



A Hybrid Short-Term Building Electrical Load Forecasting Model Combining the Periodic Pattern, Fuzzy System, and Wavelet Transform

Chengdong Li¹ · Minjia Tang¹ · Guiqing Zhang¹ · Ruiqi Wang² · Chongyi Tian¹

Received: 30 July 2019 / Revised: 6 November 2019 / Accepted: 29 November 2019 / Published online: 3 January 2020
© Taiwan Fuzzy Systems Association 2020

Abstract Accurate forecasting and scientific analysis of building electrical load can improve the level of building energy management to meet the requirements of energy saving. To further strengthen the forecasting accuracy, this study presents a hybrid model for building electrical load forecasting. The proposed method combines the fuzzy inference system and the periodicity knowledge together to generate accurate forecasting results. In this method, in order to better reflect the actual characteristic of the electrical load, the wavelet transform method is firstly utilized to filter the original building electrical load data. Then, the daily periodic pattern is extracted from such filtered electrical load data, and the residual data are obtained through removing the daily periodic pattern. Further, the residual data-driven forecasting model is constructed by the functionally weighted single-input-rule-modules connected fuzzy inference system (FWSIRM-FIS). This FWSIRM-FIS model is used to provide the compensation to the periodic component. In other words, the daily periodic component and the residual forecasting are combined to achieve the final forecasting result. Specifically, in order to assure the forecasting performance of the FWSIRM-FIS model, the subtraction clustering method is employed to construct the SIRMs while the least square estimation is utilized to optimize the parameters in the functional

weights of the FWSIRM-FIS. Finally, in this paper, two real-world experiments are made and detailed comparisons with four traditional models are given. Experimental and comparison results demonstrate that the proposed hybrid model has the smallest forecasting errors and can achieve the best performance.

Keywords Building electrical load forecasting · Fuzzy method · Periodic pattern · Wavelet transform · Least square estimation · Subtraction clustering

1 Introduction

Along with the development of China's economy, its electricity consumption gradually occupies an important proportion of the energy consumption in the whole country. The gradual increase of China's electricity demand caused a heavy burden of the power generation. As a result, it is an urgent problem to reduce the electricity consumption, especially the electricity consumption in buildings which are the terminals of energy utilization [1, 2]. On the other hand, the reasonable management and control of building equipments is conducive to effectively reducing the electricity utilization. One important thing for efficient management and control of building equipments is to accurately forecast the short-term building electrical load. Building electrical load forecasting refers to explore the internal relationship through analyzing the historical electrical load data, so as to establish a forecasting model to estimate the trend in the future. Based on the forecasting results, the electrical worker can plan power management and control scheme so as to improve the economic efficiency of power system.

✉ Chengdong Li
chengdong.li@foxmail.com; lichengdong@sdjzu.edu.cn

¹ Shandong Key Laboratory of Intelligent Buildings Technology, School of Information and Electrical Engineering, Shandong Jianzhu University, Jinan 250101, China

² State Grid Shandong Integrated Energy Services Co., Ltd., Jinan 250101, China

Existing building electrical load forecasting models can be divided into three categories: the engineering models, the statistical models, and the artificial intelligence models. In the engineering models, the physical principles are utilized to calculate the energy consumption behaviors and the thermal dynamics. Over the past 60 years, many software tools have been given to assess energy efficiency, such as the EnergyPlus [3], DOE-2 [4], BLAST, and ESP-r. Such software tools have been popularly adopted to improve building electrical load forecasting accuracy. In [5], a report was given to provide up-to-date comparison of the features and capabilities of twenty major building energy simulation programs. In [6], Westphal and Lambert analyzed the thermal loads of non-residential buildings based on some weather variables. In addition to the weather conditions, building characteristics are other important factors that can determine the forecasting performance. In [7], Yao and Steemers developed a simple method for formulating load profile. This method has been utilized to forecast the daily energy load distribution of household appliances, domestic hot water, and space heating. There is no clear boundary between the exact model and the simplified model.

The statistical methods apply the historical electrical load data to predict the future values and have been widely used in building electrical load forecasting. Among the statistical models, the AutoRegressive model with eXtra inputs (ARX), AutoRegressive Integrated Moving Average (ARIMA), AutoRegressive Integrated Moving Average with eXtra inputs (ARIMAX), and Conditional Demand Analysis (CDA) are very popular. In [8], a robust method combining two separate time-indexed ARX models was presented for hourly cooling-load forecasting. In [9], a new hybrid model combining the ARIMA and support vector regression was developed for electrical load forecasting. In [10], an ARIMAX model was given to forecast the short-term summer load. In [11], the CDA method was utilized to estimate the electrical load of various appliances.

Compared with the engineering models and the statistical models, artificial intelligence models are more popular and more suitable for accurate forecasting in dynamic and uncertain environments. Among the artificial intelligence models, the most widely used one is the neural network (NN). In [12], a cascaded NN-based hybrid forecasting model was proposed for short-term load forecasting. In [13], a new enhanced back propagation NN (BPNN) model was given to effectively predict the electrical load. In [14], Badran developed a modified forecasting model combining NN with linear regression and applied it to the electrical load demand forecasting. In [15], in order to improve the accuracy of electrical load forecasting, Li et al. presented an optimized forecasting model for building energy consumption. In this method, an improved particle swarm

optimization algorithm was applied to adjust the weight and threshold of the NN, and the forecasting results demonstrated that the accuracy of this model is higher than that of the traditional NN model. In [16], a generalized regression NN (GRNN) forecasting model was established by analyzing the load data and meteorological data to forecast the region electrical load and the forecasting accuracy of this proposed model is significantly improved compared with the traditional neural network forecasting model. Some other machine learning methods such as the support vector machine (SVM), support vector regression (SVR), and Extreme Learning Machine (ELM) are also widely used in electricity load forecasting. In [17], the forecasting results of the SVM and Random Forest Regression (RFR) were compared for power load forecasting. Comparison results showed that both SVM and RFR are excellent choices for electrical load forecasting. In [18–21], SVR was successfully applied to building energy consumption forecasting. In [22, 23], one ELM-based novel method was developed to achieve an improvement in building energy consumption forecasting accuracy.

Another popular artificial intelligence method is the fuzzy inference system (FIS), especially the fuzzy neural network (FNN) which combines the NN with FIS to aggregate their merits and overcome their shortcomings. In [24], Li et al. developed a new method named the hybrid genetic algorithm-hierarchical adaptive network-based fuzzy inference system (GA-HANFIS) to estimate the building energy consumption. In [25], a novel type-2 fuzzy set-based methodology—T2SDSA-FNN (Type-2 self-developing and self-adaptive FNN) was proposed to model the data uncertainty for electrical load forecasting. In [26], Li et al. developed a novel short-term cooling-load forecasting approach by conjunctive use of fuzzy C-mean clustering algorithm and fuzzy SVMs. In [27], a modified fuzzy logic relation-based approach was presented for future electricity demand forecasting using the dataset collected from the Central Electricity Authority in India.

According to reviewed literatures, most of the artificial intelligence-based building electrical load forecasting models are data-driven. When the number of input variables becomes large [28–30], the model will become more complicated and the structure will be very difficult to construct, especially for the fuzzy methods, which always face the curse of dimensionality [31–33]. In order to decrease the number of rules and reduce the system design difficulty, in [34, 35], Yi et al. firstly proposed the single-input-rule-modules (SIRMs)-based fuzzy inference system (FIS) to simplify the design of traditional FIS. The SIRMs connected FIS (SIRM-FIS) firstly constructs one SIRM for each input variable, and finally aggregates the outputs of all SIRMs by multiplying their weights to generate the final reasoning output [36, 37]. In order to enhance the

performance of traditional SIRM-FISs, many scholars have extended the SIRM-FIS. In [38], Seki and Nakashima proposed the SIRMs Connected Fuzzy Inference Model with Functional Weights (SIRM-FW), which replaces the constant weights of the traditional SIRM-FIS with one-variable function. In order to further improve the performance of the SIRMs method, a new kind of SIRM-FIS named the functionally weighted SIRM-FIS (FWSIRM-FIS) was proposed by Li et al. in [39]. The weights of all SIRMs in FWSIRM-FIS were replaced by the multi-variable function [39]. Thus, the importance of the corresponding SIRMs can be dynamically reflected by the values of the multiple input variables. Simulation results in [39] has proved that the FWSIRM-FIS has better performance for modeling and forecasting problems compared with the traditional SIRM-FIS.

On the other hand, building electrical load data often demonstrate obvious periodic patterns [40–43]. For example, different buildings have different monthly, daily, and hourly patterns of the electrical loads. Through considering such periodic features in the building electrical load data, not only better insights can be gained into the data, but also the forecasting accuracy can be strengthened. Despite the periodic features have great importance, to the authors' knowledge, there are few studies considering how to ensemble the periodic features into the building electrical load forecasting application. In [44], Ghaderi attempted to decompose the pattern by clustering the data for each specified period, and then to simplify and analyze the complexity of the pattern of consumption. In [45], in order to decline the periodic variable's affect and simplify the pattern of electricity consumption, Keyno decomposed the complicated pattern to a set of simple patterns by clustering the primary data and eliminating the periodic variance. However, how to utilize the periodicity knowledge to enhance the forecasting performance still needs further studies.

In order to further improve the forecasting accuracy, this study presents a hybrid model and applies it to the building electrical load forecasting. The main contributions and novelties are listed as follows:

- A hybrid model is proposed for the electrical load forecasting. In this hybrid model, the periodic pattern is extracted and the residual data obtained through removing the periodic pattern from the building electrical load data are utilized to construct the residual forecasting model. In other words, the final predicted values of this hybrid model are obtained by aggregating the periodic component and the predicted results from the residual data-driven forecasting model.
- The wavelet transform method is adopted to decompose and reconstruct the observed building electrical

load data so as to eliminate data noise and to extract the periodic pattern as accurate as possible.

- The residual forecasting model is realized by the FWSIRM-FIS which can reduce the number of fuzzy rules and has more powerful ability for forecasting and identification problems. In this study, the subtraction clustering method is employed to construct the SIRMs for the FWSIRM-FIS, while the least square method is given to learn the parameters of the FWSIRM-FIS.
- The proposed hybrid method is applied to two real-world buildings, and detailed comparisons are also given. Experimental results have verified the effectiveness of the proposed method on building electrical load forecasting. And, comparison results have demonstrated that the proposed hybrid model has the smallest forecasting errors and can achieve the best performance.

The rest of this paper is organized as follows: The wavelet transform method and the FWSIRM-FIS will be introduced in Sect. 2. The proposed hybrid model will be presented in Sect. 3. Two experiments will be done, and comparisons will be made in Sect. 4. Finally, the conclusions will be given in Sect. 5.

2 Methodologies

In this section, the wavelet transform method and the FWSIRM-FIS will be introduced.

2.1 Wavelet Transform

Fourier transform shows excellent performance in processing smooth and stationary signals, but it is not good for processing sudden changes and non-stationary signals. Wavelet transform which originates from Fourier transform is a method developed in recent years to process time series signals. Compared with the Fourier transform, the wavelet transform has the characteristics of adaptiveness and mathematical microscopy. It performs multi-scale refinement of time series through scale transformation and translation operation, and is especially suitable for processing non-stationary and non-linear signals [46].

Wavelet transform replaces the infinite triangular function basis of Fourier transform with the finite attenuating wavelet basis; in this way, not only the frequency can be obtained, but also the time can be located. The formula of wavelet transform is as follows [46]:

$$WT(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) * \psi\left(\frac{t-\tau}{a}\right) dt, \quad (1)$$

where a is the scale used to control the scaling of the wavelet function, and τ is the translation amount used to control the translation of the wavelet function.

It can be seen from this formula that, unlike Fourier transform with only one frequency variable, wavelet transform has two variables a and τ . The scale a corresponds to frequency, while the translation τ corresponds to time.

Because the wavelet transform can analyze signals simultaneously in the same time-frequency domain, it can distinguish mutation part and noise part effectively in the high-frequency part on different decomposition levels, so as to achieve the purpose of noise reduction. A time series signal with noise can be expressed as follows:

$$S = A + \varepsilon * e, \tag{2}$$

where S is the noisy signal, A is the useful signal, e is the noise, and ε is the standard deviation of the noise. Usually, we assume that e is a white noise signal. Generally speaking, the useful signal is shown in the low-frequency part, while the noisy signal has the high frequency. The decomposition of S can be computed by the wavelet decomposition structure as shown in Fig. 1.

As shown in this figure, wavelet transform is used to decompose the original noisy signal $S(k)$ layer by layer. Firstly, the first layer is decomposed into cA_1 and cD_1 . Then, continue to use wavelet decompose cA_1 into cA_2 and cD_2 . As an analogy, cA_{n-1} was decomposed into cA_n and cD_n . The amplitude of high-frequency coefficients is small, but the number is large. Through the above steps, the multi-scale decomposition of wavelet and the extraction of wavelet coefficients are completed. After obtaining the wavelet coefficients of each scale, the appropriate coefficients are selected and reconstructed by inverse wavelet transform to get filtered signal.

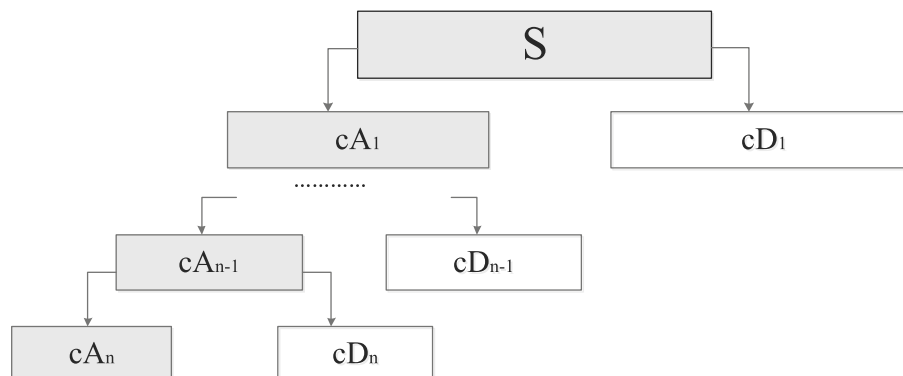


Fig. 1 Wavelet decomposition structure

2.2 FWSIRM-FIS

The FWSIRM-FIS was proposed in [39] to strengthen the performance of the conventional SIRM-FIS [34, 35]. Compared with existing conventional SIRM-FISs, the FWSIRM-FIS can not only compress the fuzzy rule base efficiently, but also improve the approximation performance greatly.

Suppose that the FWSIRM-FIS has n input variables x_1, x_1, \dots, x_n . Then, it will be composed of n SIRMs, each one of which can be regarded as a special FIS with one input and one output [34, 35, 39]. Being different from the traditional SIRMs methods with constant weights, the weights of the FWSIRM-FIS are functions of the variables x_1, x_1, \dots, x_n expressed as $w_i(x) = w_i(x_1, x_2, \dots, x_n)$.

The SIRM for the input variable x_i is as follows [39]

$$\text{SIRM-}i: \{R_i^j : x_i = \tilde{A}_i^j \rightarrow y_i = c_i^j\}_{j=1}^{m_i}, \tag{3}$$

where A_i^j s are fuzzy sets for x_i , c_i^j is the consequent parameter of rule R_i^j , and m_i represents the number of fuzzy rules in the SIRM for x_i .

With the popular singleton fuzzifier and the COS defuzzifier, we can compute the inference result of SIRM- i as [39]

$$y_i(x_i) = \frac{\sum_{j=1}^{m_i} \mu_{\tilde{A}_i^j}(x_i) c_i^j}{\sum_{j=1}^{m_i} \mu_{\tilde{A}_i^j}(x_i)}. \tag{4}$$

In the FWSIRM-FIS, we usually adopt the following functional weight for SIRM- i

$$w_i(\mathbf{x}) = w_i^{(0)} + w_i^{(1)}x_1 + \dots + w_i^{(n)}x_n, \tag{5}$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n)$.

Then, we can calculate the ultimate output of FWSIRM-FIS as

$$\begin{aligned}
 y(\mathbf{x}) &= \sum_{i=1}^n w_i(\mathbf{x})y_i(x_i) \\
 &= \sum_{i=1}^n (w_i^{(0)} + w_i^{(1)}x_1 + \dots + w_i^{(n)}x_n) \\
 &\quad \times \frac{\sum_{j_i=1}^{m_i} \mu_{\tilde{A}}(x_i)c_i^{j_i}}{\sum_{j_i=1}^{m_i} \mu_{\tilde{A}}(x_i)}
 \end{aligned} \tag{6}$$

For convenience, we give some notations for the previous results. To begin, the vector of the parameters in the functional weights is denoted as

$$\mathbf{w} = [w_1^{(0)}, \dots, w_1^{(n)}, \dots, w_n^{(0)}, \dots, w_n^{(n)}]^T. \tag{7}$$

From (6), the mapping of the FWSIRM-FIS can be expressed as [39]

$$y(\mathbf{x}) = \sum_{i=1}^n (w_i^{(0)} + w_i^{(1)}x_1 + \dots + w_i^{(n)}x_n)y_i(x_i) = \mathbf{g}(\mathbf{x})^T \mathbf{w}, \tag{8}$$

where

$$\mathbf{g}(\mathbf{x}) = [y_1(x_1), x_1y_1(x_1), \dots, x_ny_1(x_1), \dots, y_n(x_n), x_1y_n(x_n), \dots, x_ny_n(x_n)]^T. \tag{9}$$

3 The Proposed Method

In this section, the structure of the proposed method will be presented firstly. Subsequently, how to obtain the periodicity knowledge from the building electrical load data will be discussed. Finally, the construction and learning of FWSIRM-FIS which is used to predict the residual errors will be described in detail.

3.1 The Structure of the Proposed Method

As we previously discussed, the building electrical load data have the periodic characteristic. Based on the periodicity knowledge, we can compensate the uncertainties in the building electrical load data and strengthen the forecasting performance of the model. In this study, one novel

hybrid method as shown in Fig. 2 is proposed to obtain more accurate forecasting performance for the building electrical load.

More specifically, the forecasting processes in this study are listed below:

- Step 1: Utilize the wavelet transform method to filter the observed building electrical load data. Then, extract the periodic pattern from the filtered building electrical load data.
- Step 2: Generate the residual data through eliminating the periodic component. Then, construct the FWSIRM-FIS model using the residual data to realize the compensation for the periodic pattern. In this step, the clustering method and least square estimation are adopted to generate fuzzy rules and optimize the parameters in the proposed model, respectively.
- Step 3: Combine the daily periodic component and the predicted residual value to obtain the final building electrical load forecasting results.

Below, we will discuss the first and second steps in detail.

3.2 Periodic Pattern Extraction

Under the complex architectural environment, there are many systematic and random errors in the measurement process that cause measured electrical load data with high levels of noise. In order to extract the periodic pattern as accurate as possible, it is an important step to eliminate the noise from such building electrical load data. In this study, we utilize the wavelet transform method to realize this objective. Then, we will extract the periodic pattern from the filtered electrical load data.

3.2.1 Preprocessing of the Building Electrical Load Data by the Wavelet Transform Method

It is assumed that the building electrical load data have been collected for N days, and the number of sampling times is T per day. Then, the collected time series of

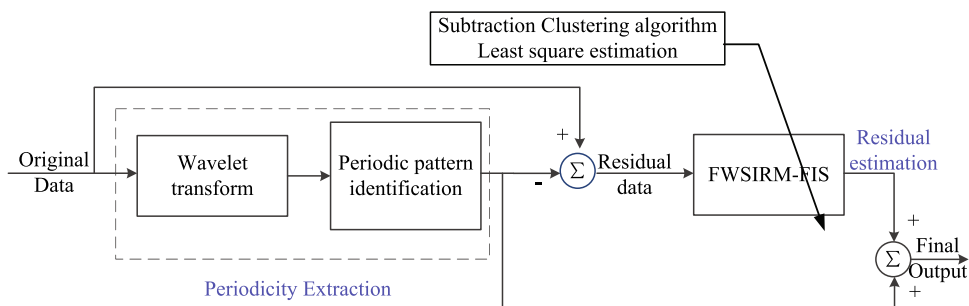


Fig. 2 The structure of the proposed hybrid method

building electrical load can be written as one-dimensional vectors as

$$S = \{S_1, S_2, \dots, S_N\} \tag{10}$$

in which S_k is a vector of the sampling electrical load data in the k th day, and can be expressed as

$$S_k = [s_k(1), s_k(2), \dots, s_k(T)], \tag{11}$$

where $s_k(j)$ is the electrical load data at time j on day k , and $j = 1, 2, \dots, T$.

For the collected original electrical load time series S , in this study, the wavelet transform is used to remove the noise in the sequence according to the following steps.

Using the wavelet transform, it can return the wavelet decomposition of the electrical load time series S at level n . The output decomposition structure consists of the wavelet decomposition vector C and the bookkeeping vector L . According to Fig. 1, the vector form of C can be expressed as

$$C = [cA_n, cD_n, cD_{n-1}, \dots, cD_1], \tag{12}$$

where A and D are line vectors. A is a low-frequency coefficient and also called approximation coefficient, D is a high-frequency coefficient, also known as detail coefficient.

After extracting the wavelet coefficients through the above steps, we need to reconstruct the extracted wavelet coefficients by using the reconstruction function; then, the reconstructed electrical load time series $A(n)$ and $D(n)$ of each decomposition coefficient are obtained.

In the decomposition process, it is found that the length of low-frequency coefficients decreases by half with the increase of scale. It can be seen that the variation of low-frequency coefficients in each layer is similar to that of the original electrical load time series, which shows that the low-frequency coefficients reflect the contour or basic information of the time series, and that the reconstructed low-frequency time series $A(n)$ is close to the original electrical load time series S . And, $D(n)$ belongs to noise time series.

In order to protect the integrity of building electrical load data while reducing the impact of noise on forecasting results, some noise can be removed while low-frequency time series $A(n)$ is kept. Consequently, the reconstructed electrical load time series Z can be obtained as

$$Z = A(n), \tag{13}$$

where n is the level of decomposition, and $A(n)$ is the reconstructed low-frequency building electrical load time series.

3.2.2 Extraction of the Periodic Pattern

According to the original electrical load time series $S = \{S_1, S_2, \dots, S_N\}$, the reconstructed electrical load time series A can be expressed as

$$Z = \{Z_1, Z_2, \dots, Z_N\} \tag{14}$$

in which Z_k is a vector of the reconstructed electrical load data in the k th day, and can be expressed as

$$Z_k = [z_k(1), z_k(2), \dots, z_k(T)], \tag{15}$$

where $z_k(j)$ is the reconstructed electrical load at time j on day k , and $j = 1, 2, \dots, T$.

Through analyzing the original building electrical load data, we can observe that this time series has a daily pattern. Thus, in this study, the building electrical load time series is divided into the periodic component and the residual part, which can be expressed as

$$Z = M + Z_r, \tag{16}$$

where Z is the reconstructed daily building electrical load time series, M is the periodic component, and Z_r is the daily residual part after removing the periodic component.

In this study, we utilize the mean value method to extract the daily electrical load pattern. Consequently, the daily periodic component of the electrical load can be computed as follows:

$$M = \left[\frac{1}{N} \sum_{k=1}^N z_k(1), \frac{1}{N} \sum_{k=1}^N z_k(2), \dots, \frac{1}{N} \sum_{k=1}^N z_k(T) \right]. \tag{17}$$

Furthermore, through removing this daily periodic component, the residual part of the electrical load data can be obtained as

$$Z_r = \{Z_1 - M, Z_2 - M, \dots, Z_N - M\}. \tag{18}$$

For simplicity, this residual part of the building electrical load data can be re-expressed as

$$Z_r = \{z_r(1), z_r(2), \dots, z_r(NT)\}, \tag{19}$$

where $z_r(j)$ is the j th data points in the residual electrical load time series of the N days, and $j = 1, 2, \dots, NT$.

3.3 Construction and Learning of the FWSIRM-FIS Model for the Residual Data

In the building electrical load forecasting model, we usually suppose that the building electrical load of the next sampling time can be affected by the building electrical load of the previous n sampling times. In other words, the proposed model constructed by the residual data has n inputs and one output and can be abstractly expressed as

$$y = \widehat{z}_r(l + 1) = f(\mathbf{x}^{(l)}) = f(z_r(l), z_r(l + 1), \dots, z_r(l + n - 1)), \tag{20}$$

where $f(\cdot)$ represents the input–output mapping of the forecasting model constructed by the residual data, and we denote $\mathbf{x} = [x_1, x_2, \dots, x_n]$, $y = \widehat{z}_r(l + 1)$.

To construct the residual forecasting model, from the residual time series, we generate the input–output training data pairs as

$$\{\mathbf{x}^{(l)}, y^{(l)}\} = \{[z_r(l), z_r(l + 1), \dots, z_r(l + n - 1)]; z_r(l + n)\}, \tag{21}$$

where $l = 1, 2, \dots, L$, and L is the number of the training data pairs.

In order to obtain satisfactory performance, the structure of the FWSIRM-FIS should be determined and its parameters need to be optimized. In this study, the subtraction clustering method is adopted to determine the structure of the FWSIRM-FIS, and the least square estimation method is utilized to learn the parameters in the functional weights of the FWSIRM-FIS model. Detailed description about the construction and learning of this fuzzy system will be given in the following subsections.

3.3.1 Construction of the SIRMs Using the Subtraction Clustering Method

In this paper, we determine the structure of the FWSIRM-FIS by the subtraction clustering method [47]. The subtraction clustering method is a fast one-time algorithm for estimating the number of clusters and the location of cluster centers from a group of data [47]. For each SIRM of the FWSIRM-FIS, the subtraction clustering method is used to achieve the fuzzy partition of the antecedent part of the SIRM and the centers of its consequent part.

The detailed steps for constructing one specified SIRM- i are given as follows [47]:

- According to the training dataset $\{(\mathbf{x}^{(l)}, y^{(l)})\}_{l=1}^L$, generate the training dataset $\{\mathbf{u}_i^{(l)} = (x_i^{(l)}, y^{(l)})^T\}_{l=1}^L$ for the SIRM- i , where $i = 1, 2, \dots, n$.
- Each two-dimensional data point $\mathbf{u}_i^{(l)}$ can be regarded as a candidate for clustering center, and the density index at the data point $\mathbf{u}_i^{(t)}$ is defined as

$$D(\mathbf{u}_i^{(t)}) = \sum_{l=1}^L e^{-\alpha \|\mathbf{u}_i^{(t)} - \mathbf{u}_i^{(l)}\|^2}, \tag{22}$$

where $t = 1, 2, \dots, L$, $\alpha = \frac{4}{r_\alpha^2}$, r_α is a positive number which defines a neighborhood of point $\mathbf{u}_i^{(t)}$. The other data points excluding r_α have less influence on the density index of this point.

- Then, the data with the largest density index are found as the first clustering center, and delete all data points near the selected center determined by the parameter r_α .
- After the $k - 1$ th cluster center has been determined, recalculate the density index for each data point according to the following formula:

$$D_k(\mathbf{u}_i^{(t)}) = D_{k-1}(\mathbf{u}_i^{(t)}) - D_{k-1}^* e^{-\beta \|\mathbf{u}_i^{(t)} - \mathbf{u}_i^{*(k-1)}\|^2}, \tag{23}$$

where $\mathbf{u}_i^{*(k-1)}$ is the location of the $k - 1$ th cluster center and D_{k-1}^* is its potential value; k means the iteration step. In order to avoid clustering centers with close distances. $\beta = \frac{4}{r_b^2}$, $r_b \approx 1.25r_\alpha$ is usually taken.

- Find the maximum density index and use this data point as the clustering center. Repeat the above steps until the density of the remaining data points is less than a threshold; then, the clustering process is over.

After finding the clustering centers of data points by clustering method, the number of fuzzy rules and membership functions of the SIRMs can be determined according to the clustering centers [47].

Suppose that m_i clusters have been obtained as $\{\mathbf{u}_i^{*(k)} = (x_i^{*(k)}, y^{*(k)})^T\}_{k=1}^{m_i}$; then, the generated fuzzy rules for SIRM- i are as follows:

$$\text{SIRM-}i: \left\{ R_i^k : x_i = \widetilde{A}_i^k \rightarrow y_i = y^{*(k)} \right\}_{k=1}^{m_i} \tag{24}$$

in which

$$\mu_{\widetilde{A}_i^k}(x_i) = e^{-\alpha(x_i - x_i^{*(k)})^2}. \tag{25}$$

3.3.2 Optimizing the Functional Weights in the FWSIRM-FIS Model

For the FWSIRM-FIS, its SIRMs have already been generated by subtraction clustering method. However, its functional weights still need to be determined. So in this part, we will optimize the parameters of the functional weights by the learning algorithms. Below, the least square method [48] will be utilized to learn such parameters.

Assume that the training dataset is $\{(\mathbf{x}^{(l)}, y^{(l)})\}_{l=1}^L$, in which $\mathbf{x}^{(l)} = [z_r(l), \dots, z_r(l + n - 1)]$. Generally, we select the following squared-error function as the learning criteria:

$$E(\mathbf{w}) = \sum_{l=1}^L (y(\mathbf{x}^{(l)}, \mathbf{w}) - y^{(l)})^2, \tag{26}$$

where $y(\mathbf{x}^{(l)}, \mathbf{w})$ is the forecasted result from the FWSIRM-FIS.

According to (8) and (23), the following equation can be obtained

$$E(\mathbf{w}) = \sum_{l=1}^L (\mathbf{g}(\mathbf{x}^{(l)})^T \mathbf{w} - y^{(l)})^2 = \|\Phi \mathbf{w} - \mathbf{y}\|_2^2 \tag{27}$$

in which

$$\begin{aligned} \Phi &= [\mathbf{g}(\mathbf{x}^{(1)}), \mathbf{g}(\mathbf{x}^{(2)}), \dots, \mathbf{g}(\mathbf{x}^{(L)})]^T, \\ \mathbf{y} &= [y^{(1)}, y^{(2)}, \dots, y^{(L)}]^T \end{aligned} \tag{28}$$

in which \mathbf{y} is a $L \times 1$ vector, and Φ is a matrix with the dimension $L \times n(n + 1)$.

Hence, the parameters in the FWSIRM-FIS’s functional weights could be calculated by solving the optimization problem below

$$\min_{\mathbf{w}} E(\mathbf{w}) = \|\Phi \mathbf{w} - \mathbf{y}\|_2^2. \tag{29}$$

This optimization problem can be solved by the least square method, and the best values of \mathbf{w} can be estimated as

$$\hat{\mathbf{w}} = [\Phi]^+ \mathbf{y}, \tag{30}$$

where $[\Phi]^+$ is the generalized Moore–Penrose inverses of the matrix Φ .

As well known in the matrix computations [49, 50], a matrix with arbitrary rows, columns, and rank has the Moore–Penrose generalized inverse. Several ways can be used to calculate the Moore–Penrose generalized inverse of a matrix, such as the modified Gram–Schmidt method, the Householder and Givens QR methods, and the singular value decomposition (SVD) method. Among all such methods, SVD is the most widely used one.

4 Experiments and Comparisons

In this section, we will present the evaluation indices and the comparative methods firstly. Then, detailed experiments and comparisons will be given. Moreover, comprehensive analysis and discussion will also be made.

4.1 Evaluation Indices and Comparative Methods

To show the superiorities of the proposed model, traditional models such as FWSIRM-FIS, ANFIS, BPNN, and MLR are selected as the comparative methods. Since the traditional FWSIRM-FIS has been introduced in detail in the previous section, the ANFIS, BPNN, and MLR will be introduced briefly here.

Adaptive neuro-fuzzy inference system (ANFIS) is an efficient FIS which combines fuzzy logic and neural network together [51, 52]. It combines them organically and not only gives full play to their advantages, but also makes up for their shortcomings. ANFIS has found lots of

applications in the research domain of modeling, forecasting, and control [51, 52].

Back Propagation Neural Network (BPNN) [53, 54] is the most basic neural network and it can be used to approximate any non-linear continuous function. Generally speaking, the more complex a network is, the more complex features it can learn and the more complex problems it can solve, but too many parameters will lead to the over-fitting phenomenon.

Multiple Linear Regression (MLR) [55] is a statistical method for analyzing the linear relationship of multiple variables. In addition to validating the correlation and causality of the variables, MLR models are mostly used for forecasting.

In this paper, to test the forecasting performance of different models, three performance measures including the mean absolute error (MAE), the root mean squared error (RMSE), and the coefficient of determination R^2 are adopted. The equations for calculating MAE, RMSE, and R^2 are shown as follows:

$$\text{MAE} = \frac{1}{K} \sum_{l=1}^K |\hat{y}^{(l)} - y^{(l)}| \tag{31}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{l=1}^K (\hat{y}^{(l)} - y^{(l)})^2}{K}} \tag{32}$$

$$R^2 = \frac{\left[\sum_{l=1}^K (\hat{y}^{(l)} - \hat{y}_{Ave})(y^{(l)} - y_{Ave}) \right]^2}{\sum_{l=1}^K (\hat{y}^{(l)} - \hat{y}_{Ave}) \sum_{l=1}^K (y^{(l)} - y_{Ave})} \tag{33}$$

in which K represents the number of training or testing data, $\hat{y}^{(l)}$ and $y^{(l)}$ represent the forecasted and real values, respectively, \hat{y}_{Ave} and y_{Ave} represent the means of the forecasted and real values, respectively.

For the first two indices, the smaller the values of the indicators are, the better the performance of the forecasting model will have. For the index R^2 which ranges from 0 to 1, the bigger its values is, the better the accuracy of the forecasting model will be.

4.2 Experiment 1

4.2.1 Applied Dataset

In this case, the building electrical load data were collected every 15 minutes in the Oak Ridge National Laboratory. The Oak Ridge National Laboratory is an Integration Center of the Campbell Creek Research House, and it is a research center for building technologies. In this experiment, we select the building electrical load data collected from March 11, 2014 to July 18, 2014 (totally 94 days) for

training and testing. The time series of building electrical load in this experiment is shown in Fig. 3. From the figure, we can clearly see that this laboratory in different days has similar electricity consumption pattern, that is to say, these electrical load data have periodicity property.

4.2.2 Data Preprocessing and Experimental Setting

In the process of de-noising the building electrical load time series by the wavelet transform method, both the choices of the wavelet basis and the decomposition level are the key factors that affect the final de-noising performance.

Because the wavelet basis function has its own characteristics in signal processing, no wavelet basis function can achieve the best de-noising effect for all types of signals. Considering factors such as support length, vanishing moment, symmetry, regularity, and similarity, in this experiments, Sym8, a special symlets wavelet basis function, is selected because of its good symmetry performance and the ability of reducing the phase distortion in signal analysis and reconstruction.

The other important step is to choose the decomposition level. On one hand, the more layers that the signals are decomposed, the better the separation of noise and signal will be. However, on the contrary hand, the more layers that the signals are decomposed, the more distorted the reconstructed signal will be, and this will affect the final forecasting accuracy. Therefore, we should pay more attention to deal with the contradiction between the above aspects and choose a suitable decomposition level. To realize this objective, cross-validation experiment was made and we found that different values of the decomposition levels have different forecasting performances. Figure 4 demonstrates the indices' values of the forecasting results corresponding to different layer numbers n in this

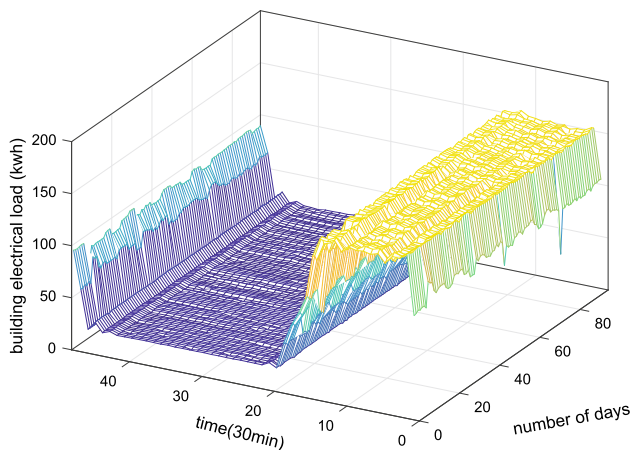


Fig. 3 The original building electrical load data in the first experiment

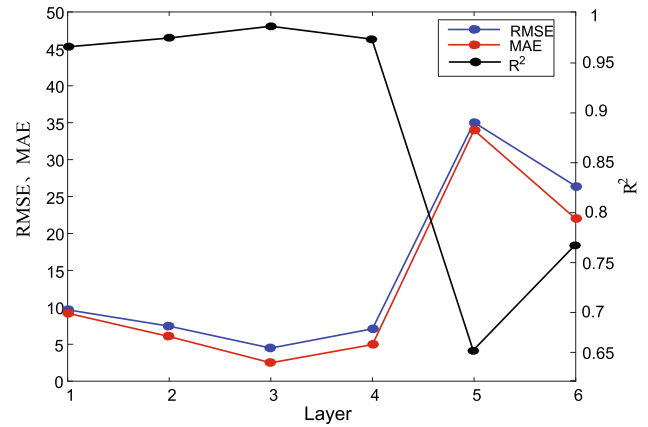


Fig. 4 Parameter selection for the layers of the wavelet decomposition in the first experiment

experiment. The experimental results show that when the decomposition level n is chosen to be 3, the values of RMSE, MAE will be the smallest and the value of R^2 will be biggest, which means that the forecasting performance achieves the best.

Consequently, the low-frequency building electrical load time series $A(3)$ is chosen to reconstruct the filtered building electrical load time series. To show more details, only the first 240 data points of the reconstructed building electrical load time series are shown in Fig. 5b. Compared with the original building electrical load time series as shown in Fig. 5a, we can observe that the noise of time series is almost entirely eliminated and the step features of

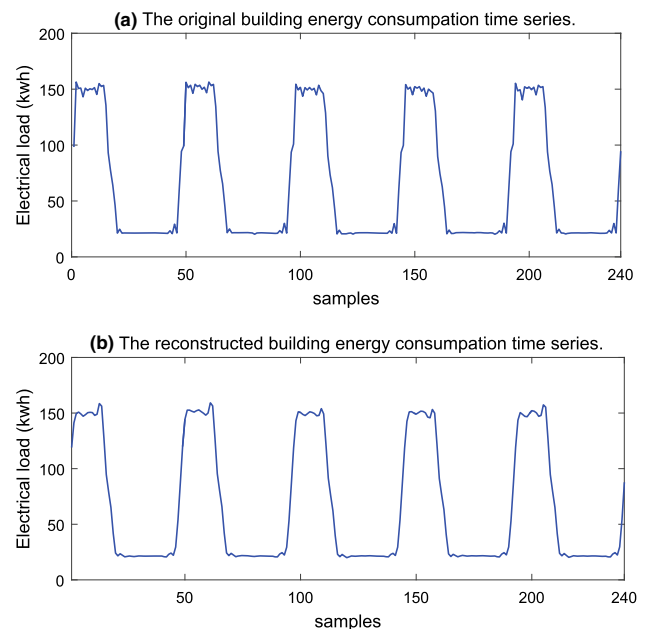


Fig. 5 The results of the wavelet decomposition in the first experiment (data in 5 days)

the original time series remain step in the reconstructed time series.

After removing the noise, we aggregate the sampling period of the time series to half an hour. Then, we use the data collected from the first 75 days as the training dataset while utilizing the data in the left days as the testing dataset.

The daily periodic component is extracted by Eq. (17) and shown in Fig. 6a. After removing the daily periodic component, the residual data are given in Fig. 6b.

In addition, when determining the best clustering radius r_α , we kept the other parameters to be constant. The clustering radius r_α is adjusted by the combination of rough and fine adjustments, the initial forecasting model is constructed, and the optimal radius parameter is determined according to the forecasting model accuracy index. Based on the above training and testing, clustering radius r_α is selected to be 0.25 in this forecasting model.

The parameters of the other four comparison models are as follows: In the FWSIRM-FIS model, the numbers of partitions and iterations are set to 3 and 10, respectively. In the ANFIS model, the numbers of partitions and iterations are set to 3 and 5, respectively. In the BPNN model, the numbers of neurons and iterations are set to 10 and 1000, respectively, while choosing the Logsig as the activation function of neuron nodes. In the MLR model, the least squares estimation is chosen to minimize the sum of squared errors.

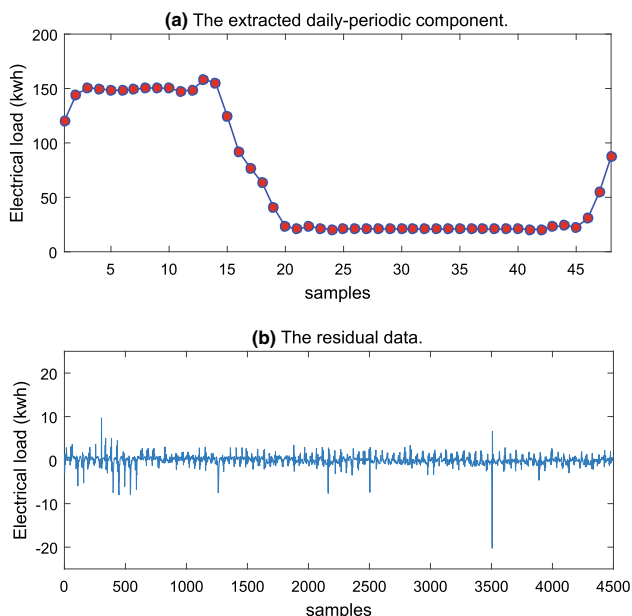


Fig. 6 The extracted daily periodicity and the residual data

4.2.3 Experimental Results and Comparative Analysis

In this experiment, we utilized the residual data to tune the proposed model, while other comparative models are trained by the original building electrical load data. After being tuned, the final forecasting results from the proposed method are shown in Fig. 7. As demonstrated in this figure, the proposed model has satisfactory forecasting performance in this building electrical load forecasting application.

Further, to highlight the improved performance of the proposed model, the MAEs, RMSEs, and R^2 of the five forecasting models in the training and testing processes are listed in Table 1. From this table, we can observe that the proposed model has the smallest values of the indices MAE and RMSE and the biggest value of R^2 , which imply that the best forecasting performance can be provided by the proposed model. Meanwhile, we also plot the forecasting error distributions of the five models in Fig. 8. Compared with other models, the mean of the forecasting errors of the proposed model is much nearer to zero. This means that the proposed model's forecasting errors are much smaller. From Fig. 8, we can also clearly observe that the normal distribution curves of the other comparative models' forecasting errors are wider and flatter than that of the proposed model, which implies that the comparative models have worse forecasting performances.

4.3 Experiment 2

4.3.1 Applied Dataset

The building electrical load dataset in the second experiment were collected from a retail store in Fremont, California, USA. The data sampling period was 15 minutes, including 34848 samples from January 2, 2010 to December 30, 2010. The electrical load time series in this experiment (totally 363 days) is shown in Fig. 9. We can clearly see that this building electrical load data in different days also have similar electricity consumption pattern.

4.3.2 Data Preprocessing and Experimental Setting

In this building electrical load data, we found that there are some missing values, which will affect the accuracy of electrical load forecasting. In order to achieve high reliability of forecasting, it is necessary to adopt scientific methods to complete or adjust these data and ensure the validity of the data. Consequently, the electrical load data at the same moment of previous day were used to fill in the missing values.

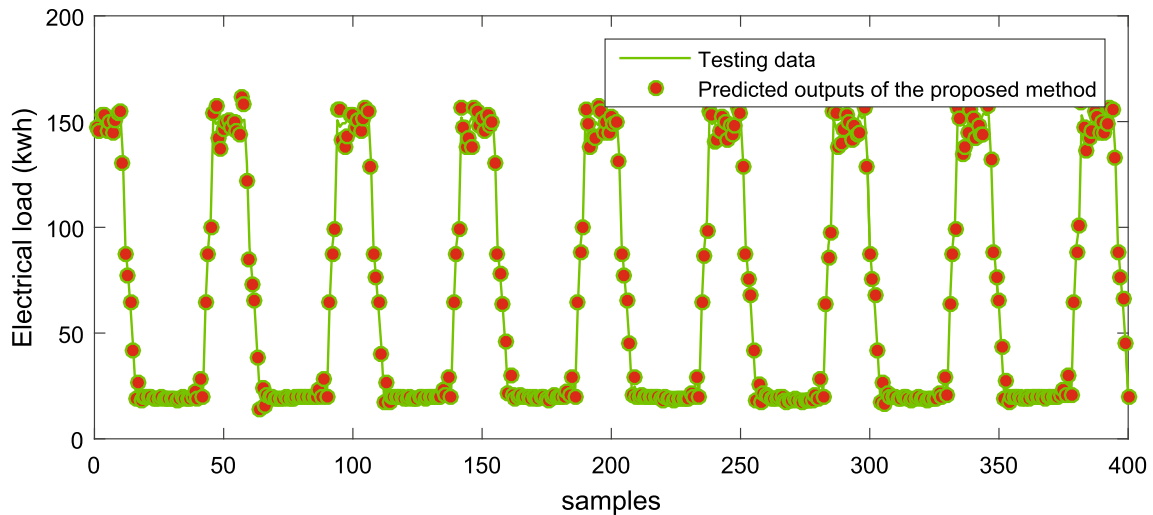


Fig. 7 Forecasting results of the proposed model in the first experiment

Table 1 Comparisons of five forecasting models in the first experiment

	Training accuracy			Testing accuracy		
	RMSE	MAE	R^2	RMSE	MAE	R^2
The proposed method	4.6108	2.7345	0.9934	4.4781	2.4925	0.9937
FWSIRM-FIS	8.2048	4.7264	0.9790	7.6962	4.0684	0.9815
ANFIS	9.2603	5.6535	0.9732	8.9496	5.8372	0.9750
BPNN	10.1691	5.4690	0.9677	9.4869	5.2924	0.9719
MLR	12.0948	7.9936	0.9544	11.6278	7.4518	0.9578

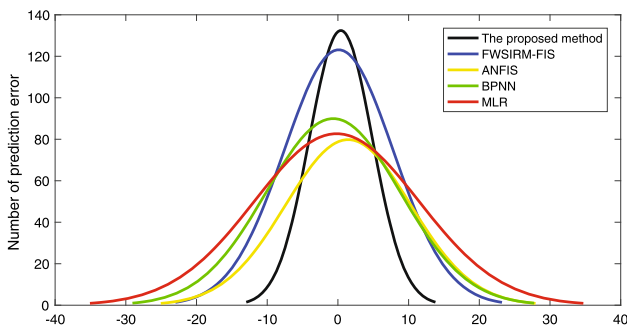


Fig. 8 Forecasting error distributions of the five predictors in the first experiment

Again, the appropriate wavelet basis function and the decomposition level should be selected before decomposing electrical load data series. In this experiment, we still choose Sym8 as the wavelet basis function. For the layers of the wavelet decomposition, it is determined by cross-validation experiment. In the second experiment, we tried different values of the decomposition levels and found that the forecasting results were significantly different too.

Figure 10 demonstrates the forecasting accuracy of the hybrid model with different decomposition levels in this

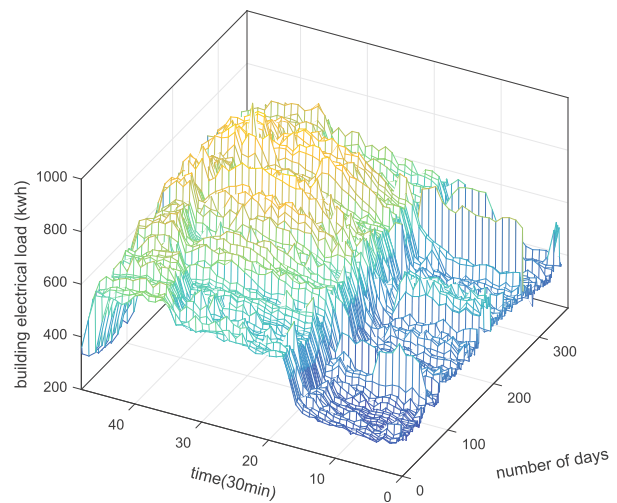


Fig. 9 The original building electrical load data in the second experiment

experiment. Experimental results show that when the decomposition level n is chosen to be 4, the values of RMSE, MAE can achieve the minimum and the value of R^2 can achieve the maximum, which means that this model achieves the best performance under this parameter. Consequently, the low-frequency building electrical load time

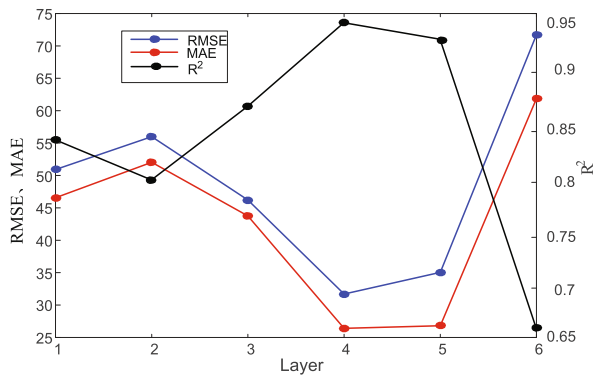


Fig. 10 Parameter selection for the layers of wavelet decomposition in the second experiment

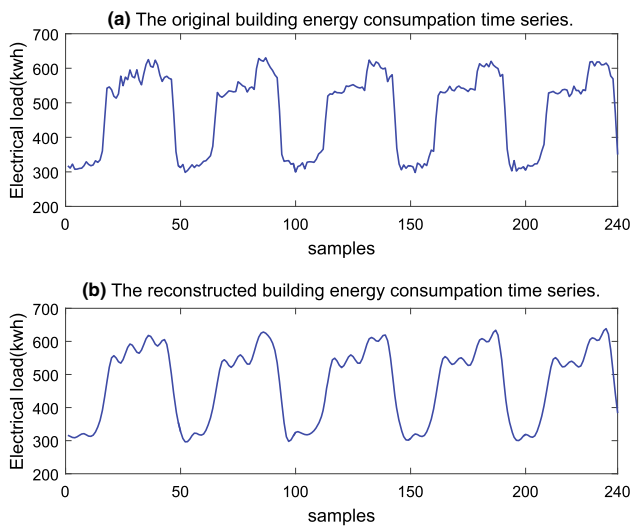


Fig. 11 The results of the wavelet decomposition in the second experiment (data in 5 days)

series A(4) is chosen to reconstruct the new building electrical load time series. The newly reconstructed building electrical load time series is shown in Fig. 11b. In order to show the more details clearly, only the first 240 samples of this data series are shown. Compared with the original building electrical load time series as shown in Fig. 11a, we can see that the reconstructed building electrical load time series can not only remove most of the noise, but also maintain the characteristics of the original time series.

In order to achieve better forecasting accuracy and comparative effect, we select the first 240-day data as the training data, while utilizing the data from the left days as the testing data. In this experiment, the periodic pattern extracted by Eq. (17) is shown in Fig. 12a, and the residual time series data used for training and testing the FWSIRM-FIS for residual forecasting are plotted in Fig. 12b.

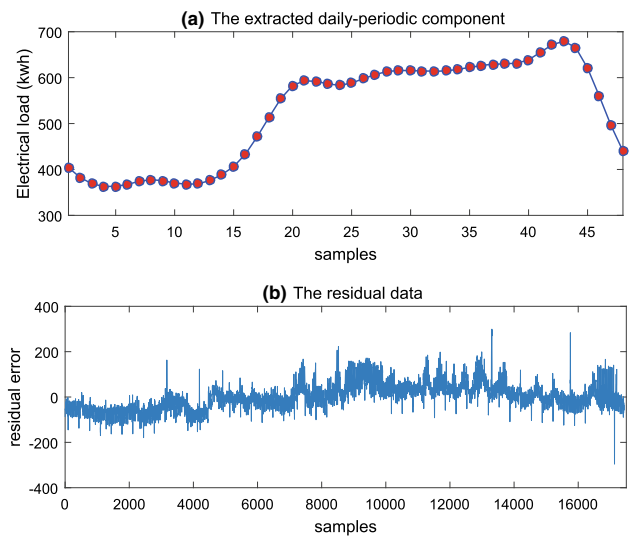


Fig. 12 The extracted daily periodicity and the residual data in the second experiment

In this experiment, the clustering radius r_x is selected to be 0.25 in the forecasting model too. The parameters of the other four comparison methods are as follows: in the FWSIRM-FIS model, the numbers of partitions and iterations are selected as 11 and 5, respectively. In the ANFIS model, the number of iterations is set to 10 and the clustering radius r_x is chosen to be 0.25. In the BPNN model, the numbers of neurons and iterations are set to 25 and 2000, respectively, while choosing the Logsig as the activation function of neuron nodes. In the MLR model, the least square estimation is also selected to minimize the sum of squared errors.

4.3.3 Experimental Results and Comparative Analysis

In this application, Fig. 13 demonstrates the building electrical load forecasting results of the proposed model. To depict more details, only 8 days' forecasting results are drawn in Fig. 13. It can be seen that the forecasting results of the building electrical load by the proposed hybrid model can satisfactorily reflect the actual fluctuation of the electrical load.

To verify the proposed model's forecasting performance, the comparisons of the five forecasting models are listed in Table 2. For the indices MAE and RMSE, the smaller their values are, the smaller the forecasting errors will be, and the better performance the forecasting model will have. For the indices R^2 , the bigger its value is, the better the forecasting performance will be. We can observe from this table that the proposed hybrid model can give the best forecasting performance again.

The forecasting error distributions of the five models are plotted in Fig. 14. This figure indicates that the normal

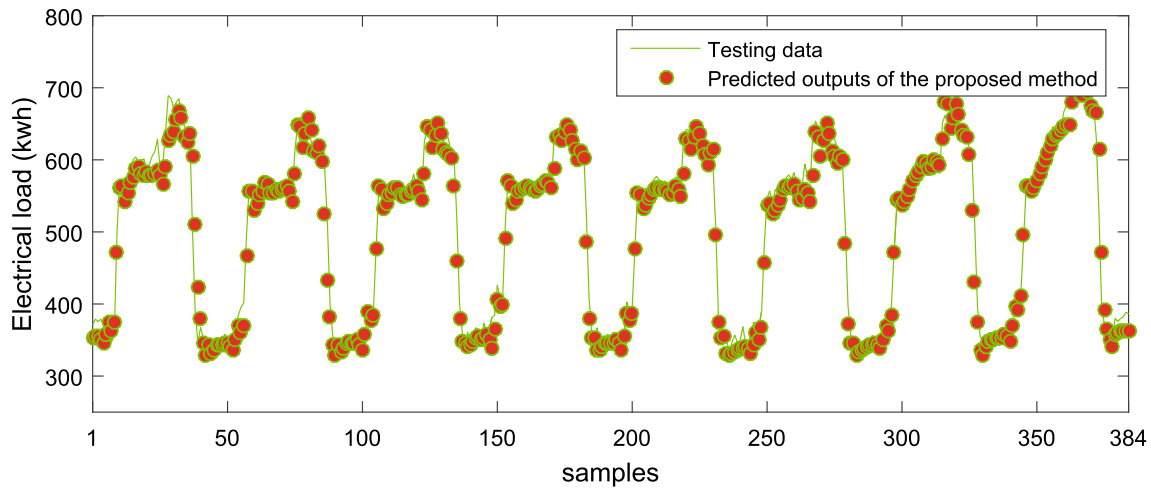


Fig. 13 Forecasting results of the proposed hybrid model in the second experiment

Table 2 Comparisons of five forecasting models in the second experiment

	Training accuracy			Testing accuracy		
	RMSE	MAE	R^2	RMSE	MAE	R^2
The proposed model	22.0767	15.5763	0.9696	30.7655	25.2439	0.9426
FWSIRM-FIS	30.5690	20.6537	0.9417	34.3639	26.3386	0.9284
ANFIS	30.4540	21.2060	0.9383	35.4395	27.2805	0.9238
BPNN	31.1027	22.2340	0.9397	36.7028	28.5681	0.9183
MLR	32.6418	21.9402	0.9326	35.6343	27.3016	0.9230

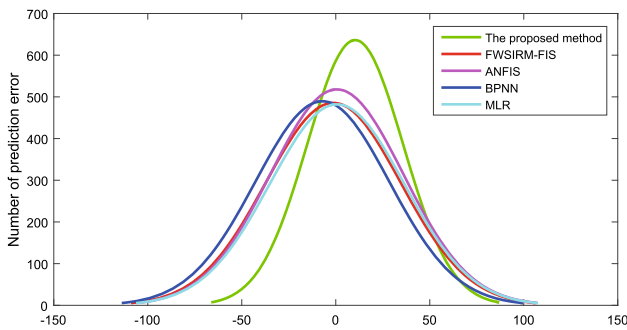


Fig. 14 Forecasting error distributions of the five predictors in the second experiment

distribution curves of the proposed model’s forecasting errors are taller and narrower than those of the other comparative models. This again means that the best performance can be achieved by the proposed model in this application.

5 Conclusions

In this paper, a new hybrid model for the building electrical load forecasting was developed. In order to strengthen the forecasting accuracy, the proposed hybrid model has been

improved from two aspects. First of all, we adopted the wavelet transform to eliminate the noise from the original building electrical load data. Secondly, we extracted the daily periodicity knowledge from the electrical load data to obtain the electrical load trend, and then we utilized the residual data to train one forecasting model as the compensation to the periodic pattern. The residual data-driven model was realized by the FWSIRM-FIS which has powerful approximation ability. To assure the forecasting performance of the FWSIRM-FIS for the residual data, the subtraction clustering method was employed to construct the SIRMs while the least square method was proposed to optimize its parameters. We have applied the proposed hybrid model to two building electrical load forecasting experiments. Experimental and comparison results indicated that the proposed model has the best performance.

In this paper, the model for the periodic pattern is data-driven and data-sensitive. Although the noise in the electrical load data has been removed in this study, the robustness of the periodic pattern model still needs to be improved through more robust methods. Hence, this will be one of our future research directions. On the other hand, more deep models may further improve the forecasting accuracy. Therefore, in the future, we will explore the deep

fuzzy model to achieve more accurate forecasting results for the short-term building electrical load.

Funding This work was financially supported by the Taishan Scholar Project of Shandong Province (TSQN201812092), the Key Research and Development Program of Shandong Province (2019GGX101072), the National Natural Science Foundation of China (61573225, 61473176), and the Youth Innovation Technology Project of Higher School in Shandong Province (2019KJN005).

Compliance with ethical standards

Conflict of interest The authors have no conflict of interest.

References

- Tian, C., Li, C., Zhang, G., Lv, Y.: Data driven parallel prediction of building energy consumption using generative adversarial nets. *Energy Build.* **186**, 230–243 (2019)
- Li, J.Q., Sang, H.Y., Han, Y.Y., Wang, C.G., Gao, K.Z.: Efficient multi-objective optimization algorithm for hybrid flow shop scheduling problems with setup energy consumptions. *J. Clean. Prod.* **181**, 584–598 (2018)
- Crawley, D.B., Lawrie, L.K., Winkelmann, F.C., Buhl, W.F., Huang, Y.J., Pedersen, C.O., Strand, R.K., Liesen, R.J., Fisher, D.E., Witte, M.J., et al.: Energyplus: creating a new-generation building energy simulation program. *Energy Build.* **33**(4), 319–331 (2001)
- Norford, L., Socolow, R.H., Hsieh, E.S., Spadaro, G.: Two-to-one discrepancy between measured and predicted performance of a low-energy office building: insights from a reconciliation based on the doe-2 model. *Energy Build.* **21**(2), 121–131 (1994)
- Crawley, D.B., Hand, J.W., Kummert, M., Griffith, B.T.: Contrasting the capabilities of building energy performance simulation programs. *Build. Environ.* **43**(4), 661–673 (2008)
- Westphal, F.S., Lamberts, R.: The use of simplified weather data to estimate thermal loads of non-residential buildings. *Energy Build.* **36**(8), 847–854 (2004)
- Yao, R., Steemers, K.: A method of formulating energy load profile for domestic buildings in the UK. *Energy Build.* **37**(6), 663–671 (2005)
- Guo, Y., Nazarian, E., Ko, J., Rajurkar, K.: Hourly cooling load forecasting using time-indexed ARX models with two-stage weighted least squares regression. *Energy Convers. Manag.* **80**, 46–53 (2014)
- Kavousi-Fard, A., Kavousi-Fard, F.: A new hybrid correction method for short-term load forecasting based on ARIMA, SVR and CSA. *J. Exp. Theor. Artif. Intell.* **25**(4), 559–574 (2013)
- Cui, H., Peng, X.: Summer short-term load forecasting based on ARIMAX model. *Power Syst. Protect. Control* **43**(4), 108–114 (2015)
- Newsham, G.R., Donnelly, C.L.: A model of residential energy end-use in Canada: using conditional demand analysis to suggest policy options for community energy planners. *Energy Policy* **59**, 133–142 (2013)
- Kouhi, S., Keynia, F.: A new cascade NN based method to short-term load forecast in deregulated electricity market. *Energy Convers. Manag.* **71**, 76–83 (2013)
- Zeng, Y.R., Zeng, Y., Choi, B., Wang, L.: Multifactor-influenced energy consumption forecasting using enhanced back-propagation neural network. *Energy* **127**, 381–396 (2017)
- Badran, S.M., Abouelatta, O.B.: Forecasting electrical load using ANN combined with multiple regression method. *Res. Bull. Jordan ACM* **2**(2), 152–158 (2011)
- Li, K., Hu, C., Liu, G., Xue, W.: Building's electricity consumption prediction using optimized artificial neural networks and principal component analysis. *Energy Build.* **108**, 106–113 (2015)
- Li, T., Guo, J., Luo, D., Tang, J.: Region load forecasting based on load characteristics analysis and GRNN. In: *Proceedings of the 9th International Symposium on Linear Drives for Industry Applications*, vol. 2, pp. 493–500. Springer, Berlin (2014)
- Huo, J., Shi, T., Chang, J.: Comparison of random forest and SVM for electrical short-term load forecast with different data sources. In: *2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, pp. 1077–1080. IEEE, New York (2016)
- Espinoza, M., Suykens, J.A., De Moor, B.: Fixed-size least squares support vector machines: a large scale application in electrical load forecasting. *Comput. Manag. Sci.* **3**(2), 113–129 (2006)
- Chen, Y., Xu, P., Chu, Y., Li, W., Wu, Y., Ni, L., Bao, Y., Wang, K.: Short-term electrical load forecasting using the support vector regression (SVR) model to calculate the demand response baseline for office buildings. *Appl. Energy* **195**, 659–670 (2017)
- Duan, P., Xie, K., Guo, T., Huang, X.: Short-term load forecasting for electric power systems using the PSO-SVR and FCM clustering techniques. *Energies* **4**(1), 173–184 (2011)
- Lei, Z., Su, W.: Mold level predict of continuous casting using hybrid EMD-SVR-GA algorithm. *Processes* **7**(3), 177 (2019)
- Naji, S., Keivani, A., Shamshirband, S., Alengaram, U.J., Jumaat, M.Z., Mansor, Z., Lee, M.: Estimating building energy consumption using extreme learning machine method. *Energy* **97**, 506–516 (2016)
- Zhang, Y., Zhang, L., Liu, Y.: Implementation of maximum power point tracking based on variable speed forecasting for wind energy systems. *Processes* **7**(3), 158 (2019)
- Li, K., Su, H.: Forecasting building energy consumption with hybrid genetic algorithm-hierarchical adaptive network-based fuzzy inference system. *Energy Build.* **42**(11), 2070–2076 (2010)
- Lou, C.W., Dong, M.C.: Modeling data uncertainty on electric load forecasting based on type-2 fuzzy logic set theory. *Eng. Appl. Artif. Intell.* **25**(8), 1567–1576 (2012)
- Li, X., Deng, Y., Ding, L., Jiang, L.: Building cooling load forecasting using fuzzy support vector machine and fuzzy c-mean clustering. In: *2010 International Conference on Computer and Communication Technologies in Agriculture Engineering*, vol. 1, pp. 438–441. IEEE, New York (2010)
- Jain, R., Jain, N., Gupta, Y., Chugh, T., Chugh, T., Hemant, D.J.: A modified fuzzy logic relation-based approach for electricity consumption forecasting in India. *Int. J. Fuzzy Syst.* (2019). <https://doi.org/10.1007/s40815-019-00704-z>
- Li, C., Ding, Z., Yi, J., Lv, Y., Zhang, G.: Deep belief network based hybrid model for building energy consumption prediction. *Energies* **11**(1), 242 (2018)
- Li, C., Ding, Z., Zhao, D., Yi, J., Zhang, G.: Building energy consumption prediction: an extreme deep learning approach. *Energies* **10**(10), 1525 (2017)
- Duan, P.Y., Li, J.Q., Wang, Y., Sang, H.Y., Jia, B.X.: Solving chiller loading optimization problems using an improved teaching-learning-based optimization algorithm. *Optim. Control Appl. Methods* **39**(4), 65–77 (2017)
- Li, C., Yi, J., Wang, H., Zhang, G., Li, J.Q.: Interval data driven construction of shadowed sets with application to linguistic word modelling. *Inf. Sci.* **507**, 503–521 (2020)
- Pal, S.S., Kar, S.: A hybridized forecasting method based on weight adjustment of neural network using generalized type-2

- fuzzy set. *Int. J. Fuzzy Syst.* **21**(1), 308–320 (2019). <https://doi.org/10.1007/s40815-018-0534-z>
33. Soto, J., Castillo, O., Melin, P., Pedrycz, W.: A new approach to multiple time series prediction using mimo fuzzy aggregation models with modular neural networks. *Int. J. Fuzzy Syst.* **21**(5), 1629–1648 (2019). <https://doi.org/10.1007/s40815-019-00642-w>
 34. Yi, J., Yubazaki, N., Hirota, K.: A proposal of SIRMs dynamically connected fuzzy inference model for plural input fuzzy control. *Fuzzy Sets Syst.* **125**(1), 79–92 (2002)
 35. Yi, J., Naoyoshi, Y., Kaoru, H.: Stabilization fuzzy control of parallel-type double inverted pendulum system using the SIRMs dynamically connected fuzzy inference model. *Artif. Intell. Eng.* **15**(3), 297–308 (2001)
 36. Peng, W., Li, C., Zhang, G., Yi, J.: Interval type-2 fuzzy logic based transmission power allocation strategy for lifetime maximization of WSNs. *Eng. Appl. Artif. Intell.* **87**, 103269 (2020). <https://doi.org/10.1016/j.engappai.2019.103269>
 37. Li, C., Wang, L., Zhang, G., Wang, H., Shang, F.: Functional-type single-input-rule-modules connected neural fuzzy system for wind speed prediction. *IEEE/CAA J. Autom. Sin.* **4**(4), 751–762 (2017)
 38. Seki, H., Nakashima, T.: Some consideration of SIRMs connected fuzzy inference model with functional weights. In: *Intelligent Decision Technologies*, pp. 471–476. Springer, Berlin (2012)
 39. Li, C., Gao, J., Yi, J., Zhang, G.: Analysis and design of functionally weighted single-input-rule-modules connected fuzzy inference systems. *IEEE Trans. Fuzzy Syst.* **26**(1), 56–71 (2018)
 40. Dordonnat, V., Koopman, S.J., Ooms, M., Dessertaine, A., Collet, J.: An hourly periodic state space model for modelling French national electricity load. *Int. J. Forecast.* **24**(4), 566–587 (2008)
 41. Xue, X., Han, H., Wang, S., Qin, C.: Computational experiment-based evaluation on context-aware O2O service recommendation. *IEEE Trans. Serv. Comput.* **PP**(99), 1–1 (2019)
 42. Xue, X., Wang, S., Zhang, L., Feng, Z., Guo, Y.: Social learning evolution (SLE): computational experiment-based modeling framework of social manufacturing. *IEEE Trans. Ind. Inform.* **15**(6), 3343–3355 (2019)
 43. Xue, X., Guo, Y., Chen, S., Wang, S.: Analysis and controlling of manufacturing service ecosystem: a research framework based on the parallel system theory. *IEEE Trans. Serv. Comput.* **1** (2019). <https://doi.org/10.1109/TSC.2019.2917445>
 44. Ghaderi, S.F., Azadeh, A., Keyno, H.S.: Forecasting electricity consumption by separating the periodic variable and decompositions the pattern. In: *2007 IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 292–296 (2007)
 45. Keyno, H.R.S., Ghaderi, F., Azade, A., Razmi, J.: Forecasting electricity consumption by clustering data in order to decline the periodic variable's affects and simplification the pattern. *Energy Convers. Manag.* **50**(3), 829–836 (2009)
 46. Zhang, D.: Wavelet transform. In: *Fundamentals of Image Data Mining*, pp. 35–44. Springer, Berlin (2019)
 47. Chiu, S.L.: Fuzzy model identification based on cluster estimation. *J. Intell. Fuzzy Syst.* **2**(3), 267–278 (1994)
 48. Nelles, O.: *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Springer Science and Business Media, Berlin (2013)
 49. Golub, G.H., Van Loan, C.F.: *Matrix Computations*, 4th edn. The Johns Hopkins University Press, Maryland (2013)
 50. Quintana-Orti, G., Quintana-Orti, E.S., Petitet, A.: Efficient solution of the rank-deficient linear least squares problem. *SIAM J. Sci. Comput.* **20**, 1155–1163 (1998)
 51. Zhang, G., Patuwo, B.E., Hu, M.Y.: Forecasting with artificial neural networks: the state of the art. *Int. J. Forecast.* **14**(1), 35–62 (1998)
 52. Güneri Tosunoğlu, N., Apaydın, A.: A new spatial algorithm based on adaptive fuzzy neural network for prediction of crustal motion velocities in earthquake research. *Int. J. Fuzzy Syst.* **20**(5), 1656–1670 (2018). <https://doi.org/10.1007/s40815-018-0483-6>
 53. Karsoliya, S.: Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture. *Int. J. Eng. Trends Technol.* **3**(6), 714–717 (2012)
 54. Wang, L., Zeng, Y., Chen, T.: Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. *Expert Syst. Appl.* **42**(2), 855–863 (2015)
 55. Shine, P., Scully, T., Upton, J., Murphy, M.D.: Multiple linear regression modelling of on-farm direct water and electricity consumption on pasture based dairy farms. *Comput. Electron. Agric.* **148**, 337–346 (2018)



Chengdong Li received the Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2010. He is currently a full Professor with the School of Information and Electrical Engineering, Shandong Jianzhu University. His major research interests include data-driven modeling and control, fuzzy logic theory and applications, fuzzy neural networks, and other computational intelligence methods. His research interests

also include the intelligent control methods with applications to intelligent buildings and smart homes. He has authored and co-authored over 100 papers in international journals and conferences. He is the associate editor of *Acta Automatica Sinica* and *IEEE/CAA Journal of Automatica Sinica*. He has also been the Program Committee Member of several international conferences and the Reviewer for several international conferences and journals.



Minjia Tang is pursuing her M.S. degree in Control Science and Engineering at the School of Information and Electrical Engineering, Shandong Jianzhu University. Her current research interests include intelligent buildings and building energy consumption prediction.



Guiqing Zhang is a Professor at the School of Information and Electrical Engineering, Shandong Jianzhu University. He received his Ph. D. degree from the Xi'an Jiaotong University in 2002. His research interest covers intelligent control methods, intelligent buildings, smart home, and Internet of Things.



Chongyi Tian received Ph.D degree from Shandong University, Jinan, China. He is now a lecture at the School of Information and Electrical Engineering, Shandong Jianzhu University. His research interests include the smart buildings, smart cities, and computational intelligence methods.



Ruiqi Wang received his MS and Ph.D degrees from Shandong University, Jinan, China. He is now with the State Grid Shandong Integrated Energy Services Co., Ltd., Jinan, China. His current research interests include the smart grid, the Internet of Things, and intelligent buildings.