Acta Technica 62 No. 2B/2017, 585–596

Environmental forecast of the solar greenhouse based on the weighted Markov chain[1]

Dongsheng Zhou[2,3,5], Yikui Bai[3,5,6], Rongfei Zhao[3,5], Yingchun Jiang[4,5]

Abstract. The prediction of Greenhouse environment is greatly important for solar greenhouse modeling. Solar irradiation, humidity and soil temperature are the main influence factors of the temperature in the greenhouse. The paper builds the weighted Markov chain model, it uses sequential cluster to realize the state classification of sample output, it calculates the coefficients and weights of each order autocorrelation, in order to simulate and predict the parameters in the short term, which include the average solar irradiation, the average humidity and soil temperature, and then it predicts the average temperature of the greenhouse. Simulation results show that environmental forecast interval is reasonable; the short term prediction modeling is accurate. It provides a theoretical basis for the greenhouse temperature prediction, which has important guiding significance in the light of the analysis of uncertainty parameter in greenhouse. It is easy to bring about, it is less time to calculate, and provides a broader prospects

Key words. Solar greenhouse, simulation, sequential cluster, the weighted Markov chain, environment prediction.

1. Introduction

Chinese solar greenhouse is the main production style of vegetables and fruits in the north of China in the winter. Solar irradiation is the main source of energy, and it

[1]This work was supported by NSFC (51508345), modern agricultural industry technology system special Funds (CARS-25-C-01), Chinese postdoctoral science foundation project(2014M561250), and general project of Education department of Liaoning province (LSNYB201618).

[2]College of Information and Electrical Engineering, Shenyang Agricultural University, Shenyang 110866, China

[3]College of Water Conservancy, Shenyang Agricultural University, Shenyang 110866, China

[4]College of Engineering, Shenyang Agricultural University, Shenyang 110866, China

[5]Engineering & Technology Research Center of Protected Agricultural Environment and Equipment of Liaoning Province, Shenyang 110866, China

[6]Corresponding author

is affected by the season and weather; humidity reduces with temperature rises, and humidity rises with temperature reduces; soil temperature rises or reduces with the temperature of greenhouse [1, 2]. So mastering the relation of temperature and the three influence factors is of great significance for the production of the greenhouse.

Research of environment of greenhouse focuses on the irradiation, temperature and humidity. Light model is established to calculate the direct solar radiation and scattered radiation [3, 4], time series method [5], nerve network [6], constrained predictive control model and self-adapting mechanism model to forecast the temperature of the solar greenhouse [7, 8]. The paper puts forward improving extreme learning machine algorithm [9], grey prediction model to forecast the temperature and humidity [10]. Daily change and math-expression method is used to analyze the temperature variation [11]. Markov chain is used to simulate and predict the PV output and wind speed in recent years, and got the better effects [12–15].

2. Materials and methods

2.1. Data analysis

The experimental greenhouse is located in the Shenyang Agricultural University scientific research base (123.4 °E, 41.8 °N). The greenhouse toward for 5 ° south by west, the length is 60 m, the span is 10 m, the angle of south roof is 30.5 °, the height of roof is 5 m, the height of back wall is 3.2 m, the wall is multilayer that is combined bricks and benzene board, the film is PE. The structure of greenhouse is main structural style in the Liaoning China.

The samples were selected from December 1st, 2015 to January 19th, 2016, a total of 50 days of data. The experiment measured the data of solar irradiation that entered greenhouse, inner temperature of greenhouse, inner soil temperature, inner humidity and outer temperature through using the wireless meteorological station made by ADCON, spectral range of pyranometer is between 400 nm and 1100 nm, operating temperature is between –30 ° and 70 °C, sensitivity is between $60\,\mathrm{V\,W^{-1}m^{-1}}$ and $100\,\mathrm{V\,W^{-1}\,m^{-1}}$, max irradiation is $2000\,\mathrm{W\,m^{-2}}$, accuracy of temperature sensor is $< \pm 0.2\,°\mathrm{C}$, measure range is between –40 °C and 60 °C. The equipment records data every 15 minutes. We selected the data from 9:00 to 15:00 and calculated the total solar irradiation, average humidity, average inner temperature and average outer temperature.

The environmental data samples in solar greenhouse are shown in Figure 1. The distribution curve shows random distribution characteristics, and there is distinct boundaries interval range. Therefore, Markov chain short-term prediction is suitable for matching the characteristics of temperature interval distribution. In the actual engineering design, the temperature specific values in solar greenhouse are undetermined, but their state boundaries can be predicted to achieve by Markov chain. So the feasible range is provided for engineering design by means of Markov chain prediction.

Inner temperature of greenhouse is a main indicator that measures greenhouse. Solar irradiation, outer temperature and soil temperature are major impact factors.

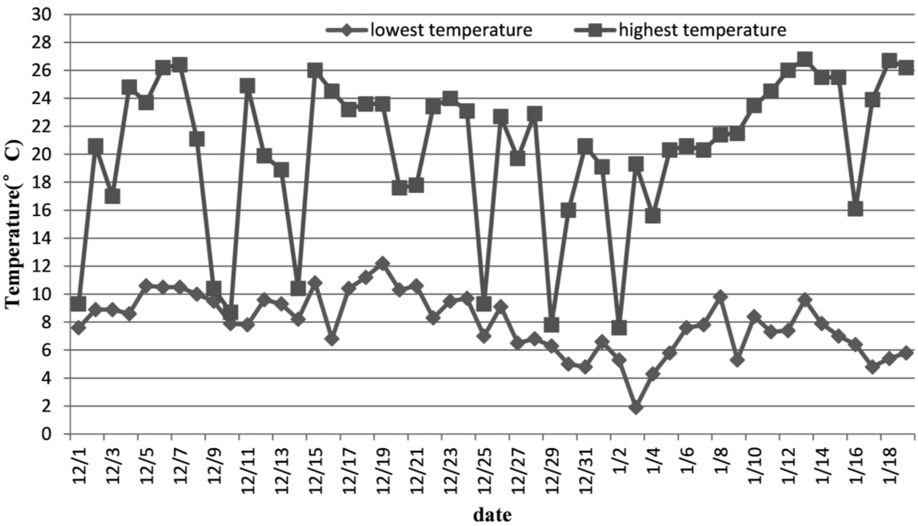

Fig. 1. Temperature range of data samples

The linear relationship analysis is based on SPSS 20.0 and significance level is less than or equal to 0.05. The analysis result is shown in Table 1.

Table 1. The linear relationship analysis of impact factors and inner temperature

Impact factor	Association	Sig.	R^2
Solar irradiation	0.711	0.000	0.769
Soil temperature	–0.790	0.000	
Humidity	0.789	0.000	

Analyzed from Table 1, the three impact factors and inner temperature exhibit a linear relationship. Multivariate regression linear equation is $y = 0.907 + 0.10x_1 - 0.65x_2 + 1.111x_3$.

For next processing, Consider the method presented by Fisher in 1958. It is used to deal with sequential data. The steps of classification are as follows: Ordinal sample is X, $b\,(n, k)$ represents the kth classifications of n samples. Symbol G represents one of the classifications and contains parameters $\left\{X_{(i),} X_{(i+1),} \ldots\ldots, X_{(j),}\right\}$ $(j > i)$. The average vector

$$\bar{X}_G = \frac{1}{j - i + 1} \sum_{t=i}^{j} X_{(t)}\,.$$

Let $D\,(i, j)$ represents diameter of the classification,

$$D(i, j) = \sum_{t=i}^{j} \left(X_{(t)} - \bar{X}_G\right)' \left(X_{(t)} - \bar{X}_G\right)\,.$$

If $m = 1$, then $D(\mathrm{i}, j) = \sum_{t=i}^{j} \left| X_{(t)} - \bar{X}_G \right|$, $\bar{X}_G$ representing the median of data. Define the error function L

$$L\left[b\left(n,k\right)\right] = \sum_{i=1}^{k} D\left(i_t, i_{t+1} - 1\right) . \tag{1}$$

If $L\left[b\left(\mathrm{n},\mathrm{k}\right)\right]$ is smaller, the sum of deviations is also smaller. The classification is then more reasonable. Then, based on the equation (1), we achieve the optimal segmentation through the iterative method.

Obtain the optimal solution. First, find j_k as the point of division, make the result of (1) the smallest possible, and obtain G_k. Now, based on the above method, obtain all classifications G_1, G_2, G_3, $\ldots$, G_k. This can be written as

$$p\left(n,k\right) = \left\{G_1, G_2, \ldots, G_{1k}\right\} . \tag{2}$$

2.2. Weighted Markov chain

Markov Chain is a random process that meets two hypothesizes as follows:

1. Probability distribution of system state at the $T + 1$th moment relates with Tth moment, and has nothing to do with phenomena before Tth moment.

2. State transition from Tth to $T + 1$th moment has nothing to do with the Tth moment. A Markov Model is represented by triple $E = (S, P, Q$, where S represents the state space, P represents the state transition probability and Q represents the initial probability density.

The weighted Markov Chain is an improved Markov chain. It can be used to describe a lot of dynamic in economic, weather forecast, etc. It can predict the current state based on the number of past states, then predict and analyze through weight on the basis of the relationship of the current state and the past states, achieve the aim of prediction adequately and reasonably. The method is suitable for prediction in short term.

3. Case study and discussion

The experimental greenhouse lies in the science and research base of Shenyang Agricultural University. Select 50 sets of data of total irradiation, average humidity and average soil temperature. Adopt the Weighted Markov Chain to predict total irradiation, average humidity and average soil temperature. Based on the three indicators, analyze the range of inner temperature of greenhouse. Explain the process based on the irradiation. The steps are as follows:

Step (1). Classify the data in accordance with the Sequential Clustering Method, generate the classification standard. In the example, the data is classified five classes, which includes total irradiation, average humidity and average soil temperature. The standard is shown in Table 2.

Table 2. Classification standard of data

State	Total irradiation	Average humidity	Average soil temperature
1	$<1000\,\mathrm{W/m^2}$	>85	$<12.5\,°\mathrm{C}$
2	$1000\sim3000\,\mathrm{W/m^2}$	$75\sim85$	$12.5\sim14\,°\mathrm{C}$
3	$3000\sim4500\,\mathrm{W/m^2}$	$65\sim75$	$14\sim15\,°\mathrm{C}$
4	$4500\sim5900\,\mathrm{W/m^2}$	$58\sim65$	$15\sim17\,°\mathrm{C}$
5	$>5900\,\mathrm{W/m^2}$	<58	$>17\,°\mathrm{C}$

Step (2). Transform the irradiation value into the corresponding state, which is shown in Fig. 2.

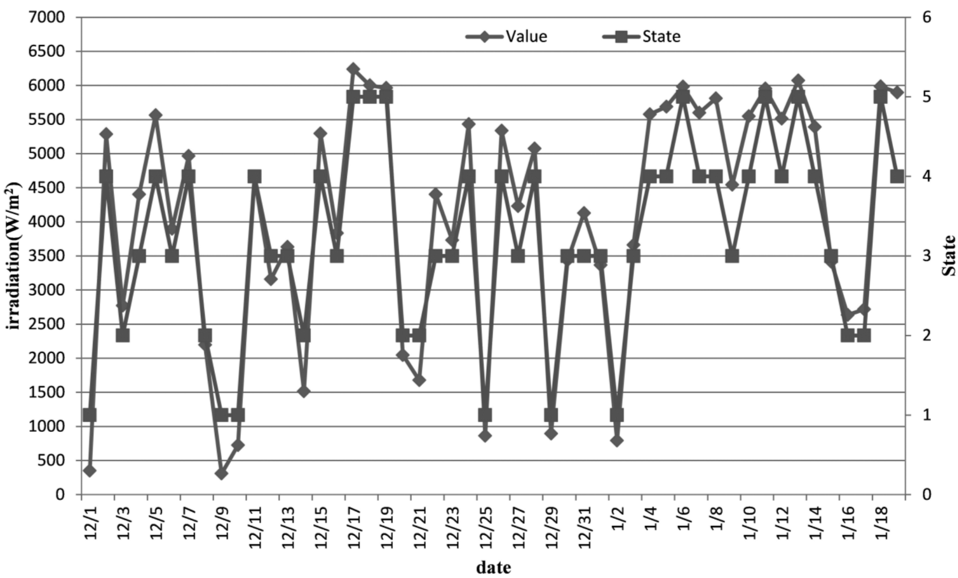

Fig. 2. Irradiation and state classification

Step (3). Calculate autocorrelation coefficient of every order $r_k,\ k \in E$.

$$r_k = \frac{\sum_{l=1}^{n-k}\left(x_l - \overline{x}\right)\left(x_{l+k} - \overline{x}\right)}{\sum_{l=1}^{n}\left(x_l - \overline{x}\right)^2}\,. \tag{3}$$

In the equation, r_k represents autocorrelation coefficient of the kth order, x_l represents irradiation of the lth moment, and $\bar{x}$ represents average irradiation. The autocorrelation coefficient is shown in Table 3.

Table 3. Autocorrelation coefficient of orders from 1 to 8

	1	2	3	4	5	6	7	8
r_k	0.9542	-1.3977	-1.3171	-2.9484	1.3758	0.9132	1.2616	0.9707
z_k	0.0857	0.1255	0.1182	0.2647	0.1235	0.0820	0.1133	0.0871

Step (4). Normalize the autocorrelation coefficient.

$$w_k = \frac{|r_k|}{\sum_{k=1}^{m} |r_k|}\,. \tag{4}$$

Here, w_k represents the weight of the kth order, m represents the max order. The weights are also shown in Table 3.

Step (5). Calculate respectively eight-order transfer matrices, the matrices representing the probability rules of state transitions.

$$P^{(1)} = \begin{bmatrix} 1/6 & 0 & 1/3 & 1/2 & 0 \\ 1/7 & 2/7 & 2/7 & 1/7 & 1/7 \\ 1/14 & 1/7 & 2/7 & 3/7 & 1/14 \\ 2/15 & 2/15 & 2/5 & 2/15 & 1/5 \\ 0 & 1/7 & 0 & 4/7 & 2/7 \end{bmatrix},$$

$$P^{(2)} = \begin{bmatrix} 0 & 1/6 & 1/2 & 1/3 & 0 \\ 1/7 & 0 & 3/7 & 2/7 & 1/7 \\ 3/14 & 3/14 & 3/14 & 3/14 & 1/7 \\ 1/15 & 1/15 & 4/15 & 7/15 & 2/15 \\ 0 & 1/3 & 1/6 & 1/6 & 1/3 \end{bmatrix},$$

$$P^{(3)} = \begin{bmatrix} 0 & 0 & 2/3 & 1/3 & 0 \\ 0 & 0 & 1/3 & 1/2 & 1/6 \\ 3/14 & 0 & 3/14 & 5/14 & 3/14 \\ 2/15 & 1/5 & 1/5 & 4/15 & 1/5 \\ 0 & 1/2 & 1/3 & 1/6 & 0 \end{bmatrix},$$

$$P^{(4)} = \begin{bmatrix} 1/3 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/5 & 0 & 1/5 & 2/5 & 1/5 \\ 1/14 & 1/7 & 2/7 & 5/14 & 1/7 \\ 1/15 & 1/15 & 4/15 & 2/5 & 1/5 \\ 0 & 1/3 & 1/2 & 1/6 & 0 \end{bmatrix},$$

$$P^{(5)} = \begin{bmatrix} 0 & 1/6 & 1/2 & 1/3 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/13 & 1/13 & 2/13 & 6/13 & 3/13 \\ 1/5 & 2/15 & 1/3 & 4/15 & 1/15 \\ 0 & 1/6 & 1/3 & 1/6 & 1/3 \end{bmatrix},$$

$$P^{(6)} = \begin{bmatrix} 0 & 0 & 1/3 & 2/3 & 0 \\ 1/5 & 2/5 & 1/5 & 1/5 & 0 \\ 3/13 & 0 & 4/13 & 3/13 & 3/13 \\ 0 & 3/14 & 2/7 & 3/14 & 2/7 \\ 1/6 & 1/6 & 1/6 & 1/2 & 0 \end{bmatrix},$$

$$P^{(7)} = \begin{bmatrix} 0 & 1/6 & 1/2 & 1/6 & 1/6 \\ 1/5 & 1/5 & 1/5 & 2/5 & 0 \\ 1/13 & 2/13 & 4/13 & 4/13 & 2/13 \\ 1/7 & 1/7 & 2/7 & 2/7 & 1/7 \\ 1/5 & 0 & 0 & 2/5 & 2/5 \end{bmatrix},$$

$$P^{(8)} = \begin{bmatrix} 1/3 & 0 & 0 & 1/6 & 1/2 \\ 1/5 & 0 & 2/5 & 2/5 & 0 \\ 0 & 4/13 & 4/13 & 4/13 & 1/13 \\ 1/13 & 1/13 & 5/13 & 3/13 & 3/13 \\ 1/5 & 0 & 1/5 & 3/5 & 0 \end{bmatrix},$$

Step (6) Calculate the state probability of the moment. $P_i^{(k)}$, $i \in E$, $k = 1, 2, \ldots, m$ Predict the irradiation state on the January 20. The results are shown in Table 4.

Table 4. Prediction table of total irradiation on January 20

Initial day	State	Hysteresis time (day)	Weight	State					Probability resource
				1	2	3	4	5	
Jan.19	4	1	0.0856645	2/15	2/15	2/5	2/15	1/5	$P^{(1)}$
Jan.18	5	2	0.1254783	0	1/3	1/6	1/6	1/3	$P^{(2)}$
Jan.17	2	3	0.1182479	0	0	1/3	1/2	1/6	$P^{(3)}$
Jan.16	2	4	0.2647003	1/5	0	1/5	2/5	1/5	$P^{(4)}$
Jan.15	3	5	0.1235196	1/13	1/13	2/13	6/13	3/13	$P^{(5)}$
Jan.14	4	6	0.0819807	0	3/14	2/7	3/14	2/7	$P^{(6)}$
Jan.13	5	7	0.1132639	1/5	0	0	2/5	2/5	$P^{(7)}$
Jan.12	4	8	0.0871448	1/13	1/13	5/13	3/13	3/13	$P^{(8)}$
P_i (the sum of weights)				0.103	0.087	0.224	0.337	0.249	

Step (7). The weighted sum of every same state prediction probability is taken as the prediction probability of the state

$$P_i = \sum_{k=1}^{m} w_k P_i^k, \ i \in E, \tag{5}$$

where $\max\{P_i,\ i \in E\}$ is the prediction probability. In Table 5, the state of total irradiation on the Jan. 20 is 4, the result agrees with the actual value. And so on, we predict the state for the next 5 day; the results are also shown in the Table 5.

The same method is used for calculation of prediction of humidity and soil temperature. The results are shown in Table 6 and Table 7, respectively.

Table 5. The results simulation and prediction of total irradiation

Date	State					Prediction state	Actual state	Actual value
	1	2	3	4	5			
Jan. 20	0.1032	0.0870	0.2235	0.3373	0.2490	4	4	5767.4
Jan. 21	0.1500	0.1198	0.2598	0.3238	0.1466	4	4	5680.0
Jan. 22	0.0921	0.2133	0.3332	0.2417	0.1198	3	3	4424.4
Jan. 23	0.0922	0.1639	0.2690	0.3165	0.1583	4	4	5803.4
Jan. 24	0.1502	0.1324	0.2677	0.3298	0.1199	4	4	5066.9

Table 6. The results simulation and prediction of humidity

Date	State					Prediction state	Actual state	Actual value
	1	2	3	4	5			
Jan. 20	0.4948	0.1450	0.2042	0.0854	0.0706	1	1	39.320
Jan. 21	0.5123	0.1360	0.2270	0.0749	0.0500	1	1	41.540
Jan. 22	0.5003	0.1576	0.2368	0.0517	0.0536	1	1	45.716
Jan. 23	0.4406	0.1725	0.2745	0.0581	0.0542	1	1	39.204
Jan. 24	0.4633	0.1535	0.2974	0.0497	0.0361	1	1	43.808

Table 7. The results simulation and prediction of soil temperature

Date	State					Prediction state	Actual state	Actual value
	1	2	3	4	5			
Jan. 20	0.2321	0.1363	0.1482	0.3286	0.1548	4	4	16.412
Jan. 21	0.3115	0.1432	0.0590	0.1474	0.3389	5	5	17.596
Jan. 22	0.1384	0.1309	0.1637	0.0743	0.4928	5	5	17.221
Jan. 23	0.1344	0.1282	0.1975	0.0337	0.5062	5	5	17.412
Jan. 24	0.1017	0.2254	0.1801	0.1930	0.2998	5	5	17.152

In the practical production, the temperature of inner greenhouse is major consideration. When the temperature is over 20 °C, the environment is suitable of the growth of vegetable, so the indicator that the temperature is over 20 °C is a very important parameter. Through predicting inner irradiation, humidity and soil temperature, we can establish a relationship of the three and temperature of inner greenhouse. The results are shown in Table 8.

Table 8. The relationship of temperature and the three (irradiation, humidity and soil temperature)

Parameter	Value	The count of temperatures >20 °C	Value	Temperature > 20 °C
Irradiation	$> 4400\,\mathrm{W/m^2}$	> 13	$< 2000\,\mathrm{W/m^2}$	0
Humidity	< 60	> 18	> 80	0
Soil temperature	> 15.5 °C	> 12	< 13.5 °C	0

From the above table, we can observe that the prediction state is matched with actual state. This indicates that the method is effective, and we can use the method to predict the environmental parameters of greenhouse.

Predicting the main impact factors of the inner environment of solar greenhouse with Weighted Markov Chain, which has been built, presents the predictive state of irradiation, humidity and soil temperature in the solar greenhouse, and shows the actual state and actual values.

In summary, the above data show that the effect of simulation and prediction of impact factors is satisfactory, the prediction agrees well with the actual situation, and the model is successful. Table 8 shows that irradiation, humidity and soil temperature are the important impact factors. When irradiation value is over $4400\,\mathrm{W/m^2}$, the count of the inner temperature over 20 °C is over 13, when irradiation value is under $2000\,\mathrm{W/m^2}$, the count is 0. When humidity value is under 60, the count of inner temperature over 20 °C is over 18 when the value is over 80, the count is 0. When soil temperature is over 15.5 °C, the count is over 12, when the value is under 13.5 °C, the count is 0. So we can predict whether the weather is suitable for the growth of plants in the short term, and provide the basis for the environment control.

4. Conclusion

In the study, the Weighted Markov Chain method was adopted to predict the major environment impact factors in the solar greenhouse. The results show that predictive accuracy of the model is high, and predictive states are according with the actual states. The predictive results are objective and reasonable, and the impact factors have a strong nonlinear relationship. Therefore, the model can be applied to forecast the indicators (including irradiation, humidity and soil temperature).

Adopted Sequential Clustering Method to classify the data, make the data groups more reasonable. With the help of MATLAB, it is easy to classify the data.

Weighted Markov Chain method is adopted to forecast the inner environment of solar greenhouse and which will be targeted to improve the management level in the solar greenhouse. It can also provide a realistic approach to environment control in the greenhouse.

References

[1] G. Tong, D. M. Christopher, T. Li, T. Wang: *Passive solar energy utilization: A review of cross-section building parameter selection for Chinese solar greenhouses.* Renewable and Sustainable Energy Reviews *26* (2013), 540–548.

[2] o. Ozgener, A. Hepbasli: *Experimental performance analysis of a solar assisted ground-source heat pump greenhouse heating system.* Energy and Buildings *37* (2005), 101–110.

[3] Y. Han, X. Xue, X. Luo, L. Guo, T. Li: *Establishment of estimation model of solar radiation within solar greenhouse.* Transactions of the Chinese Society of Agricultural Engineering *30* (2014), No. 10, 174–181.

[4] C. W. Ma, S. M. Zhao, J. Y. Cheng, N. Wang, Y. C. Jiang, S. Y. Wang, B. M. Li: *On establishing light environment model in Chinese solar greenhouse.* Journal of Shenyang Agricultural University *44* (2013), No. 5, 513–517.

[5] Z. Y. Zuo, H. P. Mao, X. D. Zhang, J. Hu, L. Han, J. Ni: *Forecast Model of Greenhouse Temperature Based on Time Series Method.* Transactions of the Chinese Society for Agricultural Machinery *41* (2010), No. 11, 173–177,182.

[6] G. A. Meehl, W. M. Washington, T. M. Wigley, J. M. Arblaster, A. Dai: *Solar and greenhouse gas forcing and climate response in the 20th century.* Journal of Climate *16* (2003) 426–444.

[7] W. Zhou, X. C. Wang: *Constrained predictive control model for greenhouse temperature.* Xinjiang Agricultural Sciences *51* (2014), No. 6, 1015–1021.

[8] P. Taalas, J. Kaurola, A. Kylling, D. Shindell, R. Sausen, M. Dameris, V. Grewe, J. Herman, J. Damski, B. Steil: *The impact of greenhouse gases and halogenated species on future solar UV radiation doses.* Geophysical Research Letters *27* (2000), No. 8, 1127–1130.

[9] P. S. Magnusson, M. Christensson, J. Eskilson, D. Forsgren, G. Hallberg, J. Hogberg, L. Larsson, A. Moested, B. Werner: *Simics: A full system simulation platform.* Computer *35* (2002), No. 2, 50–58.

[10] R. G. Sargent: *Verification and validation of simulation models.* Winter Simulation Conference, 4–7 December 2005, Orlando, FL (2005), European Journal of Operational Research *66*, (2013), No. 2, 250–258.

[11] F. Xu, C. Ma: *Daily change and math-expression method of outside temperature in winter for greenhouse environmental analysis.* Transactions of the Chinese Society of Agricultural Engineering *29* (2013), No. 12, 203–209.

[12] A. J. Hamilton, G. Hepworth: *Accounting for cluster sampling in constructing enumerative sequential sampling plans.* Journal of Economic Entomology *97* (2004), No. 3, 1132–1136.

[13] P. Jiang, Y. Huo, L. Zhang, J. Luo, H. Li: *A wind speed time series model based on advanced first-order Markov chain approach.* Automation of Electric Power Systems *38* (2014), No. 19, 22–27.

[14] Y. Li, L. He, J. Niu: *Forecasting power generation of grid-connected solar PV system based on Markov chain.* Acta Energiae Solaris Sinica *35* (2014), No. 4, 611–616.

[15] X. Yang, H. Liu, B. Zhang, Y. Xiao: *Similar day selection based on combined weight and photovoltaic power output forecasting.* Electric Power Automation Equipment *34* (2014), No. 9, 118–122.

Received September 12, 2017