Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 51-17 (2018) 790–795

Recurrent Neural Network Model for Prediction of Microclimate in Solar Greenhouse

Wang Hongkang*. Li Li*. Wu Yong*. Meng Fanjia*. Wang Haihua. N.A. Sigrimis*****

** Key Laboratory on Modern Precision Agriculture System Integration Research of Ministry of Education, China Agriculture University, Beijing 100083, China.*

*** Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture, China Agriculture University, Beijing 100083, China.*

**** Department of Natural Resources Management & Agricultural Engineering, Agricultural University of Athens, Athens 11855, Greece.*

Abstract: A recurrent neural network(RNN) based dynamic back propagation(BP) algorithm model with historical internal inputs are developed to predict the temperature and humidity of a solar greenhouse in the north of China. Climate data including air and substrate temperature、air humidity、illumination and CO_2 concentration recorded over eight days were used to build and validate models for climatic prediction. In order to compare the accuracy of predictions, different performance measures, such as average relative error (ARE), mean absolute error (AME) and root mean square error (RSME), were calculated, for using BP and untrained RNN neural networks using the same processing. For the RNN model, a context layer as the last hidden layer output , is input with the next input to the next hidden layer, which is equivalent to the state feedback. The results demonstrate that the RNN-BP model provides reasonably good predictions with the RSME for temperature 0.751 and 0.781 for humidity, including both of the $\mathbf{R}^2$ is both above 0.9, which outperforms the compared models tested in this paper.

Keywords: Recurrent Neural Network, Dynamic BP Algorithm, Prediction Model, Microclimate, Solar Greenhouse.

1. INTRODUCTION

The greenhouse microclimate is mainly affected by external radiation, temperature, humidity, and carbon dioxide. It has a close relationship with greenhouse production and crops, mainly manifested as the exchange of substances and energy between them. Thereby the further development of greenhouse requires better technologies and tools to process data at a reasonable cost and to translate the data into better decisions and actions(Tan,2016). Therefore, it is important to correctly predict greenhouse microclimate for greenhouse control and crop management. As a microclimate of agriculture, the characteristics of diurnal and seasonal changes of climate elements are affected not only by the daily and annual changes of the environmental meteorological elements but also by the influence of changes in the community of agricultural biological life. Sometimes it is also affected by the impact of surrounding facilities、 and other management factors. When the composition of greenhouse components and the weather background of the system are determined, the system's unique agricultural microclimate characteristics will be relatively stable, which is conducive to our prediction.

Because of the complex factors, large delay, strong coupling, slow time variation and non-linearity of the greenhouse microclimate, it is difficult to establish an accurate mathematical model. The current greenhouse microclimate modelling mainly has the following approaches: mechanism by design modelling method, computational fluid dynamics (CFD) model method and system identification method. The mechanism modelling method is based on the laws of physics and physiological principles. Vanthoor et al. developed and verified greenhouse climate models by designing greenhouses for various climate and economic conditions for a model-based approach. The model describes and calculates the impact of outdoor climate and greenhouse structure on indoor greenhouse climate (Vanthoor et al.,2011). However it has many limitations: the mechanism model has many unknown parameters, the measurement of which requires expensive instruments, and the test cost is high; Furthermore the measurement of some parameters takes time and effort, and is not easily mastered by the majority of farmers and the portability and the adaptability is poor. So more research focus on parameters calibration (Guzmán-Cruz et al.,2009; Hasni et al.,2011). CFD technology has powerful functions and accuracy for simulating the changes of airflow and temperature fields in the greenhouse. In the simulation process, the dynamic process of the temperature of each component of the greenhouse with time and space is taken into account (Zhou et al.,2014; Zhang et al.,2016; Hou,2016). However, its emphasis on spatial distribution and the same need for many greenhouse structural components, makes the method less applicable. System identification is a data model

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2018.08.099

based on input and output data and using internal changes as a black box. A lot of models based on regression and neural network training have been built (Patil et al.,2008;Lópezcruz et al.2010;Mustafaraj et al.2011;Fourati,2014;Taki et al.2016; Zou et al.2017). The identification of a parametric model is usually simple in structure and easy to analyse. With the popularity of the Internet of Things and cloud services, a large amount of environmental data can be saved and accessed, and the model accuracy is enhanced.

Considering the fact that effects of outdoor climate factors can be extracted from historical data but cannot show the interference caused by greenhouse management and human activities, the indoor environment data is also used as input. The impact of irrigation on the environment is also taken into consideration to add soil temperature (Zuo et al.,2010). Because of the strong coupling between temperature and humidity, both temperature and humidity are regarded as outputs (Ramos-Fernández et al.,2016). On this basis, a recurrent neural network(RNN) is used for training and identification, and parameters are optimized through a dynamic back propagation(BP) algorithm to realize the dynamic prediction of temperature and humidity in the greenhouse.

2. MATERIALS AND METHODS

2.1 Data Acquisition

The solar greenhouse is located in the teaching and testing ground of the China Agricultural University of Zhuozhou City, Hebei Province (N115.86, E39.47). The crops are strawberries and closed cultivation. In order to collect the data required by the prediction algorithm, the wireless data acquisition gateway KL-H1100 is used to acquire data through the 3G network card, and is equipped with JZH series sensors for measurement. The Illumination range is 0 to 80000 lux, the accuracy is ±5% FS, and the air temperature range is -20-60 °C, accuracy ±0.5°C, air humidity range 0%-100%RH, accuracy ±3%RH, CO_2 concentration range 0-5000ppm, accuracy 40ppm,soil temperature range -20-60°C, accuracy ±0.5°C. Sensors are placed centrally or hung in the center of the greenhouse, with a frequency of 10 minutes. For training and validation, select greenhouse climate data from March 8th to March 15th 2017. The first 6 days are used for training and the latter two days are used for verification. Schematic diagram of data acquisition is shown in Fig 1.

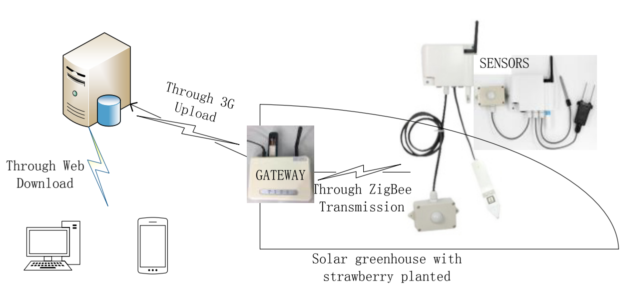


Fig. 1. Schematic diagram of data acquisition.

2.2 Recurrent Neural Network

The Elman network is a kind of recurrent neural network. It has the input layer, output layer and hidden layer similar to the standard feed-forward neural network. The difference is that there is a context layer, which is used to save the output state of the hidden layer at the current moment, to represent the historical characteristics of the object, and together with the input of the next moment as the input of the hidden layer, which is equivalent to the state feedback. The network structure is shown in Fig. 2.

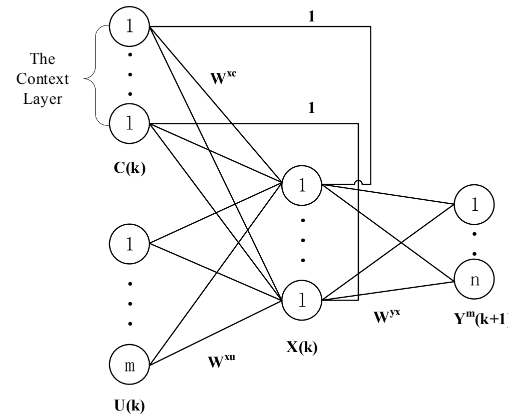


Fig. 2. Elman neural network structure.

where $\mathrm{U}(k)$ is the input vector, $\mathrm{Y}^m(k+1)$ is the output vector, W^{xu} is the weight vector between the hidden and the input layers, W^{yx} is the weight vector between the output and the hidden layers and W^{xc} is the weight vector between the hidden and the context layers. The activation function of the hidden units is a sigmoid one.

From the network structure diagram, the relationship between the input and output of the Elman network is:

$$Y^m(k+1) = W^{yx}(k) * X(k)$$

$$= W^{yx}(k) * F[W^{cx}(k)C(k) + W^{xu}(k)U(k)] \quad (1)$$

where F is the activation function of the hidden units.

In our case,$\mathrm{U}(k) = [\mathrm{T}(k), H(k), I(k), CO_2(k), T_s(k)]^T$ and $Y^m(k+1) = [\mathrm{T}^m(k+1), H^m(k+1)]^T$.

where T is greenhouse air temperature(°C), H is greenhouse air relative humidity(%),I is illumination (Lx) in greenhouse, CO_2 is greenhouse CO_2 atmospheric concentration (ppm) and T_s is substrate temperature(°C).

The learning process of the Elman neural network is the same like other neural network. In the learning step, the adaptation of the connection weight vectors is performed in the dynamic BP algorithm. W^{xu} and W^{yx} different with W^{xc} is not affected by feedback loop, so adopting the dynamic BP algorithm is similar to using the standard BP algorithm.

For ease of narrative, we assume that the network has only one output, and the error function is

$$E = \frac{1}{2}\left(y^m(k) - y(k)\right)^2 \tag{2}$$

where $y^m(k)$ is the neural network output, $y(k)$ is the realistic output of greenhouse. Take W^{xc} as an example, a certain weight w_{ij}^{xc} for the context layer to the hidden layer is expressed as follows:

$$\frac{\partial E}{\partial w_{ij}^{xc}(k-1)} = -\frac{\partial E}{\partial y^m(k)}\frac{\partial y^m(k)}{\partial x_j(k-1)}\frac{\partial x_j(k-1)}{\partial w_{ij}^{xc}(k-1)}$$

$$= -\left(y^m(k) - y(k)\right)w_j^{yx}(k-1)\frac{\partial x_j(k-1)}{\partial w_{ij}^{xc}(k-1)} \tag{3}$$

Known from the network structure,

$$C(k) = X(k-1)$$

$$= f\left(W^{xc}(k-1)C(k-1) + W^{xu}(k-1)U(k-1)\right) \tag{4}$$

and

$$C(k-1) = X(k-2) \tag{5}$$

equation (4) can be further expanded, which shows that $C(k)$ is a dynamic recursive process based on the link weights at different moments in the past.

$$\frac{\partial x_j(k-1)}{\partial w_{ij}^{xc}(k-1)} = f_j'\left[x_i(k-2) + \sum_{t=1}^{l} w_{it}^{xc}(k-1)\frac{\partial x_t(k-2)}{\partial w_{ij}^{xc}(k-1)}\right] \tag{6}$$

When the weight change for each iteration is small, equation (6) can be written in the following recursive form:

$$\frac{\partial x_j(k-1)}{\partial w_{ij}^{xc}(k-1)} = f_j'\left[x_i(k-2) + \sum_{t=1}^{l} w_{it}^{xc}(k-1)\frac{\partial x_t(k-2)}{\partial w_{ij}^{xc}(k-2)}\right] \tag{7}$$

If the second item in (7) is ignored, that is, the item due to past state feedback is ignored, the dynamic BP algorithm degenerates to the standard BP algorithm.

$$\frac{\partial x_j(k-1)}{\partial w_{ij}^{xc}(k-1)} = f_j' x_i(k-2) \tag{8}$$

3. RESULTS AND DISCUSSION

In order to effectively evaluate the predictive performance of the model, the evaluation indicators used in this paper include average relative error (ARE), mean absolute error (AME), root mean square error (RSME) and coefficient of determination (R^2). The evaluation index calculation formula is as follows:

$$ARE = \frac{1}{N}\sum_{i=1}^{N}\frac{|Y(i) - Y^m(i)|}{|Y(i)|} \tag{9}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|Y(i) - Y^m(i)| \tag{10}$$

$$RMSE = \sqrt{\frac{\sqrt{\sum_{i=1}^{N}\left(Y(i) - Y^m(i)\right)^2}}{N}} \tag{11}$$

$$R^2 = \left(\frac{\sum_{i=1}^{N}(Y(i)-\bar{Y})(Y^m(i)-\overline{Y^m})}{\sqrt{\sum_{i=1}^{N}(Y(i)-\bar{Y})^2}\sqrt{\sum_{i=1}^{N}\left(Y^m(i)-\overline{Y^m}\right)^2}}\right)^2 \tag{12}$$

3.1 Modelling results

First evaluate the model based on training set data to ensure that the model error is within the appropriate range. Fig.3 shows the training set of modelling outputs and Tabel.1 shows the evaluation indicators.

For the temperature model, the average absolute error MAE is 0.488 and the average relative error ARE is 1.7%. In other words, the prediction accuracy is ±0.48℃, and the accuracy of the sensor is considered to be ±0.5℃. The error with the actual temperature can be guaranteed within ±1℃. It is sufficient for the solar greenhouse. The root mean square error RMSE is 0.865 and the coefficient of determination R^2 is 0.953, indicating that the overall dispersion is small and does not cause a large error with the actual temperature trend.

For the humidity model, MAE is 0.473 and ARE is 2.9%. The prediction accuracy is 0.47%, and the sensor accuracy is 3%, which is about 6 times the model error, so it can be used for humidity prediction. The root mean square error RMSE is 0.699 less than the temperature and the coefficient of determination R^2 is 0.953 better than the temperature. So the model error is within the allowable range, has reliability and can be used for prediction.

Compared with the algorithm proposed in this paper, RNN neural network will produce relatively large errors without training, and its MAE、RMSE and also R^2 are greater than those after training. The BP neural network has the largest error under the same conditions. This shows that considering historical data has a positive influence on the training of the model.

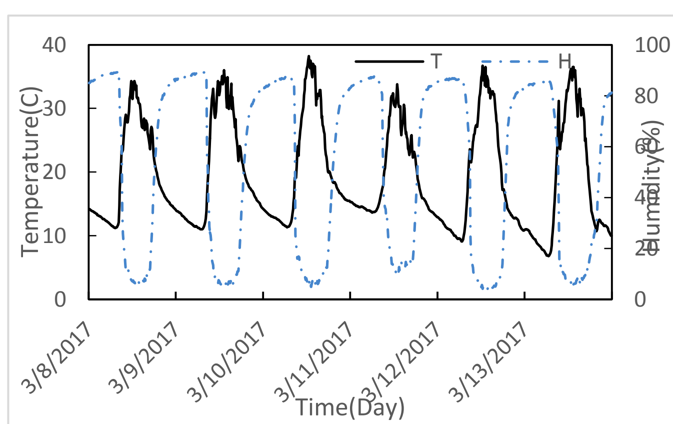


Fig. 3. Model output training set.

Table 1. Modelling evaluation indicators

Modeling	Temperature				Humidity			
	ARE	MAE	RSME	R^2	ARE	MAE	RSME	R^2
BP	0.071	0.842	1.504	0.924	0.058	1.134	1.562	0.914
RNN	0.039	0.628	1.167	0.938	0.043	0.817	1.056	0.936
BP-RNN	0.016	0.421	0.751	0.953	0.025	0.567	0.781	0.962

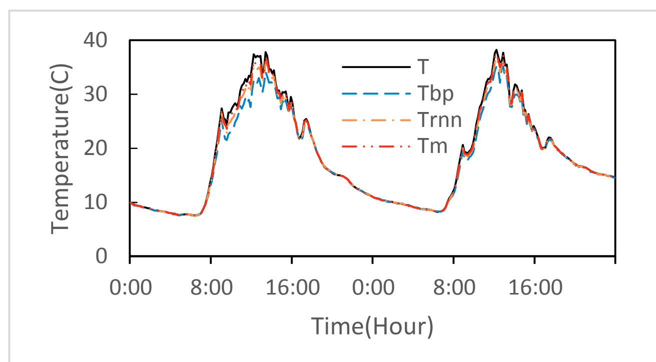
3.2 Prediction results and discussion

After the model training is completed, the two consecutive days added are predicted, and the results are analysed and discussed. Fig.4 and Tabel.2 shows the results. For temperature prediction results, the average absolute error MAE is 0.421 and the average relative error is 1.6%. It is almost the same as the model training result. The root mean square error RMSE is 0.751, which is slightly better than the training result and more consistent with the actual trend. Although the coefficient of determination R^2 is 0.953 lower than the modelling result, but it is still meets the requirements of fitness. At the same time, the model was found to be insensitive to peak values, and it was further found that at noon the sunroof was opened and the actual temperature in the greenhouse began to fluctuate. The temperature fluctuations are large, and the cycle is irregular. There are also large differences in the daily changes, showing greater randomness. Therefore, the model's prediction curve is always not higher than the actual temperature, and its fluctuation range and frequency are slightly smaller than the actual temperature.

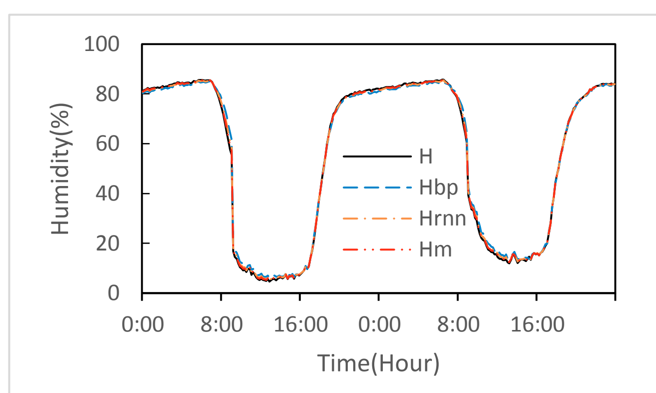
For the humidity model, the average absolute error MAE was 0.473 and the average relative error was 2.9%. In other words, the prediction accuracy is 0.47%, and the sensor accuracy is 3%, which is about 6 times the model error, so it can be used for humidity prediction. Both the root mean square error RMSE and the coefficient of determination R^2 are more in line with the actual change trend than the temperature. So the model error is within the allowable range, has reliability and can be used for prediction.

And also compared with the results of prediction used by BP neural network and untrained RNN neural network, it is obviously superior to both. When the greenhouse is closed, all three models have a good prediction curve, which is very close to the actual value. When the greenhouse is in an open state, the forecast begins to deviate with the fluctuation of the actual temperature and humidity. Specifically, the BP model is very sensitive to the opening and closing of the sunroof, but it is not very difficult to deal with actual changes. The sensitivity of the RNN model is lower than that of the BP neural network in the opening and closing of the sunroof, but it can learn the actual changes better. After the training of the dynamic BP algorithm, both the MAE and the RSME are reduced while the R^2 are improved. and the training effect is ideal.

For the problems in temperature prediction, the following thoughts are put forward: After the sunroof is opened, the greenhouse is in communication with the outside world, and heat exchange occurs. Then, the temperature in the greenhouse is basically the same as the outside temperature until the skylight is closed. During this period, the temperature does not experience excessive fluctuations. However, with the external winds forcing the air in the greenhouse to accelerate the flow and exchange with the outside world, it will also lead to temperature fluctuations. Due to the small regularity of changes in wind speed and direction in a short time, the temperature changes significantly. For model training, more data is needed for training. Then we need to increase the depth of the neural network. For the recurrent neural network, it is also the time scale. The layer of context used in this paper only contains one previous time data. It is a feasible method to introduce chaos to find the right time scale. Chaotic time series is a kind of non-linear deterministic system. The random motion appears in the system. The chaos has internal randomness and the overall stability is not local. It is stable, short-term predictable, and long-term, it contains system-rich dynamic information.



(a)



(b)

Fig. 4. Model prediction results: a) temperature; b) relative humidity.

Table 2. Prediction evaluation indicators

Prediction	Temperature				Humidity			
	ARE	MAE	RSME	R^2	ARE	MAE	RSME	R^2
BP	0.063	0.976	1.729	0.913	0.058	0.946	1.399	0.905
RNN	0.035	0.694	1.213	0.916	0.041	0.705	1.056	0.916
BP-RNN	0.017	0.488	0.865	0.925	0.029	0.473	0.699	0.937

4. CONCLUSIONS

This paper first analyses the current research direction of greenhouse microclimate modelling and selects the identification model with simple structure and easy analysis. Then, combined with the development trend of Internet of Things and big data, a recurrent neural network model is used as a deep learning algorithm for climate prediction. By analysing the input and output data of other identification models, this paper uses as input the historical environmental data from greenhouses and outputs temperature and humidity. In order to reduce the prediction error and improve the learning ability, the dynamic BP algorithm is used to modify the weights. Different from the traditional batch trained neural network, the dynamic BP method in the training process uses the output of the previous step together with the next input to the network, and the calculator outputs the weights. Compared with BP neural network and untrained RNN, as a short-term prediction, the Elman network based on dynamic BP algorithm can accurately predict the temperature and humidity in the greenhouse at the next step based on the indoor environment data because of recursive on-line training. Its RMSE remains below 0.9、MAE remains below 0.6, besides R^2 remains above 0.9, all of the parameter better than the other two forecasts. As a deep learning neural network, recurrent neural networks have already shined in language modelling and generating text, machine translation, speech recognition and image recognition, which are very suitable for time series. Therefore, when the Internet of Things and cloud computing become popular, the greenhouse-based environmental data is very impressive. The use of a recurrent neural network for prediction and control is a very effective way as it can warrant on-line model evolution and adaptation to changes.

6. ACKNOWLEDGEMENTS

This research was supported by the National Key Research and Development Program of China (2016YED0201003), the Yunnan Academician Expert Workstation (Wang Maohua, Grant No. 2015IC16) and A Traceability and Early warning system for supply chain of Agricultural Product: complementarities between EU and China (TEAP:PIRSES-GA-2013-612659). We must especially thank Academician Wang Maohua for giving us guidance and support during our research. Also Prof. Nick Sigrimis who give us a lot of information and experience.

REFERENCES

Li, T. (2016). Cloud-based decision support and automation for precision agriculture in orchards. *IFAC-PapersOnLine,*49(16), 330-335.

Vanthoor, B. H. E., Stanghellini, C., Henten, E. J. V., & Visser, P. H. B. D. (2011). A methodology for model-based greenhouse design: part 1, a greenhouse climate model for a broad range of designs and climates. *Biosystems Engineering,* 110(4), 363-377.

Zhang, X., Wang, H., Zou, Z., & Wang, S. (2016). CFD and weighted entropy based simulation and optimisation of Chinese solar greenhouse temperature distribution. *Biosystems Engineering,* 142, 12-26.

Lópezcruz, I. L., & Hernándezlarragoiti, L. (2010). Neuro-fuzzy models for air temperature and humidity of arched and Venlo type greenhouses in central Mexico. *Agrociencia,*44(7), 791-805.

Ramos-Fernández, J. C., Balmat, J. F., Márquez-Vera, M. A., Lafont, F., Pessel, N., & Espinoza-Quesada, E. S. (2016). Fuzzy modeling vapor pressure deficit to monitoring microclimate in greenhouses *. *IFAC-PapersOnLine,* 49(16), 371-374.

Hasni, A., Taibi, R., Draoui, B., & Boulard, T. (2011). Optimization of greenhouse climate model parameters using particle swarm optimization and genetic algorithms. *Energy Procedia,* 6(6), 371-380.

Guzmán-Cruz, R., Castañeda-Miranda, R., García-Escalante, J. J., López-Cruz, I. L., Lara-Herrera, A., & Rosa, J. I. D. L. (2009). Calibration of a greenhouse climate model using evolutionary algorithms. *Biosystems Engineering,* 104(1), 135-142.

Zou, W., Yao, F., Zhang, B., He, C., & Guan, Z. (2017). Verification and predicting temperature and humidity in a solar greenhouse based on convex bidirectional extreme learning machine algorithm. *Neurocomputing*. 249(2017), 72-85.

Patil, S. L., Tantau, H. J., & Salokhe, V. M. (2008). Modelling of tropical greenhouse temperature by auto regressive and neural network models. *Biosystems Engineering,* 99(3), 423-431.

Taki, M., Ajabshirchi, Y., Ranjbar, S. F., Rohani, A., & Matloobi, M. (2016). Heat transfer and MLP neural network models to predict inside environment variables and energy lost in a semi-solar greenhouse. *Energy & Buildings,* 110, 314-329.

Fourati, F. (2014). Multiple neural control of a greenhouse. *Neurocomputing*. 139(2014),138-144.

Mustafaraj, G., Lowry, G., & Chen, J. (2011). Prediction of room temperature and relative humidity by autoregressive linear and nonlinear neural network models for an open office. *Energy & Buildings,* 43(6), 1452-1460.

Hou S. (2016), Research and prediction on the control of greenhouse microclimate environment prediction based on CFD model. Thesis. *Nanjing Agricultural University*, Nanjing, China.

Zou W, Li Y., Wang X.(2014). Model Predictive Control of Air Temperature in Greenhouse Based on CFD Unsteady Model. *Chinese Journal of Agricultural Machinery,* 45(12), 335-340.

Zuo Z., Mao H., Zhang X., Hu J, Han L. &Ni J.(2010). Greenhouse Temperature Prediction Model Based on Time Series Analysis. *Chinese Journal of Agricultural Machinery,* 41(11), 173-177+182.