

Experimental study of overland flow resistance coefficient model of grassland based on BP neural network

Peng Jiao^{1,*}, Er Yang¹, and Yong Xin Ni¹

¹Yellow River Institute of Hydraulic Research, Key Laboratory of Soil and Water Loss Process and Control on the Loess Plateau of Ministry of Water Resources, Zhengzhou 450003, China

Abstract. The overland flow resistance on grassland slope of 20° was studied by using simulated rainfall experiments. Model of overland flow resistance coefficient was established based on BP neural network. The input variations of model were rainfall intensity, flow velocity, water depth, and roughness of slope surface, and the output variations was overland flow resistance coefficient. Model was optimized by Genetic Algorithm. The results show that the model can be used to calculate overland flow resistance coefficient, and has high simulation accuracy. The average prediction error of the optimized model of test set is 8.02%, and the maximum prediction error was 18.34%.

1 Introduction

The overland flow resistance is a key parameter to reflect the hydraulic characteristics of the overland flow. Studying the law of overland flow resistance is of great significance for understanding the hydrodynamic mechanism of soil erosion process. Many scholars [1-7] carried out extensive and in-depth research on the law of overland flow resistance. In the past, the research mostly focused on the experimental analysis of the overland flow resistance. For example, Keli Zhang [3] thought that the resistance coefficient in the process of rill erosion was closely related to the flow conditions and slope morphology. Peiqing Xiao [7] thought that the average resistance coefficient of slope under different covering conditions was closely related to rainfall intensity, and decreased with the increase of rainfall intensity.

Due to the difference of study objects and calibration data, the formula of overland flow resistance coefficient is quite different in form and scope of application. In this study, the neural network theory was applied to the study of overland flow resistance coefficient. In order to provide a reference for improving the calculation method of overland flow resistance coefficient and the prediction accuracy of the erosion prediction model, a BP neural network model was established based on the simulated rainfall experiment data, the genetic algorithm (GA) was used to optimize the model parameters.

2 Experiment method

2.1 Experiment equipment

The experiments were carried out in the key laboratory of soil and water loss process and control on the loess plateau of ministry of water resources. The soil tank was 5.0m long, 1.0m wide and 0.6m deep. The slope was 20°. The representative plant alfalfa in the Loess Plateau was planted on the slope by bunch planting. The distance of bunch planting was 20cm. The experiments were carried out when the grass was covered by 60% - 70%. The simulated rainfall device was a side-jet artificial rainfall device. The simulated rainfall installation was 6.0m above the ground, and had a jet height of 1.5m. The values of rainfall uniformity measured before experiments were 86% -92%.

Soil was taken from the surface layer of loess slope of mountain Mang in Zhengzhou, and the particle composition was shown in Table 1. Loess was layered paving compaction to 50cm thick, and the soil moisture content was about 15%. The soil dry bulk density is about 1.05g/cm³ above 20cm depth, and about 1.35g/cm³ below 20cm depth. Samples taken from the upper, middle and lower part of the slope by cutting ring, are used to calculate the average bulk density and dry bulk density.

Table 1. Particle distribution of the experimental soil.

Particle Size (mm)	>1.000	1.000~0.250	0.250~0.050	0.050~0.010	0.010~0.005	0.005~0.001	<0.001
Proportion (%)	0	1.05	35.45	43.40	3.20	6.40	10.50

* Corresponding author: hnzjz2010@126.com

2.2 Experiment scheme

According to the typical rainfall frequency characteristics in the hilly and gully regions of the Loess Plateau, three rainfall intensities of 45, 90 and 130mm/h were identified as the experiment conditions. Simulated rainfall of 30mm/h was applied to the experimental slope about 10 min on the day before the experiments to ensure that the soil moisture content of each experiment was consistent. Observation sections were set along the slope from top to bottom at an interval of 1m, and the 5th section was the exit section. In the measurement interval of 2min, the five sections of the experimental parameters were measured. The flow rate was determined by the dye method, and the flow width was the sum of the flow widths of each section. A runoff sample was taken at the outlet of the soil tank at each time interval. Take a sample of runoff sediment. After runoff occurred on the slope, the rainfall lasted for 60 min, and each experiment was repeated once. The overland flow resistance coefficient was calculated by the Darcy-Weisbach resistance coefficient formula.

3 Modeling method

3.1 BP neural network

BP neural network consists of input layer, hidden layer and output layer. Each layer has several neurons. Neurons in different layers are connected to each other with a certain weight. The hidden layer neurons have bias values and activation function, and the output layer neurons have bias values. The model input variables were simulated rainfall intensity, flow velocity, flow depth, slope roughness, and the model output variable was overland flow resistance coefficient. The model was a BP neural network with 3 layers of network topology, and the number of neurons in input layer, hidden layer and output layer are 4, 12, and 1, respectively. Hidden layer activation function is:

$$f(x) = \frac{2}{1 + \exp(-2x)} - 1 \quad (1)$$

3.2 Genetic algorithm

According to the topological structure and the characteristics of neuron activation function of BP neural network model, the initial weights and bias values of the model can be optimized by genetic algorithm. The weights and bias values optimization of BP neural network based on genetic algorithm, includes population initialization, construct fitness function, selection operation, cross operation, mutation operation.

Population initialization is to encode all the individuals in the population according to individual information. Each individual is a real number string, which is composed of input layer and hidden layer neuron connection weights, hidden layer neuron bias

values, hidden layer and output layer neuron connection weights and output layer bias value.

To construct fitness function is to build a BP neural network model based on information of connection weights and bias values of the neurons carried by individual, to train BP neural network using training samples, and then to simulate and predict. The sum of the absolute value of the error between the predicted output and the expected output is taken as the fitness value F_i of the individual i .

$$F_i = k \left(\sum_{i=1}^p \text{abs}(y_i - t_i) \right) \quad (2)$$

F_i is the fitness value of individual i , k is a coefficient, p is the number of neurons in the output layer of the neural network, y_i is the expected output value of the i -th neuron in the output layer of the neural network; t_i is the predicted output value of the i -th neuron in the output layer of the neural network.

The selection operation refers to selecting an individual from a population. The method is roulette method based on the selection strategy of proportion of fitness, and the selection probability P_i of an individual i is expressed as:

$$f_i = \frac{k}{F_i} \quad (3)$$

$$P_i = \frac{f_i}{\sum_{j=1}^N f_j} \quad (4)$$

f_i is the inverse of the fitness value of individual i , N is the number of individuals in the population.

The cross operation refers to using real number cross method to process real-coded individuals. The cross operation expression of the k -th chromosome a_k and the l -th chromosome a_l in the j position is:

$$\begin{cases} a_{kj} = a_{kj}(1-b) + a_{lj}b \\ a_{lj} = a_{lj}(1-b) + a_{kj}b \end{cases} \quad (5)$$

a_{kj} is the j -bit coded value of chromosome a_k after the cross operation, a_{lj} is the j -bit coded value of chromosome a_l after the cross operation, b is a parameter of the cross operation.

The mutation operation refers to using a certain algorithm to make the gene fragments of selected individual mutation. The mutation operation expression of the i -th chromosome's j -th gene a_{ij} is

$$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{\max})f(g) & r > 0.5 \\ a_{ij} + (a_{\min} - a_{ij})f(g) & r \leq 0.5 \end{cases} \quad (6)$$

$$f(g) = r^* \left(1 - \frac{g}{G_{\max}} \right)^2 \quad (7)$$

a_{\max} is the upper bound of gene a_{ij} , a_{\min} is the lower bound of gene a_{ij} , g is the current evolution generation; G_{\max} is the largest evolution generation; r^* is a random number, r is a random number within $[0,1]$, $f(g)$ is a function.

Considering the computational accuracy and computational requirements, set the population size to 20, the maximum evolutionary algebra to 50, the crossover probability to 0.7 and the mutation probability to 0.2.

* Corresponding author: hnzzjp2010@126.com

4 Results and discussions

The model of BP neural network optimized by genetic algorithm was BP-GA model. Because of the difference of the dimension and magnitude between the input and output variables can affect the accuracy of the model. Therefore, all data of the input and output matrices were transformed to the interval [-1, 1] using a normalization formula.

4.1 Model training results analysis

81 training samples of 3 simulated rainfall experiments under grassland conditions were put into BP-GA model for training. During the training of 10 BP-GA models, the model parameter optimization process of genetic algorithm was traced 5 times randomly, and the optimal fitness of all individuals in each generation was recorded. Optimization process of BP neural network model parameter using genetic algorithm was shown in Fig 1. Due to the random assignment of parameters during model initialization, the initial fitness of each population was different, the optimal fitness of the population decreased with the increase of the genetic generation, and the crossover and mutation of chromosomes led to the drastic changes of the optimal fitness, resulting in some inflection points in the optimal fitness curves. The optimal fitness of each population tended to be consistent at around 9.5, and was stable at about 9.5.

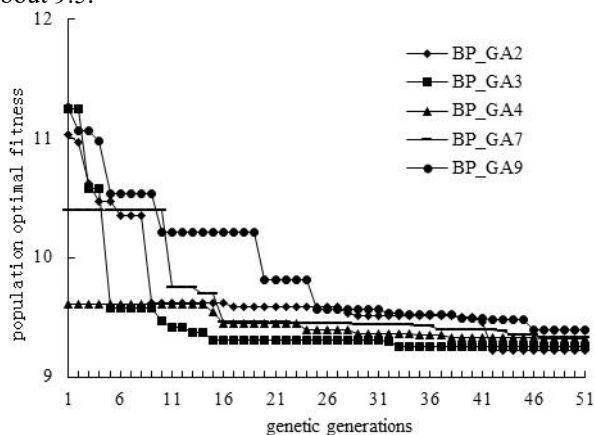


Fig. 1. Parameter optimization process of BP neural network by genetic algorithm.

The average prediction errors of 10 BP_GA models for the training samples were shown in Fig 2. When the simulated rainfall intensity is 45 mm/h, the average prediction error of all the training samples is 4.77%, and the maximum prediction error is 12.23%. When the simulated rainfall intensity is 90 mm/h, the average prediction error of all the training samples is 4.52%, and the maximum prediction error is 13.06%. When the simulated rainfall intensity is 130 mm/h, the model predicts the average error of all the training samples 6.52%, the maximum prediction error is 15.46%. The model had a high prediction accuracy. On the one hand, the grass can reduce the impact of raindrops on the surface of the slope, and on the other hand, it can increase the soil anti-corrosion ability and maintain the stability of the surface morphology.

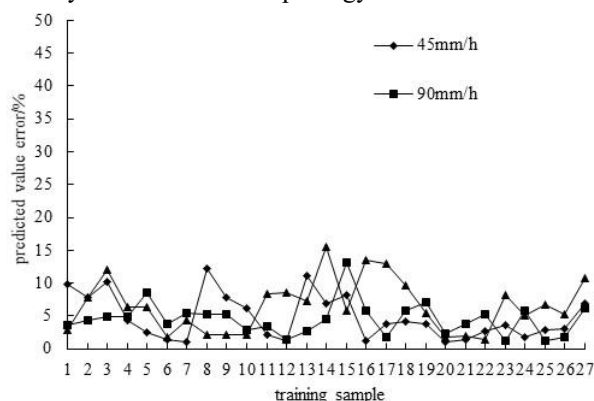


Fig. 2. Predicted value error of training samples of BP-GA model.

4.1 Model verification

9 groups of test samples of three simulated rainfall experiments under grassland conditions were respectively substituted into the BP-GA model for verification. The test samples were randomly selected in different stages of each experiment and had good representativeness. The average prediction error of the BP-GA model for the test sample set was 8.02%, and the maximum prediction error was 18.34%. The model had a high simulation accuracy and a good generalization ability.

Table 2. Model verification results of BP-GA model.

Number	Rainfall Intensity (mm/h)	Flow Velocity (cm/s)	Flow Depth (mm)	Roughness	Coverage	Measured Value	Predicted Value	Error (%)
1	45	0.025	0.5	0.1	0.5	18.30	17.26	5.68
2	45	0.036	0.5	0.1	0.5	16.00	16.41	2.56
3	45	0.034	0.8	0.1	0.5	15.60	14.78	5.26
4	90	0.064	2	0.2	0.63	12.70	13.12	3.31
5	90	0.067	2.8	0.2	0.63	10.40	12.14	16.73
6	90	0.063	2.6	0.2	0.63	12.90	13.63	5.66
7	130	0.075	3	0.3	0.9	9.62	8.37	12.99
8	130	0.077	3.1	0.3	0.9	10.98	10.80	1.64
9	130	0.093	3.9	0.3	0.9	8.56	10.13	18.34

The predicted value and the measured value calculated by BP-GA model were fitted and analyzed. The fitting relationship between the two was good, the fitting coefficient was 1.0045 and the coefficient of determination R^2 was 0.8588. Fig 3 showed the predicted and measured values have a good correlation and a one-to-one correspondence.

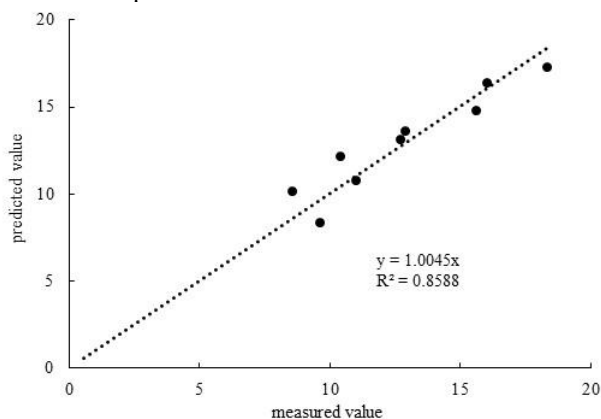


Fig. 3. Test set fitting relationship of BP-GA model.

5 Conclusions

According to the experiment data, model of overland flow resistance coefficient was established based on BP neural network optimized by genetic algorithm. The input variations of model were rainfall intensity, flow velocity, water depth, and roughness of slope surface, and the output variations was overland flow resistance coefficient. Genetic algorithm could improve the parameter optimization speed and simulation accuracy of BP neural network. The average prediction error and the maximum prediction error of the test set of the BP-GA model were respectively 8.02% and 18.34%. Because of the limitation of experiment conditions, the influences of different cover and slope on the overland flow resistance coefficient are not considered, and the prediction method of BP neural network model is only discussed by using simulated rainfall experiment data of grassland slope of 20°.

Acknowledgements

This paper is based on work supported by the National key research priorities program of China (2016YFC0402402) and the Central Nonprofit Research Institutions Basic Scientific Research Special Fund (HKY-JBYW-2014-19).

References

1. W Y. Yao. Journal of sediment research, **1** (1996): 74-82
2. A D. Abrahams, G. LI, A J. PARSON. Earth Surface Processes and Landforms, **21** (1996): 35-47
3. K L. Zhang. Yellow River, **20(8)** (1998): 13-15.
4. G H. Zhang. Advances in Water Science, **13(2)** (2002): 159-165

5. A. Knapen, J. Poesen, S. Baets de. Soil and Tillage Research, **94(2)** (2007): 425-440
6. M. Fatteta, Y. Fub, M. Ghestemc. Catena, **87(1)** (2011): 60-69
7. P Q. Xiao, W Y. Yao, L. Li, C X. Yang. Journal of sediment research, **3** (2013): 1-5

* Corresponding author: hnzzjp2010@126.com