

For publication in *Energy*

Industrial Artificial intelligence Based Energy Management System: Integrated framework for Electricity Load Forecasting and Fault Prediction

Yusha Hu¹, Jigeng Li¹, Mengna Hong¹, Jingzheng Ren², Yi Man^{2,*}

1. State Key Laboratory of Pulp and Paper Engineering, South China University of Technology, Guangzhou 510640, China
2. Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China

* Corresponding author. Email: yi.man@polyu.edu.hk, manyi@scut.edu.cn (Y. Man)

Abstract

Forecasting accuracy electricity load can help industrial enterprises optimise production scheduling based on peak and off-peak electricity prices. The electricity load forecasting results can be provided to an electricity system to improve electricity generation efficiency and minimize energy consumption by developing electricity generation plans in advance and by avoiding over or under the generation of electricity. However, because of the different informatization levels in different industries, few reliable intelligent electricity management systems are applied on the power supply side. Based on industrial big data and machine learning algorithms, this study proposes an integrated model to forecast short-term electricity load. The hybrid model based on the hybrid mode decomposition algorithms is proposed to decompose the total electricity load signal. To improve the generalisation ability of the forecasting model, a dynamic forecasting model is proposed based on the improved hybrid intelligent algorithm to forecast the short-term electricity load. The results show that the accuracy of the proposed dynamic integrated electricity load forecasting model is as high as 99%. The integrated framework could forecast abnormal electricity consumption in time and provide reliable evidence for production process scheduling.

Keywords: electricity load; dynamic forecasting model; energy system analysis; energy system optimisation; artificial intelligence

Nomenclature

| | |
|-------|----------------------------|
| a | A Lagrangian multiplier |
| A | The electricity load data |
| b | The bias |
| c | The penalty coefficient |
| c_1 | Personal learning elements |
| c_2 | Social learning elements |

| | |
|-----------|---|
| d_i | The regression coefficient |
| $F(s_q)$ | The fitness function |
| g | The regression coefficient |
| $gbest_q$ | The best place that all particles in the group have experienced |
| h | A new signal |
| $H(t)$ | The analytic signal |
| I | The identity matrix |
| IQR | Interquartile distance |
| k | The lag order |
| m | The integer representing the length of the comparison vector |
| m_1 | The mean values of the upper and lower envelope lines |
| n | The sample step length |
| N | Maximum dimension |
| $o(t)$ | The original signal |
| p | A real number representing the measure of similarity |
| $pbest_q$ | The best place that each individual has experienced |
| Q_1 | The 1st quartiles |
| Q_3 | The 3rd quartiles |
| r_k | The lag k autocorrelation coefficient |
| R_{emp} | The error control function |
| R^n | An n-dimensional vector |
| s_q | The position of the q th particle |

| | |
|-----------------|---|
| S | The population scale |
| t | Sampling time point |
| T | Time |
| $u(t)$ | The mode |
| V_i | The velocity of particle i |
| v_{in}^{kk} | The n th component of the velocity of the granule q at the k th iteration |
| w | The weight coefficient |
| w^T | Transposed matrix of w |
| $\ w\ ^2$ | The complexity of the function |
| $X(l)$ | The original signal |
| $z(l)$ | An n -dimensional time-series |
| $Z(o)$ | The n -dimensional reconstruction combination vector |
| α | Quadratic penalty factor |
| β | The kernel parameter |
| γ | The centre frequency of each IMF component |
| $\delta(t)$ | Dirac distribution |
| $\lambda(t)$ | Lagrange multiplier |
| ε_i | The quadratic term of error |
| μ | The average value |
| σ | The standard deviation |
| $\varphi(x)$ | Nonlinear function |
| ω | The centre frequency |

Abbreviations

| | |
|-------|---|
| ACF | Autocorrelation Function |
| ApEn | Approximate Entropy |
| DCS | Distributed Control System |
| ELF | Electricity load forecasting |
| EMD | Empirical Mode Decomposition |
| LSSVM | Least-Squares Support Vector Machine |
| MAPE | Mean Absolute Percent Error |
| PLC | Programmable Logic Controller |
| PSO | Particle Swarm Optimisation |
| RMSE | Root Mean Square Error |
| STELF | Short Term Electricity Load Forecasting |
| VMD | Variational Mode Decomposition |

Subscript

| | |
|------|-------------------------|
| i | A Lagrangian multiplier |
| q | The order of particles |
| k | The decomposition order |
| kk | Number of iterations |
| m | Dimension |

1 Introduction

To achieve the goal of the Paris Agreement, the United Nations Environment Programme (UNEP) reported that global carbon emissions need to be reduced by 2.7% per year between 2020 and 2030 [1]. Research shows that carbon emission is directly related to energy consumption, and the optimal use of energy can assist in controlling carbon emissions [2]. In recent years, terminal energy consumption is increasing rapidly, wherein the proportion of electricity consumption has seen a sharp rise [3]. In 2018, the proportion of electricity consumption accounted for 25.5% of the total terminal energy consumption [4]. Industrial enterprises comprise the largest proportion of the total electricity consumption, accounting for 70.3% of that in China [4]. However, China has proposed to reduce 18% and 22% of energy consumption and carbon emission for the major industrial enterprises, respectively [5]. Thus, industrial enterprises need to ensure cleaner and sustainable production.

Efficiency improvement is the most economical approach to minimize energy consumption. However, the electricity load curve of the electricity system often fluctuates greatly and frequently when a considerable number of industrial production and operation processes are unsteady. With the adjustment of the economic structure, the total electricity consumption in peak periods is increasingly higher than that in the off-peak periods. Low peak shaving capacity leads to the low efficiency of electricity system operation and a large amount of energy waste. Currently, electricity cannot be stored in large quantities because of the real-time production process. Thus, to improve the operational efficiency of the electricity system and save energy and reduce emissions, a key method is to maintain the balance between the supply and demand of electricity generation and electricity consumption at all times, thereby avoiding insufficient or excess power supply problems. To achieve the above goals, the electricity system needs to identify changes in the electricity load on the electricity consumption side in advance. Thus, it is very important to realise accurate electricity load forecasting (ELF).

Electricity load data are accumulated using some industrial enterprises with a high

level of information technology. The electricity consumption management system has been established to monitor and count electricity consumption information. However, owing to the lack of a study on the big data analysis of the electricity consumption data of industrial enterprises, the key data features are hard to extract [6]. Thus, current electricity consumption management systems cannot achieve accurate ELF and peak load shifting.

This study proposes an ELF model for the electricity consumption side. The electricity scheduling model was proposed using forecasting electricity load data and TOU electricity prices. The electricity scheduling model can reduce the operation of intermittent electricity devices in peak periods and shift the peak load. Further, reporting the accurate forecasts of the electricity load of industrial enterprises to the electricity system can help optimise the generation scheduling in advance, and help avoid over or under generation. Based on optimised generation scheduling, power generation enterprises can use efficient and clean generation devices to minimize the carbon emissions of the entire electricity system.

However, the characteristics of electricity consumption in industrial processes are complex. In particular, for the manufacturing industry, the electricity load in peak periods has a gap compared to that in the valley periods because of factors such as orders, production scheduling, and others. Many factors influence the electricity consumption characteristics of different industrial production processes, such as industry characteristics, process differences, and enterprise management.

Electricity load forecasting has obtained great results in power grids [7], renewable energy (such as wind power [8], photovoltaic power [9], etc.) generation, building energy [10], and other fields. However, the electricity loads of these industries have either a periodic character or clear influence factors. Some researches show that using environmental temperature as an input variable can help improve the accuracy of the forecasting model [11]. However, environmental factors such as environmental temperature have little impact on the electricity load of industrial enterprises. The electricity load of most industrial enterprises is non-linear, and it has no periodicity characteristics. Thus, it is difficult to select key factors as input variables to build an

electricity load forecasting model. There are few studies on building forecasting models with non-periodic electricity loads. However, the main algorithms of these forecasting models are hybrid intelligent algorithms or combination algorithms, and therefore, their accuracy is not very high [9].

Short-term electricity load forecasting (STELF) models can be divided into decomposition algorithm-based forecasting models and non-decomposition algorithm-based forecasting models [12]. The decomposition algorithm-based forecasting models can be considered as adding the decomposition algorithm to non-decomposition algorithm-based forecasting models. Many study results showed that the accuracy of the decomposition algorithm-based forecasting model is higher than that of the non-decomposition algorithm-based forecasting model [13]. Moreover, compared with non-decomposition algorithm-based forecasting models, the decomposition algorithm-based forecasting model does not suffer from a time delay. Commonly used decomposition algorithms are variational mode decomposition (VMD) [14], empirical mode decomposition (EMD) [15], and singular spectrum analysis (SSA) [16], wavelet transform (WT) [17].

Given the volatility and instability of industrial process electricity loads, the above decomposition models are not always able to accurately extract the complex feature correlations that exist in nonlinear and nonstationary data series [18]. To solve the above problems, two types of approaches have been proposed: hybrid decomposition algorithm-based models and improved classical decomposition models [19]. Although the improved classical decomposition model solves the problems of the original decomposition model, it introduces new problems. Moreover, some researches show that the multiple decomposition algorithm-based integrated models have better results than the improved classical decomposition models [20].

As illustrated in the literature, decomposition algorithms have been introduced to forecasting models to improve accuracy for many years. The hybrid decomposition algorithm-based models show better signal decomposition performance than other decomposition algorithms. Although many hybrid decomposition algorithm-based models have been introduced to ELF models, the hybrid decomposition algorithm

selection for electricity loads of the industrial process remains a concern. The electricity load of the industrial process has the characteristics of enormous fluctuations and frequent and aperiodic changes. Thus, the key to proposing an accurate ELF model for the industrial process is to choose the right hybrid decomposition algorithm to extract the whole signal characteristics. On the other hand, few ELF models have been applied in real industrial processes. One reason is that industrial enterprises lack of data preprocessing tools. Another reason is that the ELF model can not be used to realise its economic value. Thus, it is important to propose an integrated framework for industrial process to forecast electricity load accurately, improve the versatility of the ELF model and realise the application values.

This study proposes an integrated framework that can be implemented directly in industrial enterprises. The proposed integrated framework can forecast dynamically and optimise the production process including data acquisition, data preprocessing, electricity load forecasting, and production process optimisation. To propose a general-purpose for the STELF model for industrial enterprises, this study collects data from an industrial enterprise wherein the electricity load has the characteristics of enormous fluctuations and frequent and aperiodic changes. The proposed model can be applied to industrial processes with little fluctuations and periodicity in electricity loads, industrial processes with electricity fluctuations, or industrial processes with aperiodicity in electricity loads. The data preparation sub-module includes three parts: outlier elimination, data filling, and data filtering. Further, this study proposes a dynamic STELF model for industrial enterprises based on empirical mode decomposition, variational mode decomposition, particle swarm optimisation and least-squares support vector machine (EMD-VMD-PSO-LSSVM). The EMD-VMD-based combined model is used to decompose electricity loads. The PSO-LSSVM based hybrid algorithm is used to propose the forecasting model. A dynamic period is selected to propose a dynamic forecasting model. The production process optimisation sub-module includes fault forecasting and electricity consumption optimisation. The electricity consumption data from an industrial enterprise were collected to verify the proposed model. In addition, four different electricity load forecasting models are studied as contrasting

cases. The verification results show that the EMD-VMD-PSO-LSSVM based forecasting model has the best forecasting performance. Finally, this study tests the production process optimisation layer based on dynamic forecasting results.

The novelty and contributions of this study are as follows:

- The integration framework is proposed to address the issue of accurate electricity load forecasting and its application value.
- The EMD-VMD based decomposition model is proposed in the integrated framework to extract signals with stability and regularity to improve the accuracy of the ELF model.
- The integrated framework could forecast abnormal electricity consumption in time and provide reliable evidence for production process scheduling.
- The integrated framework can be applied directly in industrial process systems and can reduce production costs
- The proposed integrated framework has been implemented in a real papermaking enterprise and gets good results.

2 Methodology: Integrated Framework

It is very common for industrial enterprises to link to the Internet. Large-scale industrial enterprises collect hundreds of millions of data every day. The computational complexity of the model increases exponentially when the mathematical model develops into an integrated model. General-purpose computers cannot process hundreds of millions of data streams in real-time, and therefore, they cannot complete the calculation of the integrated model in real-time. Thus, this study proposes an integrated framework based on a cloud platform. First, a data acquisition layer is designed to collect and transmit electricity data from electricity meters to the cloud platform for storage. Then, the data pre-processing layer is built to classify and pre-process data. Next, this study proposes an ELF model based on the EMD-VMD-PSO-LSSVM hybrid algorithm. Finally, this study builds a production process optimisation layer. The flow chart of the integration framework is shown in Fig. 1.

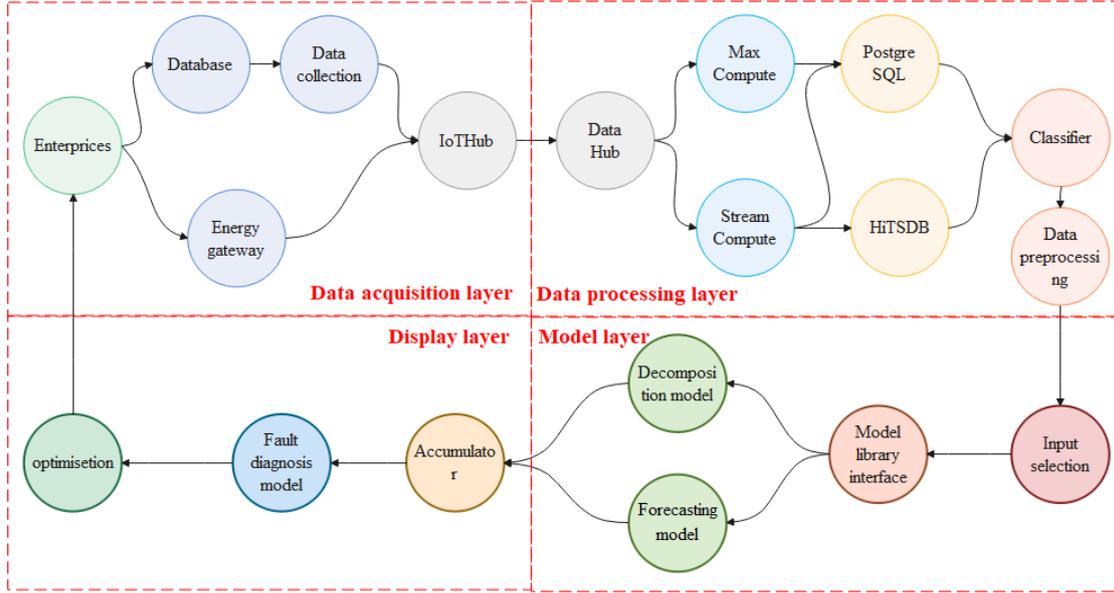


Fig. 1. Flowchart of the integration framework based on the cloud platform

2.1. Data acquisition

The study collects electricity data from the data acquisition system of industrial enterprises. The system includes three main layers: data acquisition, information transmission, and data storage. The data acquisition layer is composed of intelligent instruments and a programmable logic controller (PLC), which is used to collect real-time data from electrical devices. The information transmission layer is used to classify and transmit the collected data using a distributed control system (DCS). The data storage layer is used to save the transmitted electricity data in the database for accessing the data conveniently at any time.

2.2. Data pre-processing

For the production process, transmission signal interruption and unscheduled shutdown often occur when transmitting and storing digital signals. The STELF model is sensitive to electricity data. Owing to data quality problems such as incomplete and inconsistent data, data mining can cause serious consequences, such as increasing the uncertainty of system analysis results, which can lead to unreliable outputs, thereby reducing the efficiency of data pre-processing and affecting the performance of mining algorithms [21]. Thus, data preprocessing is very important. However, missing data is

a common problem that occurs after the collection and pre-processing of electricity load data, and the validity of forecasting results can be affected by missing data [22]. Thus, the data pre-processing layer includes four sub-modules: data acquisition, outliers and non-numerical data elimination, data filling, and data filtering. The flowchart of the data preprocessing is shown in Fig. 2. The details are as follows:

(1) Data acquisition: collecting the original electricity consumption data in the required time from the database.

(2) Outliers and non-numerical data elimination: removing outliers using the boxplot and non-numerical-type data by screening.

(3) Data filling: filling the missing data using the least-square fitting interpolation method.

(4) Data filtering: using the twice moving average filtering method to filter the electricity load data.

The boxplot states that the data distributed beyond $(Q_1 - 1.5 \times IQR, Q_3 + 1.5 \times IQR)$ can be considered outliers and should be removed [23]. Q_1 denotes 1st quartiles, Q_3 denotes 3rd quartiles, IQR denotes as interquartile distance.

The least-square fit algorithm [24] is a fitting polynomial, which is constructed by interpolating other discrete points in the interpolation interval. The polynomial coefficients are solved using the least squares method. The missing points are filled by the least-squares fit algorithm.

Moving average filtering [25] is defined as $F_t = \sum_{i=1}^n A_{t-i} / n$, and A_{t-i} is the electricity load data at time $t-i$, and n is the sample step length.

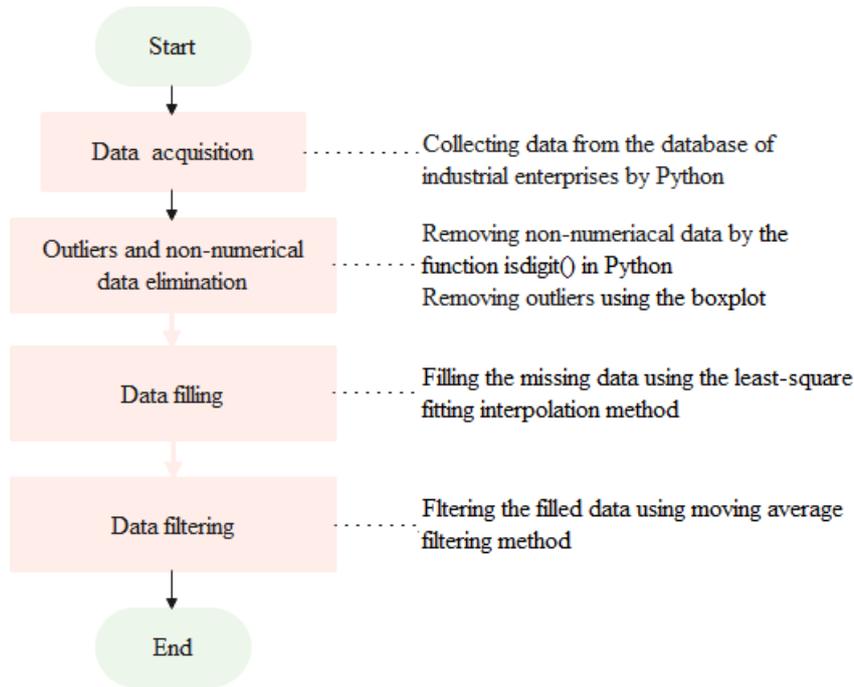


Fig. 2. Data preprocessing flowchart

2.3. Forecasting model

The accuracy of the ELF model cannot be very high when directly forecasting the electricity load data with enormous fluctuations and aperiodicity. The forecasting results can have a time delay when using key factors as input variables. To meet the accuracy demand of the forecasting model for industrial enterprises, this study proposes an EMD-VMD-based combination method to decompose the electricity load signal into electricity load components with small fluctuations and periodicity. However, after decomposing the electricity load using the EMD-VMD-based model, the increasing number of decomposed components will increase the mathematical calculation complexity and prediction error of the proposed forecasting model. The approximate entropy (ApEn) method can calculate the different complexities of the decomposed components. Thus, this study uses the ApEn method to reconstruct the decomposed components with similar complexity to reduce the dimension. Currently, a general method to select key factors affecting electricity load as input variables is lacking. However, some studies have reported that past effective electricity load can affect the present effective power [26]. Thus, this study uses the lag autocorrelation function to

select the input variables. After decomposition and reconstruction, the components are categorized as components with a single trend, components with periodicity, and components with stable fluctuation. The PSO-LSSVM algorithm can deal with high-dimensional and non-linear substances. However, the accuracy of the PSO-LSSVM based ELF model will not be high when the component of the electricity load sequence has enormous fluctuation. Thus, the study uses the PSO-LSSVM based ELF model to predict the components with small fluctuations after decomposition and reorganisation. A flowchart of the EMD-VMD-PSO-LSSVM is shown in Fig. 3.

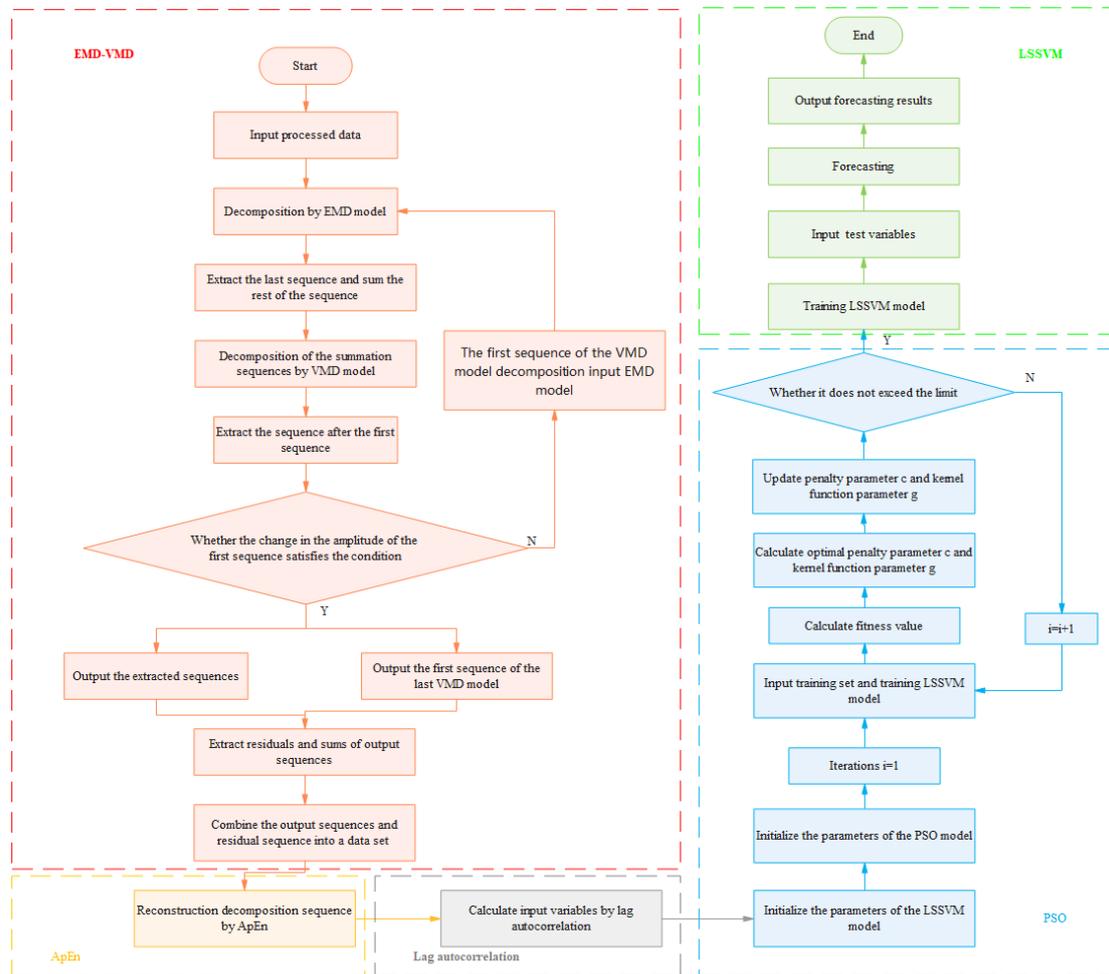


Fig. 3. Flowchart of the EMD-VMD-PSO-LSSVM based forecasting model

2.3.1. EMD-VMD based signal decomposition model

The EMD algorithm and advanced EMD algorithms are introduced into the hybrid forecasting model because of their strong adaptive capacity [27]. However, the EMD

algorithm and the advanced EMD algorithms have problems with an uncertain number of decomposed components and modal mixing, which can affect the signal decomposition results. To solve the above problems, the VMD algorithm is introduced into the hybrid forecasting model [28]. The decomposition model based on the VMD algorithm is used to decompose the electricity load signal. The decomposition results can be distorted and irregular when the selected decomposition number is large. The main trend can be extracted. However, the fluctuation frequency cannot be reduced when the selected decomposition number is too small. Thus, it is not suitable to build a forecasting model based on a single VMD algorithm. However, the VMD-based decomposition model has the advantage of decomposing signals to regular signals, except for the main trend. The EMD-based decomposition model can decompose the signal to low-frequency components, i.e., the EMD-based decomposition model has the advantage of decomposing signals to stable signals. Thus, this study combines the advantages of the EMD algorithm and the VMD algorithm to propose an EMD-VMD algorithm-based decomposition model. The proposed combined decomposition model can decompose the signal with enormous fluctuations and aperiodicity into a periodic and stable signal.

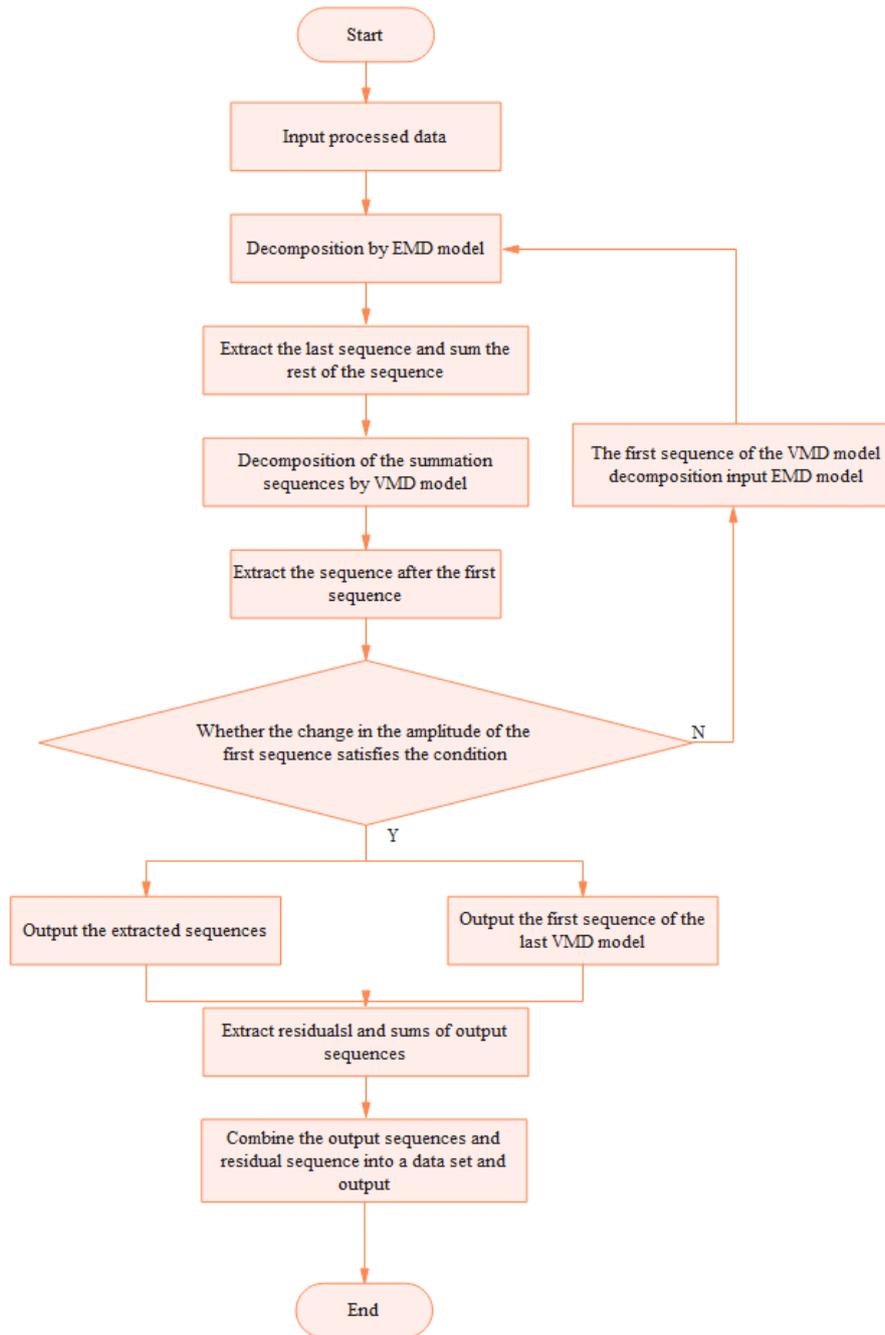


Fig. 4. Flowchart of the technical route for the EMD-VMD based decomposition model

Fig. 4 shows the flowchart of the technical route for the EMD-VMD-based decomposition model. The steps are listed below.

(1) Decompose the preprocessed electricity load signal in Section 2.2 using the EMD algorithm to extract the last decomposed component and add the other decomposed components.

(2) Initialise the parameters of the VMD algorithm (such as the decomposition number is set as 3). Then decompose the added component in step 1 and extract the second and third components.

(3) Repeat steps (1) – (2) when the maximum variation range of the first decomposed component obtained using the VMD algorithm is greater than the set value (the set value is 100 in the study), otherwise, skip to step 4.

(4) All extracted components are combined into one dataset. Extract errors between the sum of all extracted components and the preprocessed electricity load signal in step 1, and input it in the last column of the dataset.

(5) Output the dataset.

2.3.1.1. EMD algorithm

The EMD algorithm can decompose a signal with poor performance into a series of intrinsic mode functions (IMF) with good performance and residual [29]. The calculation steps of the EMD algorithm are as follows [30]:

(1) Determine all maximum and minimum points of the original signal $X(t)$, fit them with a cubic spline function, and obtain the upper and lower envelope lines of $X(t)$. The mean values of the upper and lower envelope lines are denoted as m_1 . Subtract m_1 from $X(t)$ to obtain a new signal h that filters out low-frequency signals, that is, $h = X(t) - m_1$. The above process is repeated until h is sufficiently stable to obtain the first-order IMF (c_1), which represents the highest frequency components of $X(t)$.

(2) Subtract c_1 from $X(t)$ to obtain a new signal r_1 without a high-frequency component. Decompose r_1 to obtain the second-order IMF c_2 . Repeat step (1) until the last signal r_n cannot be decomposed. Here, r_n represents the trend or mean value of $X(t)$.

2.3.1.2. VMD algorithm

Variational mode decomposition (VMD) can achieve adaptive signal decomposition by searching the optimal solution of the constrained variational model, and it can decompose the original signal into a series of modal components with sparse characteristics [31].

The calculation steps of the VMD algorithm are as follows:

(1) For each mode $(u(t))$, the Hilbert transform [32] is used to compute the related analytical signal, the formula is given as

$$H(t) = (\sigma(t) + \frac{j}{\pi \times t}) * u(t) \quad (1)$$

where $H(t)$, $\delta(t)$, t , and $*$ represent the analytic signal, Dirac distribution, sampling time point, and convolution, respectively. Further, $j^2 = -1$.

(2) Mix the centre frequency ω with the analytic signal $H(t)$, and shift the centre frequency ω pre-estimated for each modal analytical signal ω_k to modulate the frequency spectrum of the mixed signal to the corresponding fundamental band, the formula [32] is given as

$$B(t) = [(\sigma(t) + \frac{j}{\pi \times t}) * u(t)] \times e^{-j\omega_k t}. \quad (2)$$

(3) Calculate the L^2 -norm of the gradient square of the fundamental frequency band in Eq. (2), and estimate the bandwidth of each mode component. The constrained variational method [33] is equated as

$$\min\{\sum_k \left\| \partial_t [(\sigma(t) + \frac{j}{\pi \times t}) * u(t)] \times e^{-j\omega_k t} \right\|_2^2\} \quad (3)$$

where $f(t) = \sum_k u(t)$.

(4) Introduce the quadratic penalty term and the Lagrange multiplier to obtain an unconstrained problem, and finally solve the problem. The formula [33] is given by

$$L(\{u_k\}, \{\gamma_k\}, \lambda) = \alpha \times \sum_k \|\partial_t \times B(t)\|_2^2 + \|o(t) - \sum_k u_k(t)\|_2^2 + [\lambda(t), o(t) - \sum_k u_k(t)] \quad (4)$$

where $\{u_k\} = \{u_1, u_2, \dots, u_k\}$, $\{\gamma_k\} = \{\gamma_1, \gamma_2, \dots, \gamma_k\}$, $\lambda(t)$, α , and $o(t)$ represents the k IMF components obtained by decomposing the signal, centre frequency of each IMF component, Lagrange multiplier, quadratic penalty factor, and original signal, respectively.

According to the above content, the EMD algorithm can decompose components with a single trend, and VMD can decompose components with sparse characteristics. Thus, the study proposed a combined decomposition model based on the EMD and VMD algorithms to extract components with a single trend or with sparse characteristics to improve the accuracy of the ELF model.

2.3.2. Approximate entropy

ApEn is a non-negative number used to represent the complexity of a time series [35]. ApEn values are similar when time series have the same trend. Thus, this paper uses similar ApEn values to reconstruct decomposed components. The ApEn [35] is defined as

$$ApEn = \Phi^n(r) - \Phi^{n+1}(p) \quad (5)$$

where $\Phi^n(p) = (N - n + 1)^{-1} \times \sum_{i=1}^{N-n+1} \log(C_i^n(p))$, $C_i^m(r) = \frac{\text{number of } Z(o) \text{ such that } d[Z(l), Z(o)] \leq p}{N-m+1}$, and $1 \leq l \leq N-m+1$. Further, $Y(o)$ denotes the m -dimensional reconstruction combination vector, i.e., $Z(1), Z(2), \dots, Z(N-n+1)$, $Z(l) = [z(l), z(l+1), \dots, z(l+n-1)]$, $z(l)$ denotes an n -dimensional time-series obtained by equal sampled time, m denotes an integer representing the length of the comparison vector, and p denotes a real number representing the measure of similarity.

2.2.3. Lag autocorrelation function

The present effective power is affected by the past effective electricity load. To determine the influence of the past effective electricity load on the present effective power, an autocorrelation function (ACF) is used to guide the selection of information feature subsets. The lag ACF is used to choose the input variables. The historical total electricity loads are considered the input variables if the absolute lag autocorrelation coefficients exceed 0.9. The lag k autocorrelation coefficient r_k [36] is defined by

$$r_k = r(X_t, X_{t-k}) = \frac{\sum_{t=k+1}^n (X_t - \bar{X}) \times (X_{t-k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2} \quad (6)$$

where X is a data set based on a time series $X = \{X_t; t \in T\}$.

2.3.4. LSSVM algorithm

The least-squares linear system is used by the LSSVM algorithm as a loss function instead of the quadratic programming method used by the conventional SVM algorithm. The basic principle of LSSVM is to construct an optimal decision function in the

selected nonlinear mapping space. The theory of structural risk minimisation is used when constructing an optimal decision function. The kernel function of the original scope is used instead of using point multiplication in the high-dimensional feature space. Assume the sample is an n -dimensional vector where the values of a region in the sample are represented as $(x_1, y_1), \dots, (x_j, y_j) \in R^n \times R^n$.

First, the sample ($\Psi(X)$) is mapped from the initial space (R^n) to the feature space ($\Psi(X) = (\varphi(x_1), \varphi(x_2), \dots, \varphi(x_j))$) using a nonlinear map. The optimal decision function [38] in the nonlinear mapping process is constructed using

$$y(x) = w^T \times \varphi(x) + b \quad (7)$$

where w , w^T , $\varphi(x)$, and b denote the weight coefficient of the samples in the feature space, transposed matrix of w , nonlinear function, and the bias, respectively.

Based on the structural risk minimisation theory, the constrained optimisation [37] is defined as

$$R = \frac{1}{2} \times \|w\|^2 + c \times R_{emp} \quad (8)$$

where $\|w\|^2$ dominates the complexity of the function, and c denotes the penalty coefficient. Further, R_{emp} denotes the error control function, which is also called the insensitive loss function.

Widely used loss functions include linear loss functions, quadratic loss functions, and hinge loss functions. The variation in the loss function makes the form of the SVM different. As the optimisation objective, the loss function of the least-squares linear system is the quadratic term of error (ε_i). The resulting optimisation issue of the LSSVM [37] can be formulated as

$$\min_{w, b, \varepsilon} J(w, \varepsilon) = \frac{1}{2} \times w^T \times w + c \sum_{i=1}^l \varepsilon_j^2 \quad (9)$$

where $y_i = \varphi(x_i) \times w^T + b + \varepsilon_i, i = 1, \dots, l$.

The Lagrangian [37] is represented by

$$L(w, b, \varepsilon; a) = \frac{1}{2} \times w^T \times w + c \times \sum_{i=1}^l \varepsilon_i^2 - \sum_{i=1}^l a_j \times (w^T \times \varphi(x_j) + b + \varepsilon_j - y_j), \quad (10)$$

where $a_j, j=1, \dots, n$, is a Lagrangian multiplier.

According to the Lagrangian, the optimisation conditions are given as (11)

$$\frac{\partial L}{\partial w} = 0, \quad \frac{\partial L}{\partial b} = 0, \quad \frac{\partial L}{\partial a} = 0. \quad (11)$$

Using Eq. (11), the solution is found solving the system of linear equations [37] expressed in matrix form as

$$\begin{bmatrix} 0 & I^T \\ I & \Omega + I/y \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (12)$$

where $\Omega_{ij} = \varphi(x_j)^T \times \varphi(x_p) = K(x_j, x_p)$, $a=[a_1, a_2, \dots, a_j]^T$, $y=[y_1, y_2, \dots, y_j]^T$, $p=1, 2, \dots, n$; I denotes the identity matrix, and the kernel function can be described as $K(x_j, x_p) = \varphi(x_j) \times \varphi(x_p)$, and it is a symmetric function conforming to Mercer's condition.

The regression coefficients d_i and g are obtained using the least-squares method. The LSSVM regression model [37] is defined as

$$f(x) = \sum_{i=1}^l d_i \times K(x_j, x_p) + g \quad (13)$$

Because the RBF is a more widely applicable kernel function, it does not require a priori knowledge of the dataset. Therefore, this study uses the RBF as the kernel function of the LSSVM algorithm. The RBF [37] is defined as

$$K(x_j, x_p) = \exp\left(-\frac{\|x_j - x_p\|^2}{2 \times \beta^2}\right) \quad (14)$$

where β is the kernel parameter. If β is large, it is easy to classify all sample points into the same class, otherwise, it will cause an overfitting problem.

The performance of the LSSVM model largely depends on the input variables and the parameters. The regularisation parameter c and kernel parameter β are usually determined based on experience, which readily decreases the accuracy of LSSVM. Therefore, PSO is used to find the best-fit parameters of the LSSVM.

2.3.5. PSO algorithm

The elementary theory of the PSO algorithm was generated by studying the social life of fish and birds that live in groups [39]. Instead of assigning functional operations to individuals, each individual is treated as a particle (no volume) in the search space (N -dimensional), flying at a certain speed (the speed is controlled by its own experience

and social experience) in the search field. The position of the q th particle is denoted as s_q , and s_q is substituted into the fitness function $F(s_q)$ to obtain the fitness value. The best place that each individual has experienced is denoted as $pbest_q$. The best place that all particles in the group have experienced is denoted by $gbest_q$. The velocity of particle i is regarded as V_i . In general, the range of the positional variation in the nn th ($1 \leq nn \leq N$) dimension is limited to $[S_{min,nn}, S_{max,nn}]$, and the range of the speed variation is limited to $[-V_{max,nn}, V_{max,nn}]$. For each generation, its velocity and positional variation of the nn th dimension ($1 \leq nn \leq N$) is updated using [38]

$$v_{in}^k = \omega \times v_{q,nn}^{k-1} + c_1 \times rand() \times (pbest_{q,nn} - s_{q,nn}^{k-1}) + c_2 \times rand() \times (gbest_{nn} - s_{q,nn}^{k-1}) \quad (15)$$

$$s_{nn}^k = s_{q,nn}^{k-1} + v_{q,nn}^{k-1} \quad (16)$$

where s_{nn}^{kk} is the nn th segment of the position vector of particle q at the kk th iteration, v_{in}^{kk} is the nn th component of the velocity of the granule q at the kk th iteration, and c_1 , c_2 are personal learning elements and social learning elements, respectively; $rand()$ is a random function.

2.4. Production process optimisation

Frequent unscheduled downtime can affect production scheduling in process industries. Moreover, these industries do not focus on electricity consumption in peak and off-peak periods, most of which account for nearly 1:1:1. The electricity price in the peak periods is more than twice that in the off-peak periods. Thus, industrial enterprises can minimize electricity and production costs by reducing unscheduled downtime and shifting the peak load.

The production process optimisation layer forecasts faults and shifts the electricity load from peak to off-peak periods for process industries. This layer includes three sub-modules: display module for forecasting results, fault diagnosis sub-module, and peak shifting sub-module. The display sub-module shows the forecasting results and the actual results in the form of a line chart, and it sets the warning line based on the size

of the electricity load. When the electricity load is below the lower limit of the warning line, the fault diagnosis sub-module will transmit a fault warning to remind the operators to examine and repair the corresponding electricity device. When the electricity load exceeds the upper limit of the warning line, and the time is in the peak periods and day periods, and a reminder to reduce the refining and pulping time while ensuring the production process is transmitted. The upper limit of the warning line is set as μ , and the lower limit of the warning line is set as $\mu - 2 \times \sigma$ in this study. Here, μ and σ are the average value and the standard deviation of the total electricity consumption in one year under normal production, respectively.

3. Results and discussion

Many evaluation indexes are applied to assess the performance of the forecasting model. The study uses the mean absolute percent error (MAPE) and root mean square error (RMSE) to assess the performance of the developed STELF model.

3.1. Case study

Electricity load data are derived from an actual papermaking enterprise in Guangdong, China, and reserved for 60 days. The acquisition frequency was 30 min. The acquired data were processed according to the approaches explained in Section 2.2. The processed total electricity load data is illustrated in Fig. 5.

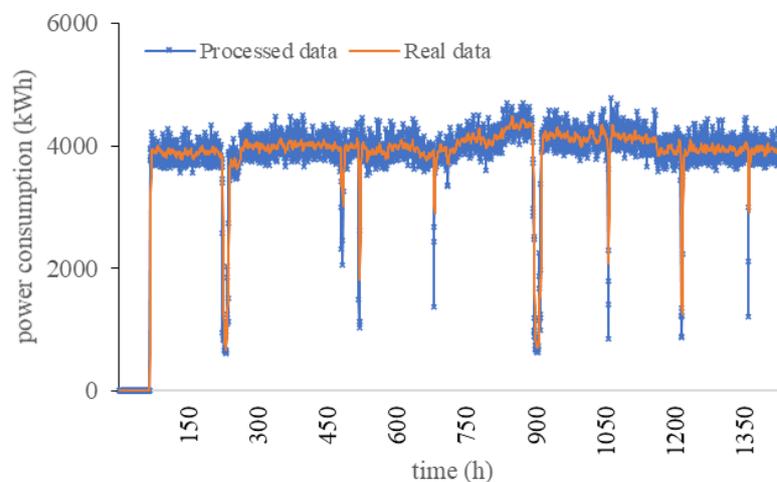
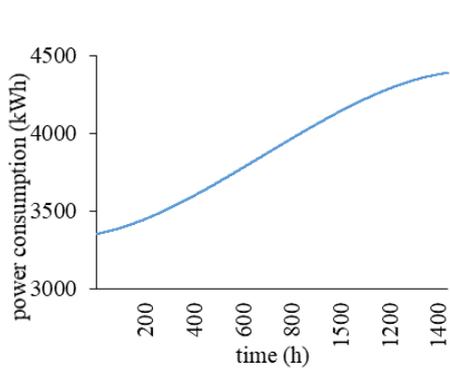
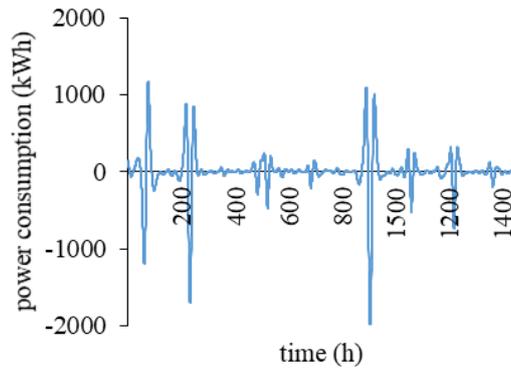


Fig. 5. Trend graph of Preprocessed total electricity load data

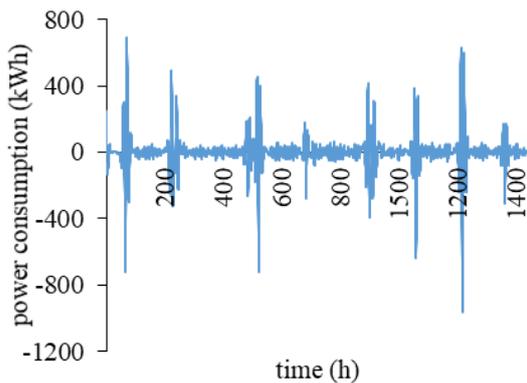
The preprocessed electricity load signal is decomposed into eight components using the EMD-VMD-based decomposition model. The decomposed components are shown in Fig. 6, where Fig. 6(a)–Fig. 6(h) show the decomposed components obtained by the EMD-VMD-based model, and Fig. 6(i) shows the residual between the preprocessed electricity load signal and the sum of the decomposed components. Fig. 6(a)- Fig. 6(f) show that the characteristic of decomposed components is regular and stable. Fig. 6(g) shows that the VMD algorithm decomposes the overall trend of the electricity load. To ensure not losing information, this study preserves the signal filtered out by the proposed EMD-VMD model. The results are shown in Fig. 6(h).



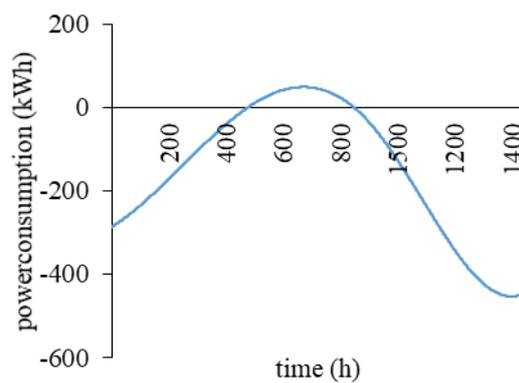
(a) First decomposed component



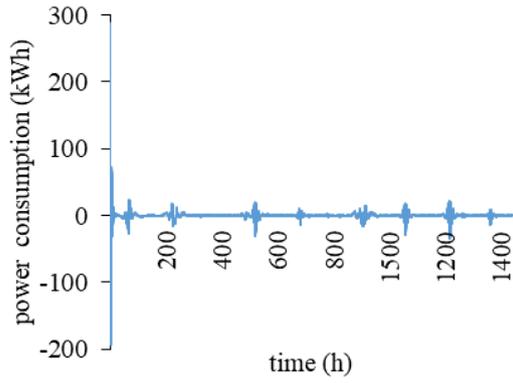
(b) Second decomposed component



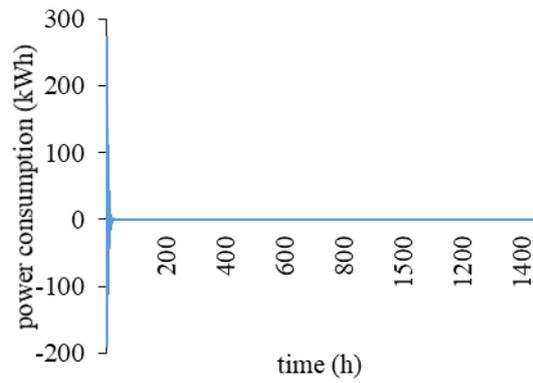
(c) Third decomposed component



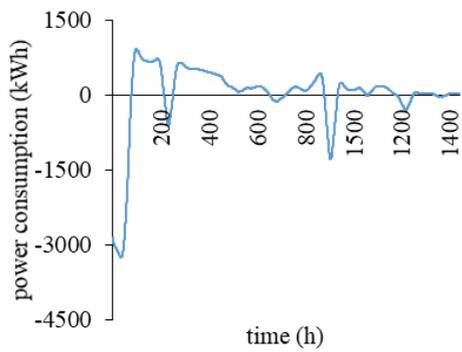
(d) Forth decomposed component



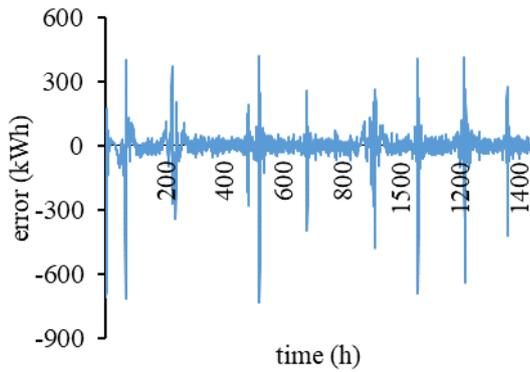
(e) Fifth decomposed component



(f) Sixth decomposed component



(g) Seventh decomposed component



(h) Eighth decomposed component

Fig. 6. Trend graph of the decomposition results

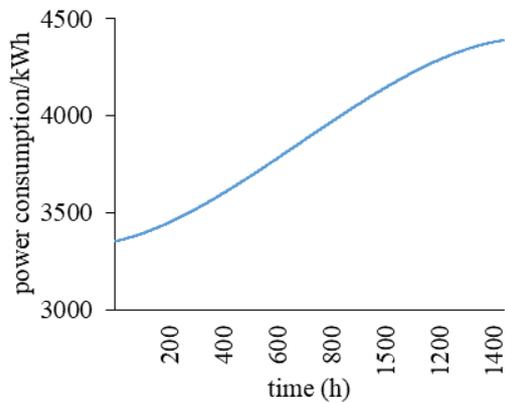
The ApEn method is used to reconstruct similar decomposed components. The ApEn values of the eight decomposed components are summarized in Table 1. Table 1 shows that the smaller the approximate entropy value is, the lower the complexity of the signal is.

Table 1. Approximate entropy values of the eight decomposed components

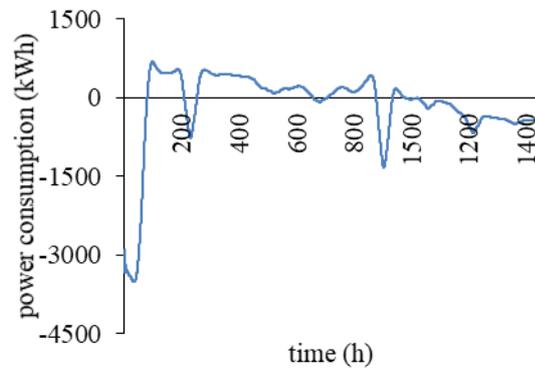
| Entropy | Order |
|----------|-------|
| 2.19E-04 | 1 |
| 1.03E-03 | 6 |
| 4.24E-03 | 4 |
| 2.52E-02 | 7 |
| 1.30E-01 | 2 |
| 2.00E-01 | 5 |

| | |
|----------|---|
| 3.18E-01 | 3 |
| 8.53E-01 | 8 |

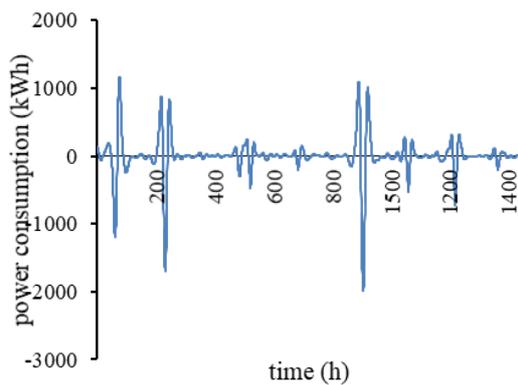
This study sets six approximate entropy intervals as follows: $[0, 0.001)$, $[0.0001, 0.05)$, $[0.05, 0.15)$, $[0.15, 0.45)$, $[0.45, 0.75)$, and $[0.75, 2]$. Similar decomposed components were added according to the six approximate entropy intervals and five reconstruction sequences were obtained. The reconstructed components are shown in Fig. 7. Fig. 7 shows that all the reconstructed components are still stable. Based on the ApEn method, this study decomposes a single trend curve, a trend curve of total electricity load without disturbance, and three more regular signals.



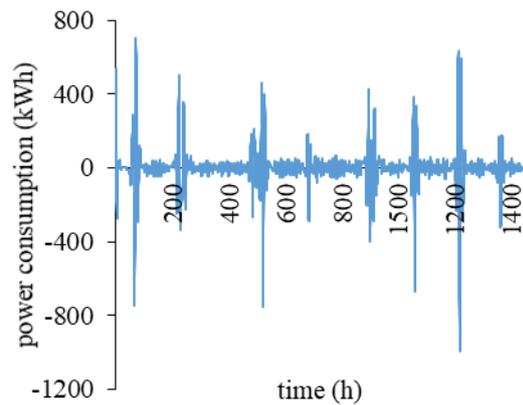
(a) First reconstructed component



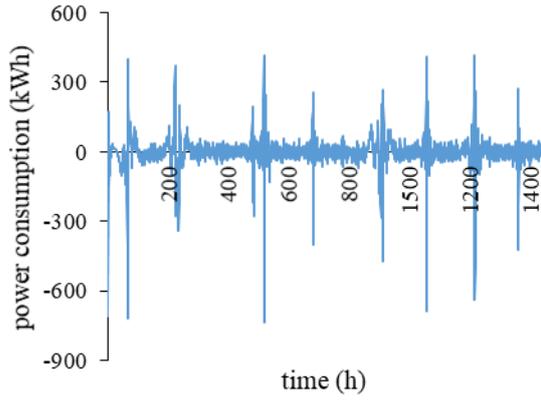
(b) Second reconstructed component



(c) Third reconstructed component



(d) Forth reconstructed component



(e) Fifth reconstructed component

Fig. 7. Trend chart of the five reconstructed components

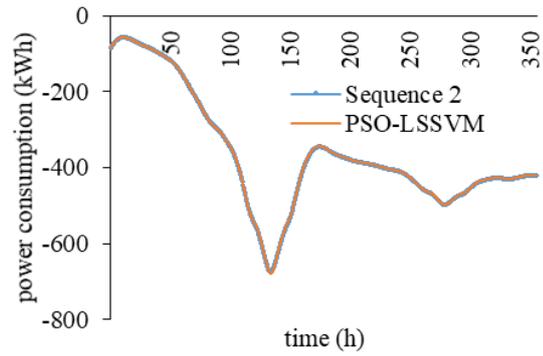
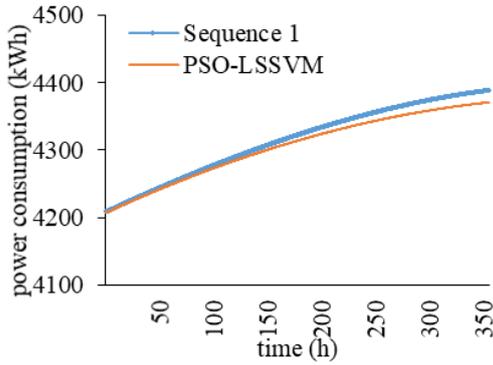
The lag autocorrelation method is adopted to choose the input variables of the ELF model for each reconstructed component. The input variables of the five reconstructed components are listed in Table 2. Table 2 shows that the lower the complexity of the component is, the greater the influence of historical information on the present status is.

Table 2. Input variables of the five reconstructed components

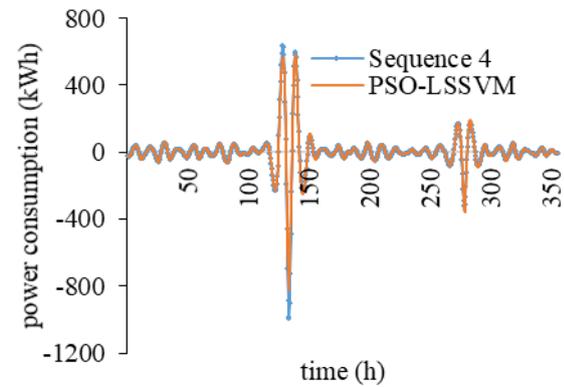
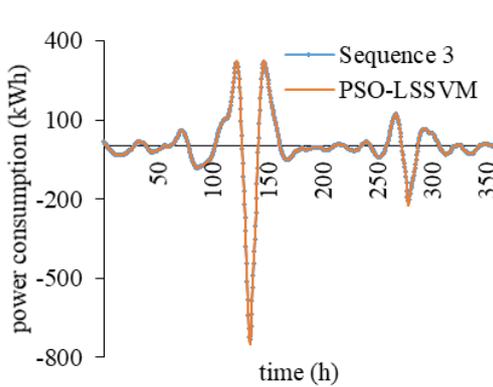
| Order | Input variables | Unit |
|---------------------------|--|------|
| Reconstructed Component 1 | $I_{t-1,1}, I_{t-2,1}, I_{t-3,1}, I_{t-4,1}, I_{t-5,1}, I_{t-6,1},$ $I_{t-7,1}, I_{t-8,1}, I_{t-9,1}, I_{t-10,1}$ | kWh |
| Reconstructed Component 2 | $I_{t-1,2}, I_{t-2,2}, I_{t-3,2}, I_{t-4,2}, I_{t-5,2}, I_{t-6,2},$ $I_{t-7,2}, I_{t-8,2}, I_{t-9,2}, I_{t-10,2}$ | kWh |
| Reconstructed Component 3 | $I_{t-1,3}, I_{t-2,3}, I_{t-3,3}, I_{t-4,3}, I_{t-5,3},$ | kWh |
| Reconstructed Component 4 | $I_{t-1,4}, I_{t-2,4}, I_{t-3,4}, I_{t-4,4}, I_{t-5,4}$ | kWh |
| Reconstructed Component 5 | $I_{t-1,5}, I_{t-2,5}, I_{t-3,5}, I_{t-4,5}, I_{t-5,5}$ | kWh |

The study divides the data set into a training set and a testing set. The ratio of the training set to the testing set is 3:1. The forecasting results of each reconstructed component are shown in Fig. 8. Fig. 8 shows that the more regular the signal is, the more accurate the forecasting results will be. For example, the error at the 149th time

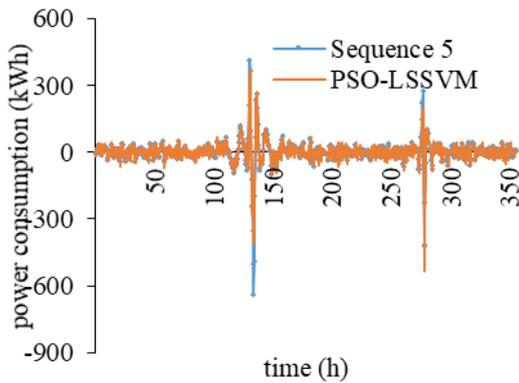
point in Fig. 8(d) is 11 times greater than that in Fig. 8(c).



(a) First reconstructed component results (b) Second reconstructed component results



(c) Third reconstructed component results (d) Forth reconstructed component results

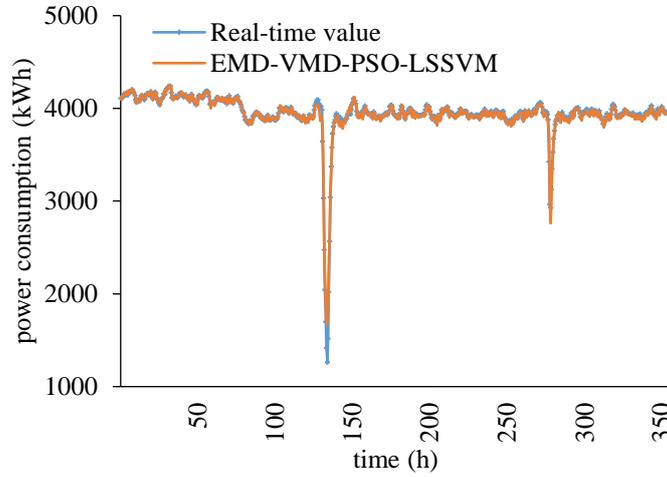


(e) Fifth reconstructed component results

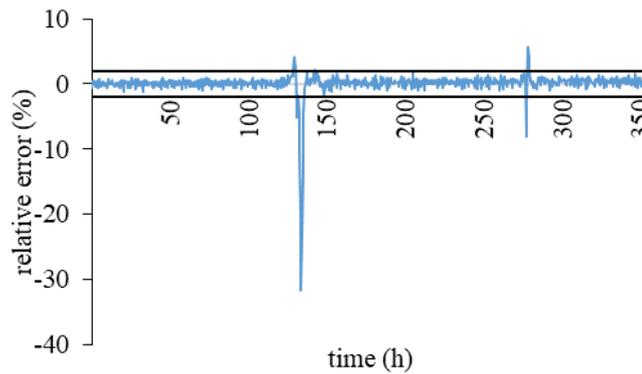
Fig. 8. Forecasting results of the reconstructed components

The parameters of the PSO algorithm are set as follows: learning coefficients (c_1 , c_2), 2; maximum particle velocity (V_{max}), 0.5; population scale (S), 30; and the maximum number of iterations, 100. The forecasting results are obtained by summing the forecasting results of the five reconstructed components. The final forecasting

results are shown in Fig. 9. Fig. 9(a) shows the trend chart of the forecasting results, and Fig. 9(b) shows the relative error. The evaluation indicators of the proposed model are listed in Table 3.



(a) Trend graph of the forecasting results



(b) relative error

Fig. 9. Final forecasting result

Table 3. Evaluation indicators of the proposed model

| Name | MAPE (%) | RMSE (kWh) |
|------------|----------|------------|
| Case study | 0.66 | 39.14 |

Fig. 9 shows that 97.7% of the forecasting results of the proposed model satisfy the errors of the industrial requirements between -2% and 2% . The maximum relative error of the proposed model is 400 kWh, and the forecasting results do not have a time

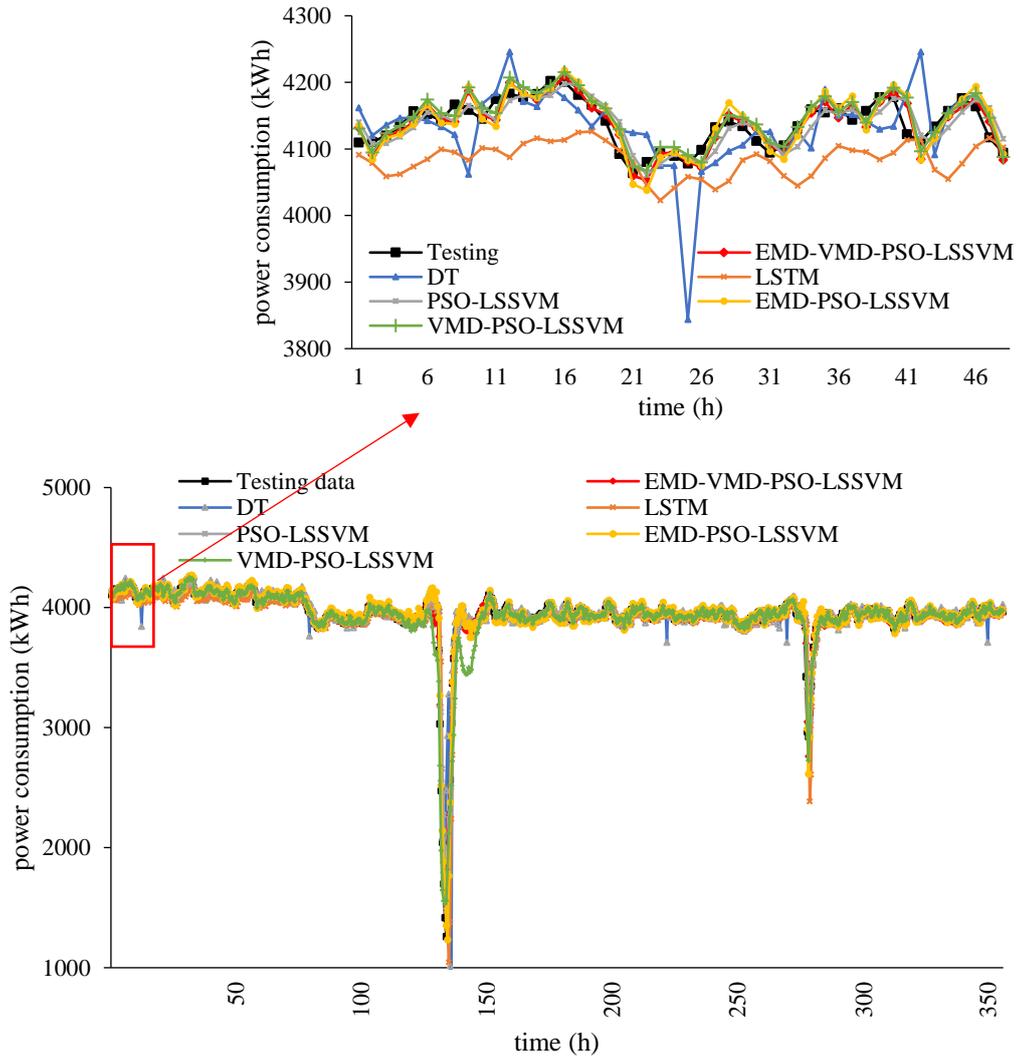
delay when forecasting the electricity load with enormous fluctuations. Thus, the test results imply that the developed EMD-VMD-PSO-LSSVM based model is suitable for forecasting industrial electricity load with enormous fluctuation and aperiodicity.

However, in Fig. 9(b), comparing the forecast results of the electricity load with the two largest fluctuations, the results show that the accuracy of the electricity load with the largest fluctuations is 10times lower than that of the smaller ones. The errors are mainly derived from Fig. 8(d) and Fig. 8(e). the reason might be that the data set has not been updated in time. As time increases, the data characteristics of the model cannot fully describe the characteristics of future electricity consumption. Thus, a dynamic ELF model has been proposed. The detailed information is shown in Section 3.3.

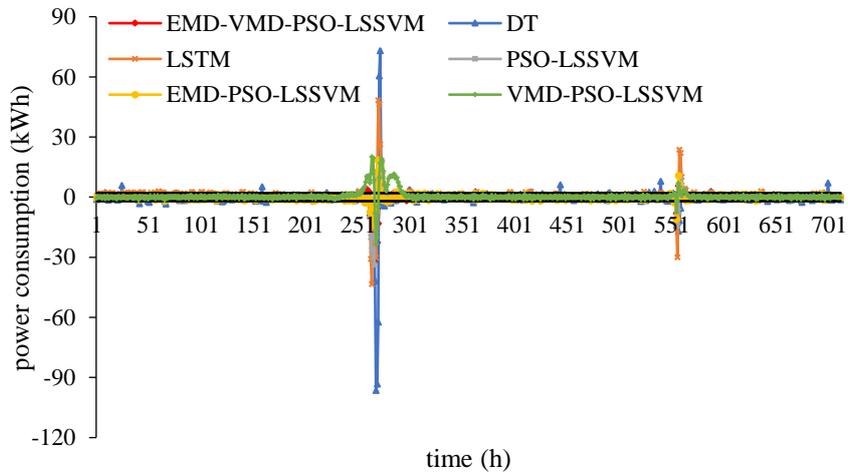
3.2. Comparative analysis

A comparative analysis of the forecasting performance of EMD-VMD-PSO-LSSVM, VMD-PSO-LSSVM, EMD-PSO-LSSVM, PSO-LSSVM, DT, and LSTM is presented in this section. The forecasting results are illustrated in Fig. 10. Fig. 10(a) illustrates the forecasting results of the four forecasting models. To illustrate the undecomposed and decomposed forecasting results in detail, Fig. 10(a) illustrates the enlarged forecasting results on the first day. Fig. 10(b) shows the relative error. A benchmark of $[-2\%, 2\%]$ of the error range is set in Fig. 10(b) to show the accuracy of the three forecasting models more intuitively.

To show the consistency of the different forecasting models, the initial parameters of VMD-PSO-LSSVM, EMD-PSO-LSSVM, and PSO-LSSVM are set to be the same as those of EMD-VMD-PSO-LSSVM.



(a) the trend graph of the forecasting results



(b) Relative error

Fig. 10. Forecasting result contrast of the forecasting model

Table 4. Forecasting performance analysis

| Case 1 | Index | | |
|-------------------|-------------------------------|-------------------------------|----------|
| | Maximum forecasting error (%) | Minimum forecasting error (%) | MAPE (%) |
| EMD-VMD-PSO-LSSVM | 23.71 | 0.00 | 0.66 |
| DT | -96.74 | 0.00 | 1.56 |
| LSTM | 48.27 | 0.00 | 1.51 |
| PSO-LSSVM | 33.88 | 0.00 | 0.85 |
| EMD-PSO-LSSVM | -18.62 | 0.00 | 0.76 |
| VMD-PSO-LSSVM | -23.49 | 0.00 | 0.98 |

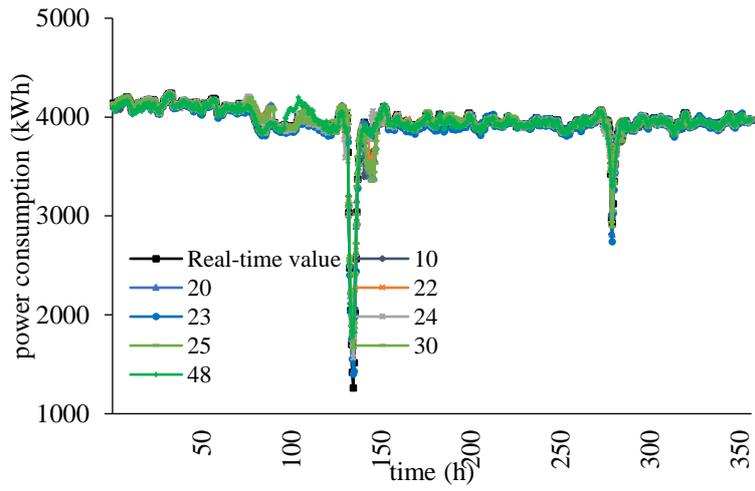
Fig. 10 displays that the proposed EMD-VMD-PSO-LSSVM based forecasting model has the most time points within the prediction error $[-2\%, 2\%]$ (97.5% time points from the entire testing time points) when compared with the other three forecasting models. The forecasting models based on VMD-PSO-LSSVM, EMD-PSO-LSSVM, and PSO-LSSVM, DT, LSTM have 96.8%, 86%, 92.7%, 91.4%, and 89.6% times points from the total test times points respectively, which lie within the prediction error $[-2\%, 2\%]$. The results indicate that the EMD-VMD-PSO-LSSVM based forecasting model has the best constant capability among all adopted models.

Table 4 lists the evaluation results of the four models. The maximum error of the EMD-VMD-PSO-LSSVM based forecasting model is greater than that of comparative decomposition algorithm-based forecasting models, but lower than that of other comparative non-decomposition algorithm-based forecasting models. Moreover, the ratio of the relative errors of the EMD-VMD-PSO-LSSVM based forecasting model without the relative error $[-2\%, 2\%]$ is 20% less than that of the VMD-PSO-LSSVM based ELF model, 82.1% less than that of the EMD-PSO-LSSVM based ELF model, 65.7% less than that of the PSO-LSSVM based ELF model, 70.9% less than that of DT based ELF model, and 75.9% less than that of LSTM based ELF model. The MAPE of

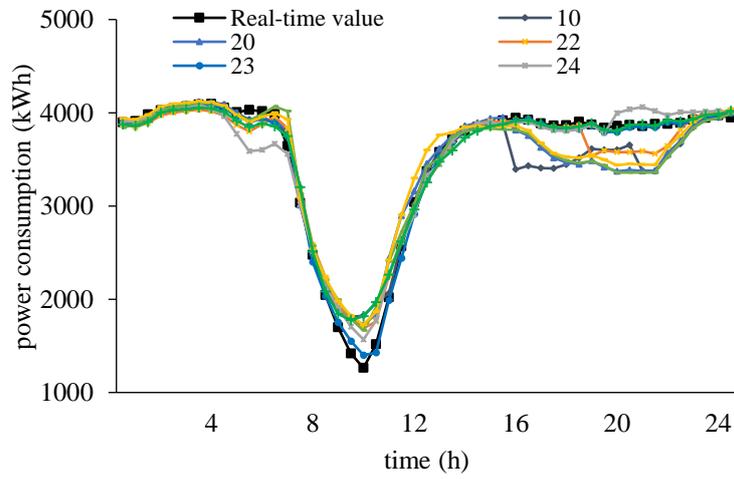
the EMD-VMD-PSO-LSSVM based ELF model is 32.6%, 13.2%, 22.3%, 57.7%, and 56.3% lower than that of the VMD-PSO-LSSVM, EMD-PSO-LSSVM, PSO-LSSVM, DT, and LSTM based ELF models, respectively. The forecasting results show that the ELF models based on decomposed algorithms do not have a time delay. However, the ELF models based on undecomposed algorithms have a time delay. Thus, the developed EMD-VMD-PSO-LSSVM based model has higher precision than the contrast models and no time delay, which is more suitable for industrial enterprises to forecast electricity load compared to the contrast models.

3.3. Dynamic analysis

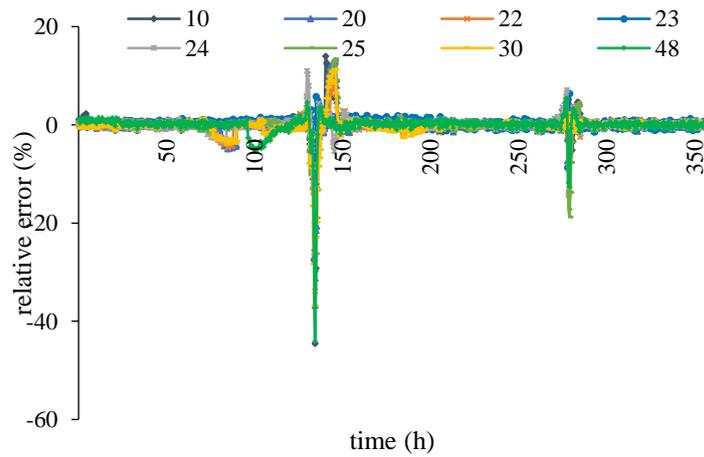
The error of the forecasting model can increase when the training dataset of the forecasting model is not updated periodically. Thus, this study selects an appropriate period to update the training dataset to ensure the stability of the long-term operation of the forecasting model. This study tests the dynamic performance of the EMD-VMD-PSO-LSSVM based forecasting model when the update periods are set as 10, 20, 22, 23, 24, 25, 30, and 48 respectively. The forecasting results are shown in Fig. 11. Fig. 11(a) is the forecasting results, Fig. 11(b) is the enlarged graph of the forecasting results with enormous fluctuation, and Fig. 11(c) is the error. Table 5 is the evaluation indicators of the proposed model with different update periods.



(a) the forecasting results



(b) the enlarged graph of the forecasting results with enormous fluctuation



(c) the error

Fig. 11. the comparative graph of the forecasting results with different periods

Table 5. the evaluation indicators of forecasting results with different periods

| Update period (30min) | MAPE (%) | Maximum forecasting error (%) | RMSE (kWh) |
|--------------------------|-------------|----------------------------------|---------------|
| 10 | 0.83 | -44.5 | 70.29 |
| 20 | 0.88 | -36.8 | 72.18 |
| 22 | 0.84 | -33.8 | 59.77 |
| 23 | 0.63 | -10.9 | 32.77 |
| 24 | 0.79 | -24.2 | 55.93 |
| 25 | 0.93 | -31.7 | 78.15 |
| 30 | 0.94 | -36.8 | 69.3 |
| 48 | 0.87 | -44.1 | 58.18 |

Fig. 11 shows that the forecasting results of all update periods are consistent with the original electricity load when the production process is stable. Fig. 11(b) shows that the results of the forecasting model with 23 update periods are closer to the original electricity load than those with other update periods. Moreover, Table 5 indicates that the accuracy of the results of the forecasting model with 23 update periods is higher than that with different update periods. The MAPE of the proposed model was 0.63% when the update period was 23, which was 31.7%, 39.7%, 33.3%, 25.4%, 47.6%, 49.2%, and 38.1% lower than the MAPE of the proposed model when the update period is 10, 20, 22, 24, 25, 30, and 48, respectively. The RMSE of the forecasting model when the update period is 10 is 114.5%, 120.3%, 82.4%, 70.7%, 138.5%, 111.5%, and 77.5% lower than the RMSE of the proposed model when the update periods are 10, 20, 22, 24, 25, 30, and 48, respectively. Thus, this study selects 23 as the update period to propose a dynamic electricity load forecasting model.

3.4. Production process optimisation analysis

This study uses data from a practical case to verify the effectiveness of the production process optimisation layer of the proposed integrated framework. Further,

this study uses the proposed dynamic model to optimise the production process optimisation layer. The electricity consumption will decrease rapidly when there is an unscheduled downtime in the production process. Thus, this study assumed that electricity device failure occurs in the production process when the difference between the electricity load of forecasting results and the current time is less than $\mu - 2 \times \sigma$. Fig. 12 shows the faults forecasting results. The forecasting results show that the fault could be forecast 30 min to 1 h in advance. According to the fault forecasting results, the process industrial enterprises could have the full time to adjust the production line, reduce the time of unscheduled downtimes, and ensure stability to minimize the production loss.

To verify the peak electricity load shaving sub-module, Guangdong Province was used as an example. The capacity of the pulping tank in this paper mill is 100 m³. The slurry consumption rate is 60 m³ per hour, and refiners and pulpers can produce the paper pulp by 70 m³ per hour. When the electricity load exceeds the upper limit of the warning line, the production process optimisation layer can stop the operation of the refiners and pulpers according to the current operation of those devices to reduce the electricity cost. The electricity cost of the optimised production schedule based on the proposed integrated framework can save 250,000 RMB per year.

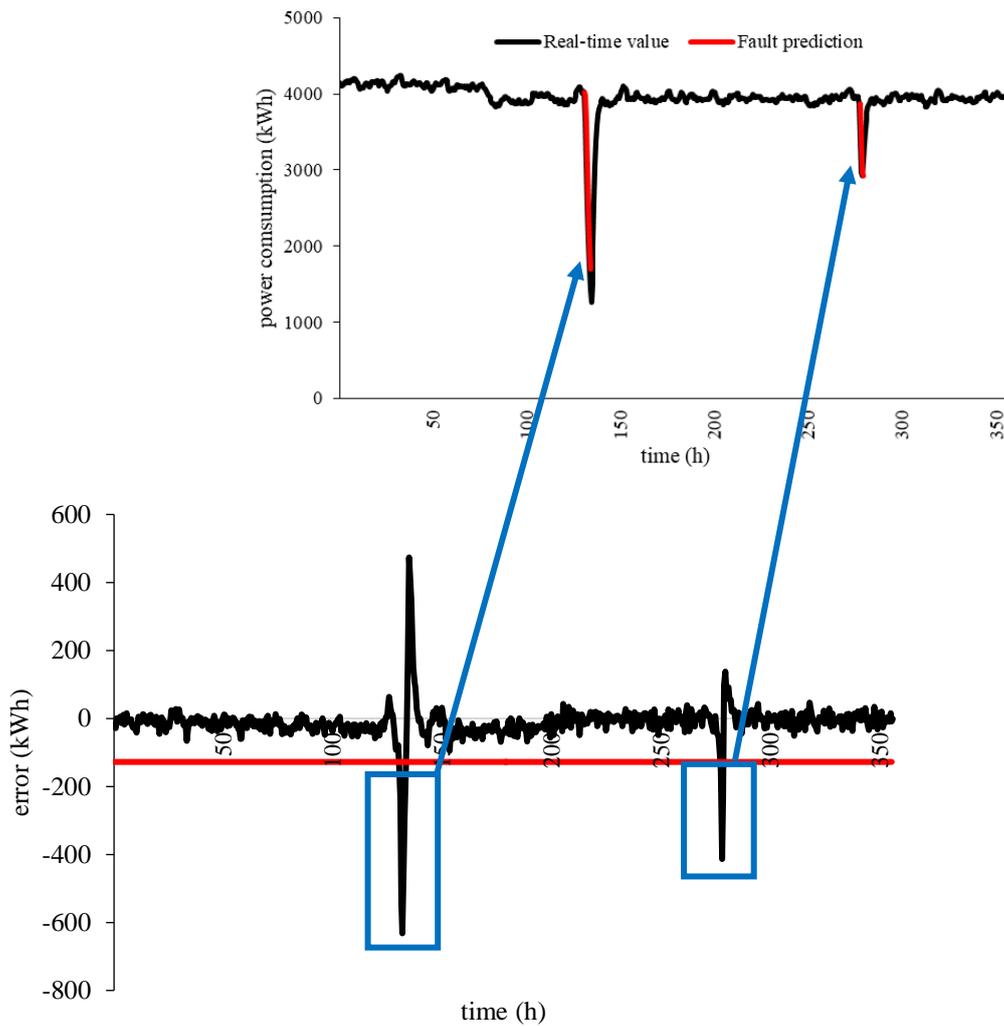


Fig. 12. Fault prediction results

4. Conclusions

Given the serious concern related to energy shortage and environmental deterioration, the establishment of electricity load prediction and production process optimisation in the electricity management system for industrial enterprises can help improve electricity consumption efficiency and achieve the goal of a cleaner and sustainable production.

In this study, an integrated framework that can be directly applied in industrial enterprises to forecast short-term electricity load and optimise the production process was proposed. This study proposed a signal decomposition model based on the EMD-

VMD algorithm that can decompose the electricity signal with aperiodicity into multiple components with stability, single trend, or periodicity. The model used an ApEn method to add components with the same trend to avoid dimension disasters and reduce the computational complexity of the forecasting model. The forecasting results showed that the dynamic STELF model based on EMD-VMD-PSO-LSSVM was superior to the ELF model based on undecomposed algorithms and the ELF model based on a single decomposition algorithm. The MAPE of the proposed dynamic STELF model was at least 4.5% and 17% lower than that of the ELF model based on undecomposed algorithms and the ELF model based on a single decomposition algorithm. The proposed ELF model has the advantages of no time delay, high accuracy, and high stability. The optimisation results showed that the times of unscheduled downtimes would be reduced based on the fault forecasting results, and the electricity cost could be minimised based on the peak shifting sub-module. As a case study, a medium-scale tissue paper enterprise could save 250,000 RMB per year by using the optimised production schedule based on the proposed integrated framework. Thus, the proposed model can be used to stabilise the operating process of industrial enterprises, improve production efficiency, and reduce electricity consumption and production costs.

The proposed integrated framework in this study opens a new direction for the study of intelligent industrial power management systems. Moreover, the proposed integrated framework will give industrial enterprises an edge in the competition of the centralised declaration of electricity consumption when the future new electricity market is formed. The proposed integrated framework can provide reliable data for the intelligent electricity consumption management of the power supply side.

Acknowledgments

The work described in this paper was supported by the grant “Postdoctoral Fellowships Scheme” from the Research Committee of The Hong Kong Polytechnic University under project ID-P0031541 (account- G-YW4Y) and this work was also

supported by Joint Supervision Scheme with the Chinese Mainland, Taiwan and Macao Universities - Other Chinese Mainland, Taiwan and Macao Universities under project ID- P0035156 (account- G-SB3R).

References

- [1] United Nations Framework Convention on Climate Change. The Paris Agreement. New York; 2018. <https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement/>.
- [2] Ghasemi-Mobtaker H, Mostashari-Rad F, Saber Z, Chau K, Nabavi-Pelesaraei A. Application of photovoltaic system to modify energy use, environmental damages and cumulative exergy demand of two irrigation systems-A case study: Barley production of Iran. *Renew Energy* 2020;160:1316–34. <https://doi.org/10.1016/j.renene.2020.07.047>.
- [3] Nabavi-Pelesaraei A, Azadi H, Van Passel S, Saber Z, Hosseini-Fashami F, Mostashari-Rad F, et al. Prospects of solar systems in production chain of sunflower oil using cold press method with concentrating energy and life cycle assessment. *Energy* 2021;223:120117. <https://doi.org/10.1016/J.ENERGY.2021.120117>.
- [4] Waheed R, Sarwar S, Wei C. The survey of economic growth, energy consumption and carbon emission. *Energy Reports* 2019;5:1103–15. <https://doi.org/10.1016/j.egy.2019.07.006>.
- [5] National Bureau of Statistics of the People's Republic of China. China Energy Statistical Yearbook. Beijing: China Statistical Publishing House;2020.
- [6] Wang Y, Sun S, Chen X, Zeng X, Kong Y, Chen J, et al. Short-term load forecasting of industrial customers based on SVM and XGBoost. *Int J Electr Power Energy Syst* 2021;129:106830.

- <https://doi.org/10.1016/J.IJEPES.2021.106830>.
- [7] Tavassoli-Hojati Z, Ghaderi SF, Iranmanesh H, Hilber P, Shayesteh E. A self-partitioning local neuro fuzzy model for short-term load forecasting in smart grids. *Energy* 2020;199:117514. <https://doi.org/10.1016/j.energy.2020.117514>.
- [8] Ogliari E, Guilizzoni M, Giglio A, Pretto S. Wind power 24-h ahead forecast by an artificial neural network and an hybrid model: Comparison of the predictive performance. *Renew Energy* 2021;178:1466–74. <https://doi.org/10.1016/J.RENENE.2021.06.108>.
- [9] Pi Z, Li X, Ding Y, Zhao M, Liu Z. Demand response scheduling algorithm of the economic energy consumption in buildings for considering comfortable working time and user target price. *Energy Build* 2021;250:111252. <https://doi.org/10.1016/j.enbuild.2021.111252>.
- [10] Kychkin AV., Chasparis GC. Feature and model selection for day-ahead electricity-load forecasting in residential buildings. *Energy Build* 2021;249:111200. <https://doi.org/10.1016/J.ENBUILD.2021.111200>.
- [11] Talaat M, Farahat MA, Mansour N, Hatata AY. Load forecasting based on grasshopper optimization and a multilayer feed-forward neural network using regressive approach. *Energy* 2020;196:117087. <https://doi.org/10.1016/j.energy.2020.117087>.
- [12] Qian Z, Pei Y, Zareipour H, Chen N. A review and discussion of decomposition-based hybrid models for wind energy forecasting applications. *Appl Energy* 2019;235:939–53. <https://doi.org/10.1016/j.apenergy.2018.10.080>.

- [13] Jaseena KU, Kovoor BC. Decomposition-based hybrid wind speed forecasting model using deep bidirectional LSTM networks. *Energy Convers Manag* 2021;234:113944. <https://doi.org/10.1016/j.enconman.2021.113944>.
- [14] Zhang Z, Hong W. Application of variational mode decomposition and chaotic grey wolf optimizer with support vector regression for forecasting electric loads. *Knowledge-Based Syst* 2021;228:107297. <https://doi.org/10.1016/j.knosys.2021.107297>.
- [15] Shao Z, Zheng Q, Liu C, Gao S, Wang G, Chu Y. A feature extraction- and ranking-based framework for electricity spot price forecasting using a hybrid deep neural network. *Electr Power Syst Res* 2021;200:107453. <https://doi.org/10.1016/j.epsr.2021.107453>.
- [16] Zhang X, Wang J, Gao Y. A hybrid short-term electricity price forecasting framework: Cuckoo search-based feature selection with singular spectrum analysis and SVM. *Energy Econ* 2019;81:899–913. <https://doi.org/10.1016/J.ENECO.2019.05.026>.
- [17] Ofori-Ntow Jnr E, Ziggah YY, Relvas S. Hybrid ensemble intelligent model based on wavelet transform, swarm intelligence and artificial neural network for electricity demand forecasting. *Sustain Cities Soc* 2021;66:102679. <https://doi.org/10.1016/J.SCS.2020.102679>.
- [18] Zhang S, Chen Y, Xiao J, Zhang W, Feng R. Hybrid wind speed forecasting model based on multivariate data secondary decomposition approach and deep learning algorithm with attention mechanism. *Renew Energy* 2021;174:688–704.

- <https://doi.org/10.1016/J.RENENE.2021.04.091>.
- [19] Shao Z, Fu C, Yang S, Zhou K. A review of the decomposition methodology for extracting and identifying the fluctuation characteristics in electricity demand forecasting. *Renew Sustain Energy Rev* 2017;75:123–36. <https://doi.org/10.1016/J.RSER.2016.10.056>.
- [20] Moreno SR, Mariani VC, dos SantosCoelho L. Hybrid multi-stage decomposition with parametric model applied to wind speed forecasting in Brazilian Northeast. *Renew Energy* 2021;164:1508–26. <https://doi.org/10.1016/J.RENENE.2020.10.126>.
- [21] Niu X, Wang J. A combined model based on data preprocessing strategy and multi-objective optimization algorithm for short-term wind speed forecasting. *Appl Energy* 2019;241:519–39. <https://doi.org/10.1016/j.apenergy.2019.03.097>.
- [22] Tawn R, Browell J, Dinwoodie I. Missing data in wind farm time series: Properties and effect on forecasts. *Electr Power Syst Res* 2020;189:106640. <https://doi.org/10.1016/j.epsr.2020.106640>.
- [23] Smiti A. A critical overview of outlier detection methods. *Comput Sci Rev* 2020;38:100306. <https://doi.org/10.1016/j.cosrev.2020.100306>.
- [24] Fong JT, Filliben JJ, Heckert NA, Marcal P V., Rainsberger R, Ma L. Uncertainty Quantification of Stresses in a Cracked Pipe Elbow Weldment Using a Logistic Function Fit, a Nonlinear Least Square Algorithm, and a Super-parametric Method. *Procedia Eng* 2015;130:135–49. <https://doi.org/10.1016/J.PROENG.2015.12.183>.

- [25] Tang X, Liu K, Lu J, Liu B, Wang X, Gao F. Battery incremental capacity curve extraction by a two-dimensional Luenberger–Gaussian-moving-average filter. *Appl Energy* 2020;280:115895. <https://doi.org/10.1016/J.APENERGY.2020.115895>.
- [26] Zhang L, Su H, Zio E, Zhang Z, Chi L, Fan L, et al. A data-driven approach to anomaly detection and vulnerability dynamic analysis for large-scale integrated energy systems. *Energy Convers Manag* 2021;234:113926. <https://doi.org/10.1016/j.enconman.2021.113926>.
- [27] Prasad SMM, Nguyen-Huy T, Deo R. Support vector machine model for multistep wind speed forecasting. In: Deo R, Samui P, Roy SS, editors. *Predictive Modelling for Energy Management and Power Systems Engineering*, Amsterdam: Elsevier; 2020, p.335–89. <https://doi.org/10.1016/B978-0-12-817772-3.00012-4>.
- [28] Niu D, Ji Z, Li W, Xu X, Liu D. Research and application of a hybrid model for mid-term power demand forecasting based on secondary decomposition and interval optimization. *Energy* 2021;234:121145. <https://doi.org/10.1016/j.energy.2021.121145>.
- [29] Bari P, Karande P. Application of PROMETHEE-GAIA method to priority sequencing rules in a dynamic job shop for single machine. *Mater Today Proc* 2021;46: 7258-64. <https://doi.org/10.1016/j.matpr.2020.12.854>.
- [30] Cheng G, Wang X, He Y. Remaining useful life and state of health prediction for lithium batteries based on empirical mode decomposition and a long and

- short memory neural network. *Energy* 2021;232:121022.
<https://doi.org/10.1016/j.energy.2021.121022>.
- [31] Zhang T, Tang Z, Wu J, Du X, Chen K. Multi-step-ahead crude oil price forecasting based on two-layer decomposition technique and extreme learning machine optimized by the particle swarm optimization algorithm. *Energy* 2021;229:120797. <https://doi.org/10.1016/j.energy.2021.120797>.
- [32] Li J, Wu Q, Tian Y, Fan L. Monthly Henry Hub natural gas spot prices forecasting using variational mode decomposition and deep belief network. *Energy* 2021;227:120478. <https://doi.org/10.1016/j.energy.2021.120478>.
- [33] Liang T, Zhao Q, Lv Q, Sun H. A novel wind speed prediction strategy based on Bi-LSTM, MOOFADA and transfer learning for centralized control centers. *Energy* 2021;230:120904. <https://doi.org/10.1016/j.energy.2021.120904>.
- [34] Lahmiri S, Tadj C, Gargour C, Bekiros S. Characterization of infant healthy and pathological cry signals in cepstrum domain based on approximate entropy and correlation dimension. *Chaos, Solitons and Fractals* 2021;143:110639. <https://doi.org/10.1016/j.chaos.2020.110639>.
- [35] Gao X, Yan X, Gao P, Gao X, Zhang S. Automatic detection of epileptic seizure based on approximate entropy, recurrence quantification analysis and convolutional neural networks. *Artif Intell Med* 2020;102:101711. <https://doi.org/10.1016/j.artmed.2019.101711>.
- [36] Kazemzadeh MR, Amjadian A, Amraee T. A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting.

- Energy 2020;204:117948. <https://doi.org/10.1016/j.energy.2020.117948>.
- [37] Kaytez F. A hybrid approach based on autoregressive integrated moving average and least-square support vector machine for long-term forecasting of net electricity consumption. Energy 2020;197:117200. <https://doi.org/10.1016/j.energy.2020.117200>.
- [38] Hu Y, Li J, Hong M, Ren J, Lin R, Liu Y, et al. Short term electric load forecasting model and its verification for process industrial enterprises based on hybrid GA-PSO-BPNN algorithm—A case study of papermaking process. Energy 2019;170:1215–27. <https://doi.org/10.1016/j.energy.2018.12.208>.