaluation index parameter statistics

NO	RMSE		MAPE	
	BP	GA-BP	BP	GA-BP
1	6.3081	0.22726	10.592 %	0.41234 %
2	1.7392	0.555	3.2466 %	1.283 %
3	9.0197	0.42413	19.6638 %	0.8252 %
4	2.445	0.79798	3.5562 %	1.8244 %
5	4.2637	0.81116	9.3364 %	1.6981 %
6	4.0055	0.28749	7.6028 %	0.57294 %
7	2.1669	0.4809	4.2679 %	1.0654 %
8	6.8948	0.68498	13.8876%	1.5158%
9	6.5857	0.66154	10.0203%	1.3527%
10	5.0258	0.46632	10.9858 %	0.97458 %
average value	4.84544	0.539676	9.31594 %	1.152446 %

It can be seen from the table 4 that the average RMSE and MAPE of BP neural network forecasting model are 4.84544 and 9.31594%, while the average RMSE and MAPE of GA-BP neural network forecasting model are 0.53967 and 1.152446%, respectively. After cross-validation, it can be seen that the prediction of BP neural network fluctuates greatly, while the prediction result of GA-BP neural network is relatively stable. Ten experiments show that its accuracy is better than that of BP neural network. Only one random prediction is used to evaluate the accuracy of the model, and the random error is large. After cross-validation, it is more reasonable to use the mean value of RMSE and MAPE as the evaluation standard of the model.

Discussion

Advantages of GA-BP neural network regression model. In this paper, the GA-BP prediction model is used to simulate and predict the daily average soil moisture through eight kinds of meteorological factors obtained from the monitoring of ecological slope protections in Taizhou area as independent variables, and the following conclusions are drawn. The advantages are as follows: (a) Soil moisture is closely related to the surrounding meteorological factors. Because of the randomness and non-linearity of meteorological factors, soil moisture is uncertain. Using meteorological factors to predict soil moisture can obtain high-precision prediction results. Therefore, it is feasible to analyze and predict the change of soil moisture on ecological slope protection s by establishing related mathematical models, which can provide theoretical basis for the selection of vegetation on ecological slope protections. (b) Soil moisture is greatly affected by meteorological factors, and has a certain lag, especially rainfall and environmental humidity. This lag should be considered when using meteorological factors to

predict and simulate soil humidity. Considering the lag, the prediction accuracy has greatly improved. (c) The established GA-BP neural network prediction model has high accuracy, and the deviation between soil moisture prediction results and actual values is small, which shows that the model can accurately predict the change of soil moisture with meteorological factors. (d) Comparing the prediction results of BP neural network and GA-BP neural network, it is found that GA-BP neural network is superior to BP neural network in predicting soil moisture of ecological slope protection s, and the optimization of GA plays a very important role.

Shortcomings of GA-BP neural network regression model In this paper, the artificial neural network algorithm optimized by GA is applied to the prediction of soil moisture in ecological slope protection for the first time, and a more accurate prediction result have been obtained. However, there are some defects in this paper. For example, all meteorological factors are not fully considered, and meteorological factors such as sunshine time and daily average pressure can be introduced in future research. In addition, GA is used to optimize neural network, and the optimized initial weights and thresholds are obtained. Besides GA, particle swarm optimization and ant colony algorithm can also be used to optimize the initial weights and thresholds of neural network.

References:

1. Li, X., Huo, Z. & Xu, B. Optimal Allocation Method of Irrigation Water from River and Lake by Considering the Field Water Cycle Process. *Water-Sui*. **9**, <https://doi.org/10.3390/w9120911> (2017).
2. Liao, R., Yang, P., Wang, Z., Wu, W. & Ren, S. Development of a Soil Water Movement Model for the Superabsorbent Polymer Application. *Soil Sci. Soc. Am. J.* **82**, 436 (2018).
3. Liang, Y., Kang, S. & Zhang, C. 14 the Effects of Soil Moisture and Nutrients On Cropland Productivity in the Highland Area of the Loess Plateau. *aci-ar.gov.au*. (2002).
4. Chang, J. H. *Climate and Agriculture: An Ecological Survey*, Climate and Agriculture: An Ecological Survey, (2017).
5. Rockstr M, J. et al. Sustainable Intensification of Agriculture for Human Prosperity and Global Sustainability. *Ambio*. **46**, (2017).
6. Ngo, H. T. T. & Cavagnaro, T. R. Interactive Effects of Compost and Pre-Planting Soil Moisture On Plant Biomass, Nutrition and Formation of Mycorrhizas: A Context Dependent Response. *Sci. Rep.-UK*. **8**, <https://doi.org/10.1038/s41598-017-18780-2> (2018).
7. Chauhan, Y. S., Ryan, M., Chandra, S. & Sadras, V. O. Accounting for Soil Moisture Improves Prediction of Flowering Time in Chickpea and Wheat. *Sci. Rep.-UK*. **9**, <https://doi.org/10.1038/s41598-019-43848-6> (2019).
8. Martinez-Fernandez, J., Gonzalez-Zamora, A., Sanchez, N., Gumuzzio, A. & Herrero-Jimenez, C. M. Satellite Soil Moisture for Agricultural Drought Monitoring: Assessment of the SMOS Derived Soil Water Deficit Index. *Remote Sens. Environ.* **177**, 277-286, <https://doi.org/10.1016/j.rse.2016.02.064> (2016).
9. Chukalla, A. D., Krol, M. S. & Hoekstra, A. Y. Green and Blue Water Footprint Reduction in Irrigated Agriculture: Effect of Irrigation Techniques, Irrigation Strategies and Mulching. *Hydrol. Earth Syst. Sc.* **19**, 4877-4891, <https://doi.org/10.5194/hess-19-4877-2015> (2015).
10. Feki, M., Ravazzani, G., Ceppi, A., Milleo, G. & Mancini, M. Impact of Infiltration Process Modeling on Soil Water Content Simulations for Irrigation Management. *Water-Sui*. **10**, (2018).
11. Feng, H. Individual Contributions of Climate and Vegetation Change to Soil Moisture Trends

- Across Multiple Spatial Scales. *Sci. Rep.-UK*. **6**, <https://doi.org/10.1038/srep32782> (2016).
12. Cai et al. Research On Soil Moisture Prediction Model Based On Deep Learning. *PLoS One*. (2019).
 13. Hummel, J. W., Sudduth, K. A. & Hollinger, S. E. Soil Moisture and Organic Matter Prediction of Surface and Subsurface Soils Using an NIR Soil Sensor. *Comput. Electron. Agr.* **32**, 149-165 (2001).
 14. Joo, J. Y. & Lee, S. O. Prediction of Soil Moisture Using Hydrometeorological Data in Selmacheon. *Journal of the Korean Society of Civil Engineers B*. **30**, 437-444 (2010).
 15. Gill, M. K., Asefa, T., Kemblowski, M. W. & McKee, M. Soil Moisture Prediction Using Support Vector Machines. *J. Am. Water Resour. As.* **42**, 1033-1046 , <https://doi.org/10.1111/j.1752-1688.2006.tb04512.x> (2006).
 16. Prakash, S., Sharma, A. & Sahu, S. S. Soil Moisture Prediction Using Machine Learning., 2018.
 17. Zhu, Q., Wang, Y. & Luo, Y. Improvement of Multi-Ayer Soil Moisture Prediction Using SVM +nKF Coupled with Remote Sensing Soil Moisture Datasets Over an Agriculture Dominant Basin in China. *Hydrol. Process.* (2021).
 18. Xu, J., Zhao, J., Zhang, W., Hu, Z. & Zheng, Z. Mid-Short-Term Daily Runoff Forecasting by ANNs and Multiple Process-Based Hydrological Models. *2009 IEEE Youth Conference on Information, Computing and Telecommunication*, 2010.
 19. Xu, J., Zhu, X., Zhang, W., Xu, X. & Xian, J. Daily Streamflow Forecasting by Artificial Neural Network in a Large-Scale Basin. *IEEE Youth Conference on Information, Computing & Telecommunication*, 2009.
 20. Ji, R., Zhang, S., Zheng, L., Liu, Q. & Saeed, I. A. Prediction of Soil Moisture with Complex-Valued Neural Network. *Chinese Control and Decision Conference*, 2017:1231-1236.
 21. An, X., Lu, K. & Cui, G. Prediction of Soil Moisture Based On BP Neural Network Optimized by Adaptive Salp Swarm Algorithm. *Journal of Chinese Agricultural Mechanization*. (2019).
 22. Cao, J., Chen, L., Wang, M., Shi, H. & Tian, Y. A Parallel Adaboost-Backpropagation Neural Network for Massive Image Dataset Classification. *Sci. Rep.-UK*. **6** , <https://doi.org/10.1038/srep38201>(2016).
 23. Feng, W., Zhu, W. & Zhou, Y. The Application of Genetic Algorithm and Neural Network in Construction Cost Estimate. *proceedings of the international symposium on electronic commerc.*
 24. Qiu, M., Song, Y. & Akagi, F. Application of Artificial Neural Network for the Prediction of Stock Market Returns: The Case of the Japanese Stock Market. *Chaos Solitons & Fractals the Interdisciplinary Journal of Nonlinear Science & Nonequilibrium & Complex Phenomena*. **85**, 1-7 (2016).
 25. Abbassi, N. M., Aghaei, M. A. & Fard, M. M. An Integrated System Based On Fuzzy Genetic Algorithm and Neural Networks for Stock Price Forecasting: Case Study of Price Index of Tehran Stock Exchange. *International Journal of Quality & Reliability Management*. **31**, 281-292 (2014).
 26. Yang, H. J. & Hu, X. Wavelet Neural Network with Improved Genetic Algorithm for Traffic Flow Time Series Prediction. *Optik - International Journal for Light and Electron Optics*. **127**, 8103-8110 (2016).
 27. Agarwal, M. & Kaushik, B. Text Recognition From Image Using Artificial Neural Network and Genetic Algorithm. *International Conference on Green Computing & Internet of Things*, 2016.
 28. Nasri, M. et al. Cascade-System Genetic Algorithm: Multilayer Neural Network for a Supervised Classification of Texture Images. *Proceedings of SPIE - The International Society for Optical Engineering*. **5132**, 169 (2003).
 29. Li, P. et al. Orthodontic Treatment Planning Based On Artificial Neural Networks. *Sci. Rep.-UK*.

- 9, <https://doi.org/10.1038/s41598-018-38439-w>(2019).
30. Taye, G. T., Hwang, H. & Lim, K. M. Application of a Convolutional Neural Network for Predicting the Occurrence of Ventricular Tachyarrhythmia Using Heart Rate Variability Features. *Sci. Rep.-UK*. **10**, <https://doi.org/10.1038/s41598-020-63566-8> (2020).
 31. Zhang, Q., Deng, D., Dai, W., Li, J. & Jin, X. Optimization of Culture Conditions for Differentiation of Melon Based On Artificial Neural Network and Genetic Algorithm. *Sci. Rep.-UK*. **10**, <https://doi.org/10.1038/s41598-020-60278-x> (2020).
 32. Witten, I. H. & Frank, E. Data Mining : Practical Machine Learning Tools and Techniques. *Acm Sigmod Record*. **31**, 76-77 (2011).
 33. Allen, D. M. The Relationship Between Variable Selection and Data Augmentation and a Method for Prediction. *Technometrics*. **16**, 125-127 (1974).
 34. Geisser & Seymour. The Predictive Sample Reuse Method with Applications. *J. Am. Stat. Assoc.* **70**, 320-328 (1975).

Acknowledgements

The research has been supported by the Initiation fund for postdoctoral research of Central South University, grant number 228697.

Author Contributions

All authors contributed to the current work. D.L. devised the study plan and led the writing of the article. C.L. conducted the experiments and collected the data and conducted the analysis, and Y.T. and C.G. supervised the whole process and gave constructive advice.

Additional Information

Competing financial interests: The authors declare no competing financial interests.