ICIC Express Letters ICIC International ©2017 ISSN 1881-803X
Volume **11**, Number **1**, January **2017** pp. **95–100**

PRACTICAL DATA-DRIVEN ANALYSIS ON STOCHASTIC PROPERTY OF POWER ASSUMPTION IN SMART GRIDS

Yiwen Le and Jinghan He

School of Electrical Engineering
Beijing Jiaotong University
No. 3, Shangyuancun, Haidian District, Beijing 100044, P. R. China
leyiwen13117376@sina.com; jhhe@bjtu.edu.cn

Received June 2016; accepted September 2016

Abstract. *Smart Grid is an emerging power system to provide efficient electricity management by leveraging the advanced information technologies. Power optimized distribution is a main concern in Smart Grid. This letter first makes a practical analysis for the stochastic features of power assumption based on the practical electricity consumption dataset. In particular, we build a stochastic theory model for power assumption based on the Markov chain model, and derive a stochastic assessment method in theory. Then, we analyze how the power assumption dataset satisfies the stochastic feature. According to the statistics analysis, the results show that practical power assumption may not fit the Markov stochastic model. Therefore, the inaccuracy can be introduced if directly adopting the straightforward Markov based prediction without any amendment. To our best knowledge, this study is the first attempt to test the accuracy of the stochastic property of power assumption in Smart Grids.*
Keywords: Stochastic property, Markov chain, Power assumption, Data collection, Smart Grids

1. **Introduction.** Electrical power industry is undergoing rapid developments and moving to the Smart Grid [1]. Smart Grid is believed as a next generation electric power grid system with many modern information technologies, such as Internet, automated control, power metering, and modern energy management [2]. In general, a Smart Grid includes a variety of operational and energy measures for power prediction, such as smart meters, smart appliances, and smart sensing components. Accurate forecasting provides benefits in the real-time power generation, efficient energy management, and economic cost saving. However, the prediction accuracy is highly related with the geographical locations as well as the time periods. Stochastic property of power assumption greatly increases the power prediction difficulty in Smart Grids [3,4].

The power prediction models can be roughly divided into four types: regression method, time series method, expert-based method and neural network [5,6,10]. The prediction accuracies rely on the distribution of original series as well as a large amount of observed data. Recently, some studies conduct the power assumption predictions using the Markov-based models. For example, Yang et al. [6] presented a support vector machine enhanced Markov model for short-term wind power forecast. Zhao et al. [7] proposed a high-order Markov chain based time-varying weighted average method to predict the monthly electricity consumption. Zhou et al. [8] described a fuzzy probability-based Markov chain model for forecasting regional long-term electric power demand. Ardakanian et al. [9] proposed parsimonious Markovian reference models of home load. Kumar and Jain [10] proposed a Grey-Markov model to forecast the consumption of conventional energy. He et al. [11] motivated data-driven tools to perceive the complex grids in high dimension and proposed a big data architecture design for Smart Grids based on random matrix theory. However, most of them only give an assumption considering that power assumption follows a Markov process without a detailed analysis of the stochastic properties of power assumption. Any impractical artificial assumption may have a great impact on the prediction results.

This letter focuses on an in-depth investigation on the stochastic properties of power assumption based on a set of practical electricity consumption. The main contributions are summarized as follows.

(1) This letter collects a set of first-hand practical electric power data in a community, which may provide benefits to more researchers on power assumption analysis.
(2) This letter builds a stochastic theory model for power assumption based on the Markov chain model, and derives a simple stochastic assessment method. It can be used to evaluate how power assumption fits the Markov randomness.
(3) This letter makes experimental analysis of practical power assumption dataset and finds that power assumption may not fit the Markov stochastic model. Finally, this letter proposes some guides and suggestions for the power assumption prediction.

The rest of this letter is organized as follows. Section 2 introduces the theory modeling for the stochastic assessment method. Section 3 elaborates the experiments details. Section 4 presents the statistics analysis and results. Finally, Section 5 concludes this work.

2. **Theory Modeling.** Markov process is usually used to model a random system in which the next state only depends on the current state. In this section, we build a stochastic theory model for power assumption based on the Markov chain model.

Let $P_{ij}(t) = P\{X_{t+s} = j/X_s = i\}$ denote the probability that the state being in state i presently and in state j a time t later. A Markov chain holds the following natures:

$$P_{ij}(t) \geq 0,\ i, j \in E \tag{1}$$

$$\sum_{j\in E} P_{ij}(t) = 1 \tag{2}$$

$$P_{ij}(t+s) = \sum_{k\in E} P_{ik}(t)P_{kj}(s) \tag{3}$$

According to the Taylor's formula,

$$\begin{aligned} P_{ij}(\Delta t) &= P_{ij}(0) + P'_{ij}(\xi)\Delta t + o(\Delta t) \\ &= P_{ij}(0) + q_{ij}\Delta t + o(\Delta t), \qquad \xi \in (0, \Delta t) \end{aligned} \tag{4}$$

where q_{ij} is element in the transition rate matrix $\mathbf{Q}$, which is called the infinitesimal transition rate. It can be calculated by:

$$q_{ij} = \lim_{\Delta t\to 0} \frac{P_{ij}(\Delta t) - P_{ij}(0)}{\Delta t} \tag{5}$$

According to the Formula (5), when $\Delta t \to 0$, q_{ij} equals the derivative of transition probability function.

Consider

$$P_{ij}(0) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{6}$$

Then, we have

$$q_{ij} = \begin{cases} \lim\limits_{\Delta t\to 0} \dfrac{P_{ij}(\Delta t) - 1}{\Delta t}, & i = j \\ \lim\limits_{\Delta t\to 0} \dfrac{P_{ij}(\Delta t) - 0}{\Delta t}, & i \neq j \end{cases} \tag{7}$$

Let stochastic process $N_i(t)$ $(t \geq 0)$ denote the total number of state transitions after the Markov chain jumps into the state i. It is immediate that $N_i(t)$ is a counting process. The stochastic process $N_i(t)$ $(t \geq 0)$ is a Poisson process.

Let $P_{ii}(\Delta t)$ denote the probability that the state being in state i presently and in state i a time Δt later. When $t \to 0$, we have

$$\begin{aligned} P_{ii}(\Delta t) &= P\{X(t+\Delta t) = i, N_i(\Delta t) = 0/X(t) = i\} \\ &\quad + P\{X(t+\Delta t) = i, N_i(\Delta t) \geq 2/X(t) = i\} \\ &= e^{-\lambda_i \Delta t} + o(\Delta t) \\ &= 1 - \lambda_i \Delta t + o(\Delta t) \end{aligned} \tag{8}$$

Based on the theoretical analysis, we derive a simple stochastic assessment method for power assumption. We denote the lifetime of state i as a random variable T. Let $P_i(T \leq \Delta t)$ denote the probability that the Markov process at state i stays for no longer than Δt. Then, $P_i(T \leq \Delta t)$ is the complementary probability of $P_{ii}(\Delta t)$. From Formula (8), we can get

$$P_i(T \leq \Delta t) = 1 - P_{ii}(\Delta t) = 1 - e^{-\lambda_i \Delta t} \tag{9}$$

The lifetime of state i follows the exponential distribution and λ_i can be acquired by the curving fitting with the exponential distribution during the practical analysis.

3. **Experiments.** In this section, we make a practical analysis of power assumption in Smart Grids based on a large amount of power assumption data. To analyze the electricity consumption stochastic properties, we select a piece of power equipment of a community and collect the first-hand practical electric power data (this dataset can be obtained with the permission of the authors). The data is collected every half an hour and with duration of a month (form August 10th, 2015 to September 9th 2015). The dataset has about 1440 data samples. Figure 1 shows the change of the electric power within 14.5 days in a community. As we can see, electric power distributes between 200 kW to 400 kW; clearly, it presents periodic changes. Figure 2 further shows the change of electric power in one cycle, which is about one day.

In order to further explore its periodic variation, we calculate the average value of the electric power within a month at each moment of the day, and analyze the rules of change. The results are shown in Figure 3. We can see that the change trend of Figure 3 is basically the same as Figure 2, with two peaks of electric power in 12:00 and 19:30 around and a valley in current 5:00 each day. It shows that the electric power first decreases before 5 am and then increases until 10 am. From 13 pm, it further starts increasing until 21 pm, then the electric power decreases greatly until 0 am.

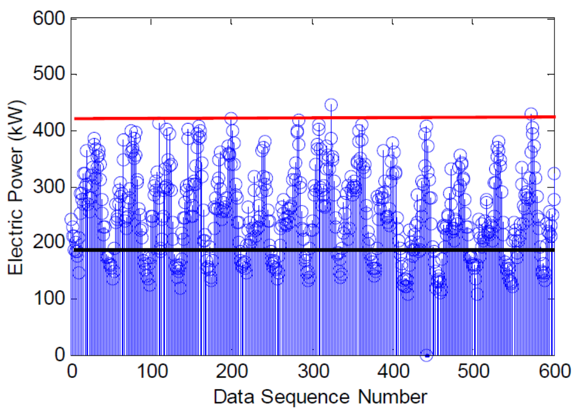

FIGURE 1. Electric power within 14.5 days

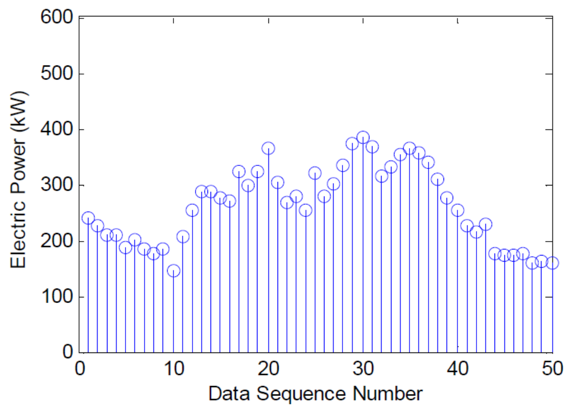

FIGURE 2. Electric power in one cycle

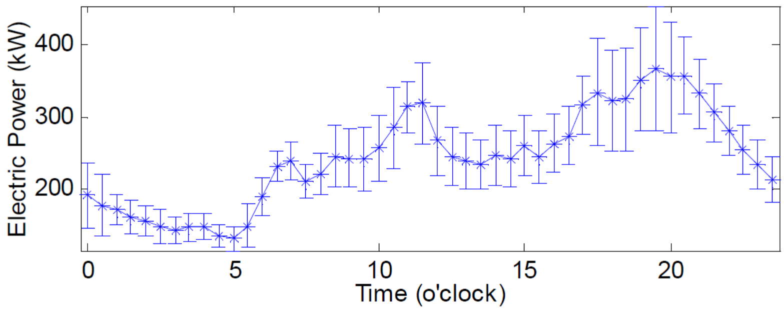

FIGURE 3. The change of average electric power

In order to analyze and process the data more convenient, as well as further explore the distribution rule followed by the duration of different electric power sections, we discretize the electric power. The formula of discretization is as follows:

$$S = \frac{P}{\max(P) + 1} \cdot DR + 1 \tag{10}$$

where S is the state, P is the electric power, and DR is the discretization range.

Figure 4 shows the change of average electric power after the discretization ($DR = 6$) in a day. We can see that the states 3 and 4 appear many times in a row. However, other discretization states appear rarely. Specifically, the state of 3 appears eleven times in total, and the state 4 occurs more often. For example, between 6:00 to 10:00, the state 4 appears eight times, and between 12:00 to 17:00, this state appears ten times. However, other states only appear a couple of times. The state 5 occurs nine times in total, but there is no continuous emergence of this state, which is not like the state three and four.

Figure 5 shows the statistical and fitting results of the state 4 ($DR = 6$). The histogram is the result of statistics, while the asterisk-marked curve is a fitting result with the negative exponential distribution. Clearly, we find that the fitting results are very poor.

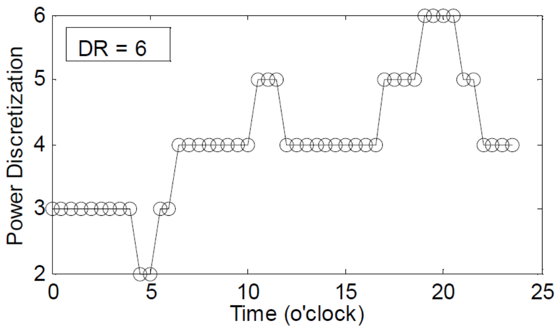

FIGURE 4. The power discretization ($DR = 6$)

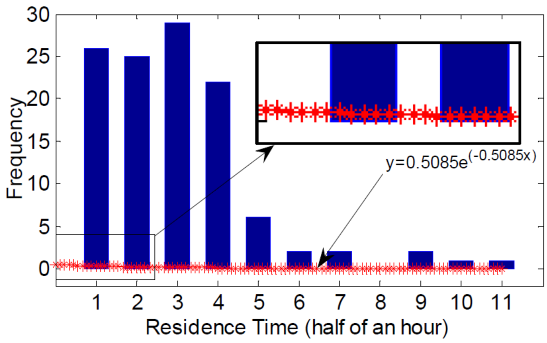

FIGURE 5. The statistics results ($DR = 6$)

4. **Statistics Analysis.** To further test the accuracy of results, we further analyze other cases with different DR values based on the collected datasets. For example, when $DR = 7$ and $DR = 8$, the results of statistics are as Figures 6-9.

Figure 6 shows the change of electric power average after discretization ($DR = 7$). In this situation, we find that the state 7 occurs in the least number of times, and the numbers are relatively bigger in other states. Furthermore, the state 3 appears very continuously. Figure 7 shows the statistical and fitting results of the state 4 ($DR = 7$). Figure 8 shows that the state 8 appears in the least number of times, and the occurrences of other states are relatively frequent, especially in the states 5 and 7. Figure 9 shows

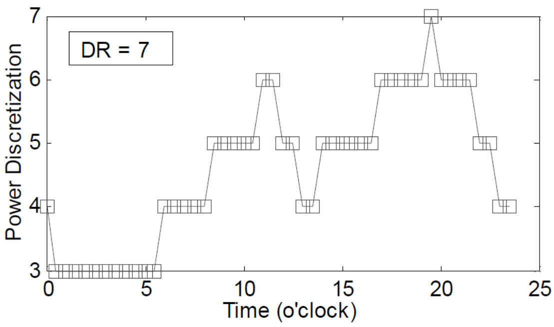

FIGURE 6. The power discretization ($DR = 7$)

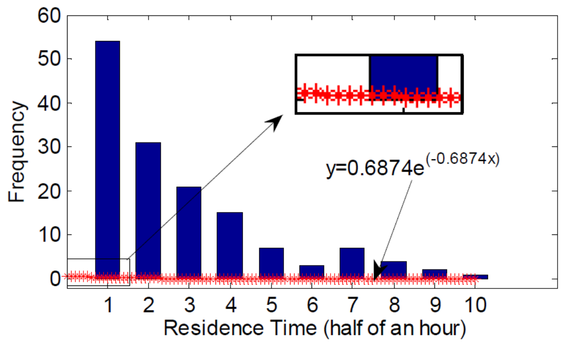

FIGURE 7. The statistics results ($DR = 7$)

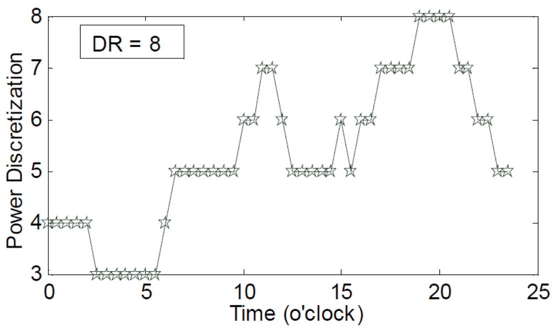

FIGURE 8. The power discretization ($DR = 8$)

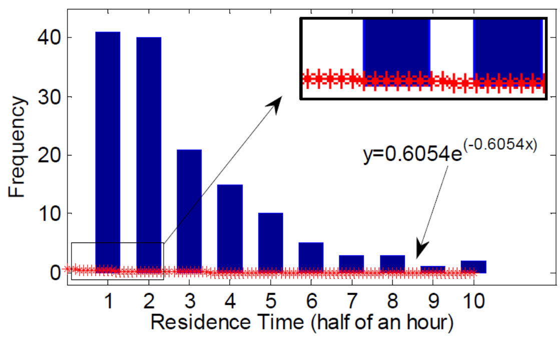

FIGURE 9. The statistics results ($DR = 8$)

TABLE 1. Statistics analysis results

Deviation	$\boldsymbol{DR = 6}$	$\boldsymbol{DR = 7}$	$\boldsymbol{DR = 8}$	$\boldsymbol{DR = 9}$
$\boldsymbol{S = 3}$	30.57	44.90	50.21	19.22
$\boldsymbol{S = 4}$	42.95	82.43	64.41	50.43
$\boldsymbol{S = 5}$	39.60	59.65	65.48	108.7

the statistical and fitting results of the state 4 ($DR = 8$). We further find that the fitting effect is very poor from the diagram.

Table 1 summarizes all statistics results. Through the above analysis, we can find that it is very poor when the duration of Electric Power State (EPS) is fitted by negative exponential distribution. Therefore, it is not strict that the duration of EPS is in accord with negative exponential distribution mentioned in previous studies. Therefore, the inaccuracy can be brought naturally if directly adopting the straightforward Markov based prediction without any amendment.

5. **Conclusions.** This letter makes a detailed analysis of stochastic property of power assumption through practical data sampling, averaging, status discretization, probability computing and model fitting. In particular, it analyzes how the power assumption satisfies the Markov stochastic features. In future work, we will further analyze power assumption data and explore its stochastic properties to build accurate power prediction solutions.

Acknowledgements. This work is supported by the Joint Project with National Energy Administration [grant number NY20150302].

REFERENCES

[1] K. Mets, J. A. Ojea and C. Develder, Combining power and communication network simulation for cost-effective smart grid analysis, *IEEE Communications Surveys & Tutorials*, vol.16, no.3, pp.1771-1796, 2014.

[2] H. Liang, A. K. Tamang, W. Zhuang and X. Shen, Stochastic information management in smart grid, *IEEE Communications Surveys & Tutorials*, vol.16, no.3, pp.1746-1770, 2014.

[3] H. Liang, A. Abdrabou and W. Zhuang, Stochastic information management for voltage regulation in smart distribution systems, *IEEE INFOCOM*, pp.2652-2660, 2014.

[4] M. Ghomi, M. Moeti and M. Azimi, Introduction of neural network model and fuzzy regression method for short term load forecasting in center regional of IRAN, *IEEE ICNEE*, pp.90-95, 2011.

[5] Y. He et al., A two-stage electricity demand forecasting model in the smart grid, *Power System Protection and Control*, vol.38, no.21, pp.167-172, 2010.

[6] L. Yang et al., Support vector machine enhanced Markov model for short term wind power forecast, *Spatio-Temporal Data Analytics for Wind Energy Integration*, Springer, pp.35-57, 2014.

[7] W. Zhao, J. Wang and H. Lu, Combining forecasts of electricity consumption in China with time-varying weights updated by a high-order Markov chain model, *Omega*, vol.45, pp.80-91, 2014.

[8] X. Zhou et al., A fuzzy probability-based Markov chain model for electric power demand forecasting of Beijing, China, *Energy & Power Engineering*, vol.5, no.4, pp.488-492, 2013.

[9] O. Ardakanian, S. Keshav and C. Rosenberg, Markovian models for home electricity consumption, *ACM SIGCOMM Workshop on Green Networking*, pp.31-36, 2011.

[10] U. Kumar and V. K. Jain, Time series models to forecast energy consumption in India, *Energy*, vol.35, no.4, pp.1709-1716, 2010.

[11] X. He, Q. Ai, R. C. Qiu et al., A big data architecture design for smart grids based on random matrix theory, *IEEE Trans. Smart Grid*, no.99, pp.1-13, 2015.