Prediction Model of Greenhouse Temperature of Multiple-factor Correlation Based on Error Back Propagation Algorithm

Dong Wang[a], Xue Pei, Xuanke Cui

College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling 712100, Shaanxi, China

[a]wangdong510@163.com

Abstract

Greenhouse can create a small environment suitable for growing plant by adjusting various environmental factors, and temperature is the first essential factor, however, because it is correlative with many factors, how to establish greenhouse temperature model based on multi-factor correlation has become a hot research topic. Different control methods cause differences of greenhouse complex models to be big, while closed greenhouse models are the basis of complex model, so this paper uses closed greenhouse as the research object modeling. The main influencing factors are the temperature outside the greenhouse, the temperature inside the greenhouse and the intensity of the solar light outside the greenhouse, and then three types of monitoring data are obtained through the monitoring system, a total of 1296 groups of test samples are selected to construct the training set, and the prediction model of greenhouse temperature based on error back propagation algorithm are established. Experiment shows that the predicted root mean square error is 0.42 ℃; maximum error is 1.86 ℃, with high accuracy, and it can provide a good theoretical basis for establishment of greenhouse control system.

Keywords

Greenhouse, temperature model, multiple-factor correlation, BP network, solar light intensity.

1. Introduction

The growth of crops is limited by environmental conditions, and the greenhouse can create a microclimate to promote crop growth. The microclimate system has the time varying, nonlinearity, strong coupling, great inertia and other characteristics, temperature is the first factor that affects crop growth in the many factors of greenhouse environment [1], and the greenhouse temperature is related to light and heat radiation and other factors [1-2], how to establish the temperature model is of great significance to improve the control accuracy and efficiency.

In recent years, the research on greenhouse temperature model is mainly the physical mechanism model [3] and greenhouse model of intelligent algorithm identification [4]. However, the structure of mechanism model is complex in and parameters are more, and many parameters are difficult to determine, which are obtained by experience alone [5], it is not suitable for intelligent control. The accuracy of the identification model is directly dependent on the choice of parameters and the structure of the model, because the main influencing factors are not screened; and cause the prediction results of this model have bigger error and increases input dimension and other problems.

In view of the above analysis, considering differences of stable impact factors of greenhouse control way are vary greatly, the screening relevant impact factors is difficult, and closed greenhouse analysis is the foundation of the greenhouse complex model analysis. Therefore, this paper takes the closed greenhouse model as research object, based on the analysis of the classical model, puts forward outdoor temperature, indoor temperature and outdoor sunlight intensity three factors as the main influencing factors of greenhouse temperature, then based on model building method of (Error Back Propagation Algorithm Network, BP network), prediction model of greenhouse temperature is established through the network training, and the model is verified by using XOR checkout method, and provide an effective solution for achieving efficient regulation of the greenhouse temperature.

2. Materials and Methods

2.1 Acquisition System of Environmental Information

The temperature monitoring system in greenhouse uses the DS18B20 greenhouse sensor, the single bus characteristics of DS18B20 are used, the system access to six DS18B20 temperature sensors; the monitoring device with same structure is used outside the greenhouse to monitor the temperature outside the greenhouse, meanwhile outdoor increase QY-150B light sensor for monitoring outdoor light intensity. The above-mentioned distributed temperature monitoring system and the monitoring system outside greenhouse all have a clock chip, collect and store data once every 5 minutes synchronously.

2.2 Experimental method

The experiment is made in located in the Venlo modern glass research greenhouse C10 in the South Campus of Northwest A & F University in Yangling Demonstration Zone, Shaanxi Province, it is located at 34.2833°N and 108.0617°E, mean altitude is 500m, the greenhouse is oriented south and north, and east-west length is 8.0m, north-south width is 4.0m, there are 2 steps along the east-west direction, each step is 4.0m wide, greenhouse eaves is 4.0m high, the top is 4.8m, the east of each ridge of greenhouse has a skylight along the north and south, length of skylight is 2.0m , the width is 0.9m, a fan is installed on the outside of east side of the greenhouse, the structure as shown in Fig.1.

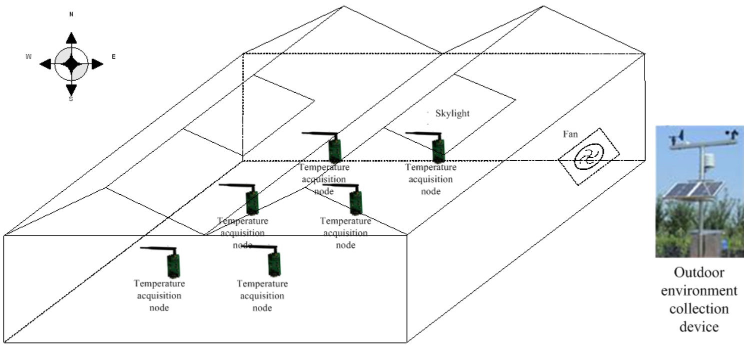


Fig.1 The schematic diagram of greenhouse structure and monitoring node distribution

The experimental environment which meets fully enclosed state, dry ground, no plant and heating, forced cooling was constructed in the above-mentioned greenhouse, and achieve the basic data acquisition of the temperature inside the greenhouse, the temperature outside the greenhouse and the outdoor light intensity three kinds of factors. Its data acquisition uses environmental monitoring systems by independent research and development, including the distributed temperature monitoring system inside greenhouse and temperature outside the greenhouse and light intensity monitoring system.

The temperature monitoring system inside greenhouse uses DS18B20 greenhouse sensors among them, the system access to six DS18B20 temperature sensors, they are deployed in the central position of greenhouse 1 m away from the ground, and get the average value as the temperature inside greenhouse, reduce the measurement error caused by sensor accuracy. The distributions temperature monitoring nodes are shown in Figure.1. In this paper, DS18B20 temperature sensor and QY-150B light sensor of this system are used to achieve information acquisition and storage of temperature and light intensity outside the greenhouse. Based on the above monitoring system, the experiment was made from August

14, 2016 to September 14, 2016, and accumulatively measure 4320 sets of experimental data including indoor temperature, outdoor temperature and light intensity.

2.3 Sampling Data Processing

2.3.1 Eliminate Coarse Data

A total of 12 days of data was selected in the above-mentioned one month in the experiment, there are 1728 groups, and sample set (X, T_a) of the P group was constructed. $X = (X^1, X^2, \dots, X^p, \dots, X^P)$, $T_a = (T_a^1, T_a^2, \dots, T_a^p, \dots, T_a^P)$, $X^p = (X_1^p, X_2^p, X_3^p)$, X_1^p, X_2^p, X_3^p , each parameter represents the temperature outside the greenhouse, the temperature inside the greenhouse, the sunlight intensity outside the greenhouse at time k, respectively, and T_a is the output of each sample set. The coarse errors in the data are eliminated before establishing the model to ensure the accuracy of the measurement data. This experiment uses the 3d criterion [6-7] to eliminate coarse errors. The basic idea is to obey the normal distribution; the absolute values of the error are mainly concentrated around the mean (0), if the residual error $v_i = x_i - \bar{x}$ ($\bar{x}$ is the average value of measurement) of the measured value x_i, then:

$$P(|v_i| > 3d) = 0.027 \quad (1)$$

Therefore, the probability of the data error exceeding 3d is small, and it is considered that x_i is eliminated due to coarse error, d represents the standard deviation, and the calculation formula is as shown in formula (2).

$$d \approx \sqrt{\frac{1}{n-1} \sum v_i^2} \quad (2)$$

This rule is reused until there are no coarse errors in the retained data.

After measured 1728 groups of data eliminate coarse error after 12 days and obtain average value of (for the indoor temperature, because the temperature of 6 points is measured at the same time, 3δ is used to eliminate coarse errors for these 6 points, and the average value as the model indoor temperature value of model input), and calculate the variance of the 1728 data. The calculation formula of variance is shown in formula (3). Through the calculation, the variance values of the three measured inpu value are: the indoor temperature is equal to 0.25, the outdoor temperature is equal to 0.21, and the solar light intensity is 0.37, all of which are not greater than 0.40, respectively.

$$S^2 = \frac{1}{n}[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2] \quad (3)$$

2.3.2 Data Normalization Processing

Because the above processed 3D network input data volume are different, the numerical differences are obvious, according to the characteristics of network, if directly input network, which will become unusually large after weighted by accumulator, it makes the network converge hardly, in order to improve the training and the convergence rate of the network, input data of network need normalization processing [8]. This paper uses linear function transformation, the data in the same dimension are normalized within the interval [0.2 0.9], X and Ta generate input array T and the output array Td after the normalization processing, thus completing the multi-dimensional data sampling set of normalized multi-factor correlation greenhouse temperature modeling, the normalization formula as shown in formula (12) :

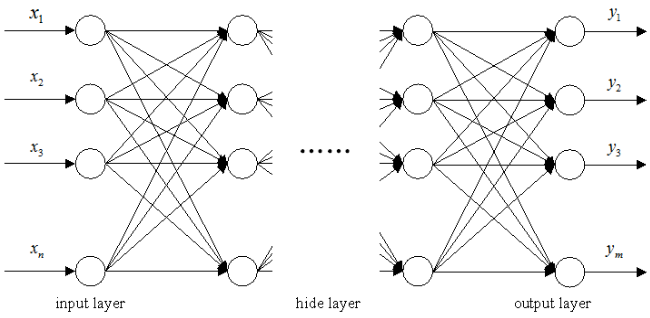
$$t = z_1 + (z_2 - z_1) * (x - x_{min})/(x_{max} - x_{min}) \quad (4)$$

[z1, z2] are the normalized data scope in the formula, z1 = 0.2 in this paper, z2 = 0.9; x is the data to be normalized; xmin, xmax are the minimum and maximum value of the data sequence in the same dimension, respectively.

3. Prediction Model of Greenhouse Temperature Based on BP Network

3.1 BP Neural Network

BP network is one of the most widely used identification models, which is a multi-layer feed-forward network trained by error back propagation algorithm, including input layer, hide layer and output layer. The topology structure of BP neural network model is shown in Figure.2.



Note: $x_i(i = 1,2,3, \ldots, n)$ is BP neural network input factor, i- is the number of input factors: $y_i(i = 1,2,3, \ldots, m)$ —BP is the number of output factors.

Fig.2 The topology structure of BP neural network model

Based on the above structure, the BP network can complete the weight adjustment of multi-layer feed-forward neural network of nonlinear continuous function by the error back propagation algorithm under the premise that mathematical equation describing the mapping relationship is unrevealed, and obtain the square-error and minimum network model, thus learning and storing a large number of input-output mode mapping relationship, which can achieve fitting and prediction of variable nonlinear complex function.

3.2 Structure Design of BP network

The network uses a single hidden layer in this paper, the number of first hidden layer nodes use the formula (5) initially identifies:

$$L = (m + n)/2 + c \tag{5}$$

Among them, m is the number of input nodes, n is the number of output nodes, c is a constant between 1 and 10, and the number of hidden layer nodes is determined to be 8 based on experience by repeated training with trial and error. Therefore, the topology structure of BP network in this paper is 3-8-1. On this basis, the transfer function of network hidden layer neuron adopts the S type tangent function Tansig, the transfer function of output layer adopts the linear function Purelin, and the Trainlm function with the fastest convergence rate is selected as the training function.

Among them, the system randomly sets weight vector V from the input layer to the hidden layer, and the weight vector W from hidden layer to the output layer; the network learning rate is 0.1, the dynamic factor is set to 0.9, and the target error takes 0.0001.

3.3 Network Training Methods

According to the BP neural network established by above structure, input X'=(X1' , X2', …, X5')T , the output signal represents the calculated greenhouse temperature value by network computing, and the corresponding measured greenhouse temperature value of each group is the teacher signal Td. The analysis process of BP gradient training method as shown in Fig. 3, the greenhouse temperature model Td '(X') is established.

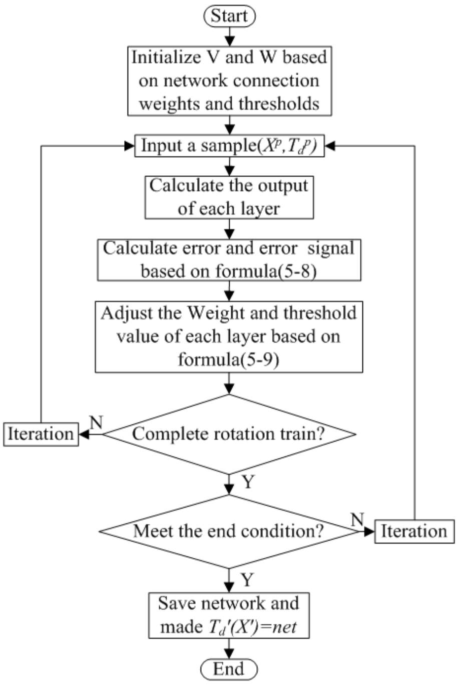


Fig.3 The gradient train procedure of BP neural network

When BP neural network program is running, first of all, a small enough initial value according to the network connection value and threshold set V and W; then input a pair of training set samples (Xp, Tdp) which have been processed; calculate the output y i of each layer, and trigger the following process:

The total system error is calculated based on the teacher signal and the network output signal, as shown in formula (6):

$$E_{RES}=\sqrt{\frac{1}{p}\sum_{p=1}^{p}\sum_{k=1}^{l}(T_{dk}^{p}-T_{ok}^{p})^{2}}\ \text{k=1,2,3,...,l} \tag{6}$$

The error signal in output layer and the neuron error signal are calculated based on the teacher signal, the network output signal, the weight vector from hidden layer to the output layer, and the output component of the hidden layer. The error signal of output layer is shown as formula (7), and the neuron error signal is shown as formula (8)

$$d_k^o=(T_{dk}-T_{ok})o_k(1-T_{ok}) \tag{7}$$

$$d_j^y=\left(\sum_{k=1}^{l}d_k^o w_{jk}\right)y_j\left(1-y_j\right) \tag{8}$$

In the formula, w is weight vector from hidden layer to the output layer; y_j is first j output component of hidden layer.

The weights from the input layer to the hidden layer and the weight s from hidden layer to the output layer based on the input component, the output component of hidden layer, the error signal of output layer, the error signal of neuron and the learning rate. The weight from the input layer to the hidden layer is shown in formula (9), and the weight from hidden layer to the output layer is shown in formula (10):

$$v_{ij}=v_{ij}+hd_j^y x_i \tag{9}$$

$$w_{jk}=w_{jk}+hd_k^o y_j \tag{10}$$

In the formula, v is the weight vector from input layer to the hidden layer; h is the learning rate.

Finally, when one training in rotation is completed for all samples and meets the end condition, and save the trained network.

3.4 Network Training Effect

According to the network structure, transfer function and parameter settings of above nerve, a total of 1296 sets of data (account for 75% of the total sample) are selected in the experimental data to form the training set data in this paper, the pre-processed training set data are input in MATLAB to train the network model of BP artificial neuron for predicting temperature inside greenhouse, the error performance of the BP neural network training is shown in Fig. 4, the model does not show any concussion and local flat areas, and only reaches The desired error level is reached only after six iterations, it shows that the convergence rate of the model is very fast, the model design meets the requirements.

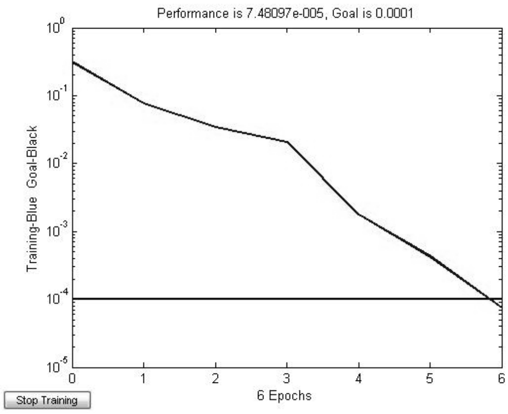


Fig.4 curves of the error variation

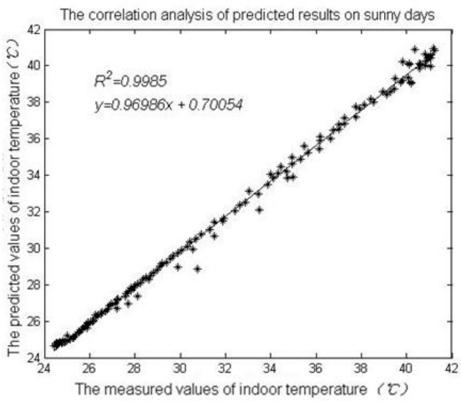


Fig.5 The correlation analysis of predicted results

4. Verification and Discussion

The advantages and disadvantages of BP network performance mainly depends on whether it has good generalization ability, and the experiment of generalization ability should use test data outside the training set to test, the XOR method is used to verify the accuracy of the model in this paper, 25% of the total samples outside the test data are selected to constitute a test set in the test data of this paper, the above-mentioned prediction model of greenhouse temperature based on BP neural network is verified, the correlation analysis between measured and predicted values of temperature inside the greenhouse is shown in Fig. 5, where the horizontal axis is the measured value and the vertical axis is the predicted value.

It can be found from Figure.5 that the determination coefficient of the correlation analysis between the measured and predicted values of greenhouse temperature is 0.9985, the slope of the straight line is 0.96986 and the intercept is 0.70054. The above data shows that the measured value and the predicted value have a good correlation, and the overall similarity is high.

It can be seen from further error analysis that the mean square root error of measured value and predicted values of greenhouse temperature is 0.43 ℃, the maximum error of 1.87 ℃.On the whole, the predicted effect of BP network is very good, which shows that the construction process of the model is highly reliable.

5. Conclusion

The temperature outside the greenhouse, temperature inside the greenhouse and sunlight intensity outside the greenhouse three types of major impact factors are selected in this paper, BP neural network is used to establish prediction model of greenhouse temperature. 432 groups of data are selected to verify, it can be found from the experimental verification that root mean square error of prediction is 0.43 ℃, the maximum error of 1.87 ℃. The prediction results of model for temperature show that the prediction effect of model for greenhouse temperature is good, which conforms to actual requirement; and it provides an efficient solution for greenhouse control requirements which orient actual production.

Acknowledgements

Fund Project: Promotion Program of Scientific and Technological Achievements of Experimental Demonstration Site (Base) of Northwest A&F University (TGZX2016-31)

Xi'an Science and Technology Project (NC1504 (3))

References

[1] Du Shangfeng, Xu Lihong, Ma Chengwei, et al. Research on modeling, simulation and control for controlled environment production systems[J].Scientia Sinica Informationis, 2010, 40: 54-70.(in Chinese with English abstract).

[2] Gieling T H, Janssen H J J, Van Straten G, et al. Identification and simulated control of greenhouse closed water supply systems[J]. Computers and electronics in agriculture, 2000, 26(3): 361-374.

[3] Li Shuhai, Ma Chengwei, Zhang Junfang, et al. Thermal model of multi-span greenhouses with multi-layer covers[J]. Transactions of the Chinese Society of Agricultural Engineering, 2004, 20(3): 217-221. (in Chinese with English abstract).

[4] Xu Lihong, Su Yuanping, Liang Yuming. Requirement and current situation of control-oriented microclimate environmental model in greenhouse system[J]. Transactions of the Chinese Society of Agricultural Engineering, 2013,(19):1-15. (in Chinese with English abstract).

[5] Roy J C, Boulard T, Kittas C, et al. PA—Precision Agriculture: Convective and Ventilation Transfers in Greenhouses, Part 1: the Greenhouse considered as a Perfectly Stirred Tank[J]. Biosystems Engineering, 2002, 83(1): 1-20.

[6] He Shi-biao,Yang Shi-zhong.Applied of 3σ-rule in Reducing Noise in Signal by Wavelet analysis[J]. Journal of Chongqing University,2002,25(12):58-61. (in Chinese with English abstract).

[7] Jiang Jun,Fan Wei-hua,GuoJian,et al.aveket filtering algorithm for gyro signals in moving satellite communication system[J]. Systems Engineering and Electronics,2010,32(4):825-828.(in Chinese with English abstract).

[8] Sun Bai-qing,Pan Qi-shu,Feng Ying-jun,et al.Improment of BP network training rate[J]. Journal of Harbin Institute of Technology,2001,33(4):439-441. (in Chinese with English abstract).