Minimum temperature prediction of solar greenhouse based on evaluation model established by principal component analysis

Lina Wang[1,2,3], Haihui Zhang[1,2,3*], Zhihang Liu[1]

[1]College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi, P. R. China

[2]Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture, Northwest A&F University, Yangling, Shaanxi, P. R. China

[3]Shaanxi Key Laboratory of Agriculture Information Sensing and Intelligent Service, Northwest A&F University, Yangling, Shaanxi, P. R. China

* Corresponding author, e-mail: lina_wang@nwafu.edu.cn

Abstract

Principal component analysis (PCA) was adopted to establish the evaluation model of solar greenhouse. Seven climatic varieties outside the greenhouse were analyzed in this paper. Illustrative results showed there was intimate correlation between minimum temperature and several indexes that include inside temperature, outside temperature and solar radiation. The three components were extracted to for establishing the PCA model. A simple model was built between the three indexes and minimum temperature. Verification results showed that mean absolute error (MSE) between theoretical and experimental values were less than 0.7℃, average relative error (ARE) were less than 12.8%, which indicated these models can forecast the minimum temperature inside solar greenhouse effectively. The application of models was more convenient and efficient for predicting minimum temperature for greenhouse microclimate.

Key words: Solar greenhouse; temperature prediction; evaluation model; minimum temperature; principal component regression; facility vegetables

1. Introduction:

Facilities horticulture through modern agricultural engineering, machinery, technology and management technology is to improve the local environment, to provide crop to adapt to the growth temperature, light, water, gas, fertilizer and other microclimate environment in a controlled environment industry. China has more than 90 percent of the total area of the world's facilities. It has become an important part of modern agriculture in our country. A typical solar greenhouse in northwestern China where in plants are grown and utilized as collectors to absorb solar radiation and produce heat inside. In the winter, the temperature of northwestern greenhouse is very low, even falling to below -10°C. These solar greenhouses can extend the planting time and solve the problem of insufficient vegetables and fruit wages in winter (Wang, et al., 2016). An accurate estimation of crop environmental temperature in greenhouses is critical for precise management, because it can provide the better conditions for plat growth. The greenhouse system is characterized by multivariate, nonlinear, multimodel, model coupling, multi-time scale and conflict multi-objective. It is a complex system that includes the interaction of crops, facilities, environment and management technologies. In order to improve the yield and quality of crops, the greenhouse should create a microclimate environment that is suitable for different stages of crop development. The relationship between crops and the environment is very strong and very interactive. The growth and physiological processes of crops affect the environment, and the environment, in turn, directly affects the growth of crops. As a factor of greenhouse environment control, temperature plays an important role. Accurately predicting temperatures in solar greenhouse has been a focus of research because it can influence crop photosynthesis greatly (Li, et al., 2013). Both higher and lower temperature may influence the quality and quantity of crop in greenhouse, even make the plant suffer seriously damaged or diseases when extreme temperature happen without necessary protective measures (Su, et al., 2012). Therefore, this result will decrease the yield of greenhouse and economic losses. The main objective in this paper is using automatic control strategies to get the higher quality and quantity of crop.

Solar energy is the source of heat in the greenhouse. Greenhouse microclimate environment is effected by outside weather conditions. In recent years, some authors have studied different computational approaches for various aspects of environment control of solar greenhouse. A photosynthesis prediction model of the tomato plant in the seedling stage is proposed by applying SVM to determine optimal CO_2 concentration (Li, et al., 2017). Zou presented a novel temperature and humidity prediction model based on convex bidirectional extreme learning machine to predict the solar greenhouse temperature and humidity(Zou, et al., 2017). It was a model of the elements of the meteorological elements inside and outside the greenhouse, without considering the elements of the greenhouse in the previous day. They used multi-purpose stepwise regression method for establishing models, without sufficient consideration for the error of the correlation between meteorological elements. The temperature predictive methods can be divided into two aspects: the first research method based on the mechanism model using mathematical theory; the second ones based on data driven model using some intelligent methods (Paniagua Tineo, et al., 2011; Van Beveren, et al., 2015). He and Ma established physical model which need to determine many parameters (He and Ma, 2010). Van Beveren presented the numerous temperature predictive method in his paper (Van Beveren, et al., 2015). But lots of parameters could be decided is very difficult. Sometimes during the model derivation some parameters have to be assumed, this will cause great uncertainty and low prediction accuracy in the numerical model. Over the last decades, some intelligent 'black box' control methods like neural network have been used in the greenhouse environment temperature prediction. Patil built linear neural network auto regressive method in Thailand that don't need to set parameters inside (Patil, et al., 2008). But its drawback is the model noise due to the sequential effect of rain. There has been some more computer intelligent methods that be used in temperature control in greenhouse, such as fuzzy control (Lu, et al., 2014), artificial neural network and support vector regression algorithm method. Dariouchy et al. (2009) established a neural network model to predict the internal temperature. This method performs better than other methods, however, it will easily have low convergence, over-fitting and poor stability. By means of small subset of training points, SVM proposes enormous computational advantages and avoids into global minimum or local optimum effectively (Tian, et al., 2017). Coelho predicted the greenhouse environment behavior using SVM models to describe indoor climate process dynamics. (Coelho， et al., 2010) But they only considered temperature or some other climatic conditions as model inputs to predict the next temperature value, not consider the reason of using these parameters.

In prediction analysis, the principal component analysis (PCA) has been applied widely in many field. W Li proposed two recursive PCA algorithms to updating the correlation matrix recursively for adaptive process monitoring.(Li, et al., 2000, Cao, et al., 2003). PCA model for recognizing human actions learning is consisted of two components in. First, features are learned respectively from the RGB and depth channels, using sparse auto-encoder with convolutional neural networks. Second, the learned features from both channels is concatenated and fed into a multiple layer PCA to get the final feature.(Li, et al., 2015). This work studies the fault classification issue using PCA focused on complicated industrial processes for multi-fault classification purpos.(Chen, et al., 2015).

This study focuses on the study of temperature changes in winter solar greenhouse. According to the weather prediction results and the meteorological elements of the previous day in the greenhouse, the minimum temperature prediction model of the solar greenhouse in winter is established by using the method of principal component regression. In the hope that before the cold air comes, the weather forecast will be released in time to warn against low temperatures.

The paper is organized as follows. Section 2 describes the theory of principal component analysis. In Section 3, the solar greenhouse materials and method model are presented. Use PCA method to stablish low temperature prediction model and show results and the experimental process. Finally, conclusions are in Section 4.

2. Principal Component Analysis

The PCA is a statistical method, which can present the most important variabilities information and results in a graphical plot. This algorithm can transform the original principal components (PCs) that are relevant into uncorrelated variabilities with each other. Also it can decrease the high dimension data to low dimension and save as much as possible of the total variation. In this paper, the PCA algorithm is used to choose the initial variables obtained by greenhouse sensors defined *X,* which is a m×n matrix(m variables and n samples). The variables acquired by greenhouse sensors have different units and scales. But, here we can use PCA analysis to change each variable to the same weight when in the data analysis. The dimension data can be reduced as X_i, which is defined by (Destefanis, et al., 2000):

$$X_i = \frac{\lambda_i - \widehat{\lambda_i}}{S_i} \tag{1}$$

in above equation, λ_i defines the i^{th} row vector, $\widehat{\lambda_i}$ is the man of this vector, S_i is the standard deviation. The PCA model parameters can be defined by the eigenvalues $\xi_1, \xi_2, \cdots, \xi_t$ and eigenvectors $\gamma_1, \gamma_2, \cdots, \gamma_t$. The PCs can be expressed by Y_i as follows:

$$Y_i = X_i \gamma_i \tag{2}$$

in which γ_i denotes the i^{th} eigenvectors of the correlation matrix.

Through PCA analysis we can obtain the each variability percentage whose occupancy among all the measured variabilities (Wang, et al., 2017). The percentage of variability can be estimated by the equation below:

$$w_{pc}(i) = \frac{\xi_i}{\sum \xi} \times 100 \tag{3}$$

The coefficients of correlation can be calculated by eq.(4) as follows:

$$\mu_i = \gamma_i \sqrt{\xi_i} \tag{4}$$

The scope of μ_i vary between -1 and 1, which are obtained by the multiplications between each eigenvector and the square root of the eigenvalue associated.

Using the coefficients of correlation make the PCA achieve a graphic information representation (Nair, et al., 2013). Through studying the projections of the coefficients of correlation between two correlated variables can decide the correlation of these two variables and the new space of visualization. The space is a circle axes called “factorial space” or “correlation circle”. If the circle between the vector are close, the two variabilities are correlated (Chen, et al., 2010). The angle about the correlation circle between these variables can be measured by eq.(5) as follows:

$$\cos\alpha = angle(\mu_i, \mu_j) \tag{5}$$

If these two variables are both closed to themselves, they are correlated. By opposition, if α is 90°, the two variables are not correlated. Through the analysis of PCA model, we can see the correlation of many variables.

3. Materials and methods

3.1 Greenhouse installations and data acquisition

The data used in this paper were acquired by the management system of vegetable base in Yanliang district, Shaanxi Province of China. Until now this agricultural Internet of things monitoring platform has managed more than 100 greenhouses in Shaanxi province. Typical solar greenhouses in the northwestern China have a plastic film covering the slanted front, as shown in Fig.1. It can keep light and temperature (Fourati, et al., 2014). At the top of greenhouse there is a curtain rolling machine which can maintain the heat inside at night. At the top of greenhouse there is a thick soil wall against the cold in winter. Our greenhouse data were collected over several months at approximately longitude 109°08′54″, latitude 34°35′11″. As shown in figure 2, the air temperature and humidity sensors and light radiation sensors are dispersed in the greenhouse and the outer partition. The sensor will be uploaded and saved per five minutes after the required experimental data, which will eventually be available for indoor and external air temperature and light intensity (Yu, et al., 2016, Frausto, et al., 2004). A small weather station was set up outside the greenhouse to detect the influencing factors, concluding air humidity, solar radiation, wind speed, outside maximum temperature, outside minimum temperature, outside mean temperature, average sunshine time. The data comes from the winter of vegetable base from 17th Dec. 2016 to 19th Jau. 2017. Its meteorological elements are in line with the real time meteorological forecast project issued by the meteorological station. Using SPSS20.0 to realize principal component analysis and multiple linear regression. All indexes were either of analytical grade or from Kaiser Meyer Olkin (KMO) and Bartlett inspection.

Figure 1. Solar greenhouse in China

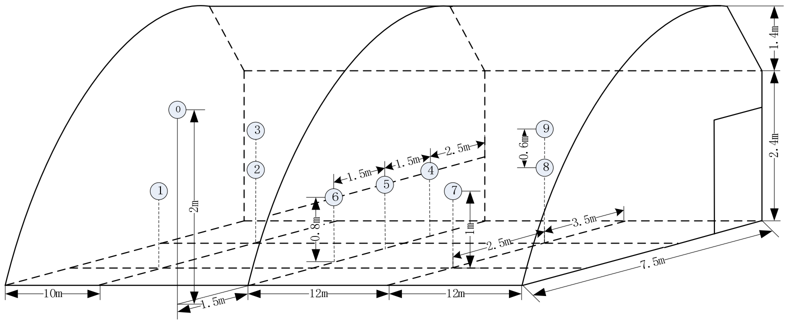

Figure 2. Solar greenhouse structure diagram

3.2 Preparation of microclimate conditions outside the solar greenhouse

The detail minimum temperature variables of solar greenhouse are shown as Fig.3. According to the classification standard of sunshine percentage of meteorological, the percentage of sunshine is greater than 60%, 20% < S < 60% and 0 < S < 20%, respectively, as sunny, cloudy and low sunny weather as shown in Fig.4. From the characteristics of temperature change in solar greenhouse, we can see that during sunny days and cloudy conditions, the temperature in the solar greenhouse has a distinct daily change. During the day, the temperature rises and the night temperature decreases. Under the conditions of low sunny weather, the daily change is not obvious and the temperature of the whole day is lower (Note: Sunny days are 2016.12.27-2016.12.28, 2017.01.15, 2017.01.18-2017.01.19; cloudy days are 2016.12.17-2016.12.26, 2017.01.12-2017.01.14 and low sunny are 2016.12.29-2016.12.31, 2017.01.01-2017.01.11, 2017.01.16-2017.01.17). The maximum temperature and minimum temperature data is shown in Fig.5. The temperature is high at sunny day owing to abundant sunlight. Variation of the average illumination outside solar greenhouse is also higher. At cloudy day the maximum temperature is lower than sunny day. The temperature is the lowest at low sunny weather. We can see that different weather condition will influence the greenhouse temperature and the crop growth.

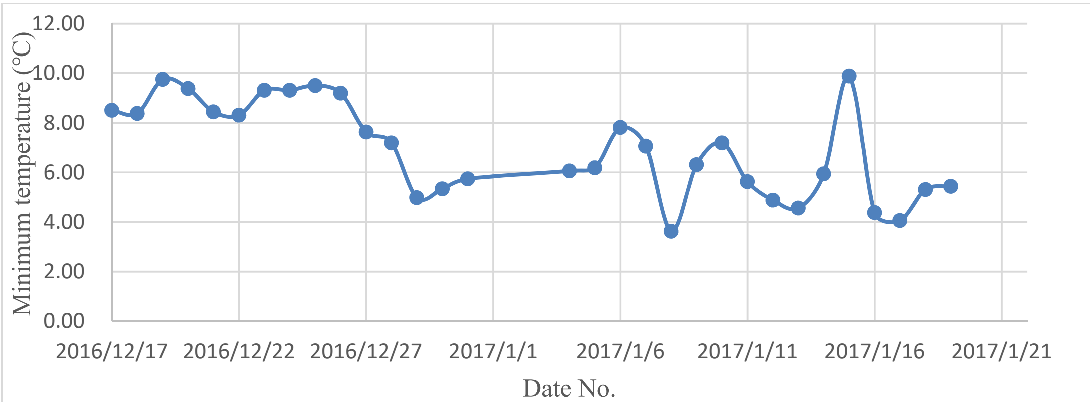

Figure 3. Variation of solar greenhouse of minimum temperature

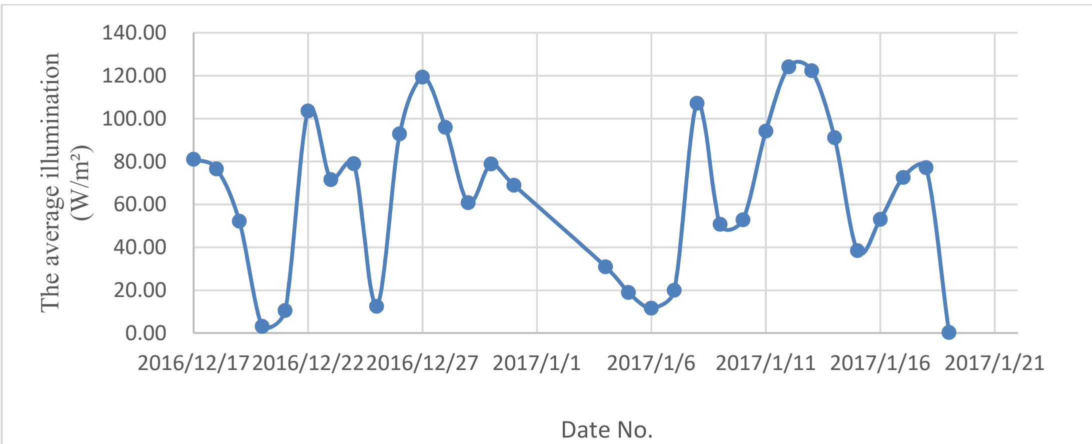

Figure 4. Variation of the average illumination outside solar greenhouse

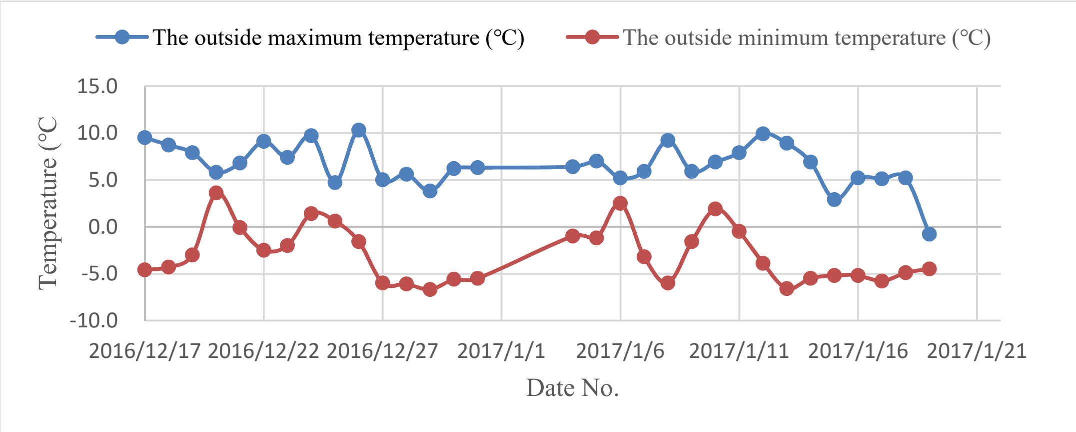

Figure 5. The variation curve of maximum temperature and minimum temperature outside greenhouse

Take tomatoes as example, tomatoes is for warm-season vegetables, it can grow up well from the range 15℃ to 33℃, in 22 ~ 25℃ during the day, in 15 ~ 18℃ at night. Influence of low temperature is bigger on growth of warm-season vegetables. Germination period is so slow, early seeding stage will easily have deformity fruit and the color of fruiting stage is very slow. We can see the minimum temperature outside greenhouse is -4.06℃. If extreme low temperatures weather can be predicted, the farmers can do some protective things to defect low temperature disaster.

In this paper, select several meteorological factors affecting weather condition concluding the maximum temperature and minimum temperature outside solar greenhouse, mean temperature outside solar greenhouse, mean humidity outside solar greenhouse, maximum wind velocity outside solar greenhouse, illumination and mean illumination outside solar greenhouse. Use PCA to analyze the relationship between the temperature inside solar greenhouse and meteorological factors outside solar greenhouse. The variation curve of maximum wind velocity is in Fig.6. The mean illumination outside solar greenhouse is in Fig.7.

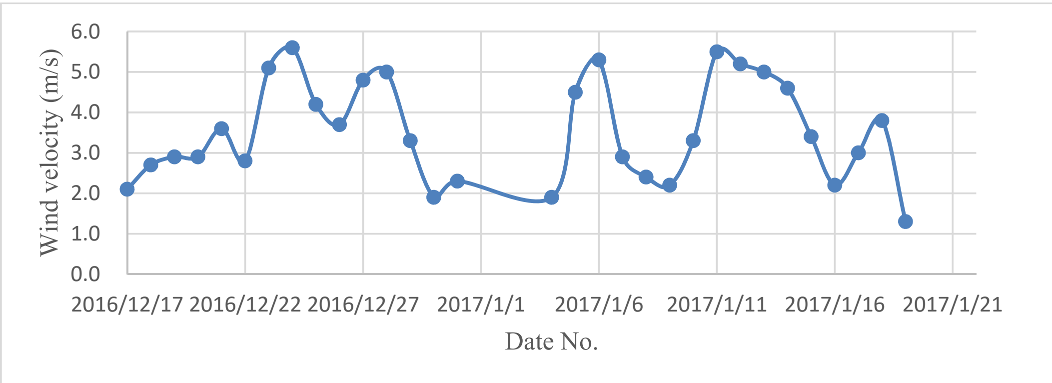

Figure 6. The variation curve of maximum wind velocity outside solar greenhouse

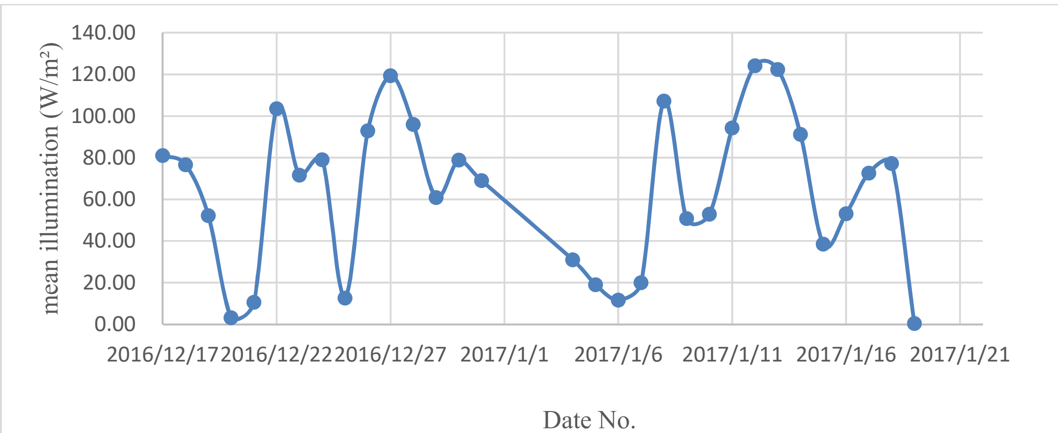

Figure 7. The variation curve of mean illumination outside solar greenhouse

3.3 Establish low temperature prediction model using PCA

The 7 meteorological elements and minimum temperature in greenhouse was selected to principal component analysis, the result is shown in Tab.1. The results show, the factor eigenvalues of the first three principal components are greater than 1. The accumulative contribution rate of the 3 major factors to total variation accounted for over 86%, maintaining most of information of 7 characters. In other words, 86.6% of the total variance in the 7 considered variables can be condensed into three PCs.

Table 1. The total variance explained by principal component

Component	Initial eigenvalue		
	Eigenvalues	Percentage of variance (100%)	Cumulative variance (100%)
1	2.595	37.075	37.075
2	2.432	34.738	71.813
3	1.04	14.863	86.676
4	0.527	7.523	94.199
5	0.296	4.236	98.435
6	0.082	1.178	99.613
7	0.027	0.387	100

Select the first PCs to calculate the loading matrix and show the result in Table 2. Note: T_{tmax} is the maximum temperature outside solar greenhouse, T_{tmin} is the minimum temperature outside solar greenhouse, T_{tmean} is mean temperature outside solar greenhouse, H is mean humidity outside solar greenhouse, W is maximum wind velocity outside solar greenhouse, R is mean temperature outside solar greenhouse, R_{mean} is illumination and mean illumination outside solar greenhouse. The greater the absolute value of the coefficient, the relationship between the principal component and the index indicates the closer. Through the loading matrix and eigenvalues, unit eigenvector can be acquired. The relationship between meteorological elements and minimum temperature can be expressed as follows below：

$$y_1=0.783T_{tmax}+0.006T_{tmin}+0.267T_{tmean}-0.684T_{tmean}+0.654W+0.778R+0.640R_{mean} \quad (6)$$
$$y_2=0.268T_{tmax}+0.94T_{tmin}+0.933T_{tmean}+0.448T_{tmean}+0.137W-0.553R+0.284R_{mean} \quad (7)$$
$$y_3=0.413T_{tmax}-0.219T_{tmin}-0.034T_{tmean}+0.421T_{tmean}-0.657W+0.168R+0.428R_{mean} \quad (8)$$

On the basis of the overall results analyses, PCA proved to be a useful and effective method to point the relationship between minimum temperature and meteorological elements.

Table 2. Loading matrix of 3 principal component factors

	Component		
	1	2	3
T_{tmax}(℃)	0.783	0.268	0.413
T_{tmin}(℃)	0.006	0.94	-0.219
T_{tmean}(℃)	0.267	0.933	-0.034
H(%RH)	-0.684	0.448	0.421
W(m/s)	0.654	0.137	-0.657
R(W/m^2)	0.778	-0.553	0.168
R_{mean}(h)	0.64	0.284	0.428

3.4 Test prediction model

Using the above minimum temperature prediction model to test winter temperature data from 2016 to 2017. The result between prediction value and actual value in different weather conditions is shown in Table 3. The established model has satisfactory fitting goodness and can be employed to forecast the temperature at sunny, cloudy and low sunny weather day. The results of the test data indicate that the prediction model is reliable and the root of mean square error is 1.49℃ in 2016, 1.65℃ in 2017. Average absolute error is 0.6℃ at sunny day and low sunny day, which is better than cloudy day. Average relative error is 11% at sunny day, which is better than cloudy day and low sunny day. The prediction accuracy using PCA method in this paper is higher. The error is smaller and satisfies the low temperature prediction need in solar greenhouse.

Table 3. The minimum temperature prediction model test of solar greenhouse

	Root mean squared	Average absolute error(℃)	Average relative error (100%)

		error(°C)		
2016	Sunny day	1.49	0.6	11.4
	Cloudy day	1.49	0.7	12.6
	Low sunny day	1.49	0.6	11.7
2017	Sunny day	1.65	0.6	11.9
	Cloudy day	1.65	0.7	12.8
	Low sunny day	1.65	0.7	12.4

4. Conclusions

This paper explains a PCA applied to the complex temperature prediction in solar greenhouse. Simplify the model variables but keep an efficient prediction model. We can observe this method is efficient to define a set of relevant variables to uncorrelated variables. The results presented show the relationship between minimum temperature inside solar greenhouse and meteorological weather elements outside greenhouse. The work research performed on prediction modeling, propose a significant result in solar greenhouse.

Acknowledgements

This paper was supported by the Chinese Universities Scientific Fund (2452016154), a Test Demonstration Station Technology Achievements Promotion project (TGZX2017-33), the China Postdoctoral Science Foundation (2017M623253) and the Fundamental Research Funds for the Central Universities (2452017128).

Reference

Coelho, J. P.， Cunha, J. B.， Oliveira, P. D. M.， Pires, E. S., (2010). Greenhouse heat load prediction using a support vector regression model, Advances in Intelligent & Soft Computing, 73:111-117

Chen, J.， Jian, H., (2015). SVM and PCA based fault classification approaches for complicated industrial process, Neurocomputing, 167 (C):636-642

Chen, X., Wang, L., Hu, B., Guo, M. S., Barnard, J., Zhu, X. F., (2010). Pathway-based analysis for genome-wide association studies using supervised principal components, Genetic Epidemiology, 34(7), 716-724.

Cao, L. J.， Chua, K. S.， Chong, W. K.， Lee, H. P.， Gu, Q. M., (2003). A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine, Neurocomputing, 55(1):321-336

Dariouchy, A., Aassif, E., Lekouch, K., Bouirden, L., Maze, G., (2009). Prediction of the intern parameters tomato greenhouse in a semi-arid area using a time-series model of artificial neural networks, Measurement, 42(3), 456-463

Destefanis, G., Barge, M. T., Brugiapaglia, A., Tassone, S., (2000). The use of principal component analysis (PCA) to characterize beef, Meat Science, 56 (2000) 255-259

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

Frausto, H., Pieters, H., (2004). Modelling greenhouse temperature using system identification by means of neural networks, Neurocomputing, 56(56):423-428

He, F., Ma, C., (2010). Modeling greenhouse air humidity by means of artificial neural network and principal component analysis, Comput. Electron. Agric, 71, 19-23

Nair, A., Mohanty, U. C., Acharya, N. (2013). Monthly prediction of rainfall over India and its homogeneous zones during monsoon season: a supervised principal component regression approach on general circulation model products, Theoretical and Applied Climatology, 111(1-2), 327-339.

Li, W.， Yue, H. H.， Valle-Cervantes, S.， Qin, S. J., (2000). Recursive PCA for adaptive process monitoring, Journal of Process Control, 10(5):471-486

Li, S. Z.， Yu, B.， Wu, W.， Su, S. Z.， Ji, R. R., (2015) Feature learning based on SAE–PCA network for human gesture recognition in RGBD images, Neurocomputing, 151(151):565-573

Li, N., Shen, S., Li, Z., Qing, Y., Li, C., Liu, S., (2013). Forecast model of minimum temperature inside greenhouse based on principal component regression, Chinese J. Agrometeorol, 3, 306-311

Li, T.， Ji, Y. H.， Zhang, M.， Li, M. Z., (2017) Determining optimal CO_2 concentration of greenhouse tomato based on PSO-SVM, Applied Engineering in Agriculture, 33(2):157-166

Lu, W., Pedrycz, W., Liu, X., Yang, J., Li, P., (2014). The modeling of time series based on fuzzy information granules, Expert Syst. Appl, 41(8), 3799-3808

Paniagua, T. A., Salcedo, S. S., Casanova M. C., Ortiz, E.G., Cony, M. A., Hernández, M. E., (2011). Prediction of daily maximum temperature using a support vector regression algorithm, Renewable Energy, 36(11), 3054-3060

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

Su, Z.G., Wang, P.H., Shen, J., Li, Y.G., Zhang, Y.F., Hu, E.J., (2012). Automatic fuzzy partitioning approach using variable string length artificial bee colony (VABC) algorithm, Appl. Soft Comput. 12, 3421-3441.

Tian, Z. D., Zhang, C., Li, S. J., Wang, Y. H., (2017). Time-delay prediction based on phase space reconstruction and least squares support machine, Acta Electronica Sinica, 45(5), 1044-1051

Van, Beveren, P. J. M., Bontsema, J., Straten, G., Van, Henten, E.J., (2015). Minimal heating and cooling in a modern rose greenhouse, Appl. Energy, 137, 97-109

Wang, L. L., Hong, Z., Liu, L., Wang, Q., Li, S. R., Li, Q. Z., (2017). Prediction of peanut protein solubility based on the evaluation model established by supervised principal component regression, Food Chemistry, 218(1), 553-560

Wang, C. J., Liu, C. Z., Wang, J. Z., (2016). Climate change may threaten habitat suitability of threatened plant species within Chinese nature reserves, PeerJ, 4, e2091

Yu, H. H., Chen, Y. Y., Hassan, S. G., Li, D. L., (2016). Prediction of the temperature in a Chinese solar greenhouse based on LSSVM optimized by improved PSO, Computers and Electronics in Agriculture, 122, 94-102

Zou, W.， Yao, F.， Zhang, B.， He, C.， Guan, Z., Verification and predicting temperature and humidity in a solar greenhouse based on convex bidirectional extreme learning machine algorithm, Neurocomputing, 249(2), 72-85