Research of Soil Moisture Content Forecast Model Based on Genetic Algorithm BP Neural Network

Caojun Huang , Lin Li, Souhua Ren, Zhisheng Zhou

College of Information and Technology of Heilongjiang Bayi Agricultural University, Daqing 163319, China

Abstract. Soil moisture forecast model based on genetic neural network is established because the soil moisture forecasting is nonlinear and complex. The weights and threshold value of BP network are optimized according to the total situation optimization ability of genetic algorithm, which can avoid effectively that BP network is vulnerable to run into the local minimum value as its poor total optimization ability. The model is applied to Hongxing farm in Heilongjiang Province to predict the soil moisture. The forecasting result shows that the model has favorable forecasting precision, which indicates that the genetic neural network model is feasible and effective to predict the soil moisture.

Keywords: GA-BP, soil in the field, moisture forecasting, modeling, forecast precision

1 Introduction

With the increasing development of industrialization, the human environment and the global extreme climate deteriorates more and more, the drought and flood disaster happen continually. All of these restrict the agricultural production and the healthy development of social economy badly. To ensure food safety, fertility and get in timely, the accurate forecast information of soil moisture of farmland is needed to be achieved in time and exactly. Therefore, establishing the soil moisture prediction model is conducive to science management of farming and has important significance for the food fertility and stable grain.

The different soil moisture prediction models were established by experts and scholars, such as: empirical formula of the model, the water balance model, the dynamic model of soil water, time series model, remote sensing model and neural network model. These models are always restricted in the practical application [1-6]. For example, although the empirical model is simple and useful, the model parameters have limited applied scope. Other models such as the water balance model, the dynamic model of soil water, time series model etc. have better suitability, but they need a mass of measurement data. The remote sensing model has poor stability. Neural network model get into the local minimum values easily, which makes the network cannot search the global optimal solution [7].

According to the complexity and the shortages of forecast of soil moisture model, a suitable algorithm for processing the non-linear problem and with global optimization ability of Genetic Algorithm (GA) is used. It optimizes neural network coefficients and threshold and form a hybrid GA-BP Algorithm, which has the practical value of the soil moisture prediction model of the soil moisture forecast.

2 Genetic BP Network Model

The genetic BP network utilizes the overall situation optimizing ability of genetic algorithm to search optimization for weights and threshold value of BP network within large scope, then the optimum solution is endued to the weights and threshold value of BP network, finally BP network is used to search precisely to obtain the optimal weight and threshold value [8]. The specific implementation process is as follows [9-10]:

2.1 Population initialization and coding

Population initialization is used to refers to initialize the BP network's weights and threshold. According to population design experience, the population A with M individual is produced, $A=\{X_1, X_2, X_3, \cdots, X_m\}^T$, individual $X_i=\{x_1, x_2, x_3, \cdots, x_m\}$

M genes in $X_i=(x_1, x_2, x_3, \cdots, x_m)$ are composed of BP network's weights and threshold. The length computation is as follows:

$$m = r \cdot S_1 + S_1 \cdot S_2 + S_1 + S_2 \tag{1}$$

Where r is the input node numbers, S_1 is the hidden nodes, and S_2 is the output layer nodes.

As the soil moisture prediction is more complicated, the real number coding scheme is chosen, which can effectively avoid weight step change. Next, with real coding method, the vector of three-layers BP networks weights and threshold is:

$$Xi=[\omega_{11}, \cdots, \omega_{1S_1}, \cdots, \omega_{1S_1*r}, \omega_{21}, \cdots, \omega_{2S_2}, \cdots, \omega_{2S_2*S_1}, B_{11}, \cdots, B_{1S_1}, B_{21}, \cdots, B_{2S_2}] \tag{2}$$

Where ω_1 is the weights value between input and hidden layer, ω_2 is the weights value between output and hidden layer, B_1 is the hidden layer's threshold value, and B_2 is the output layer's threshold value.

2.2 Fitness function

The fitness function is used to evaluate individual. Each individual is evaluated according to the fitness function value, then BP neural network input sample is obtained by decoding corresponding individual. Finally, neural network output error is calculated, the fitness function is used:

$$\mathrm{F(i)} = \frac{1}{E} \tag{3}$$

Where E is BP network output error, its computation formula is:

$$\mathrm{E} = \frac{1}{2}\sum (y - Y_b)^2 \tag{4}$$

Where y is network actual output, Y_h is network standard output.

2.3 Choose operation

According to the calculation results of fitness, the highest fitness groups are retained. High fitness individuals are selected to pass down offspring directly. The other individuals are chosen by roulette selection method, so the small fitness of individuals also have the opportunity to pass down offspring. This ensures the diversity of the individuals in the group, and avoids the local optimum. Select formula is as follows:

$$P_{is} = \frac{F(i)}{\sum_{i=1}^{M} F(i)} \tag{5}$$

Where P_{is} is probability for selection: M is the population size: i is the first i individual in group M .

2.4 Crossover operation

Since the weight coefficients are obtained by using real number coding, arithmetic crossover is used. If $\mathrm{X}_{\mathrm{a}}^{\tau}$ and X_b^{τ} are two crossover individuals, then the new individuals are generated after operation:

$$X_a^{\tau+1} = \alpha X_b^{\tau} + (1-\alpha) X_a^{\tau} \tag{6}$$

$$X_b^{\tau+1} = \alpha X_a^{\tau} + (1-\alpha) X_b^{\tau} \tag{7}$$

Where α is linear combination coefficient, and usually $\alpha \in (0,1)$.

2.5 Mutation operator

The non-uniform mutation operation can improve the algorithm accuracy, which makes the final search closer to the optimal solution. So, non-uniform mutation is selected. Mutation operating is as follows:

x_m is variation point of variance $X = x_1 x_2 x_3 \cdots x_m \cdots x_n$, $x_m \in [a,b]$, the new genetic $\tilde{x}_m$ of variation point can be obtained by:

$$\tilde{x}_m = \begin{cases} x_m + (b - x_m)(1 - r^{(1-\frac{t}{T})b}) \\ x_m - (x_m - a)(1 - r^{(1-\frac{t}{T})b}) \end{cases} \tag{8}$$

Where $r \in [0,1]$, and it has the uniform probability distribution; t is the current evolution algebra, T is the termination algebra, b is shape parameters of the system.

According to (8), search scope is large in the early stage of algorithm operation(when t is small).While it is small in the late stage(when t is close to T) ,which is equivalent to local search.

2.6 Generate new population

A new population will generates when new individuals are inserted into the primary populations, then the connection weights of individuals in the new population are given neural network, and the new individual fitness function is calculated, if it meets predefined criteria, then goes to the next step, otherwise continues to genetic operation.

2.7 Optimal operation

The optimal individuals which have been decoded are endowed to neural network's weights and threshold value as the net's initial value. BP neural network is trained until the error square sum reach the specified accuracy, or reach the maximum iteration times. The flow chart of Genetic neural network algorithm is shown in Figure 1:

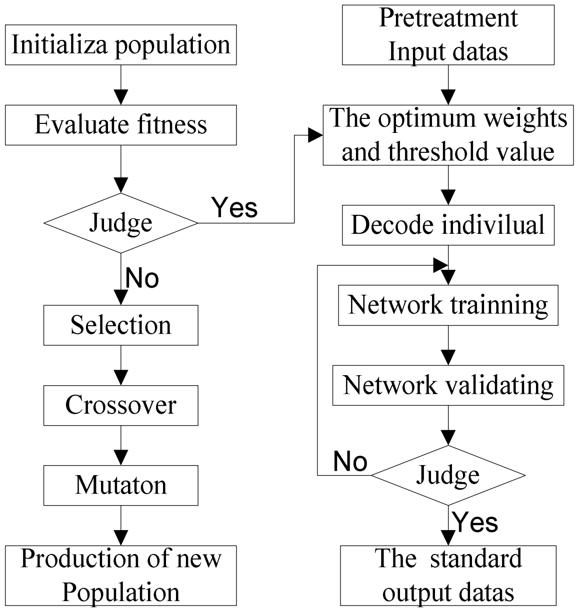

Fig.1. Realizing flow chart of GA-BP

3 Implementation of soil moisture forecast

This research is carried out in Heilongjiang Hongxing farm, which is located in the northern of Heilongjiang province Songnen Plain and Xiaoxing'an Mountain zone, belongs to cool transitional continental monsoon climate. Location: latitude 48°02'~ 48°17', longitude 126 ° 47 '15 ° ~ 127. Annual average temperature is 0.8 ℃, the average frost-free period is117 days, annual rainfall is about 553.0mm. The temperature is below 0 ℃ in five months every year. Hongxing farm has plenty of rain and heat than adequate, long sunshine hours, high soil organic matter content, strong water storage capacity, which provide more advantageous conditions for soybeans, wheat and other crops of high economic value.

3.1 Feature extraction of Model input and output

The actual soil moisture data, under10cm, 20cm, 30cm, 40cm and 50cm from 2004 to 2009, is used in the model. The 10 factors effecting soil moisture remarkably, such as temperature, rainfall, evaporation, relative humidity and sunshine hours and other meteorological factors, are applied to the input of the model.

The model output is soil moisture data. According to the training results, it can be known that single output factor has higher precision and faster convergence speed. Therefore, we select single soil moisture data as the model output. Research shows that with the crops growing, the main water area of rhizome has downward trend [11-14]. Therefore, we focus on this layer of the surface 10cm of soil moisture conditions in April and May, in June, layer of the surface 20cm, in July, August, September, layer of the surface 30cm.

3.2 Hidden layer and hidden layer nodes determination

As the three-layer neural network can approximate any function, we use a single hidden layer BP network structure [15]. Based on the above analysis, there are ten input factors and one output factor. According to the calculation formula of the hidden nodes [16], the optimal hidden nodes number is 6 to 23. According to the training results, the optimal hidden nodes is 22. Therefore, the network topology is :10-22-1.

3.3 Description of training samples

Firstly, the input and output sample data are normalized for network training [17]. Input samples and the corresponding output data are inputted to the network with the recurrence relations, thus completing the genetic neural network model training, modeling process, the input and output data table as follows:

Table1. Correspondence of the input and output data

Input sample data	Output sample data
$X_1, X_2, \cdots\cdots, X_n$	$Y_1, Y_2, \cdots\cdots, Y_n$
$X_{n+1}, X_{n+2}, \cdots\cdots, X_{n+n}$	$Y_{n+1}, Y_{n+2}, \cdots\cdots, Y_{n+n}$
……	……
$X_{kn+1}, X_{kn+2}, \cdots\cdots, X_{kn+n}$	$Y_{kn+1}, Y_{kn+2}, \cdots\cdots, Y_{kn+n}$

Where X_1 to X_n and Y_1 to Y_n are input and output sample data in the first year arranged by time serial. X_{kn+1} to X_{kn+n} and Y_{kn+1} to Y_{kn+n} homoplastically. The input sample data includes: 10 ~ 50cm of five levels of soil moisture data, temperature, rainfall, evaporation, relative humidity and sunshine duration, etc; output sample data is a certain level of soil moisture data

According to its requirement by Heilongjiang province of soil moisture report, soil moisture data should be reported to Land Reclamation Bureau monthly from April to October each year on the 8th, the 18th, the 28th. To this end, each of the input sample data arranged in chronological order: April 8 sample data (X_1), April 18 sample data

(X_2), April 28 sample data(X_3), ..., October 18 sample data X_n; output sample data arranged in chronological order: April 18 sample data (Y_1), April 28 sample data (Y_2), May 8 sample data (Y_3), ..., October 28 sample data (Y_n).

3.4 Forecast implementation

The genetic neural network parameters are set up firstly, then the input and output data are put into the network for training. When the training result is in accordance with established standards, the training stop, the genetic neural network structure and parameters are saved. The trained neural network is used to forecast the soil moisture from April to October in 2009, and the forecast results are compared with BP neural network forecast model results, the detailed as follows:

Table 2. Comparison of forecasted results by two models

Date（day/month）	Measured results（%）	GA-BP network model		BP network model	
		Forecasting results（%）	Error（%）	Forecasting results（%）	Error（%）
18/4	34.98	36.253	1.273	36.676	1.696
28/4	29.15	29.775	0.625	29.255	0.105
8/5	17.49	17.363	0.127	15.386	2.104
18/5	22.79	23.711	0.921	23.627	0.837
28/5	23.64	24.672	1.032	25.235	1.595
8/6	33.74	30.381	3.359	29.525	4.215
18/6	35.52	37.585	2.065	38.61	3.090
28/6	42.62	37.128	5.492	35.444	7.176
8/7	40.25	44.842	4.592	43.704	3.454
18/7	45.51	47.498	1.988	49.681	4.171
28/7	49.37	50.62	1.25	53.164	3.794
8/8	43.99	42.98	1.01	39.015	4.975
18/8	30.54	32.409	1.869	32.227	1.687
28/8	35.22	32.587	2.633	31.683	3.537
8/9	38.14	39.48	1.34	40.63	2.490
18/9	36.62	37.356	0.736	38.034	1.414
28/9	34.87	32.51	2.36	31.503	3.367

It can be seen from the table that BP network prediction of soil moisture average absolute error is 2.92%; GA-BP network prediction of soil moisture average absolute error is 1.92%. Obviously, GA-BP network prediction accuracy increases one percentage point. The results show that the genetic algorithm of BP neural network has better mapping capabilities and better generalization ability comparing to the single BP neural network. It satisfies modeling requirements and is applicable to the soil moisture forecast.

4 Conclusion

The soil moisture prediction model was established based on the combination of genetic algorithm and neural network. This model offsets the shortcomings that the neural network for global searching capability is weak and easily fallen into the local minimum. It also gives full play to the neural network nonlinear mapping capability. The model has been applied to Heilongjiang Hong Xing Farm on soil moisture prediction and has achieved good results. Compared to the BP neural network model, the genetic BP neural network model has better prediction accuracy and application value, and favorable application prospect in the soil moisture forecast.

References

1. Chen Hongwei : Research on Soil Moistrue Prediction Models of Inadequate Irrigation Rice Field in South China . Yangzhou: Yangzhou University (2004)
2. Mahmood R , Hubbard K G.: An analysis of simulated long-term soil moisture date for three land uses under contrasting hydroclimatic conditions in the northern greet plains.J. Hydrometeoroloy, vol.5, pp . 160--179(2004)
3. Zhou Liangchen: Study on Estimation of Soil-water Content by Using Soil-Water Dynamics Model.J. Water Saving Irrigation, vol.3, pp . 10--13(2007)
4. Zhang Hexi, Yang Jing, Fang Xiaoyu, Fang Jun, Feng Cheng, et al.: Application of Time Series Analysis in Soil Moisture Forecast . J . Research of Soil and -Water Conservation, , vol.15, pp . 82--84(2008)
5. loess soil are . J. Remote Sensing, 1, vol. 19(2), pp . 237--243(1998)
6. Ma Xiaogang: Forecast of Soil Moisture Content during Critical Period of Spring Sowing Based on Precipitation in Last Autumn. J . Chinese Journal of Agrometeorology, vol. 29(1), pp. 55--57(2008)
7. Luo Chengfeng, Liu Zhengjun, Wang Changyao, et al.: Optimized BP neural network classifier based on genetic algorithm for land cover classification using remotely-sensed data. J . Transactions of the CSAE, vol. 22(12), pp. 133--137(2006)
8. Zhou Shiguan; Li Zhongxia: Genetic Algorithm for Optimization of Neural Network Structure and Weight Distribution. J. Ordnance Industry Automation, vol.23(4), pp. 48--49(2004)
9. Zhi Jing, Zhang Dongmei,Jiang Pengfei: Based on PCA of genetic neural network prediction of stock index. J . Computer Engineering and Applications, vol. 45(26),pp. 210-212(2009)
10. Vittorio M. Genetic evolution of the topology and weight distribution of neural networks . J. IEEE Transactions on Neural Networks,vol. 5, pp. 39--53(1994,)
11. Zhang Xiying, Yuan Xiaoquan: A Field Study on the Relationship of Soil Water Content and Water Uptake by Winter Wheat Root System. J. Acta Agriculturae Boreali-Sinica, vol. 10(4),pp. 99--104(1995)

12.Zhang Yingshan: Croproot system and Soil water application. China Meteorological Press, Beijing(1999)
13.Cao Minjian: Principle of cropping system and soil management. China Agriculture Press, Beijing(2002)
14.Cui Lijie，Xu Baolin: Experimental study of Soil moisture content in the Midwest of Jilin province. J . Jilin Water Resources, vol. 1, pp. 14--15(2003)
15.R.Hecht-Nielsen: Counterpropagation Networks. J . Applied Optics, Vol. 26, pp. 4976--4984(1987)
16.Paola J D, Schowengerdt R A. : A review and analysis of backpropagation neural networks for classification of remote-sensed multi-spectral imagery . J . International Journal of Remote Sensing, vol. 16(16), pp. 3033--3058(1995,)
17.Wang Ling, Zheng Dazhong**:**Two hybrid learning strategies of feedforward network. J . Journal of Tsinghua University(Science and Technology), vol., 38(09), pp. 95--97(1998)