

A Method to Forecast Photo-voltaic Power Outputs Based on Markov Chain

ZHANG Hua^{1,a}, HAN Xiaojuan^{1, b}, DOU Jin-yue^{2,c}

1 School of Control and Computer Engineering, North China Electric Power University,
Beijing, China, 102206

2 School of Economic and Management, North China Electric Power
University, Beijing, China, 102206

^a656305355@qq.com, ^bwmhxj@163.com, ^c719086129@qq.com

Keywords: Photo-voltaic power prediction; Markov chain; State space; Prediction step

Abstract: Large-scale grid-connected operation of photo-voltaic power station will influence the stable and economic operation of the grid. Effective forecast method of photo-voltaic power outputs is benefit for the whole scheduling and reasonable operating of the grid. A prediction method based on Markov chain absolute distribution is proposed to realize the short-term forecast of photo-voltaic power outputs. The impacts of different state spaces numbers and different prediction steps on the prediction accuracy are analyzed. The simulation results show the validity of this method using history data at a photo-voltaic station.

Introduction

With continuous development of solar photo-voltaic power generation technology and the rapid growth of global photo-voltaic installed capacity, solar power connected to the grid has become an important developing direction of solar photo-voltaic power generation[1]. However, photo-voltaic power generation has the characteristics of randomness, volatility and intermittent, large-scale grid-connected operation of photo-voltaic power station will lead to a great negative impact on the security, stability and economy of power grid operation[2,3]. Therefore, an accurate predicting method of photo-voltaic power outputs can effectively reduce this kind of negative impact[4], which can also help dispatching departments of the grid to adjust the scheduling plan in real time and integrally arrange the coordination operation between photo-voltaic power generation and other power supplies in order to achieve reasonable operation requirements.

At present, prediction methods of photo-voltaic power at home and abroad are mainly focused on neural network, support vector machine and Markov chain. A short-term prediction method of photo-voltaic power based on BP neural network is proposed in [5], combining additional momentum method with self-adaptive learning rate to improve the speed and accuracy of prediction. But the prediction accuracy of this method in the case of low light intensity is undesirable. A support vector machine (SVM) regression model of photo-voltaic system output is proposed in [6] by analyzing non-linear relationships between photo-voltaic system and its influence factors to achieve the short-term prediction of photo-voltaic power. A generalized regression neural network (GRNN) power prediction model based on the improved similar day algorithm is proposed in [7]. The prediction accuracy of this method is superior to that of traditional BP neural network obviously. As described in [8,9], Markov chain method is used to correct the errors of BP neural network prediction model. The prediction accuracy after the correction is higher than that of BP neural network model. But the prediction deviation in morning and evening time is quite large. Short-term forecasting method of photo-voltaic power outputs based on Markov chain is proposed in [10,11].

In this paper, a prediction model based on Markov chain absolute distribution is proposed to realize short-term forecasting of photo-voltaic power outputs, according to history data at a photo-voltaic power station. The impact of different state space numbers and different prediction steps on the prediction accuracy is analyzed in this paper. The simulation results show that the prediction model

based on Markov chain absolute distribution is feasible in short-term prediction of photo-voltaic power outputs.

Prediction principles of Markov chain absolute distribution

Definition of Markov chain. One of the most notable feature of Markov chain is the “no memory”, which means “the state in the future” is only related to “the state at present” and unrelated to “the state in the past”. Assume that a random process $\{X(t), t \in T\}$ meets the following conditions.

- 1) The time set is a set of nonnegative integers $T = \{n=0, 1, 2, \dots\}$. For each moment, the state space is a discrete set S . $S = \{S_0, S_1, \dots, S_n, \dots\}$.
- 2) For any positive integer m, k and any nonnegative integer $L_n > \dots > L_2 > L_1$ (of which $m+k > m > L_n$), as well as the corresponding state at the moment $X(t)$, $S_j, S_i, S_{L_n}, \dots, S_{L_2}, S_{L_1}$, there is the following formula established,

$$P\{X(m+k) = S_j \mid X(m) = S_i, X(L_n) = S_{L_n}, \dots, X(L_2) = S_{L_2}, X(L_1) = S_{L_1}\} \\ = P\{X(m+k) = S_j \mid X(m) = S_i\} \quad (1)$$

where the state of $X(t)$ is $X(m+k) = S_j$ in the moment $m+k$, the probability is only related to the state in the moment m ($X(m) = S_i$) and unrelated to the state before the moment m .

Predicting process of Markov chain absolute distribution. Using Markov chain prediction model whose prediction step is 1 and the initial vector distribution to forecast the absolute distribution of a list of mutually dependent random variables in the next period of time is the prediction method based on Markov chain absolute distribution. The specific steps of this method are as follows.

- 1) Choose the model data, then pretreat and correct the data.
- 2) Divide the state space of the processed data.

3) Get the transfer frequency matrix and the one-step transfer probability matrix through the statistics and calculation. Assume that $X(t) (t=0, 1, 2, \dots)$ is a sample sequence of Markov chain, the state space of which is $S = \{S_0, S_1, \dots, S_n\}$. f_{ij} is the transfer frequency when $X(t)$ transfers from the state i to the state j through one step, $i, j \in S$. The $n \times n$ matrix composed of f_{ij} is the transfer frequency matrix. f_i is the total frequency when the state of $X(t)$ is i . That is the row sum of line i of the transfer frequency matrix. Then the transfer probability can be calculated as follows.

$$p_{ij} = \frac{f_{ij}}{f_i} \quad (2)$$

k -step transfer probability matrix $P(k)$ is

$$P(k) = \begin{bmatrix} p_{00}(k) & p_{01}(k) & p_{02}(k) & \dots & p_{0n}(k) \\ p_{10}(k) & p_{11}(k) & p_{12}(k) & \dots & p_{1n}(k) \\ p_{20}(k) & p_{21}(k) & p_{22}(k) & \dots & p_{2n}(k) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n0}(k) & p_{n1}(k) & p_{n2}(k) & \dots & p_{nn}(k) \end{bmatrix} \quad (3)$$

- 4) Examine the Markov property. Calculate the marginal probability p_j .

$$p_j = \frac{\sum_{i=1}^n f_{ij}}{\sum_{j=1}^n \sum_{i=1}^n f_{ij}} \quad (4)$$

When n is quite large, χ^2 statistical magnitude is

$$\chi^2 = 2 \sum_{i=1}^n \sum_{j=1}^n f_{ij} \left| \log \frac{p_{ij}}{p_j} \right| \quad (5)$$

If χ^2 is more than $\chi_{\alpha}^2((n-1)^2)$, it can be assumed that this sequence has the Markov property, or this sequence can not be processed and analysed as a Markov chain.

5) Get the initial probability vector through the initial datas. Combined with one-step transfer probability matrix, calculate and predict as the following formula.

$$P_1 = P_0 P(1) \quad (6)$$

6) Calculate the predicted value. The state space of the predicted value is the corresponding state space of the maximum in the probability distribution got in the last step. The predicted value is the average in this state space.

7) In order to verify the prediction effect, the error of prediction results need to be analysed. In this paper, root-mean-square error (RMSE), mean absolute error (MAE) and mean relative error (MRE) are used to analyse the prediction results.

Example of photo-voltaic power prediction using the Markov chain

Selection of photo-voltaic power. The history data of a 40 MW photo-voltaic power station in a year are selected in this paper. The sampling interval is 1 minute. In order to achieve the short-term prediction of photo-voltaic power whose prediction time interval is 15 minutes, the raw data of photo-voltaic power are processed. Then select the power data of 29 days in June as a data sample to build the prediction model, the 7 days of this data sample is shown in Fig.1(a)

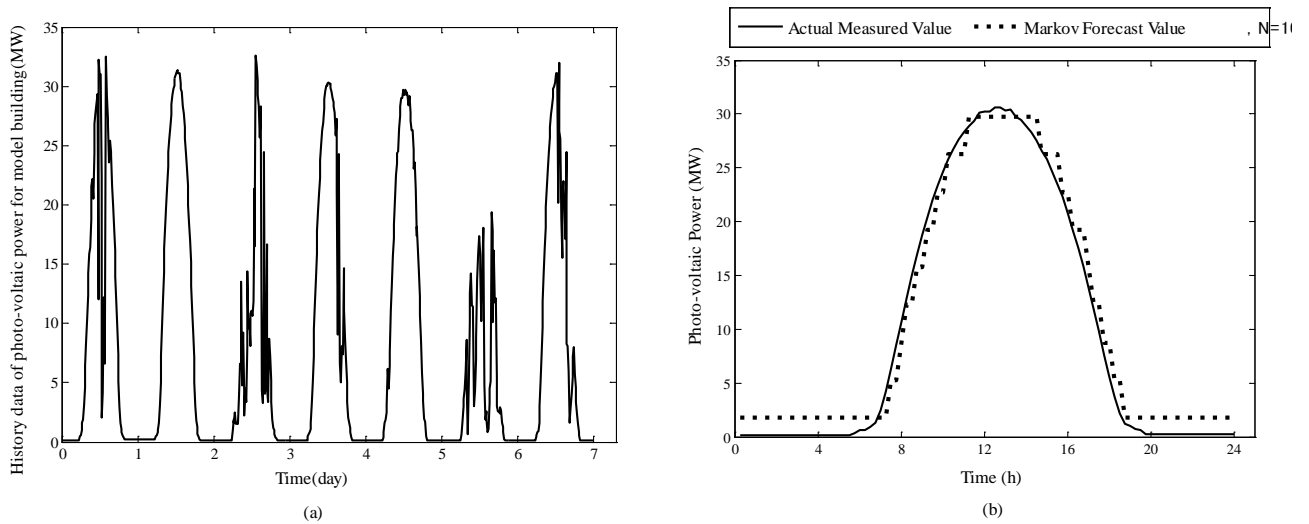


Fig.1. (a) 7 days' actual measured value of photo-voltaic power in June (b) Prediction results of photo-voltaic power when N=10

State division of the Markov chain. Due to the large fluctuation of photo-voltaic power, the equisection method is adopted to divide the history power data into N status. The upper limit P_{max} is the maximum of the history power data. Here, the P_{max} is 35MW. The interval $[0, P_{max}]$ is divided into N equal interval. The length of each interval is $l = P_{max}/N$. When $N=10$, $l=3.5$. The state division is shown in Table 1 and the prediction process when $N=10$ is described in detail as follows.

Table1

The state division when N=10

State sype	S_1	S_2	S_3	S_4	S_5
Interval range	[0,3.5]	[3.5,7]	[7,10.5]	[10.5,14]	[14,17.5]
State sype	S_6	S_7	S_8	S_9	S_{10}
Interval range	[17.5,21]	[21,24.5]	[24.5,28]	[28.5,31.5]	[31.5,35]

Calculation of one-step transfer probability matrix. Count the frequency that data sample in each 10 status and get the transfer frequency matrix. According to the Eq.2, the one-step transfer probability matrix can be calculated as follows.

$$P = \begin{bmatrix} 0.9664 & 0.0294 & 0.0024 & 6.0060 \times 10^{-4} & 6.0060 \times 10^{-4} & 0 & 6.0060 \times 10^{-4} & 0 & 0 & 0 \\ 0.2110 & 0.4083 & 0.2431 & 0.0780 & 0.0229 & 0.0138 & 0.0138 & 0.0046 & 0 & 0.0046 \\ 0.0223 & 0.2626 & 0.3687 & 0.1788 & 0.0894 & 0.0279 & 0.0168 & 0.0056 & 0.0112 & 0.0168 \\ 0.0229 & 0.1374 & 0.1985 & 0.3130 & 0.1908 & 0.0382 & 0.0305 & 0.0305 & 0.0076 & 0.0305 \\ 0.0167 & 0.0500 & 0.0833 & 0.1833 & 0.2833 & 0.2083 & 0.0677 & 0.0583 & 0.0167 & 0.0333 \\ 0 & 0.0112 & 0.0787 & 0.0449 & 0.2360 & 0.3258 & 0.2135 & 0.0562 & 0.0225 & 0.0112 \\ 0.0102 & 0.0306 & 0.0714 & 0.0510 & 0.0816 & 0.1327 & 0.3571 & 0.2143 & 0.0408 & 0.0102 \\ 0 & 0.0159 & 0.0238 & 0.0238 & 0.0556 & 0.0317 & 0.1349 & 0.0528 & 0.1746 & 0.0159 \\ 0 & 0.0157 & 0.0236 & 0.0157 & 0.0157 & 0.0157 & 0.0630 & 0.1417 & 0.6929 & 0.0157 \\ 0 & 0.0333 & 0 & 0.1333 & 0.0333 & 0.1000 & 0 & 0.1000 & 0.2000 & 0.4000 \end{bmatrix}$$

Checkout of the Markov property. According to the Eq.4 and Eq.5, the marginal probability p_j and statistical magnitude χ^2 can be calculated. $\chi^2=2235.5$. The prescribed significance level is $\alpha=0.05$. Lookup the table we can get $\chi^2(81)=102.7$. That means $\chi^2 > \chi^2((n-1)^2)$. So this state sequence of photo-voltaic power meets the Markov property.

The determination of the initial state distribution. When to predict the first data of the 30th day, the initial vector is the last data of the 29th day (the 2784th data point). The initial state S_0 is decided by the state of the current moment. The 2784th power data is $p_w=0.1000$ and is in the state S_1 . So the initial state vector is $P_0=[1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$.

Prediction calculation. According to the Eq.6, the state of the next point can be predicted. Choose the maximum of the vector P_1 as the distribution probability of the next point. Choose the mean value of the state interval corresponding to this distribution probability as the output value. The forecast value and the actual measured value are compared in Fig.1(b).

Error analysis. According the forecast value, we can calculate that $RMSE=0.0457$, $MAE=1.4313$, $MRE=0.0409$. Thus, the forecasting results exist some problems that the error is quite large at some point. The division of the state space is related to the prediction accuracy, so it is supposed to adjust the state spaces numbers N and do further analysis. The power prediction results when $N=10, N=40, N=80, N=120$ are compared in Fig.2, and the three error evaluation indicators are shown in Table 2.

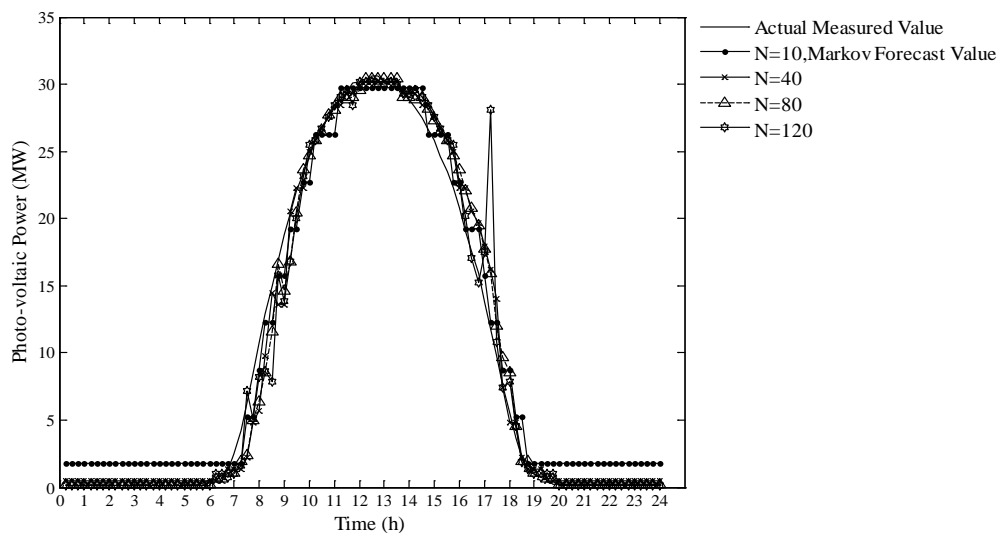


Fig.3. Comparison of the power prediction results when $N=10, N=40, N=80, N=120$

Table2

Three error evaluation indicators of different N						
Indicators	N=10	N=30	N=50	N=80	N=100	N=120
<i>RMSE</i>	0.0457	0.0460	0.0457	0.0447	0.0379	0.0631
<i>MAE</i>	1.4313	1.0443	0.9497	0.8842	0.7174	0.9136
<i>MRE</i>	0.0409	0.0298	0.0271	0.0253	0.0205	0.0261

A closer look at Table 2 reveals that as N increases, root-mean-square error (RMSE), mean absolute error (MAE) and mean relative error (MRE) are all firstly decreased and then increased when N increases to a certain extent.

Single-point photo-voltaic power prediction with more prediction steps.The single-point Markov chain prediction method based on absolute distribution with 1 step can forecast the data of next 15 minutes every time ,with 2 steps can forecast the data of next 30 minutes every time and 4 steps can forecast the data of next 1 hour every time. The increase of the step length can extend the prediction time span, but at the same time ,will change the prediction accuracy. The detailed changing is shown in Fig.3.

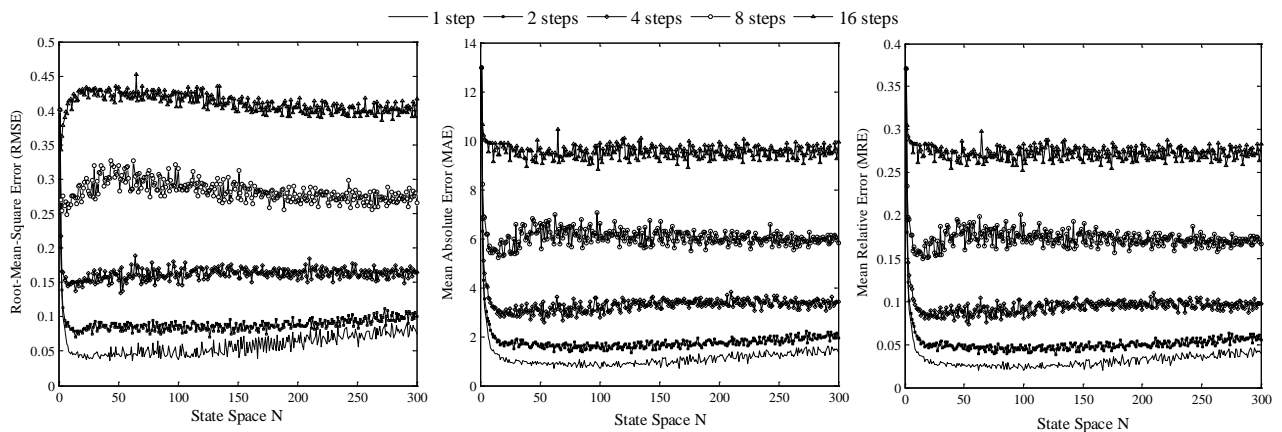


Fig.3. Comparison of error evaluation indicators with different prediction steps

When the prediction step length changes from 1 to 2,4,8,16 and even higher, the prediction accuracy is gradually reduced although the prediction time span is extended. This is because the prediction accuracy of the single-point Markov chain prediction method based on absolute distribution is largely depends on the similarity between the prediction data and the measured data used to generate the prediction model. The longer the step length, the larger the prediction data span every time, the lower the accuracy. In consequence, applied to the super short-term photo-voltaic power prediction, the Markov chain prediction model based on absolute distribution has quite great prediction effect. But in long-term photo-voltaic power prediction, the prediction effect reduces greatly and it need to combine with other prediction methods to achieve better prediction effect.

Summary

In this paper, according to the historical power data of a photo-voltaic power station,a Markov chain prediction model based on absolute distribution is built. According to single-point photo-voltaic power prediction model with one prediction step, the impact of the state spaces number Non the prediction accuracy is analyzed. According to single-point photo-voltaic power prediction model with more prediction steps, the impact of different prediction steps on the prediction accuracy is analyzed. The simulation results show that the prediction method based on Markov chain absolute distribution has higher prediction accuracy in super short-term photo-voltaic power prediction. But in long-term prediction, the forecast error of this method is too large to meet the standards for grid connected system and still need further study.

References

- [1] Li Fen, Chen Zhenghong, He Mingqiong, et al. Review of Status and Prospect of Solar Photo-voltaic Power Generation [J]. Water Resources and Power, 2011,29(12): 188-192.
- [2] Zhang Zuoyu, Wang Jingyan, Li Yanqing. The current situation and prospect of photo-voltaic power generation [J]. Science & Technology Information,2010,2(17): 352-352.

- [3] European Photo-voltaic Industry Association (EPIA).Global Market Outlook for Photo-voltaics until 2015[R].Brussels: EPIA,2011.
- [4] Lu Jing, Zhai Haiqing, Feng Shuanglei, et al. Physical Method for Photo-voltaic Power Prediction[J]. East China Electric Power, 2013, 41(2): 380-384.
- [5] Ding Ming, Wang Lei,Bi Rui. A short-term prediction model to forecast output power of photo-voltaic system based on improved BP neural network[J]. Power System Protection and Control, 2012, 40(11):93-99.
- [6] Li Ran, Li Guangmin . Photo-voltaic power generation output forecasting based on support vector machine regression technique[J] .Electric power, 2008,41(2):74-78.
- [7] Yang Lei, Dong Haiying, Ma Bo, et al. Power forecasting of photo-voltaic plant based on improved similar day [J]. Chinese Journal of Power Sources,2013,37(7):1176-1179.
- [8] Jiang Qiaona, Chen Zhong. The study of solar power forecasting based on BP-Markov method[J]. Power Demand Side Management, 2011, 13(6): 21-24.
- [9] Wu Xuelian, Du Hongji. Photo-voltaic Power Generation Forecast Based on BP Neural Network and Markov Chain[J]. Electrotechnics Electric (Jiangsu Electrical Apparatus), 2014,(3):22-27.
- [10]Li Yingzi,Niu Jincang,Ru Luan,el al.Research of multi-power structure optimization for grid-connected photo-voltaic system based on Markov decision-making model [C]//International Conference on Electrical Machines and Systems,Wuhan,China,2008;2607-2610
- [11]Ding Ming. A Method to Forecast Short-Term Output Power of Photo-voltaicGeneration System Based on Markov Chain[J]. Power System Technology, 2011,35(1): 152-157.