Application of BP-ANN Based on RAGA to Forecasting the Relation between Rice Accumulated Temperature and Yield

Qiang Fu[1,2], Ying Qi[2], Wei Zu[3]

1. Doctoral Working Station of Beidahuang Company, General Bureau of Agricultural in Heilongjiang, Harbin, Heilongjiang 150040, China; fuqiang100@371.net
2. School of Water Conservancy & Civil Engineering, Northeast Agricultural University, Harbin, Heilongjiang 150030, China;
3. School of Agriculture, Northeast Agricultural University, Harbin, Heilongjiang 150030, China

Abstract: The yield of rice is related to accumulated temperature. So, the yield of rice can be predicted according to accumulated temperature. The Error Back-Propagation artificial net based on real coded accelerating genetic algorithm (RAGABP) was put forward to predict the rice harvest. It was shown that RAGABP has several advantages such as well whole optimum and powerful non-linear mapping. The model has more accuracy and fast calculation velocity, and it can be used in other problems related to predicting. [The Journal of American Science. 2005;1(2):28-32].

Key Words: genetic algorithm; BP-ANN; rice accumulated temperature; yield; forecasting

Introduction

Rice is one of the main crops in china. The yield of rice changed largely in different years, which influences the total output of crop. Some studies showed that there is close relationship between the yield and accumulated temperature of more than 20℃ during the period of growth. When the accumulated temperature more than 20℃ was below 2100℃, the phenomenon of cold disaster will be happened, then the yield will reduce seriously. Only when the accumulated temperature more than 20℃ is right, and the illumination and irrigation is befitting, rice will get high yields. Thus, it can be seen that the number of the accumulated temperature ≥20 ℃ has great effect to the yield. It is very important to give exact forecasting of the accumulated temperature ≥20℃ to arrange the agricultural production and keep a dependable crop.

1 Basic Condition

At present, the mean annual temperature has been used often. Because the accumulated temperature changed greatly in different year, if the farming has been arranged according to routine climate index, the accumulated temperature is more in high-temperature year and less in low-temperature year.

The relationship between accumulated temperature more than 20℃ and rice yield has been researched (Xu, 1995). According to the date from 1951 to 1995 between rice yield (kg/hm^2) and the accumulated temperature more than 10℃ and 20℃ in Shenyang of China, the author used the method of moving average to analyze the weather yield. Through calculating, the correlation coefficient between weather yield and accumulated temperature more than 20℃ is 0.4446 ($\gamma = 0.4446$, $\alpha = 0.05$). The correlation coefficient between weather yield and accumulated temperature more than 20℃ is 0.264 ($\gamma = 0.264$). It indicates that the accumulated temperature more than 20℃ is rather to rice yield. Thus, according to the original date, the accumulated temperature more than 20℃ can be divided into three types (Xu, 1995). The first is normal but on the high side, the second is normal but on the low side, and the third is cold disaster.

Table 1. Category Table (1955~1990)

Year	Category	Year	Category	Year	Category
1955	3	1967	2	1979	2
1956	2	1968	3	1980	2
1957	4	1969	4	1981	1
1958	1	1970	2	1982	2
1959	2	1971	2	1983	1
1960	3	1972	4	1984	1
1961	1	1973	4	1985	2
1962	2	1974	3	1986	2
1963	1	1975	2	1987	2
1964	2	1976	4	1988	1
1965	2	1977	2	1989	3
1966	4	1978	2	1990	3

Correlation analysis has been used in original reference (Xu, 1995). The method is complex and poor accuracy. The artificial neural net has powerful function mapping, each given net can express quite complex function relationship, which cannot be expressed by mathematics formula, especially the non-linear relationship (Fu, 2003; Xing, 2003). The neural net is in poor convergence speed, and getting into the part optimal point easily. Because the Genetic Algorithm has the capability to handle these problems, the real coded accelerating genetic algorithm linked with BP-ANN (RAGABP) has been applied to forecasting the rice yield according to accumulated temperature more than 20℃ (Fu, 2003).

2 Error back Propagation Artificial Neural Net (BP-ANN)

Single implicit layer net showed in Figure 1 has been used commonly in multiplayer feedforward BP-ANN. Those three lawyers include input layer, implicit layer and output layer.

In three layer feed forward net, the input sample is $X=(x_1,x_2,\cdots,x_r)^T$. Let $x_0=1$, the implicit layer has a threshold value. In implicit payer, the output vector is $Y=(y_1,y_2,\cdots,y_m)^T$. Let $y_0=1$. In output layer, the output vector is $O=(o_1,o_2,\cdots,o_k,\cdots,o_l)^T$; the expected output vector is $d=(d_1,d_2,\cdots,d_k,\cdots,d_l)^T$.

Weight matrix from output layer to implicit layer is $V=(V_1,V_2,\cdots,V_j,\cdots,V_m)$, where the column vector V_j is weight vector of the j^{th} neuron corresponding in implicit layer. Weight matrix between implicit layer and output layer is $W=(W_1,W_2,\cdots,W_k,\cdots,W_l)$, where the column vector W_k is weight vector of the k^{th} neuron corresponding in implicit layer. The type of S function has been used as in implicit layer, which can avoid the phenomenon of super saturation. Linear function has been often used in output layer in order to handle the practical problem.

Defining network error as follows (Figure 1).

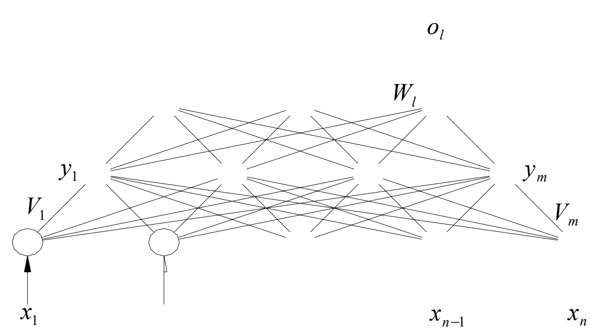

Figure 1. Three layers of BP-ANN

$$E = \frac{1}{2}(d-o)^2 = \frac{1}{2}\sum_{k=1}^{l}(d_k - o_k)^2 \qquad (1)$$

According to formula (1), where

$$E = \frac{1}{2}\sum_{k=1}^{l}[d_k - f(\sum_{j=0}^{m} w_{jk} y_j)]^2 \qquad (2)$$

According formula (2), where

$$E = \frac{1}{2}\sum_{k=1}^{l}\{d_k - f[\sum_{j=0}^{m} w_{jk} f(\sum_{i=0}^{n} v_{ij} x_i)]\}^2 \qquad (3)$$

According to formula (3), the input error of network is a function of weight value w_{jk} and v_{ij} in each layer. By the way of adjusting weight value constantly, error E will be changed in order to attain the scheduled accuracy.

Adjusting weight value formula of BP algorithm is as follow.

$$\begin{cases} \Delta w_{jk} = \alpha \delta_k^{\,o} y_j = \alpha (d_k - o_k) o_k (1 - o_k) y_j \\ \Delta v_{ij} = \alpha \delta_j^{\,y} x_i = \alpha (\sum_{k=1}^{l} \delta_k^{\,o} w_{jk}) y_j (1 - y_j) x_i \end{cases} \qquad (4)$$

In formula (4), α is studying speed.

2.1 Error back propagation flow of BP algorithm

The process of error back propagation is to calculating output layer e_k. According formula, the δ_{ki} can be calculated. Because it is not expected output in implicit layer, the Δw_{ki} can be calculated by δ_{ki} after that, $e_i = \sum_{k=1}^{l} \delta_{ki} w_{ki}$. According to e_i , the δ_{ij} can be calculated. Then, the Δv_{ij} can be calculated also.

If there has implicit layer former, following with the method above-mentioned, till making output error e_k of each layer back reckon to first layer. Direct course described as Figure 2.

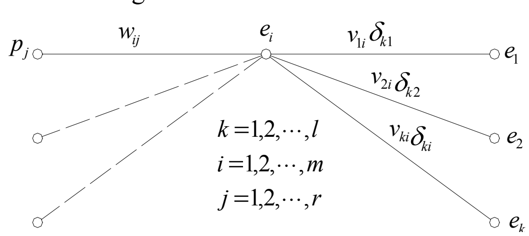

Figure 2. Errors back propagation flow

2.2 Limit and shortage of BP algorithm

Error back propagation algorithm has been applied extensively, but it has shortage itself, it is uncertainty in training process (Fu, 2003), such as follows：

(1) It will waste long time in training. For some special problem, runtime may need few hours even longer. Poor studying velocity is the main reason of these, so the adaptive studying velocity will be used to improve it.

(2) Far from training. Weights data adjust too much, so that excitation function arrives at saturation in training course, adjusting net weight data nearly stagnate. In order to avoid this problem, the first method is to choose less initial value. The second method is to adopt less studying velocity.

(3) Getting into part minimum. BP algorithm could make net weight value astringe final value, but it could not ensure that the answer gotten is optimal of error hyperplane, sometimes, it maybe a part minimum. BP algorithm adopt gradient method, training begin from some initial point, along the bevel of error function to get error minimum gradually. The minimum varies in different initial point, so the optimal answer cannot be calculated. In order to improve precision of answer, multilayer net and more nerve cell often be used when the training answer cannot arrive at precision, which increase training time and complexity of net at the same time.

3 Error Back-propagation Artificial Net Based on Real Coded Accelerating Genetic Algorithm (RAGABP)

Genetic algorithm has favorable full-searching ability, and BP algorithm has power part-searching ability, there is the error back-propagation artificial net based on real coded accelerating genetic algorithm (RAGABP) when combining the two algorithms.

The modeling steps of RAGABP as follows:

(1) Producing many neural net configuration by some probability of experience formula, getting initial weight data and threshold value of different neural net structure, and taking energy function (error function) of neural net as aim function, through choosing, hybridization and variation (Xing, 2003), the initial weight data and threshold value of optimal net structure will be calculated.

(2) Based on step (1), taking initial weight data and threshold value of optimal net structure as input parameter, entering BP algorithm, it will get stable neural net weight value and net structure by much training.

4 Application Example

4.1 The processing of the input and output data

Define the input nerve cell is 1, which means the value of accumulated temperature. When the input value is 1, because the function of sigmoid is sensitive to zone of [0,1], dealing with these data normalization is as follows:

Changing the data into [0,1].

$$\overline{x} = (x_i - x_{\min})/(x_{\max} - x_{\min}) \qquad (5)$$

where, x_i is input or output data, $x_{\min}$ is min variation, $x_{\max}$ is max variation.

4.2 Selecting of modeling parameter and net structure

There are 4 neurons in output layer, which indicate the four different temperature types. There is the right yield while input value is 1 and the irrelevancy yield while input value is 0, such as （1，0，0，0)，(0，1，0，)，(0，0，1，0)，(0，0，0，1）.

Genetic algorithm have 400 colonies, accelerating 3 times, entering BP algorithm, training 1368 times, concluding the data of energy function is 4.5137×10^{-4}, and average error is 3.1345×10^{-6}. Now topological structure describe as chart 1-5-4 and training answer showed as Table 2.

Table 2. Training result of RAGABP according to different yield

Year	Output	Category	Year	Output	Category	Year	Output	Category
1955	(0,0,1,0.04)	3	1967	(0,1,0,0)	2	1979	(0,1,0,0)	2
1956	(0,0.98,0,0.01)	2	1968	(0,0,1,0)	3	1980	(0,1,0,0)	2
1957	(0,0.01,0,1)	4	1969	(0,0,0.1,1)	4	1981	(1,0,0,0,)	1
1958	(1,0,0,0.0.1)	1	1970	(0,1,0,0)	2	1982	(0,1,0,0)	2
1959	(0,0.99,0,0)	2	1971	(0,1,0.1,0)	2	1983	(1,0.0,0,0)	1
1960	(0,0,1,0)	3	1972	(0,0,0,1)	4	1984	(0,1,0,0)	1
1961	(1,0,0,0)	1	1973	(0,0,0,0.98)	4	1985	(0,1,0,0)	2
1962	(0,1,0,0)	2	1974	(0,0,0.1,1)	3	1986	(0,0.97,0,0)	2
1963	(1,0,0,0)	1	1975	(0,0.99,0,0)	2	1987	(0,1,0,0)	2
1964	(0,1,0,0.1)	2	1976	(0,0,0,1)	4	1988	(0,1,0,0.1)	1
1965	(0,1,0,0)	2	1977	（0.1,1,0,0）	2	1989	(0,0,1,0)	3
1966	(0,0,0,1)	4	1978	(0,1,0,0)	2	1990	(0,0,0.98,0)	3

Table 3. Forecasting the yield of 1991~1994 by RAGABP

Year	1991	1992	1993	1994
Practical yield	1	1	2	1
Forecasted yield	（0.98,0,0,0）	(0.97,0,0,0.1)	(0,1,0,0)	(1,0,0,0.1)

From Table 2, the fit precision of RAGABP model is rather high. Using the BP net structure to forecast the yield of 1991 to 1994, and training answer are showed in Table 3.

From Table 3, the practical yield of rice is the same as the forecasting by RAGABP net structure. Forecasting the yield is in high accuracy and credibility.

4 Conclusions

(1) RAGABP can describe non-linear relationship between the yield and accumulated temperature more

than 20℃ during the growth period accurately.

(2) The model of RAGABP can be used in calculating course easily, and it has higher accurate than the algorithm of reference.

(3) The model of RAGABP could not be influenced by artificial factor and the net input vector will be improved with the predicting data fully, forecasting yield is more accuracy.

(4) The model of RAGABP can be widely used in some other forecasting problems.

Acknowledge

The financial support is provided by Chinese National "863" High-Technique Programme (No. 2002AA2Z4251-09); China Postdoctoral Science Foundation (No. 2004035167); Heilongjiang Province Youth Foundation (No. QC04C28); Natural Science Fund in China (No. 30400275).

Correspondence to:

Qiang Fu，Wei Zu
School of Water Conservancy & Civil Engineering
Northeast Agriculture University
Harbin, Heilongjiang 150030, China
Telephone: 01186-451-55190298 (Office)
Cellular phone: 01186-13936246215
E-mail: fuqiang100@371.net

References

[1] Cong S. Application and theory of neural net to face MATLAB toolbox. Science and Technology Publishing Company, China, 1998.

[2] Fu Q, Liang C. Modeling and optimize technique of water saving irrigation system. Sichuan Technology Publishing Company, Chengdu, Sichuan, China. 2003;22（5）: 56-62.

[3] Fu Q, Li XQ, Xiao JM. Application of Based of RAGA to ascertain the relation of flux and depth. Northeast Hydraulic and Hydroelectric Publishing Company. 2001;19（10）:1-4.

[4] Fu Q, Song YF, Xiao JM. Application of ANN Forecasting water demand of rice irrigation. Northeast Hydraulic and Hydroelectric Publishing Company 2005;5:38-40.

[5] Fu Q, Fu H. Applying PPE Model Based on RAGA in the Investment Decision-Making of Water Saving Irrigation Project. Nature and Science 2003;1(1):57-61.

[6] Fu Q, Zu W. Using Multi-Dimension Dynamic Planning Based on RAGA to Optimize Irrigation System under Non-Sufficient Irrigation Condition. The Journal of American Science 2005;1(1):68-74.

[7] Ma XD, Fu Q. Applying Self-Organizing Competition Artificial Neural Networks to Classify the Soil. Nature and Science 2003;1(1):75-81.

[8] Xing ZX, Fu Q. Application of Based of RAGA to nature riverway surface line. Transaction of Irrigation and Drainage. 2003:60-63

[9] Xu FL, Jiang XY, Li KX. The relationship between rice accumulate temperature more than 20℃ during growth period and yield and the forecasting method. Liaoning Weather 1955（3）: 22-23.