

# Enhanced Markov-Difference Based Power Consumption Prediction for Smart Grids

Yiwen Le<sup>†</sup> and Jinghan He<sup>\*</sup>

**Abstract** – Power prediction is critical to improve power efficiency in Smart Grids. Markov chain provides a useful tool for power prediction. With careful investigation of practical power datasets, we find an interesting phenomenon that the stochastic property of practical power datasets does not follow the Markov features. This mismatch affects the prediction accuracy if directly using Markov prediction methods. In this paper, we innovatively propose a spatial transform based data processing to alleviate this inconsistency. Furthermore, we propose an enhanced power prediction method, named by Spatial Mapping Markov-Difference (SMMD), to guarantee the prediction accuracy. In particular, SMMD adopts a second prediction adjustment based on the differential data to reduce the stochastic error. Experimental results validate that the proposed SMMD achieves an improvement in terms of the prediction accuracy with respect to state-of-the-art solutions.

**Keywords:** Smart grid, Power prediction, Markov-difference, Spatial mapping

## 1. Introduction

Electricity power has become an indispensable part in our daily life [1-5]. In recent years, increasing electricity consumption has caused terrible power shortage and brings a lot of inconvenience to the citizens. One reason is that traditional grid uses a rigid system with many deficiencies [6, 7], including: (i) the power transmission process is lack of flexibility; (ii) the self-recovery capability is low and its duration is long; (iii) the electric services only support simple user operations; (iv) information sharing is largely limited. These defects are becoming more and more obvious so that the grid is hard to constitute a real-time and efficient power system.

Smart Grid is a new type of modern power grid, which is formed by advanced information communication technology, sensing technology, analysis technology, decision technology and automatic control technology [8-10]. Different from the traditional ones, Smart Grid is featured by using two-way digital communications to control electric appliances at homes. Hence, it can save energy, reduce costs and increase reliability and transparency when delivering electricity from suppliers to consumers [11-13]. The concept of Smart Grid is originated in the United States, which announced that they would invest more than 40 billion dollars to promote grid modernization. Nowadays, Smart Grids have been promoted by many governments as a way of addressing energy independence, global warming and emergency resilience issues [37]. Hence, smart grids become more and more ubiquitous. Fig. 1 illustrates the

ubiquitous smart grids applications.

Electricity supply management is a key factor to guarantee the stable electricity in Smart Grids. The electricity prediction is the foundation for the electricity power management in Smart Grids [14-18]. The prediction accuracy will directly affect the effectiveness, reduce electricity costs, ensure the normal production, and improve economic benefits effectively. There are various available forecasting methods for power data in Smart Grid, such as trend extrapolation forecasting (TEF), regression analysis forecasting (RAF), artificial neural network forecasting (ANNF), gray theory forecasting (GTF), time series forecasting (TSF), wavelet analysis forecasting (WAF) and so on. Recently, many researches have studied different kinds of mathematics models for the electricity prediction [32]-[36]. These works make great contributions to improving prediction accuracy in Smart Grids. However, due to complex stochastic property of power consumption, real-time power monitor and prediction is still a challenging task in Smart Grids.

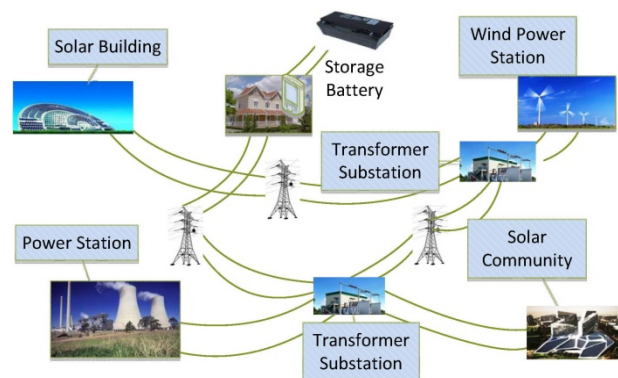


Fig. 1. Ubiquitous applications in smart grids

<sup>†</sup> Corresponding Author: School of Electrical Engineering, Beijing Jiaotong University, China. (leyiwen13117376@sina.com)

<sup>\*</sup> School of Electrical Engineering, Beijing Jiaotong University, China. (jhhe@bjtu.edu.cn)

Received: July 20, 2016; Accepted: February 24, 2017

Markov prediction model is widely used to predict a process satisfies the Markov property. However, with an investigation of practical power consumption datasets, our former work found that not all power consumption data is well fitted with the Markov stochastic property [19]. If directly using the Markov model, the prediction accuracy cannot be guaranteed. In this paper, we further investigate this interesting phenomenon and propose an enhanced power prediction method, named by Spatial Mapping Markov-Difference (SMMD) prediction, to improve the prediction accuracy. As far as we know, SMMD is the first work to adopt a second prediction calibration for the data difference to reduce the stochastic error.

The novelty and contributions of this paper are summarized as follows:

- We analyze and compare several kinds of prediction models and collect a set of practical electricity consumption data (with duration of one month). Based on the realistic data collections, we make a detailed analysis of the stochastic property of electricity consumption, and highlight that not all power consumption data fit the Markov property completely.
- We make a set of tests to adjust each set of electricity consumption. Spatial mapping with different parameters are first adopted to alleviate the inconsistency between the Markov model and the data stochastic property. This spatial mapping is reversible. Based on this preprocess, it allows us to use the Markov prediction model for the practical power data.
- We explore an enhanced Markov-difference prediction model. In particular, the power data generate two new data sequences. The first sequence is built by discretizing data into many different ranges; the second one is built by the data difference. Then, a Markov prediction is conducted for the data range sequence to predict the range. The gray prediction model is further considered to establish an infinite differences information model for the power fluctuation dataset based on the data difference sequence. This process aims to fine-adjust the primary predict result and can make up for the limitations of only using Markov prediction model.
- We test the performance of SMMD method in practical scenarios and carry out a set of experiments to analyze the prediction accuracy by comparing with the traditional Markov prediction method. Practical electricity consumption data are used as the experiment samples. Experimental results demonstrate that SMMD method can improve the prediction accuracy with respect to state-of-the-art solutions.

The rest of this paper is organized as follows. Section 2 summarizes the background and related works. Section 3 details the basic theory models and workflow of the proposed SMMD solution. Section 4 presents the experimental results and analysis. Finally, section 5 concludes this paper.

## 2. Background and Related Works

In this section, we make a brief survey on existing prediction methods, and summarize the related researches on power consumption in Smart Grids.

### 2.1 Traditional prediction methods

Traditional forecasting methods mainly include the trend extrapolation forecasting method, regression analysis forecasting method and time series forecasting method.

#### 2.1.1 TEF method

TEF is a prediction method using the data fitting [20]. Due to the randomness property, it analyzes data trends using the fitting equation and determines whether the tendency obeys a linear or nonlinear law. Power load forecasting usually has a certain trend, for example, the residential electricity consumption can be fitted by using trend analysis. TEF method is suitable for the prediction which has less data and short forecast period. It also determines the fitting tendency through analyzing data trends [4]. However, this method cannot change along with dynamic load, and it will increase the chance of calculation errors.

#### 2.1.2 RAF method

RAF method is used to set up the regression equations between different variables by analyzing the relation of independent variables and dependent variables. It predicts dependent variable according to independent variable, thereby determining relevant relationship. Therefore, RAF method is a high effective and practical method for short-term forecasts [22]. This method can be adopted to predict future development level, as long as we have the main factors affecting the variables. Regression analysis forecasting method can be adopted when we know the main factors affecting the market [20, 23]. This method is an effective and practical method for short-term forecasts. However, this method is not suitable for a long-term experiment.

#### 2.1.3 TSF method

TSF method is widely used in electric power load analysis [6]. It takes power load data as random sequence which changes along with the time, that is, present load data has inevitable connection with the past load data. Therefore, we can use the random model to analyze the statistical data, get the relevant parameters, and complete the model of future power load forecasting.

Time series analysis is a statistical method for dynamic data processing [22]. Based on stochastic process theory and mathematical statistics, it is used to study the statistical law of random data sequence to solve practical problems.

In other words, electric power load is a stochastic model with continuation of time. Then, a random model can be used to analyze the statistical data.

**2.2 Advanced prediction methods**

Modern forecasting method mainly includes artificial neural network forecasting method, wavelet analysis method and gray theory forecasting method.

**2.2.1 ANNF method**

ANNF method is a nonlinear and adaptive information processing system which is composed of a large number of processing units [21, 31]. It is put forward on the basis of modern neuroscience, and tries to process the information by simulating the mechanism of brain neural network and the way of memory. ANNF is configured with the following characteristics: (i) information processing occurs at many single elements called nodes, or neurons; (ii) signals are passed between nodes through connection links; (iii) each connection link has an associated weight that represents its connection strength; (iv) each node typically applies a nonlinear transformation called an activation function to its net input to determine its output signal.

**2.2.2 WAF method**

WAF method is carried out by wavelet transform which decomposes load into different frequency components [6, 8, 26]. It can be used to decompose the complex load sequence into a number of relatively simple sub sequences. The ability of variable metric analysis and local time interval information of wavelet transform makes it widely used in load forecasting and other aspects of power system. Extensive work have been done for more models, for example, Liang et al. presented a clear description of wavelet analysis forecasting method for Smart Grids [10].

**2.2.3 GTF method**

GTF is a deepening development of the forecasting

system [26-29]. It changes the original data into a regular pattern of the formation and then sets a model. The essence of the gray model is: with limited information to establish an infinite differences information model. Typically, GM (1,1) model is one of the most common gray modeling methods [23].

In fact, the gray model is a generation series model, and usually described by differential equations. In [26, 28, 29], detailed descriptions elaborate the practical application in our lives. This method has a wide application range, such as the forecast of land-using change and total freight of ocean shipments. What’s more, this method can be applied with a high accuracy. Some improved GM(1,1) forecast models are also proposed based on the Markov theory [27]. Therefore, the gray prediction can become the main means for power prediction. However, the original data is often not well applicable to the gray theory model [24]. How to deal with the data is also a key point for GTF method.

Through a careful comparison, we make a brief summary of the advantages and disadvantages of the aforementioned methods, which are concluded in the Table 1. All these methods provide available tools to forecast power consumption. However, each solution has its limitations highly dependent on the practical dataset. Each of them is difficult to achieve a required accuracy for the complex power consumption data, which have various and undiscovered stochastic properties for different data samples. Therefore, it is of great significance to further research some advanced, integrated and refined power consumption prediction approaches.

Our solution is motivated based on two main considerations:

(1) With an investigation of practical power consumption datasets, we found that power consumption is not well fitted with the Markov model. Therefore, it will degrade the accuracy of power prediction if directly using traditional Markov-based methods. Thus, we use a spatial mapping to preprocess the original data, to make them follow the Markov features.

(2) Markov chain forecasting model is suitable for

**Table 1.** A brief summary of power forecasting methods

Types		Advantages	Disadvantages
Traditional forecasting method	TEF	a) small demand of data b) short-term forecast	a) applicable to static load b) increase calculation error
	RAF	a) simple structure b) high prediction accuracy	a) not suit complex situations b) high data requirements
	TSF	a) little calculating work b) does not need much historical data	a) cannot fully consider the factors affecting the load data b) cannot be summed up
Modern forecasting method	ANNF	a) have the ability of self-learning, optimization calculation b) fully consider the impact of the load changes c) high prediction accuracy	a) long learning time will affect the learning objectives b) slow network convergence c) poor adaptability to unexpected events
	WAF	a) have adaptive capability b) high prediction accuracy	a) affected by weather, temperature, and humidity
	GTF	a) simple calculation b) does not consider the distribution law and the change tendency c) high accuracy of short-term prediction	a) only can be applied to load index with exponential growth trend b) not suitable for long-term prediction

describing the large random volatility of the prediction problem. However, due to the non-memory property of Markov chain, it easily misses some relations hidden in the power data sequences, hence, brings some prediction errors. To reduce this error, we use the gray prediction model to establish an infinite differences information model, according to the power fluctuation dataset. It can make up for the limitations of only using Markov prediction model.

Based on above considerations, we propose an enhanced prediction method named SMMD to guarantee the predict accuracy.

### 3. Enhanced SMMD Solution

#### 3.1 Basic models for SMMD

To facilitate the analysis, a set of power consumption data can be seen as a continuous time stochastic process. We use  $\{X(t), t \geq 0\}$  to denote a continuous time stochastic process and its countable state space is defined as  $E = \{i, j, i_1, \dots, i_k\}$ . We suppose this stochastic process is a Markov process. The state transition probabilities should follow the condition:

$$P\{X(t+s) = j | X(s) = i\} = P\{X(t+s) = j | X(s) = i, X(s_k) = i_k, \dots\} \quad (1)$$

where  $i, j, i_1, \dots, i_k \in E; s > s_k \geq \dots \geq s_1 \geq 0; t, s > 0; k \in N$ .

As we can see, the state  $j$  in time  $t+s$  only associates with its former state  $i$  in time  $s$ .

Furthermore, let  $p_j(t) = P\{X(t+s) = j | X(s) = i\}$  denote the transition probability that the current process being in state  $i$  and then in state  $j$  in a time  $t$  later. Then, all transition probabilities at the time  $t$  can make up a transition matrix  $P$  as shown follows:

$$P = \begin{bmatrix} P_{11} & P_{12} & \dots & \dots & P_{1j} & \dots \\ P_{21} & P_{22} & \dots & \dots & P_{2j} & \dots \\ \vdots & \vdots & \dots & \dots & \vdots & \dots \\ P_{i1} & P_{i2} & \dots & \dots & P_{ij} & \dots \\ \vdots & \vdots & \dots & \dots & \vdots & \dots \end{bmatrix} \quad (2)$$

Based on this, the lifetime of each state should follow the exponential distribution as for a Markov process [19]. We make a detailed analysis of the stochastic property of electricity consumption. We find that the lifetime of each state does not well follow the exponential distribution in the practical power consumption data. It means that not all power consumption data fit the Markov property completely. Therefore, the prediction accuracy will be reduced if directly using traditional Markov prediction methods. Fig. 2 shows the fitting results by the exponential

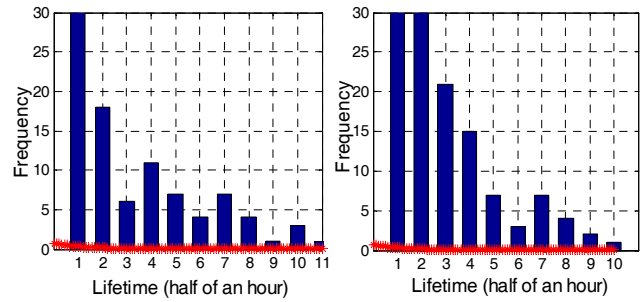


Fig. 2. Fitting by the exponential distribution

---

#### Algorithm: Enhanced SMMD Prediction

---

- 1:  $A_i (i=1, 2, \dots, a)$ : a series of original data;
  - 2: Determine  $N$  by an exhaust algorithm;
  - 3:  $k = A_{\max} - A_{\min}$  as the Interval range  $K$ ;
  - 4: Divided the interval equally into  $N$  parts  $k_1, k_2, k_3 \dots k_N$ ;
  - 5: **for**  $u$  from 1 to  $N$  **do**
  - 6:     **for**  $w$  from 1 to  $N$  **do**
  - 7:          $Num_{uw}$  = the number of points from  $K_l$  to the next one;
  - 8:     **end for**
  - 9: **end for**
  - 10: **for**  $u$  from 1 to  $N$  **do**
  - 11:     **for**  $w$  from 1 to  $N$  **do**
  - 12:          $P_{uw} = Num_{uw} \div (\sum_{i=1}^N \sum_{r=1}^N Num_{ir})$ ;
  - 13:     **end for**
  - 14: **end for**
  - 15: Generate a transfer matrix  $P = \begin{matrix} P_{11} & \dots & P_{1N} \\ \vdots & \dots & \vdots \\ P_{N1} & \dots & P_{NN} \end{matrix}$
  - 16: Investigate  $A_a$  in which interval;
  - 17: Use  $P$  and  $A_a$  to determine the predicted interval;
  - 18:  $Q_I$  = the median of the predicted interval;
  - 19: **for**  $u$  from 1 to  $(a-1)$  **do**
  - 20:      $J_u = A_{u+1} - A_u$ ;
  - 21: **end for**
  - 22:  $J_u (i=1, 2, \dots, a-1)$  is a series of jitter data;
  - 23: Apply gray prediction model in step 6 we get  $Q_2$ ;
  - 24: Set  $Q$  as the final predicted value for power consumption.
  - 25:  $Q = Q_1 + Q_2$ ;
- 

distribution.

To resolve this problem, we adopt a spatial mapping to make the statistical property of data sequence satisfy the Markov property better. This spatial mapping is simply defined as follows:

$$\begin{aligned} \mathbb{T}' &= F(\mathbb{T}) \\ \mathbb{T} &= F^{-1}(\mathbb{T}') \end{aligned} \quad (3)$$

where  $\mathbb{T}$  and  $\mathbb{T}'$  denote two different spaces,  $F(\cdot)$  denotes the mapping function, and  $F^{-1}(\cdot)$  is the inverse mapping. It is noted that this mapping may be very

different according to the practical dataset. The mapping should be reversible. Its function is to preprocess the dataset to make the data follow the Markov stochastic property well. Based on this spatial mapping, it allows us to use the Markov prediction model to predict the practical power data.

Markov chain forecasting is suitable for describing the large random volatility of the prediction problem. Therefore, we use the Markov chain model to predict the primary result interval with a long term observation. Then, we take the gray prediction as a short observation in order to further optimize prediction for the jitters in each interval.

Based on the above consideration, given time data sequence  $\{x^0(k), k=1, 2, \dots, n\}$ , which is the jitters of the power data in each interval. Then, we can accumulate the original data of prediction  $x^0(1), x^0(2), \dots, x^0(n)$ , that is,

$$x^{(1)}(i) = \sum_{m=1}^i x^{(0)}(m) \tag{4}$$

Then, we establish the differential equation,

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = u \tag{5}$$

Utilizing least square method to compute the values of  $a, u$ , we get the structural data matrix  $N$ .

$$B = \begin{bmatrix} -\frac{1}{2}[x^{(1)}(1) + x^{(1)}(2)] & 1 \\ -\frac{1}{2}[x^{(1)}(2) + x^{(1)}(3)] & 1 \\ \vdots & \vdots \\ -\frac{1}{2}[x^{(1)}(n-1) + x^{(1)}(n)] & 1 \end{bmatrix}, \tag{6}$$

Let  $y_n$  be a column vector,

$$y_n = [x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)]^T. \tag{7}$$

The parameter identification  $a, u$ , can be calculated by

$$\hat{a} = \begin{bmatrix} a \\ u \end{bmatrix} = (B^T B)^{-1} B^T y_n. \tag{8}$$

Then, we get the GM (1,1) gray prediction model for the jitters in each interval:

$$\hat{x}^{(1)}(i+1) = \left(x^{(0)}(1) - \frac{u}{a}\right) e^{-ai} + \frac{u}{a}, \tag{9}$$

$$\begin{cases} \hat{x}^{(0)}(1) = \hat{x}^{(1)}(1) \\ \hat{x}^{(0)}(i) = \hat{x}^{(1)}(i) - \hat{x}^{(1)}(i-1), i=2, 3, \dots, n \end{cases} \tag{10}$$

All above present the basic theory models used in enhanced SMMD solution. In the next section, we will detail the workflow of SMMD solution.

### 3.2 SMMD workflow

Based on the above analysis, we propose an enhanced prediction mechanism to improve the prediction accuracy. In particular, the prediction method mainly includes eight steps:

**Step 1:** Check the raw power consumption data whether follow the Markov stochastic property and conduct a spatial mapping.

**Step 2:** Disposal of data discretization and mapping the data series into different intervals.

**Step 3:** Calculate the probability of data locating in each interval, and build a matrix.

**Step 4:** Calculate the transition probability between any two intervals, and build the transfer matrix.

**Step 5:** Compute the predicted interval value.

**Step 6:** Construct the gray prediction model for the jitters of each interval.

**Step 7:** Compute the predicted jitter value.

**Step 8:** Obtain an optimal predicted value based on the Step 5 and Step 7.

As for the step 3, we calculate the original probability of each interval with a method of difference. Suppose we get a series of data  $A_1, A_2, A_3, \dots, A_a$  and divide them into  $N$  intervals. We record the number of data in each section  $K_1, K_2, K_3, \dots, K_N$  and then divided by the total number  $A$  to calculate the original probability of each interval  $P_1, P_2, P_3, \dots, P_N$ .

As for the step 4, we apply the traditional Markov method to  $(A_1, A_2, A_3, \dots, A_a)$  to get transfer matrix  $P$ .  $P$  is produced by  $(A_1, A_2, A_3, \dots, A_a)$ . The element  $P_{ab}$  of  $P$  is the probability that data in range  $a$  jumping to the next hop range  $b$ . Then, we can compute the predicted interval value.

As for the step 5, we investigate  $A_a$  in which interval,

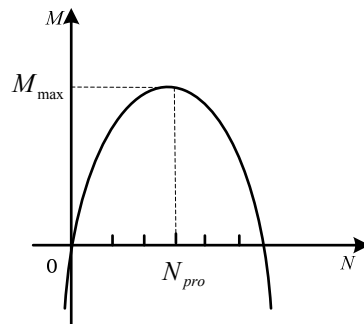


Fig. 3. Optimization analysis of  $N$

then we use  $P$  and  $A_a$  to determine the predicted interval, then we take the median of the predicted interval as the predicted value  $Q_1$ .

As for the *step 6*, we can get a new series of jitter data  $J_1, J_2, J_3 \dots J_{L-1}$ , where each element is computed by  $J_1 = A_2 - A_1, J_2 = A_3 - A_2, J_3 = A_4 - A_3 \dots J_{a-1} = A_a - A_{a-1}$ . Then, we can the transition probability matrix  $(P'_1, P'_2, P'_3, \dots, P'_N)$ .

As for the *step 7*, we construct the gray prediction model for the jitter data  $J_{L-1}$ , and obtain the predict value  $Q_2$ .

As for the *step 8*, we can obtain the predicted value with two groups of data. From the steps above, we obtain  $(P_1, P_2, P_3, \dots, P_N)$ ,  $(P'_1, P'_2, P'_3, \dots, P'_N)$ ,  $A_a$  and  $J_{L-1}$ . Then, we can get two predicted value  $Q_1$  and  $Q_2$ . Take an overall consideration about the two predicted value, we get will the final predicted value calculated by  $Q = Q_1 + Q_2$ .

### 3.3 Optimized data discretization

How many parts to be divided will be better? The prediction accuracy is determined by division patterns. Therefore the first question we care about is how to divide the series of data. We can coarsely divide the collected data into equal parts. The more parts we divide, the smaller each range is.

In order to analyze the optimized division pattern, we define two variables for the accuracy. One is the interval accuracy  $M_1$ , and the other one is the self-accuracy  $M_2$ . If the discretization interval is too big, the predicted value is easier to fall in this interval, thus the range accuracy  $M_1$  is high, while the self-accuracy  $M_2$  is low, or vice versa. To this end, we take the range between the highest and lowest values as the segmented region.

Assuming the number of parts we divide is  $N$ . As  $N$  increases, the range will become smaller,  $M_1$  will be smaller, and  $M_2$  will be bigger. If  $N$  is small enough,  $M_1$  will take a dominant position, and the overall accuracy  $M$  will be near to 0. If  $N$  is large enough,  $M_2$  will take the dominant position, and the overall accuracy  $M$  will also be near to 0. Therefore, there must be an optimal  $N$  to make the result more accurate.

The optimization analysis between  $M$  and  $N$  is shown in Fig. 3. To find the optimal value for  $N$  ( $N_{PRO}$ ), we can adopt an exhaust algorithm. According to the collected data, we let  $N$  increases from 1 to a larger number. In the beginning,  $M$  will rise along with the increasing of  $N$ . When  $N$  is greater than a certain number,  $M$  will decrease with the increase of  $N$ . Then, we can obtain  $N_{PRO}$ , which is corresponding to the maximum value of  $M$ .

## 4. Experimental Evaluations

### 4.1 Data collections and test model

To analyze the stochastic property of power consumption, we first collect an amount of practical electric consumption

data. The practical location of the specific observation area is shown in the Fig. 4. The gray circle section indicates the location of the substation in the village. This substation equipment introduces a total of 10 lines connected to different areas. Each line is connected to a community.

We randomly selected a cell as the experimental community (marked with “experimental line”), and selected a family in the cell to test their electricity consumption.

In this observation area, an electricity meter is deployed to collect the real-time power consumption data. In order to reduce the amount of data sample, we conducted a data collection every 30 minutes. The whole collection process is with duration of one month (from August 10, 2015 to September 9, 2015). The collected power consumption data includes: voltage, current, zero-sequence current, actual power and so on. The voltage and current are three phases. Finally, we obtained  $1459 \times 6$  data entries in total. Some descriptions of collected dataset are listed in Table 2. With these practical data collections, we will further the prediction performance.

In our experiments, we first make a study on the periodic variation law of the Electric Power. We calculate the average value of the Electric Power at each hour of every day in a month, and analyze the rules of change. Then, we use the power consumption data in the first 27 days as the training data samples, and predict the next three power consumption by using two prediction methods. One is the Markovian prediction model proposed in [38], and another one is the enhanced SMMD prediction method proposed in this paper. In detail, a Discrete Time Markov Chain model (DTMC) is used as our comparison solution and the test model.

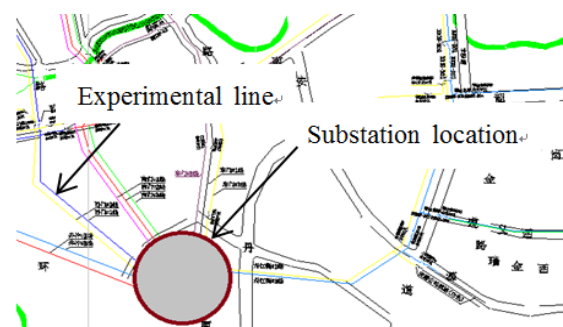


Fig. 4. Experimental observation area

Table 2. Dataset description

Parameter	Descriptions
Duration	August 10, 2015-september 9, 2015
Object	Village, Community, House
Voltage	Three phases
Current	Three phases
Zero-sequence current	0~36A
Time interval	30 minutes
Total number of data entries	$1459 \times 6$

### 4.2 Experimental analysis

Fig. 5 shows the variation law of electric power in one period, which is about one day. As we can see, the electric power distributes between 200 to 400 Kw, and there are two peaks in around 12:00 and 19:30. The lowest point is at 5:00 am each day, which shows the electric power first decreases before 5:00 am and then increase until to 11:00 am. From 1:30 pm, the electric power further starts increasing until 9:00 pm, then it decreases greatly until to 0:00 am.

In order to analyze and process data more convenient, we discretize electric power to further explore the distribution rule followed by the duration of each Electric Power State.

Similar with the reference [19], the formula of discretization is as follows:

$$S = \left\lfloor \frac{A}{\text{Max}(A)+1} \cdot DR \right\rfloor + 1 \tag{11}$$

where  $S$  is the Electric Power State (abbreviated as **state**),  $A$  is the electric power, and  $DR$  is the discretization range.

Fig. 6 shows the status change of Electric Power Consumption after discretization ( $DR=9$ ) within a day. We can see that the state of 6 ( $S=6$ ) accounts for the majority since they appear many times in a row. It means that there is a relative big probability for the state 6.

With a careful dataset analysis, we find that the duration of Electric Power State is poorly fitted by negative exponential distribution (as the Fig. 2 shows). Therefore, we cannot use the Markov based prediction without any amendment scheme. To this end, we introduce a simple solution to make the power prediction suit for the Markov model.

We further discover that the Electric Power State can be well fitted by negative exponential distribution according to a special spatial mapping. In other words, through a spatial mapping, the Electric Power Consumption will greatly present the Markov properties. Therefore, before using Markov-based prediction, we should first conduct a transformation. Then, we calculate the fit parameter for Markov model, and thus the Markov prediction model is constructed. Next, we can conduct the prediction process based on Markov model to obtain the predictive data in the transformed space. Finally, through an inverse transformation, we make the predictive data return to the actual space, so as to obtain the actual power consumption. By this way, we can enhance the accuracy of Markov prediction.

In this paper, we adopt a spatial mapping defined as follows:

$$\mathbb{T}' = \alpha \cdot \mathbb{T} \tag{12}$$

where  $\mathbb{T}$  and  $\mathbb{T}'$  denote two different spaces, and  $\alpha$  is the

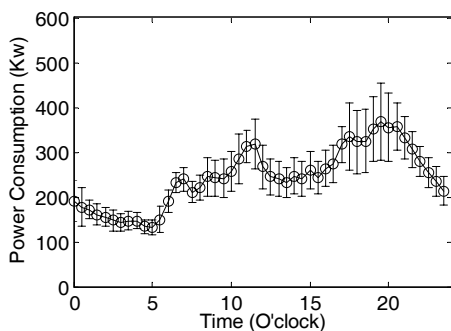


Fig. 5. Power consumption in a period

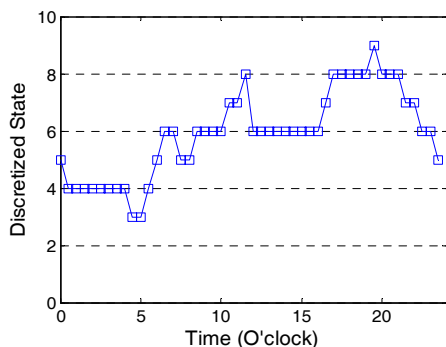


Fig. 6. Discretized electric power

Table 3. The values of  $\alpha$  in different situation

	DR=6	DR=7	DR=8	DR=9
S=3	0.007	0.007	0.009	0.014
S=4	0.006	0.006	0.006	0.007
S=5	0.016	0.008	0.006	0.005

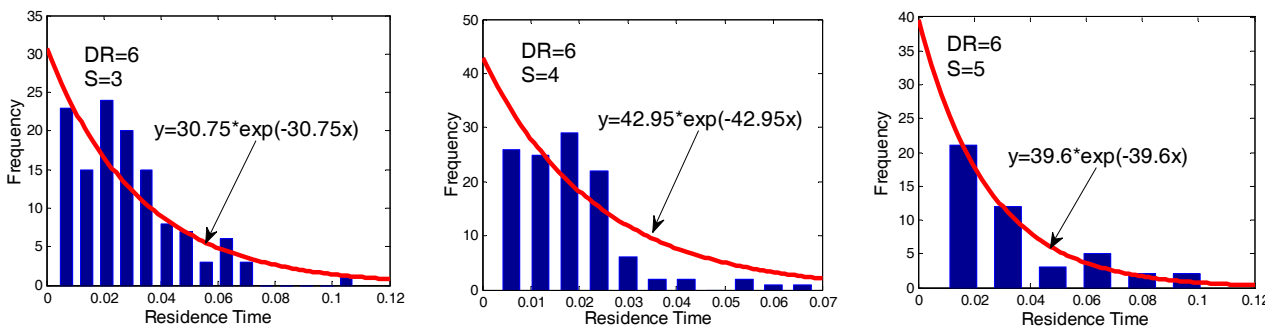


Fig. 7. Frequency fitting with DR=6

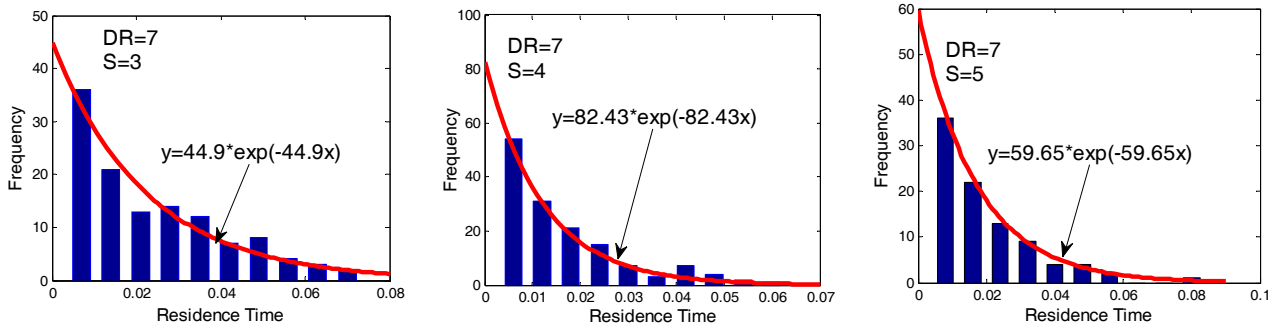


Fig. 8. Frequency fitting with DR=7

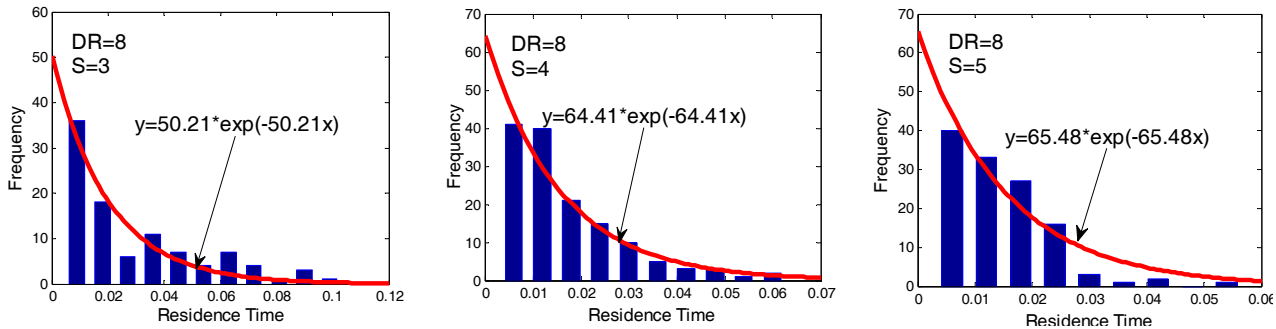


Fig. 9. Frequency fitting with DR=8

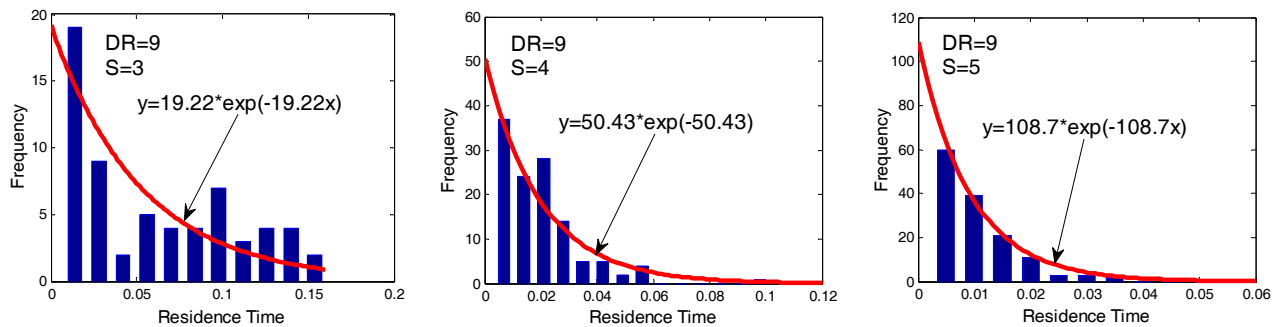


Fig. 10. Frequency fitting with DR=9

coordinate change parameter. The values of  $\alpha$  in our experiment are shown in the Table 3.

After the defined spatial mapping, we found that the residence time of each state can meet a negative exponential distribution well. When DR=6 and DR=7, the results of statistics and simulation are shown in Fig. 7 and Fig. 8, we can see that Residence Time of Electric Power State conforms to the negative exponential distribution, and the fitting result of negative exponential distribution is quite good. The maximum of Frequency in all cases when DR=6 is between 20 to 30, and the range of Frequency maximum when DR=7 is wide, from 35 to 60. In most cases, results of statistics and simulation are consistent with overall change rules of negative exponential distribution. For example, frequency shows tendency of decrease all the time, and a sudden drop will occur in a certain moment. However, there are also abnormal situations of some statistical data. For example, some frequency

statistics drop first and rise late when DR=6 and S=3, and there is a case that the frequency substantially unchanged when Residence Time is small. The causes of these abnormal phenomena are mainly the inadequate discretization state values, but it is also possible that there is a sudden electricity situation.

As shown in Fig. 9 and Fig. 10, we can see that the results of statistics and simulation are basically same as the above when DR=8 and DR=9, and the abnormal situation mentioned above also appears. From the overall pictures contrast, it is the best when DR=7, and there is no abnormal statistics. The root-mean-square-error (RMSE) of statistics and simulation in all cases are shown in Table 4. We can see that it is the best when DR=7, and the worst when DR=8. When DR=7 and S=5, the results of negative exponential distribution is best, and the RMSE is 0.97. When DR=6 and S=4, the results of negative exponential distribution is the worst, and the RMSE is 5.78. The  $\lambda$

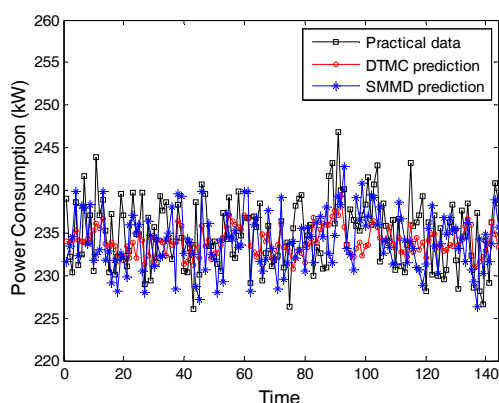


**Table 4.** RMSE of statistics and simulation

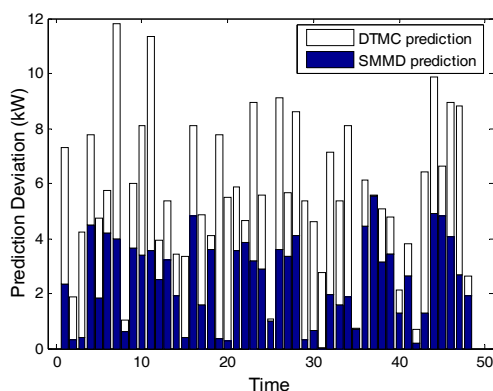
	DR=6	DR=7	DR=8	DR=9
S=3	3.67	2.52	3.41	3.17
S=4	5.78	2.59	3.66	3.29
S=5	1.67	0.97	4.46	2.08

**Table 5.**  $\lambda$  fitting values

$\lambda$	DR=6	DR=7	DR=8	DR=9
S=3	30.75	44.90	50.21	19.22
S=4	42.95	82.43	64.41	50.43
S=5	39.60	59.65	65.48	108.7



**Fig. 11.** Power consumption predictions (3 days)



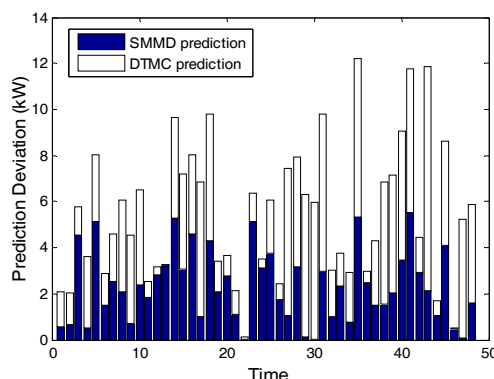
**Fig. 12.** Prediction deviation (1st day)

fitting values are listed in the Table 5.

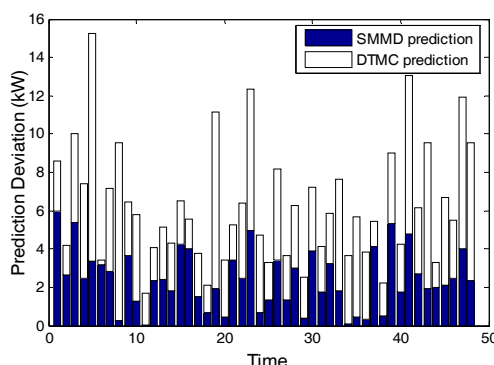
Through the above analysis, we can find that Residence Time of Electric Power State conforms to the negative exponential distribution after the suitable spatial mapping process.

### 4.3 Prediction analysis

The prediction comparisons are shown in Fig. 11. In this figure, we can see the prediction results in both solutions are approximate to the practical data, where the value of power consumption jitters in the range of 227~247. However, the varying pattern in SMMD prediction is more similar to the practical one.



**Fig. 13.** Prediction deviation (2nd day)



**Fig. 14.** Prediction deviation (3rd day)

In order to further compare the performance of the SMMD prediction method with the traditional one, we analyze the prediction deviation of both methods. The prediction deviation comparisons are illustrated in Fig. 12, Fig. 13 and Fig. 14. The white column chart represents the prediction deviation of the state-of-the-art solution, and the blue column chart represents the prediction deviation of our proposed SMMD method.

In these figures, we can see that the deviation of SMMD prediction method is smaller than the traditional Markov one. For example, it keeps smaller than 4 in most cases of SMMD. The deviation of traditional method is mostly about 7, and several cases are even more than 12. Besides, the deviation of traditional method fluctuates sharply, but the new proposed solution does not. We know that the volatility of the prediction deviation will have a great impact on the stability of the prediction. On the other hand, we can also see that the deviation of traditional method generally grows over the forecast time for the traditional Markov method. For example, the prediction deviation in the 3rd day is much bigger than the one in the 1st day. However, the new SMMD prediction method basically remains unchanged, which proves the stability of the SMMD prediction.

Based on above analysis, we conclude that the proposed SMMD prediction method has great improvement in terms of prediction accuracy, prediction stability and predictable term.

## 5. Conclusion

Power prediction is a critical concern in smart grids to improve power efficiency. In this paper, we propose an enhanced power consumption prediction method, named SMMD. In particular, we first analyze and compare several kinds of prediction models. Furthermore, we collect a set of practical electricity consumption data in a certain community and make a detailed analysis of its stochastic property. We explore a special spatial mapping approach to make the power data conform to the Markov process. Then, an enhanced Markov-difference power consumption prediction method is proposed to improve the prediction quality. Finally, a simulation campaign has been carried out, demonstrating that SMMD achieves improved prediction accuracy with respect to state-of-the-art solutions in terms of prediction accuracy, stability and predictable term.

## Acknowledgements

This work was supported by the National High-Tech Research and Development Program of China (863 Program) [grant number 2015AA0501954].

## References

- [1] X. Liu, "The Research of Power Load Forecasting Based on Empirical Mode Decomposition-Grey Markov Model," *Master Degree Thesis, Hefei University of Technology*, Apr. 2013.
- [2] Z. Ziadi, S. Taira, M. Oshiro, T.a Funabashi, "Optimal Power Scheduling for Smart Grids Considering Controllable Loads and High Penetration of Photovoltaic Generation," *IEEE Transactions on Smart Grid*, vol. 5, no. 5, pp. 2350-2359, Sep. 2014.
- [3] G. Koutitas, "Control of Flexible Smart Devices in the Smart Grid," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1333-1343, Sep. 2012.
- [4] G. Carpinelli, G. Celli, S. Mocci, et al, "Optimal Integration of Distributed Energy Storage Devices in Smart Grids," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 985-995, Jun. 2013.
- [5] L. Hernandez, C. Baladron, J. M. Agular, et al, "A Multi-Agent System Architecture for Smart Grid Management and Forecast in of Energy Demand in Virtual Power Plants," *IEEE Communication Magazine*, vol.51, no. 1, pp. 106-113, Jan. 2013.
- [6] J. Cao, Y. Cao, "Research on Common Information Model and its Key Applications for Smart Grid," *Doctor Degree Thesis, Zhejiang University*, Oct. 2012.
- [7] Z. Wan, G. Wang, Y. Yang, S. Shi, "SKM: Scalable Key Management for Advanced Metering Infrastructure in Smart Grids," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 12, pp. 7055-7066, Dec. 2014.
- [8] L. Sun, "The study of Evaluation System of Smart Grid Management and Multi-attribute Analysis Model," *Doctor Degree Thesis, North China Electric Power University*, 2015.
- [9] L. Zhou, J. Rodrigues, L. M. Oliveira, "QoE-Driven Power Scheduling in Smart Grid: Architecture, Strategy, and Methodology," *IEEE Communication Magazine*, vol. 50, no. 5, pp. 136-141, May. 2012.
- [10] J. Liang, G. K. Venayagamoorthy, R. G. Harley, "Wide-Area Measurement Based Dynamic Stochastic Optimal Power Flow Control for Smart Grids With High Variability and Uncertainty," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 59-69, Mar. 2012.
- [11] N. Saputro, K. Akkaya, S. Uludag, "A Survey of Routing Protocols for Smart Grid Communications," *Computer Networks*, vol. 56, no. 11, pp. 2742-2771, Jul. 2012.
- [12] H. Liang, A. Tamang, W. Zhuang, and X. Shen, "Stochastic Information Management in Smart Grid," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, vol. 16, no. 3, pp. 1746-1770, Aug. 2014.
- [13] T. Logenthiran, D. Srinivasan, T. Shun, "Demand Side Management in Smart Grid Using Heuristic Optimization," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1244-1252, Sep. 2012.
- [14] K. Mets, J.A. Ojea, C. Develder, "Combining Power and Communication Network Simulation for Cost-Effective Smart Grid Analysis," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, 1771-1796, pp. 2014.
- [15] H. Liang, A. K. Tamang, W. Zhuang, X. Shen, "Stochastic Information Management in Smart Grid," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1746-1770, 2014.
- [16] H. Liang, A. Abdrabou, W. Zhuang, "Stochastic Information Management for Voltage Regulation in Smart Distribution Systems," *IEEE INFOCOM*, pp. 2652-2660, 2014.
- [17] 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 International Publishing*, 35-57, 2014.
- [18] W. Zhao, J. Wang, H. Lu, "Combining forecasts of electricity consumption in China with time-varying weights updated by a high-order Markov chain model," *Omega*, vol. 45, no. 2014, pp. 80-91, 2014.
- [19] Y. W. Le and J. H. He, "Practical Data-Driven Analysis on Stochastic Property of Power Assumption in Smart Grids," *ICIC Express Letters*, vol. 11, no. 1, pp. 95-100, Jan. 2017.
- [20] W. Zhang, Y. Zhang, "Power Load Forecast under Intelligent Grid Environment," *Master Degree Thesis. Southwest Jiaotong University*, Jun. 2011.

- [21] J. Abbot, J. Marohasy, "Application of Artificial Neural Networks to Rainfall Forecasting in Queensland, Australia," *Advances in Atmospheric Sciences*, vol. 29, no. 4, pp. 717-730, 2012.
- [22] S. Yang, "Study and Application of Time Series Forecasting Based on Rough Set and Kernel Method," *Report, North China Electric Power University*, Dec. 2008.
- [23] L. Zhang, "The Research and Realization of Electric Load Forecast Method," *Master Degree Thesis, Northeast Forestry University*, Jun. 2006.
- [24] X. Wang, S. Yang, H. Wang, "Application Research on Grey Orthogonal Prediction Model Based on Markov Chain for Electricity Consumption Forecasting of Smart Grid," *Chinese Journal of Management Science*, vol. 18, no. 6, pp. 46-51, Nov. 2010.
- [25] C. Shi, P. Wang, "The Improvement and Application of the Grey Markov Chain," *Master Degree Thesis, Lanzhou University*, Jun. 2014.
- [26] W. Ma, "The Mechanical Fault Diagnosis Research Based on Wavelet Analysis Method," *Master Degree Thesis, Inner Mongolia University of Science and Technology*, Jun. 2008.
- [27] F. Liang, W. Zhang, F. Ju, "Forecast of Land-using Structure Change Based on Markov-gray Model," *Journal of Shenyang University*, vol. 20, no. 6, pp. 110-113, Dec. 2008.
- [28] W. Gao, "An Improved GM(1,1) Forecast Model Based on the Markov Theory," *Computer Engineering & Science*, vol. 33, no. 2, pp. 159-163, Dec. 2011.
- [29] W. Feng, J. Yang, Y. Feng, "Forecast of the Total Freight of Chinese Ocean Shipments Based on Gray Markov Model," *Ship Electronic Engineering*, vol. 34, no. 11, pp. 124-127, Feb. 2014.
- [30] B. Li, "Research on GM(1,1) Predicted Effects Based on Grey Generation Technology and Grey-Markov Forecasting Model Optimization," *Master Degree Thesis, Agricultural University of Henan*, Jun. 2014.
- [31] H. Zhou, Z. Li, "Clustering based Short Term Load Forecasting using Artificial Neural Network," in *Proceedings of 2009 IEEE PES Power Systems Conference Exposition (PSCE)*, pp. 1-7, March 2009.
- [32] Y. He, A. Dai, et al, "A Two-stage Electricity Demand Forecasting Model in The Smart Grid," *Power System Prediction and Control*, vol. 38, no. 21, pp. 167-172, Nov. 2010.
- [33] X. Zhou et al, "A Fuzzy Probability-based Markov Chain Model for Electric Power Demand Forecasting of Beijing," *Energy & Power Engineering*, vol. 5, no. 4, pp. 488-492, 2013.
- [34] O. Ardakanian, S. Keshav, and C. Rosenberg, "Markovian Models for Home Electricity Consumption," in *Proceedings of the 2nd ACM SIGCOMM workshop on Green networking*, pp. 31-36, Aug. 2011.
- [35] C. Liu, Q. Zeng, Y. Liu, "A Dynamic Load Control Scheme for Smart Grid Systems," *Energy Procedia*, vol. 12, no. 1, pp. 200-205, 2011.
- [36] X. Liu, H. Zhang, Q. Xiang, et al, "Taming Uncertainties in Real-Time Routing for Wireless Networked Sensing and Control," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 288-301, Mar. 2013.
- [37] F. Xi, G. Xue, D. Yang, "Smart Grid – The New and Improved Power Grid: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 944-980, 2012.
- [38] M. K. Haider, A. K. Ismail and I. A. Qazi, "Markovian Models for Electrical Load Prediction in Smart Buildings," *ICONIP, LNCS 7664, Springer*, 2012.



**Yiwen Le** He is a Ph.D. student with school of Electrical Engineering, Beijing Jiaotong University. His research interests include modeling, data aggregation and power prediction technologies in smart grid.



**Jinghan He** She is the dean and a professor (full) now with school of Electrical Engineering, Beijing Jiaotong University. Her major research areas are on-line monitoring, protection and control of power system, power quality, new energy and intelligent power grid, and electrical rail transportation.