An improved method for prediction of tomato photosynthetic rate based on WSN in greenhouse

Ji Yuhan[1], Jiang Yiqiong[1], Li Ting[2], Zhang Man[1*], Sha Sha[2], Li Minzan[1]

(1. *Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education, China Agricultural University, Beijing 100083, China*; 2. *Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture, China Agricultural University, Beijing 100083, China*)

Abstract: In order to improve the efficiency of CO_2 fertilizer and promote high quality and yield, it is necessary to precisely control CO_2 fertilizer by wireless sensor network based on a model of photosynthetic rate prediction in greenhouse. An experiment was carried out on tomato plants in greenhouse for photosynthetic rate prediction modeling combined rough set and BP neural network. In data acquiring phase, plants growth information and greenhouse environmental information that may have influences on photosynthetic rate, including plant height, stem diameter, the number of leaves and chlorophyll content of functional leaves, air temperature, air humidity, light intensity, CO_2 concentration and soil moisture, which were measured. And LI-6400XT photosynthetic rate instrument was used for obtaining net photosynthetic rate of functional leaf. After preliminary processing, 135 sets of data were obtained. And twelve of them were used for model test of neural network, while the others were used for modeling. All of the data were normalized before modeling. Two models were built to predict photosynthetic rate based on BP neural network. One had total nine input parameters. The other had six input parameters, chlorophyll content, air temperature, air humidity, light intensity, CO_2 concentration, and soil moisture, which were reducted from original nine based on attributes reduction theory of rough set. Both two models have one output parameter, the net photosynthetic rate of single leaf. The genetic algorithm was adopted to reduct attributes. Since continuous data cannot be processed by rough set, the *K*-mean cluster method was used to discretize the data of nine input parameters before attributes reduction. The prediction results of two models showed that the model with six input parameters had a mean absolute error of 0.6958, an average relative error of 7.28%, a root-mean-square error of 0.7428, and a correlation coefficient of 0.9964, while the other model respectively had 0.4026, 4.53%, 0.3245 and 0.9965, which proved that the model with minimum attributes had higher prediction accuracy. On the other hand, the number of iterations was used to represent the neural network train speed. The result showed that the model with six input parameters had an iteration of 544, while the other had 1038. Hence, the reduction model was applied to controlling CO_2 concentration. The net photosynthetic rates at different CO_2 concentrations were predicted at a certain condition. The results had the same curve trend with theory analysis, and a high prediction accuracy, which proved that the model was useful for CO_2 concentration control.

Keywords: tomato, photosynthetic rate, wireless sensor network, greenhouse, rough set, BP neural network
DOI: 10.3965/j.ijabe.20160901.1243

Citation: Ji Y H, Jiang Y Q, Li T, Zhang M, Sha Sh, Li M Z. An improved method for prediction of tomato photosynthetic rate based on WSN in greenhouse. Int J Agric & Biol Eng, 2016; 9(1): 146－152.

1 Introduction

Wireless sensor network (WSN), as a new technology of information acquisition, has been widely used in agriculture, especially in greenhouse[1-3]. Greenhouse environment is convenient to control for providing better

Received data: 2014-10-25 **Accepted data:** 2015-11-20

Biographies: Ji Yuhan, Master, Major in agricultural informatization, Email: jyh1009845844@163.com; **Jiang Yiqiong**, Master, Major in agricultural informatization, Email: 550931538@qq.com; **Li Ting**, Master, Major in agricultural informatization, Email: 15201423238@163.com; **Sha Sha**, Master, Major in agricultural informatization, Email: shashaok2010@163.com; **Li Minzan**, PhD, Professor, Major in precision agriculture, Email: limz@cau.edu.cn

***Corresponding author: Zhang Man,** Associate Professor, Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education, China Agricultural University, Beijing 100083, China. Tel: +86-10-62737914, Email: cauzm@cau.edu.cn

crop growth conditions. Applying WSN in greenhouse is to benefit production management, thus to promote high quality and yields[4]. The large amounts of data acquired by WSN can be used for further suggestion and decision of production management.

Carbon dioxide (CO_2) is one of the most important raw materials for plant growth. Research results show that an appropriate supply of CO_2 air fertilizer will improve crop yields and quality[5-7]. However, crop growth information and greenhouse environmental information will influence the photosynthetic rates of crop, and further influence the requirements of CO_2 fertilizer[8]. Therefore, how to determine the optimal CO_2 demand has become a key of precision fertilization[9,10]. Many crop growth models have been studied and applied, and photosynthetic rate prediction models have been developed. However, most of these models were based on mechanism of photosynthesis, the parameters of the models could not be measured automatically[11-14]. Wang et al. used a WSN system to collect environmental information automatically, and extracted the information for photosynthetic rate. However, the crop growth information was not considered[15,16].

The objective of this research was to build a single leaf net photosynthetic rate prediction model on tomatoes, on the purpose of CO_2 concentration control in greenhouse. In this study, the greenhouse environmental parameters were automatically collected by WSN, the plants growth information was measured manually during the whole tomato growth stage, both of this information data were considered as the input parameters of prediction model.

2 Materials and methods

2.1 Experimental design

The system architecture is shown in Figure 1. The source of data consists of environmental data collected by WSN, crop growth data measured manually, and photosynthetic rate data measured by LI-6400XT device. Based on the above information, the photosynthetic rate prediction model and CO_2 requirement model are built. According to these models, CO_2 concentration can be adjusted in the greenhouse through controlling the executive unit.

The experiment was conducted in an experimental greenhouse of the College of Water Conservancy and Civil Engineering, China Agricultural University. The experiments were carried out from 12 July to 16 August, 2013, and tomato was in the seeding stage during this period. Tomatoes were planted under 3 different soil moisture conditions, including 35%-45%, 55%-65% and 75%-85%. The soil moisture was controlled by weighing and watering quantificationally. The environmental parameters were automatically collected by WSN every 30 min, including air temperature and humidity, CO_2 concentration and light intensity. The crop growth information was measured manually once a week, including plant height, stem diameter, the number of leaves and chlorophyll content of functional leaf. Among which, chlorophyll content was measured by SPAD502 chlorophyll meter. In order to reduce measurement errors, the value of chlorophyll was averaged by several measurements. Function leaf (the third phyllotaxy from top to bottom) of tomatoes was selected as the object leaf during the experiment. Function leaf has a special position, excellent morphology and physiological function in plants, and plays an important role in the course of crop yield formation.

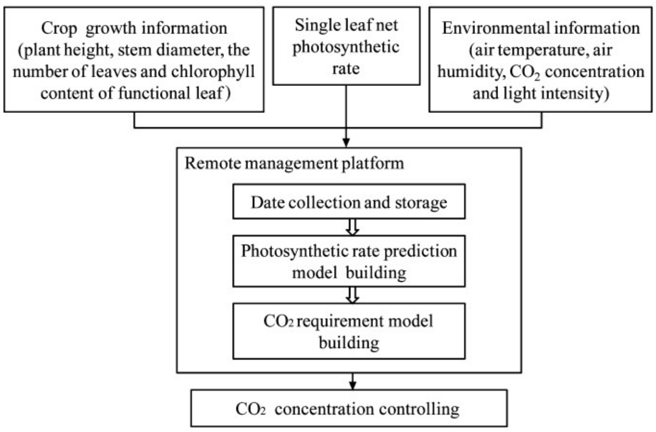

Figure 1 Architecture of CO_2 control system

Experiments about photosynthetic rate were carried out on 2, 9 and 16 August when the weather was sunny. During the experiment, the single leaf net photosynthetic rate was measured by the LI-6400XT. In order to obtain a wide range of experimental data, the Red/Blue LED

Light Source and CO_2 injection system were used to artificially control the leaf chamber temperature, humidity, light intensity, and CO_2 concentration. The whole experiments included three steps. First of all, a set of light-response curves were measured in order to find the light saturation point of photosynthetic rate, and a set of CO_2 concentration response curves under the saturation light intensity were measured to estimate the saturation CO_2 concentration. Secondly, a set of temperatures response curves were measured at the near saturation light intensity and CO_2 concentration, in which the measured temperature was controlled in plus or minus six degrees environment temperature. A set of humidity response curves were measured at saturation light intensity, CO_2 concentration and optimum temperature, in which gas flow rate was used to control the humidity. Four groups of curves and a total of 30 groups of data were acquired in this step. Finally, three healthy plants were selected from different groups of soil moisture to measure single leaf net photosynthetic rate with controlled light intensities and CO_2 concentrations. In this process, light intensity was set at 300 μmol/(m^2·s), 600 μmol/(m^2·s), 900 μmol/(m^2·s), respectively. The CO_2 concentration of greenhouse air is about 400 μmol/mol, therefore, the CO_2 concentration range were set from 400 μmol/mol, to slightly beyond the position of CO_2 concentration saturation point, at an interval of 200 μmol/mol. A total of 105 sets of data were obtained. Besides, in order to test the model, a CO_2 response curve was measured at a certain condition, which obtained eight sets of data. Based on the above experiments, 143 sets of data were obtained, which included the growth information and the photosynthetic rate data of nine plants under different environment. Among which, eight sets of data were used for the CO_2 response curve forecast.

2.2 Data analysis methods

2.2.1 The attributes reduction theory of rough set

Rough set is a symbol method that analyzes the correlation and dependence between data[17]. It can simplify data and reduce redundancies, improve the speed of neural network training and then optimize the neural network model.

Rough set can be regarded as an information system, which consists of a set of objects. The objects are represented by a set of attributes, including condition attributes and decision attributes. Attributes reduction is to remove the unnecessary condition attributes and obtain the minimum condition attributes, at the same time, maintain the same classification capability of the information system[18]. In the information system of tomato photosynthetic rates, single leaf net photosynthetic rate is the decision attribute and the rest of the parameters are the condition attributes.

2.2.2 BP neural network modeling

Artificial neural network is a non-linear statistical data modeling tool with high nonlinear mapping capability. It is mainly used in pattern recognition and function approximation, etc. Back-propagation neural network (BPNN) is the most common kind of neural networks[19]. The drawback of BPNN is difficult to interpret the result, so a good assessment is required [20-21]. In order to effectively evaluate the prediction performance of BPNN, several evaluation parameters were used in this paper, including R (Relative coefficient), ARE (Average relative error), MAE (Mean absolute error) and RMSE (Root-mean-square error).

Figure 2 is the flow process of building photosynthetic rate prediction model.

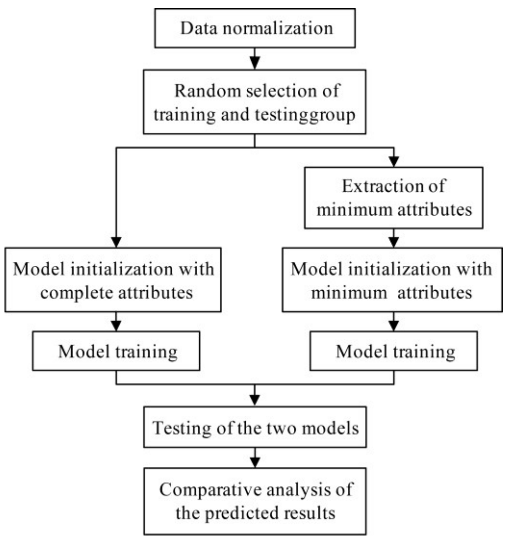

Figure 2 The modeling chart of photosynthetic rate prediction

Firstly, the data were normalized to make the index of each input parameter in the same dimension. Secondly, the normalized data were randomly divided into training group and testing group. Thirdly, according to the results of rough set reduction, the data with minimum condition attributes were extracted from both training and testing group, and then the neural network model was developed. In order to evaluate the performance of the developed models, the model based on the data with total condition attributes was also built. Fourthly, the two models were tested by testing group data respectively. Finally, the training accuracy and training speed of the two models were analyzed and compared by the evaluation parameters.

3 Results and analysis

The environmental parameters in greenhouse and plant growth parameters were used as the input parameters of the photosynthetic rate prediction model. And the single leaf net photosynthetic rate of tomato was used as the output parameter. Before building the model, the data collected by WSN and the photosynthetic rate device needed to be matched firstly.

3.1 Attributes reduction

Among the nine condition attributes, only the data of soil moisture was discrete (three levels of soil moisture). The rest of attributes were continuous data. The continuous data needed to be discretized before attributes reduction by rough set. In this paper, K-means clustering method was adopted to discretize the data. K-means clustering method takes distance as the evaluation index of similarity. In other words, the closer the two objects are, the more similar they are. The specific implementation procedures were listed as follows. Firstly, the cluster centers were initialized according to the specified number of clusters K. Secondly, for each object, it decided which class to belong to, and for each class, it recalculated its cluster center. Finally, repeated the second step until convergence. The SPSS software was used to cluster the continuous input data of each attribute. The output cluster centers of each attribute were shown in Table 1. All objects of each attribution were represented by the cluster number, therefore, a discrete information system was formed and could be processed by rough set.

Genetic algorithm is one of the common methods of attributes reduction in rough set. It initializes population by initializing binary codes with a fixed length. The length of the binary code is as same as the number of condition attributes. Every bit in binary codes has two values. Value “1” or “0” represented “selected” or “not selected” state of the corresponding condition attribute respectively. The individual fitness is evaluated by the less number of condition attributes and the higher resolution. The operations of selecting, intersecting and variation are executed to get the best individual until final convergence. The best individual represents the minimum attributes. In this paper, the genetic algorithm in Rosetta software was used to reduct condition attributes of discrete information system. From the reduction result, the minimum condition attributes had six attributes, including chlorophyll content, soil moisture, air temperature, air humidity, light intensity and CO_2 concentration. Both the data with minimum attributes and complete attributes were used to establish prediction models of photosynthetic rate respectively.

Table 1 Cluster centers of the attributes

Attribute	Cluster centers
Plant height/cm	23.8, 19.2, 13.5
Stem diameter/cm	2.36, 3.02, 3.57
The number of leaves	6.8, 8.5, 9.5
Chlorophyll content	44.4, 49.4, 53.6
Air temperature/°C	31.87, 35.97, 41.46, 45.87
Air humidity/%	25.41, 30.93, 44.37, 57.47
Light intensity/μmol·m^{-2}·s^{-1}	1200.55, 906.29, 599.99, 299.98, 56.42
CO_2 concentration/μmol·mol^{-1}	127.32, 392.99, 595.34, 800.11, 1000.31, 1199.99, 1399.43
Photosynthetic rate/μmol·CO_2 m^{-2}·s^{-1}	−0.84, 1.44, 4.10, 6.07, 8.54, 11.14, 13.59, 16.01, 18.91, 21.37

3.2 BPNN model building

The Matlab software was used to establish BPNN model. Data were randomly divided into training group and testing group. The training group was used to train the neural network, and the testing group was used to verify the effectiveness of the models. Among the 135 sets data obtained from the experiment, twelve sets were randomly selected as the testing group, and the rest 123 sets were used as the training data.

In this study, a three-layer neural network was used to build the prediction model, including input layer, hidden layer and output layer. Logsig and Purelin functions were used as transfer functions, wherein Logsig was used for the input layer to the hidden layer, and Purelin for the hidden layer to the output layer. Bayesian normalization method was used as the learning method, and Trainbr was used as the analytic function. The number of nodes in the hidden layer was determined according to the empirical Equation (1).

$$n_h = \sqrt{n_i + n_o} + l \tag{1}$$

where, n_h, n_i, and n_o represent the number of hidden layer neurons, input layer neurons and output layer neurons respectively; and l is a constant with value fluctuated from 1 to 10. After several trials, the number of hidden layer nodes was finally set to 15 in the model with the complete attributes and to 12 in the model with the minimum attributes.

3.2.1 Photosynthetic rate prediction

After the models were built, twelve sets of test data randomly selected were used to evaluate the performance of the BPNN models. Figure 3 shows the predicted value and observed value of 12 sets data. The performance comparison of two models was analyzed in Table 2.

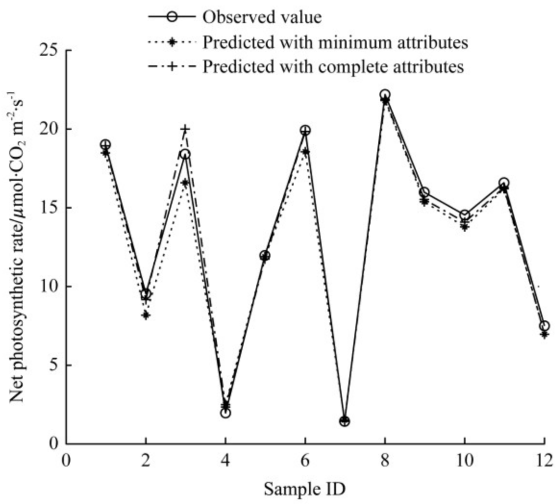

Figure 3 The prediction results of BPNN

Table 2 Performance comparison of two models

Prediction samples ID	Observed value/ μmol $CO_2 \cdot m^{-2} \cdot s^{-1}$	Complete attributes prediction			Minimum attributes prediction		
		Predicted value	Absolute error	Relative error/%	Predicted value	Absolute error	Relative error/%
1	18.9719	18.4711	−0.5008	2.6396	18.9326	−0.0393	0.2071
2	9.5274	8.1729	−1.3545	14.2172	9.1337	−0.3937	4.1325
3	18.3465	16.5671	−1.7794	9.6988	19.9629	1.6163	8.8099
4	1.9143	2.4533	0.5390	28.1562	2.3372	0.4230	22.0951
5	11.9241	11.8578	−0.0663	0.5562	11.6992	−0.2249	1.8865
6	19.9079	18.5598	−1.3481	6.7716	19.8318	−0.0761	0.3822
7	1.4210	1.4940	0.0729	5.1324	1.4279	0.0069	0.4852
8	22.1368	21.7687	−0.3681	1.6629	21.8516	−0.2852	1.2883
9	15.9336	15.3247	−0.6089	3.8214	15.4653	−0.4683	2.9389
10	14.4928	13.7250	−0.7678	5.2975	14.0242	−0.4686	3.2333
11	16.5767	16.1446	−0.4321	2.6066	16.2756	−0.3011	1.8164
12	7.4770	6.9656	−0.5114	6.8400	6.9492	−0.5278	7.0595
MAE			0.6958			0.4026	
ARE/%			7.28			4.53	
RMSE			0.7428			0.3245	
R			0.9964			0.9965	
The number of iterations			1038			544	

From Table 2, the prediction model established by the minimum attributes group was better than the complete attributes group in absolute and relative errors. According to the model with minimum attributes, the R, MAE, ARE and RMSE, were 0.9965, 0.4026, 4.53%, and 0.3245 respectively. The results of evaluation indices showed the model with minimum attributes had a more accurate prediction and a faster training speed. The above results prove that the BPNN is an effective method to predict the net photosynthetic rates. And the rough set attributes reduction can optimize the BPNN model.

3.2.2 CO_2-photosynthetic rate prediction

The real time environmental information in greenhouse and the plants growth information could be used to predict the photosynthetic rates and find the optimal CO_2 concentration with the established BP neural network model. Under a certain condition, the maximum photosynthesis rate could be found by prediction of CO_2-photosynthetic rate curve. The CO_2 concentration which corresponds to the maximum photosynthesis rate is the optimal level of CO_2 concentration. Figure 4 shows the relationship between

CO_2 concentration and photosynthetic rate according to the prediction model with minimum attributes. The other environmental factors included soil moisture set to 75%-85%, air temperature set to 38°C, air humidity set to 41.5%, light intensity set to 900 μmol/(m^2·s), and chlorophyll content set to 44.4. The prediction curve was similar to the observation curve with small errors, which proved the model could use for predicting net photosynthetic rate of single leaf.

In addition, Figure 4 demonstrates that the maximum rate of photosynthesis was obtained at the CO_2 concentration of 800 μmol/mol. Under this condition, through controlling the switch of CO_2 air source, application of the appropriate amount of CO_2 concentration could make the plant grow better.

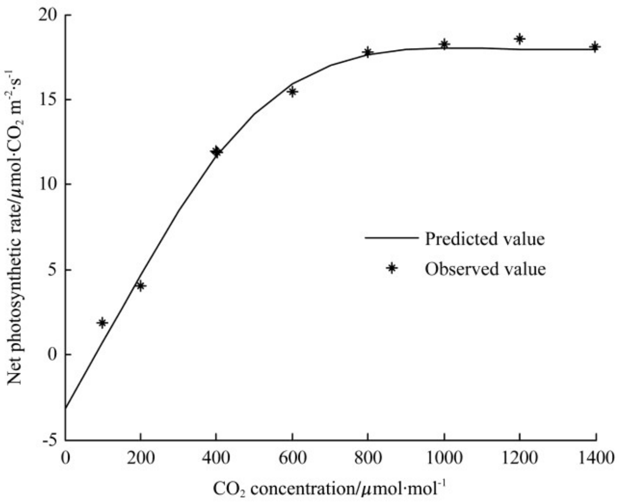

Figure 4 The prediction of photosynthetic rates at different CO_2 concentrations

4 Conclusions

In order to achieve the precision control of CO_2 fertilizer, a photosynthetic rate prediction method was designed. Experiments about the net photosynthetic rate of single leaf were carried out on tomatoes at the seeding stage. The experiments extracted nine attributes, including greenhouse environment information and plants growth information, as the input parameters of photosynthetic rate prediction model, and the single leaf net photosynthetic rate of tomatoes as output parameter. The model was established based on a BP neural network. In order to further improve the prediction accuracy and efficiency of the model, the nine input parameters were reducted by rough set. The reduction results contained six input parameters, which were used to build another model of BP neural network. The comparative results of the two models showed that the model with six input parameters was faster in training speed and better in prediction accuracy. Finally, the model with six input parameters was used to predict the photosynthetic rates with varying CO_2 concentrations under a certain environmental condition. Therefore, the optimal CO_2 concentration was obtained under this condition. The prediction results showed that the BPNN model was effective on photosynthetic rate prediction. In order to further verify the building model, the experiments should be implemented at other growing stage of tomato plants. The values of plants growth parameters should be extended in relative wide ranges. And the controlled range of air temperature and humidity should also be extended. Therefore, more experiments would be carried out in the further.

Acknowledgements

This work was supported by the National Natural Science Fund (Grant No. 31271619) and the Doctoral Program of Higher Education of China (Grant No. 20110008130006).

[References]

[1] Li L, Li H X, Liu H. Greenhouse environment monitoring system based on wireless sensor network. Transactions of the CSAM 2009; 9(40): 228−231. (in Chinese with English abstract)

[2] Guo W C, Chen H J, Li R M, Liu J, Zhang H H. Greenhouse monitoring system based on wireless sensor networks. Transactions of the CSAM, 2010; 41(7): 181−185. (in Chinese with English abstract)

[3] Li Y H, Ji G F, Han J Y. Application of the wireless sensor network in environment monitoring system of greenhouse. Instrument and Meter for Automation, 2010; 31(10): 61−64. (in Chinese with English abstract)

[4] Gonda L, Cugnasca C E. A proposal of greenhouse control using wireless sensor networks. ASABE Publication Number: 701P0606, Florida, USA, 2006.

[5] Nederhoff E M. Effects of CO_2 concentration on photosynthesis, transpiration and production of greenhouse fruit vegetable crops. Doctoral thesis. Agricultural University, Wageningen, Netherlands, 1994.

[6] Vance P, Spalding M H. Growth, photosynthesis, and gene expression in chlamydomonas over a range of CO_2 concentrations and CO_2/O_2 ratios: CO_2 regulates multiple acclimation states1. Canadian Journal of Botany, 2005; 83(7): 796−809.

[7] Chen S C, Zou Z R, He C X, Zhang Z B, Yang X. Rules of CO_2 concentration change under organic soil cultivation and effects of CO_2 application on tomato plants in solar greenhouse. Acta Bot. Boreal.-Occident. Sin, 2004; 24(9): 1624−1629.

[8] Hou J L. Study on model to greenhouse tomato growth and development. Doctoral thesis. Beijing: China Agricultural University, 2005. (in Chinese with English abstract)

[9] Li T L, Yan A D, Luo X L, Qiu J Q, Li D, Yao Z K. Temperature modified model for single-leaf net photosynthetic rate of greenhouse tomato. Transactions of the CSAE, 2010; 26(9): 274−279. (in Chinese with English abstract)

[10] Luo X L, Li T L, Li G C, Liu Z Y, Diao J, Gu J G. Relations between leaf net photosynthetic rate of greenhouse tomatoes and meteorological factors. Jiangsu Agricultural Sciences, 2007; 4: 89−92. (in Chinese with English abstract)

[11] El-Sharkawy M A. Overview: Early history of crop growth and photosynthesis modeling. BioSystems, 2011; 103(2): 205−211.

[12] Gary C, Jones J W, Tchamitchian M. Crop modeling in horticulture: state of the art. Scientia Horticulture, 1998; 74: 3−20.

[13] Jones J W, Dayan E, Allen L H, Keulen H. A dynamic tomato growth and yield model (TOMGRO). Transaction of ASAE, 1991; 34(2): 663−672.

[14] Zhang J, Wang S X. Simulation of the canopy photosynthesis model of greenhouse tomato. Procedia Engineering, 2011; 16: 632−639.

[15] Wang W Z, Zhang M, Jiang Y Q, Sha S, Li M Z. Photosynthetic rate prediction of tomato plants based on wireless sensor network in greenhouse. Transactions of the CSAM, 2013; 40(Supp 2): 192−197. (in Chinese with English abstract)

[16] Wang W Z, Zhang M, Liu C H, Li M Z, Liu G. Real-time monitoring of environmental information and modeling of the photosynthetic rate of tomato plants under greenhouse conditions. Applied Engineering in Agriculture, 2013; 29(5): 783−792.

[17] Li M, Zhang H G. Research on the method of neural network modeling based on rough sets theory. Acta Automatica Sinica, 2002; 28(1): 1−7. (in Chinese with English abstract)

[18] Yao H X, Huang Z L, Liu Z G. Comparison of two methods used to data preprocessing for one kind of neural network. Microcomputer Development, 2003; 13(6): 77−79. (in Chinese with English abstract)

[19] Han L, Li R, Zhu H L. Comprehensive evaluation model of soil nutrient based on BP neural network. Transactions of the CSAM, 2011; 42(7): 109−115. (in Chinese with English abstract)

[20] Ehret D L, Hill B D, Raworth D A, Estergaard B. Artificial neural network modeling to predict cuticle cracking in greenhouse peppers and tomatoes. Computers and Electronics in Agriculture, 2008; 61(2): 108−116.

[21] Xiang M J. Optimization and regulation of greenhouse environmental factors based on information fusion technology. Doctoral thesis. Jiangsu University, China, 2009. (in Chinese)