

**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**

Yusha Hu¹, Jigeng Li¹, Mengna Hong¹, Jingzheng Ren², Ruoju Lin², Mengru Liu¹, Yi

Man^{1,2*}

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, China

* Corresponding author:

Yi Man, PhD, Email: manyi@scut.edu.cn, yi.man@polyu.edu.hk

Abstract

Process industry consumes tremendous amounts of electricity for production. Electric load forecasting could be conducive to managing the electricity consumption, determining the optimal production scheduling, and planning the maintenance schedule, which could improve the energy efficiency and reduce the production cost. This paper proposed a short term electric load forecasting model based on the hybrid GA-PSO-BPNN algorithm. The GA-PSO algorithm is used in a short-term electric load forecasting model to optimize the parameters of BPNN. The forecasting model avoids the shortcoming that the prediction result is easy to fall into local optimum. The papermaking process, as one of the most representative process industries, is selected as the study case. The real-time production data from two different papermaking enterprises is collected to verify the proposed model. Besides the proposed GA-PSO-BPNN model, the GA-BPNN and PSO-BPNN based electric load forecasting models are also studied as the contrasting cases. The verification results reveal that the GA-PSO-BPNN model is superior to the other two hybrid forecasting models for future application in the papermaking process since its MAPE is only 0.77%.

Keywords: electric load forecasting; modeling and simulation; papermaking process; energy saving; energy consumption

Nomenclature

α	Crossover variable
σ	The standard deviation
μ	Mean value
ω	The inertia weight
b_k	Bias value
c_1, c_2	Acceleration constants
d	Population size
f'	The partial derivative of f
f	The activation function of the hidden layer
$f(x_i)$	The fitness value of the particle
$gbest_n$	The global best position
i	The iteration number
k	The particle's index
L	The number of training samples size
l	The dimension of the output variables
n	The length of the sequence
o_{ik}	The input value of the hidden layer
p	The position of the particle in the search space
$pbest_{in}$	The local best position
P_i	the probability of selecting the particle
r	Correlation coefficient
$rand()$	The random variables in the range of [0, 1]
t	Time
v_{in}^k	The particle's velocity
w_{ij}	The connection weights of the input and hidden layers
w_{jk}	The connection weights of the hidden and output layers
w'_{kj}, b'_j	The updated connection weight and bias
\bar{X}, \bar{Y}	The average value

X_i, Y_i	Two different sequences
x	A data sequence
x_i	The input value of the input layer
x_{in}^k	The particle's position
x_{t+1}^i, x_{t+1}^j	New individuals
y_k	The output value of the output layer
y_{ij}	The forecasting value
\hat{y}_{ij}	The real value
BPNN	Back Propagation Neural Networks
GA	Genetic Algorithm
MAPE	Mean absolute percent error
MAFRE	Maximum forecasting relative error
MIFRE	Minimum forecasting relative error
PSO	Particle Swarm Optimization
RE	Relative error

1. Introduction

Process industry consumes massive amounts of electricity for production. In China, the electricity consumption of process industrial enterprises accounts for about 70% of the total electricity consumption of the whole society^[1]. Forecasting future electric load for large-scale process industrial enterprises could be conducive to managing the electricity consumption, determining the optimal production scheduling, and planning the maintenance schedule, etc., which could improve the energy efficiency and reduce the production cost. The electricity load for the large-scale process industrial enterprises could be affected by many different kinds of factors, such as the production scale, product types, the number and characteristics of electric equipment, and regional policies, etc.^[2]. The accuracy and robustness of electric load forecasting is difficult to ensure due to these complicated impact factors.

The paper manufacturing is a typical process industry, and it is the fourth largest energy-intensive industrial sector in the world as well. The energy consumption of papermaking industry accounts for nearly 7% of the total industrial energy consumption in the world^[3]. In the papermaking process, the electricity consumption accounts for about 40%-50% of the total energy consumption in the papermaking process^[4]. For industry users, the electricity prices during peak and off-peak period are different. The price for peak period is almost two times higher than that for off-peak period in most areas in China^[5]. Since there are a large number of intermittent electric equipment in the papermaking process, optimizing the startup and shutdown plans of these intermittent electric equipment could shift the electricity load from peak to off-peak

period, which could reduce the production cost. In China, the industrial enterprises have to purchase the electric load quota before consuming. Currently in practice, the papermaking enterprises usually purchase electric load quota by experience. This method often results in the problem of excessive or insufficient electric load plan, which increases the production cost with the needless wastage. Accurate short-term electric load forecasting method could guide papermaking process to determine the optimal and rational production scheduling, such as setting up the day-ahead power generation plan, the electricity purchase plan, the rational production scheduling that shift the intermittent electric equipment from peak to off-peak period, and the electricity consumption anomaly detection, etc. [6-7]. Those scheduling plans help to reduce unnecessary energy consumption. Therefore, to forecast the electric load with an efficient and accurate method could help optimizing the production scheduling and electricity consumption for the papermaking enterprises. It is a convenient way to improve the economic benefits and enhance production stability by adopting the electric load forecasting during production process.

In recent years, many studies on electric load forecasting have been presented. The first category is linear forecasting method based on time series, such as trend extrapolation [8], linear regression (LR) [9], and autoregressive integral moving average model (ARIMA) [10]. However, the embedded drawback is that they could hardly deal with the complicated nonlinear characteristics of electric load series, as a result, the forecasting performance is usually unsatisfied for these complicated industrial cases [11]. To overcome these defects of the linear forecasting methods, intelligent algorithms with

learning capabilities are applied: Jetcheva *et al.* proposed a neural network-based integration model for daily construction-level electric load forecasting, which reduced 50% of the error in comparison with the ARIMA model ^[12]. Ding *et al.* established a new gray forecasting model to forecast the total electricity consumption and industrial electricity consumption in China from 2015 to 2020, and achieved satisfactory results ^[13]. Egrioglu *et al.* added seasonal trend effects to linear models, and proposed a new hybrid approach based on SARIMA algorithm and partial high order bivariate fuzzy time series forecasting model ^[14]. Tarsitano *et al.* used two-stage seasonal ARIMA to forecast hourly electricity loads in six macro regions of Italy ^[15]. However, the model based on single algorithm still has some defects that needs to be improved. For example, single algorithm based on forecasting models often fall into local optimum, and the convergence time is much longer than the requirement of the industry. Therefore, the hybrid forecasting model based on multiple algorithms has been widely studied and applied in recent years ^[16-18]. For example, Safari *et al.* proposed a hybrid forecasting model based on the exponential smoothing model (ESM), ARIMA model and nonlinear autoregressive (NAR) neural network to forecast oil prices. This hybrid model solved the problems in accurate diagnosis of linear and nonlinear patterns in economic and financial time series ^[19]. Oliveira *et al.* proposed a bagging ARIMA and exponential smoothing method to forecast the mid-long term electric energy consumption in different countries. The error of their forecasting results is much better than the single algorithm based forecasting model ^[20].

Recently, a large number of metaheuristic methods have been developed to provide

more accurate and realistic estimation model through altering the way by which the weighting factors are determined [21]. genetic algorithm (GA) [22], particle swarm optimization (PSO) [23], rain-fall optimization algorithm (RFO) [24], gene expression programming (GEP) [21][25], artificial corporative search (ACS) [26], backtracking search algorithm (BSA) [27] and other metaheuristic optimization algorithms have been satisfactorily applied in the field of power, petroleum, steel and chemical industry. To improve the accuracy of the model, combinatorial metaheuristic methods have been developed to eliminate the disadvantages of the single optimization algorithm. For example, Particle Swarm Optimization and Ant Colony Optimization (ACO-PSO) [28], sub-gradient combined with harmony search algorithms (MSG-HS) [29], GA-PSO [30] and etc.

All these works of course facilitate the development of the electric load forecasting methods. However, the current research mainly aim at the power demand of the electricity grid with relatively stable electricity load, or the industrial sectors such as steel and petrochemical with relatively stable electricity consumption. For industrial processes with enormous fluctuations and aperiodicity in electric loads, there still lacks the adaptable forecasting model.

In order to solve those problems, this study proposes a short-term electric load forecasting model based on GA-PSO-BPNN (Genetic Algorithm - Particle Swarm Optimization - Back Propagation Neural Networks), combining the local search advantage of GA and the global search advantage of PSO to eliminate the drawback that BPNN is easy to fall into local optimum. The novelty of this study is that the

proposed model fills the gap of electric load forecasting applications for process industry, especially papermaking process. Since the electric load of papermaking process has the characters of enormous fluctuation ranges and frequency and aperiodic changes, the proposed electric load forecasting model will have highly extensive application scenarios.

2. Short term electric load forecasting model

In this study, a hybrid BPNN model with GA and PSO algorithm is established for enhancing the accuracy of short-term electric load forecasting. The technical route of the proposed model includes four sections: data collection, data preprocessing, model training, and forecasting and model evaluation. The moving average filter method is applied in the data preprocessing section. The hybrid optimization model based on GA and PSO algorithm is adapted to optimize the weights and thresholds of the BPNN algorithm. The hybrid PSO-BPNN model and the hybrid GA-BPNN model are adopted as the contrasting cases for the proposed GA-PSO-BPNN electric load forecasting model. Fig. 1 shows the flow chart of the technical route for this work.

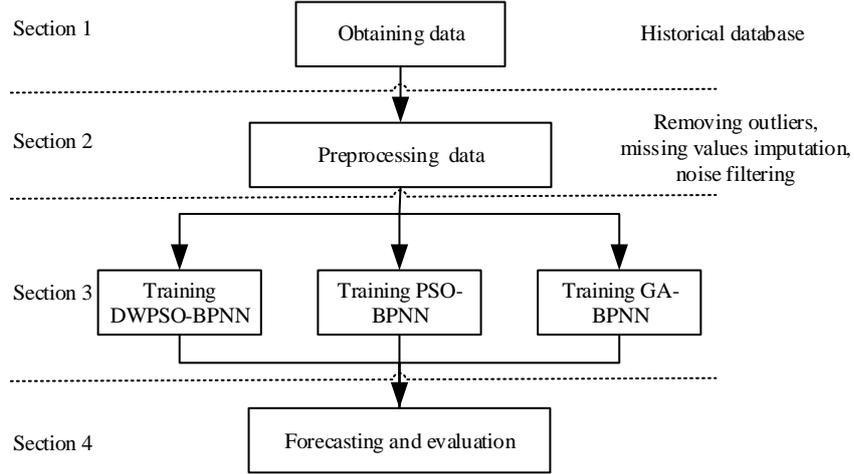


Fig. 1. The flow chart of the proposed GA-PSO-BPNN electric load forecasting model

Section 1 introduces data acquisition and input variable selection for the forecasting model. The real-time load data of the electric equipment that have high correlation with the total electric power are selected as the input variables, as well as the temperature and the relative humidity of the production environment. Section 2 is data preprocessing. All the production data are derived from online real-time data collection database. Due to the interruption of communication in the online collection, the interaction between the electric equipment and the impact loads generated during the intermittent equipment booting, the electric load data could exceed the actual range, namely outliers. These outliers could not reflect the actual production process and would decrease the accuracy of the forecasting model. Thus, the outliers need to be preprocessed to filter the abnormal data before modeling. The raw data are firstly classified by different production scheduling, then the outliers are removed by the 3σ method. The removed outliers could be filled up by the nearest neighbor interpolation method. Finally, the moving average filter is used for data de-noising to remove the high frequency component of electric load data. In Section 3, three different algorithms,

GA-PSO, PSO and GA, are adopted to optimize the weights and thresholds of BPNN. PSO-BPNN and GA-BPNN models are used as the contrasting case to analyze the accuracy and convergence rate of the proposed GA-PSO-BPNN model. In Section 4, the forecasting results from the GA-PSO-BPNN model are verified in two different industrial cases. And the feasibility of a prediction model is verified by the economic analysis of a papermaking enterprise in Hubei province.

2.1. Data preparation and preprocessing

In this work, two different papermaking processes are taken as the study cases. All the real-time production data are collected from these two papermaking mills. As mentioned above, the electric load data could sometimes exceed the actual range because of the interruption of data communication, the equipment interaction, and the impact loads generated during the intermittent equipment booting. However, apart from these outliers, the unscheduled paper-machine downtime due to paper break, the voltage fluctuations, and the shifting of paper products factors. These kinds of the data reflect the actual electric load and they cannot be regarded as the outliers. Therefore, the working conditions of the production process need to be analyzed before preprocessing data. Different products, products shifting time, scheduled and unscheduled paper-machine downtimes are recognized by shutdown signals from the DCS (Distributed Control System) and production schedules of the papermaking enterprises. After that, data preprocessing is performed to identify invalid data.

The most widely used method to remove outliers is the 3σ method. The first step is to determine whether the data satisfies the normal or approximate normal distribution

by using Eq. (1)^[31].

$$f(x) = \frac{1}{\sqrt{2 \times \pi} \times \sigma} \times e^{-\frac{(x-\mu)^2}{2 \times \sigma^2}} \quad (1)$$

where μ is the mean value of the data sequence and σ is the standard deviation. For the data sequence that satisfies the normal or approximate normal distribution, the 3σ method defines that the probability of a numerical distribution in $(\mu-3\sigma, \mu+3\sigma)$ is 0.9974^[31], and the possibility of exceeding this interval is less than 0.3%. That means the data distributed beyond $(\mu-3\sigma, \mu+3\sigma)$ could be considered as outliers and need to be removed.

After removing the outliers, the remaining data sequence faces a problem of data missing in the time series. The incomplete data sets could affect the accuracy of forecasting models^{[32] [33]}. Therefore, the missing data needs to be re-filled before modeling. The details of the missing data re-filling method are shown in the Appendix A.

For the production process, it is inevitable to introduce erroneous data, redundant data and measurement noise when converting electrical signals of power consumption equipment into digital signals^[34]. These kinds of data also have great influence on the accuracy of electric load forecasting model. Since the low-quality data could not be effectively identified by the 3σ method, data filtering method is introduced. In this work, the preprocessed data are filtered by the moving average filtering method and the Kalman filtering method. A brief description of two filtering methods is provided in the Appendix B.

2.2. Electric load forecasting model based on hybrid GA-PSO-BPNN algorithm

There are two key elements of the proposed hybrid GA-PSO-BPNN algorithm

based electric load forecasting model:

(1) The basic short-term electric load forecasting model is developed by BPNN algorithm.

(2) The weights and thresholds of BPNN algorithm are optimized by a hybrid GA-PSO optimization algorithm.

The electricity consumption of papermaking process is irregular and not cyclical. There are many factors that could affect the total electricity consumption, such as the electricity consumption of high-power electric equipment, production scheduling, etc. If the traditional prediction model is adopted, the prediction accuracy cannot satisfy the requirement of papermaking process. Therefore, a multiple inputs and single output forecasting model needs to be developed based on the electric equipment load and the production scheduling. BPNN algorithm is an optimal option to tackle the large amounts of non-linear real-time production data in papermaking process.

However, the forecasting model based on BPNN algorithm needs to be optimized because of the overfitting problem. The electricity consumption of papermaking process changes rapidly. It is suitable to directly train and test the electric load data instead of discretizing and encoding them. The PSO algorithm is sophisticated for real-value processing ^[35]. However, the PSO algorithm could not achieve complete global optimization. It is easy for the results of multi-peak problems to fall into local optimum ^[36]. To solve this problem, GA, a classic global optimization algorithm, is selected in the forecasting model to enhance the global search ability. The hybrid forecasting model introduces GA into the PSO algorithm by the selection, crossover and mutation

processes. The hybrid GA-PSO algorithm is used to optimize the weights and thresholds of BPNN, which could improve the global search ability and avoid the problem of local optimum.

The crossover process is the operation of replacing the partial structure of two parent individuals to generate new individuals, thus improving the search ability of the genetic algorithm. The algorithm used in this study is real-valued encoding. Therefore, crossover in genetic operations is achieved by arithmetic crossover operators. The formula of crossover at time t is shown as follows [37]:

$$x_{t+1}^i = \alpha \times x_t^j + (1 - \alpha) \times x_n^i \quad (2)$$

$$x_{t+1}^j = \alpha \times x_t^i + (1 - \alpha) \times x_n^j \quad (3)$$

where α is a crossover variable, x_{t+1}^i , x_{t+1}^j are two new individuals generated at time $t+1$ after crossover calculation at time t , $i \neq j$.

The core formula of variation process is [37]:

$$\Delta x_{max,t}^i = \Delta x_{max,t-1}^i + (\Delta x_{max,t}^i - \Delta x_{max,t-1}^i) / n \quad (4)$$

Combine the Eqs. (2) - (4), the updated formula of the GA-PSO algorithm can be written as [37]:

$$\Delta x_{max,t+1}^i = \Delta x_{max,t}^i + c_1 \times rand() \times (\Delta x_{max}^i - x_{in}^{k-1}) + c_2 \times rand() \times (\Delta x_{max}^j - x_{in}^{k-1}) \quad (5)$$

$$x_{in}^k = x_{in}^{k-1} + \Delta x_{max,t+1}^i \quad (6)$$

The framework of the proposed hybrid forecasting model in this study is given in Fig. 2.

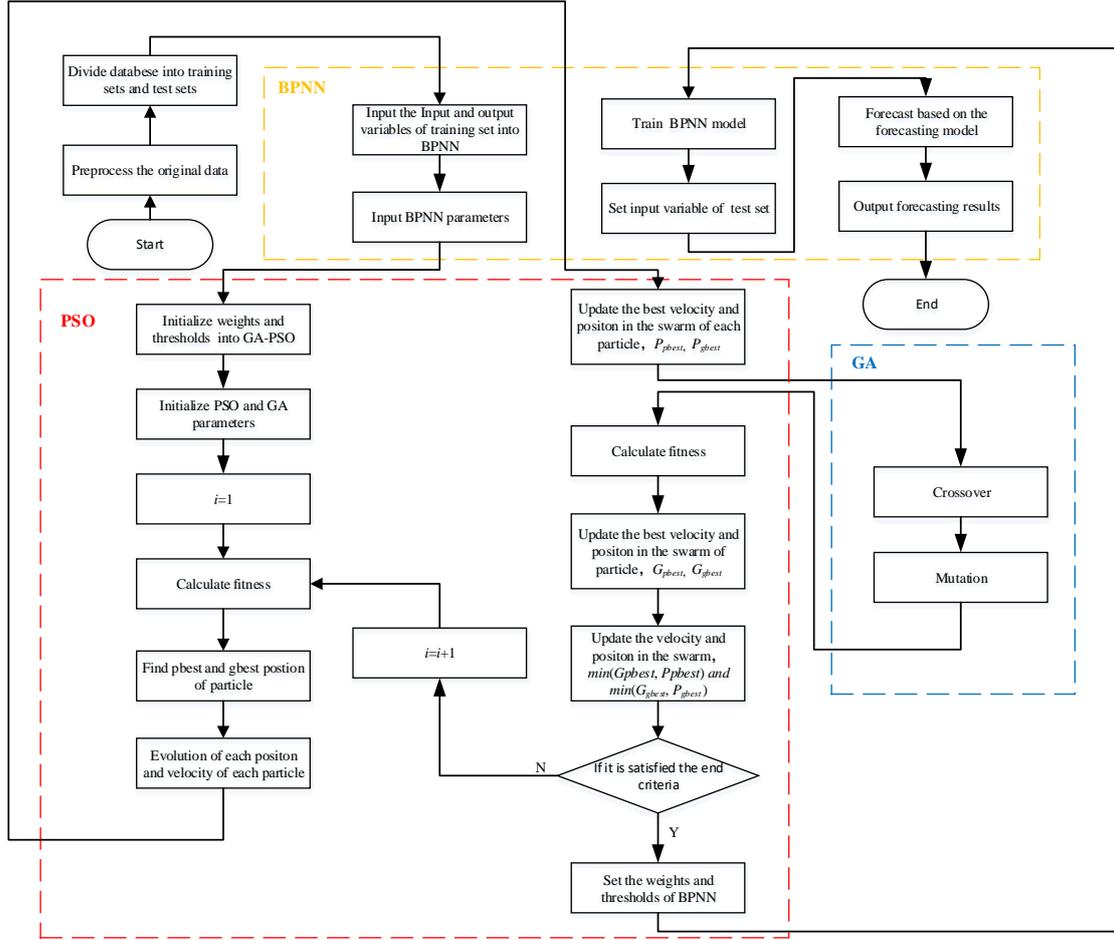


Fig. 2. Flow chart of the proposed GA-PSO-BPNN based forecasting model

The detailed steps of the hybrid algorithm are shown as follows to:

(1) Select the input and output variables by the correlation analysis, and then preprocess the raw data.

(2) Divide the input and output data sets into training and testing data set.

(3) Initialize the weights and thresholds of the BPNN. The parameters for the BPNN include the neural number of input layer, the neural number of hidden layers, the neural number of output layers, the maximum number of training (1000), BPNN convergence value (0.001), and learning rate (0.1).

(4) Initialize the parameters for the GA-PSO, including learning coefficients c_1 and

c_2 (2), population scale N (30), inertia weights w_{max} and w_{min} , (0.9 and 0.4), maximum number of iteration (100), maximum particle velocity V_{max} (0.5), crossover rate (0.5), mutation rate (0.1), upper bound of position B_{max} , and lower bound of position B_{min} .

(5) Train BPNN and calculate the fitness of each particle. Sort the fitness size, and use each particle as the local best of the current population in the particle group, denoted as P_{pbest} . The particle with the least fitness is regarded as the global optimal in the global group, denoted as P_{gbest} .

(6) Set the current iteration number i , and update the velocity and position of the particle by using Eq. (5), (6).

(7) Encode the resulting population, and perform the population selection, the genetic crossover and mutation operations on the particles.

(8) Recalculate the fitness value of the particles, and sort the fitness size. The particles with the best fitness in the current particle is taken as the local optimum, denoted as G_{pbest} . The particle with the best fitness in the particle population is taken as the global optimum, denoted as G_{gbest} .

(9) Compare the size of P_{pbest} and G_{pbest} , P_{gbest} and G_{gbest} , select the best local optimum, denoted as G_{Ppbest} , and the best global optimal, denoted as G_{Pgbest} .

(10) Update the velocity and position of the particle according to Eq. (5), (6) and determine them whether satisfy the end condition (the end condition judgement is whether the upper and lower bounds are exceeded). If not, the number of iterations will be increased by 1 and repeat Step (4)-(8) until the end condition is satisfied.

(11) Set the optimized weights and thresholds of the BPNN, and training the BPNN

model.

(12) Input the processed testing variables, and output the forecasting results.

2.2.1. Variable Selection

The accuracy of the input variables can directly affect the accuracy of the forecasting model. The total effective electric power is the sum of the electric equipment used in the whole papermaking process. How to select input variables from these large number of factors is an important issue. According to the electricity consumption characteristics of papermaking process, the input variables are divided into two parts: one is from the papermaking production process, and the other one is from external environment. For the production process, the correlation analysis is conducted between the total effective electric power and the electric equipment by Eq. (7) [38]. The electric equipment is recognized as the input variables when their absolute correlation coefficient is higher than 0.6. For the external environment, the temperature and relative humidity during production process are selected as the input variables.

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X}) \times (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \times \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (7)$$

where r is the correlation coefficient, X_i and Y_i are the i th number of the two different sequences, \bar{X} and \bar{Y} are the average of two different sequences, and n is the length of the sequence.

2.2.2. The algorithm descriptions

The core algorithms used in the electric load forecasting model are described.

BPNN algorithm

BPNN is a multi-layer feed forward network trained by the back propagation of

error to optimize artificial neural networks. Its learning rule is to use the steepest descent method to continuously adjust the weights and thresholds of the neural network by the back propagation of the error until the square of the error of the network cannot be further minimized. The objective function is [39]:

$$E = \frac{1}{2 \times L} \sum_{i=1}^n \sum_{j=1}^l (y_{ij} - \hat{y}_{ij})^2 \quad (8)$$

where L is the number of training samples size, l is the dimension of the output variable y , y_{ij} is the forecasting value, and \hat{y}_{ij} is the real value.

The BP neural network model topology includes input layers, hidden layers and output layers. The construction of BPNN is shown in Fig. 3.

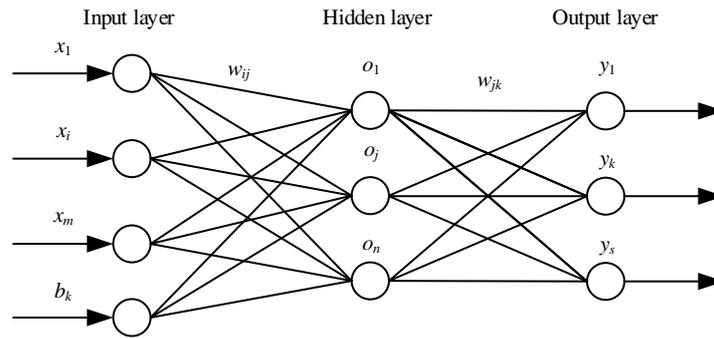


Fig.3 The construction of BPNN

This study chooses the typical BPNN algorithm, which has three-layer structure, including one output layer, one input layer and one hidden layer. The input layer, the hidden layer, and the output layer are supposed to have m neurons, n neurons, and s neurons, respectively. Then the output function is as follows:

$$y_k = f(w_{1k} \times o_1 + \dots + w_{jk} \times o_j + \dots + w_{nk} \times o_n + b_k) \quad (9)$$

where x_i represents the input value of the input layer, y_k is the output value of the output layer, f represents the activation function of the hidden layer, sigmoid function is used as the activation function in this study, o_{ik} represents the input value of the

hidden layer, w_{ij} represents the connection weights of the input and hidden layers, and w_{jk} represents the connection weights of the hidden and output layers, b_k represent bias value.

The functions of weights of hidden layers and the biases are as follows:

$$w'_{jk} = w_{jk} - \alpha \times \frac{1}{s} \times \sum_{k=1}^s \Delta_k \times o_j \quad (10)$$

$$b'_j = b_j - \alpha \times \frac{1}{s} \times \sum_{k=1}^s \Delta_k \quad (11)$$

where w'_{jk} , b'_j is the updated connection weight and bias respectively, and $\Delta_k = (y_k - \hat{y}_k) \times f'$, f' is the partial derivative of the sigmoid function.

PSO algorithm

The PSO algorithm moves the individuals in a group to find the best areas based on the fitness of the group. It can be illustrated as a particle in the swarm. Each particle moves in the search space to look for the most favorable flight path. Therefore, each particle is specified by its position and velocity in the search space which updates them based on its personal and its neighbor experiences. The velocity and positional variation are updated using the following equations ^[40]:

$$v_{ip}^k = \omega \times v_{ip}^{k-1} + c_1 \times rand() \times (pbest_{ip} - x_{ip}^{k-1}) + c_2 \times rand() \times (gbest_p - x_{ip}^{k-1}) \quad (12)$$

$$x_{ip}^k = x_{ip}^{k-1} + v_{ip}^{k-1} \quad (13)$$

where, x_{ip}^k and v_{ip}^k represent the particle's position and velocity, $pbest_{in}$ is the local best position, $gbest_n$ is the global best position, c_1 and c_2 adjust the maximum step size of learning, which are the constants called 'acceleration constants', $rand()$ are the random variables in the range of $[0, 1]$, ω is used to evaluate the effect of the previous

velocities on the current velocity, which is called ‘the inertia weight’, p indicates the position of the particle in the search space, k is the particle’s index, and i is the iteration number.

GA

Genetic algorithm is a kind of random search method that learns from the natural biological evolution (the survival of the fittest, Genetic mechanism of superiority)^[41]. The basic steps of the genetic algorithm are expressed as follows:

(1) Initialization of the optimization parameters: crossover rate, mutation rate, population size, etc.

(2) Calculation of fitness function: The fitness of each chromosome in the generation is assessed by the different of original and forecasting electricity load.

(3) Selection: Calculate the selection rate by Roulette Wheel Selection. As shown in Eq. (14)^[42]. And the chromosomes are chosen according to their fitness and used as parents. The larger the fitness value of the particle is, the greater the rate would be selected.

$$P_i = \frac{f(x_i)}{\sum_{j=1}^d f(x_j)} \quad (14)$$

Where P_i is the probability of selecting the i th particle, d is population size, $f(x_i)$ is the fitness of the i th particle.

(4) Crossover: exchanging partial data of two different chromosomes to produce a new chromosome, and the new chromosome contains the characteristics of the two original exchanged chromosomes,

(5) Mutation: a randomly change in one gene value of a chromosome from its

initial state to produce a new chromosome,

(6) Output: it is decided whether the end condition is satisfied. If not, repeat step (3)-(5). Otherwise, end the program and output the optimization goal.

3. Industrial verification of the model

The mean absolute percent error (MAPE) and relative error (RE) are used to evaluate the performance of proposed forecasting model. RE and MAPE are defined as follows ^{[43][44]}:

$$RE_i = (y_{ij} - \hat{y}_{ij}) / y_{oi} \times 100\% \quad (15)$$

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{y_{ij} - \hat{y}_{ij}}{\hat{y}_{ij}} \right|}{n} \times 100\% \quad (16)$$

The number of electric equipment and the installation of the meter in different papermaking enterprises could be totally different. In order to broad the applicability of the proposed model, the data from two different papermaking enterprises are selected to verify the forecasting model. Furthermore, the performances of the proposed forecasting model are compared with the other two hybrid forecasting models, GA-BPNN algorithm and PSO-BPNN algorithm based model.

3.1. Case 1

In the experiment of Case 1, the data are obtained from a real-world paper mill in Guangdong, China. Thirty three sampling locations in this paper mill are selected and remained for 60 days. The acquisition frequency is one per 30 min. The raw data is preprocessed by the method described in Section 2.1. The processed data of the effective electric load for Case 1 is shown in Fig. 4. The input variables are selected by

the method described in Section 2.2.1. The input variables include the real-time electric load data of 7 sampling positions, the production planning and the relative humidity, as shown in Table 1.

Table 1. Input variables list

No.	Input variables	Unit
1	The effective electric power of refiner	kWh
2	The effective electric power of pulper	kWh
3	The effective electric power of vacuum pump	kWh
4	The effective electric power of drive side	kWh
5	The effective electric power of lightening	kWh
6	The total effective electric power of pulp production	kWh
7	The total effective electric power of paper production	kWh
8	Production planning	T
9	Relative humidity	%

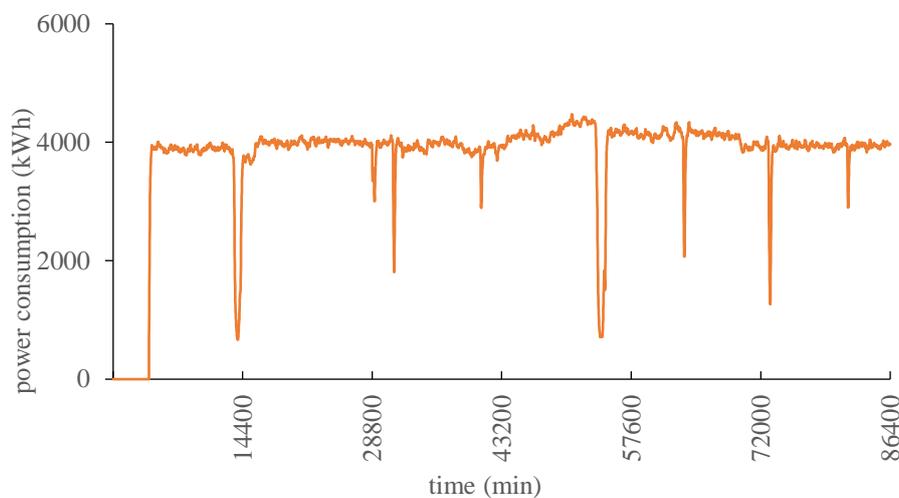
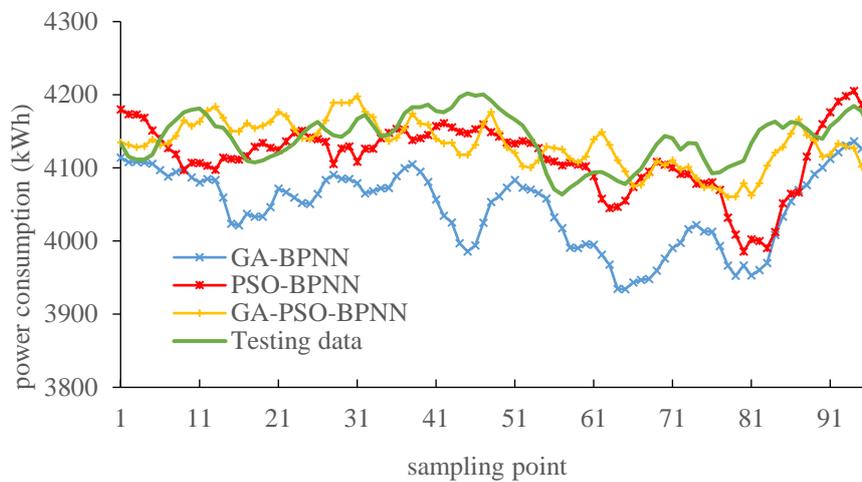


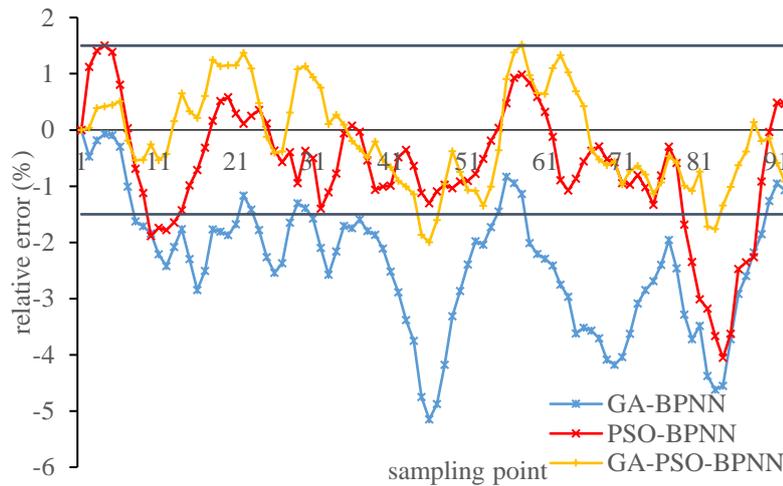
Fig. 4. The preprocessed original data

The first 59 days of the processed data are used for training, and the data on the 60th day are used to verify the forecasting performance. The parameters for GA-PSO-BPNN are set as follows: position upper limit, i.e., $\max(P_{load})$; position lower limit, i.e., $\min(P_{load})$; the neural number of input layer, i.e., 9; the neural number of hidden layer,

i.e., 15; and the neural number of output layer, i.e., 1. The parameters of GA-BPNN and PSO-BPNN are set as the same as the GA-PSO-BPNN. The forecasting results are shown in Fig. 5. Fig. 5(a) reveals a comparison of the forecasting performance between three different models, and Fig. 5(b) shows the relative error for them. The forecasting result shows that the GA-PSO-BPNN based forecasting model achieve the highest accuracy and the smallest error when compared with the other two employed models in Case 1.



(a) Forecasting results



(b) Relative error

Fig. 5. Forecasting results comparison of the three models for Case 1

3.2. Case 2

In the experiment of Case 2, the data are obtained from a paper mill in Hubei, China. Seventeen sampling locations in this paper mill are selected and remained for 60 days. And the data acquisition frequency is one per 30 min. The raw data is preprocessed by the method described in Section 2.1. The processed data of effective electric load for Case 2 is shown in Fig. 6. The input variables are selected by the method described in Section 2.2.1. The input variables include the real-time electric load data of 5 sampling positions, the production planning and the relative humidity, as shown in Table 2.

Table 2. Input variables list

No.	Input variables	Unit
1	The effective electric power of refiner	kWh
2	The effective electric power of pulper	kWh
3	The effective electric power of duster	kWh
4	The effective electric power of drive side	kWh
5	The effective electric power of air compressor	kWh
6	Production planning	T
7	Relative humidity	%

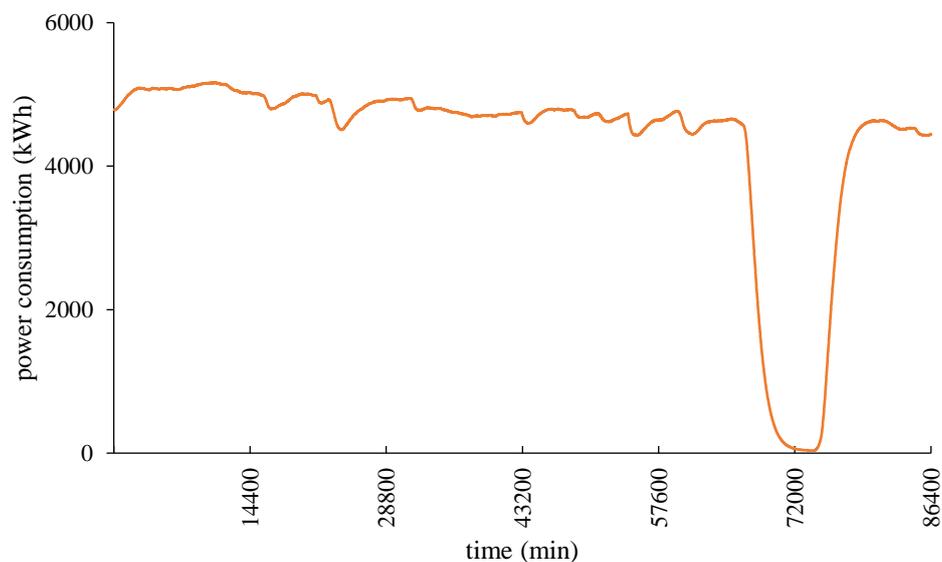
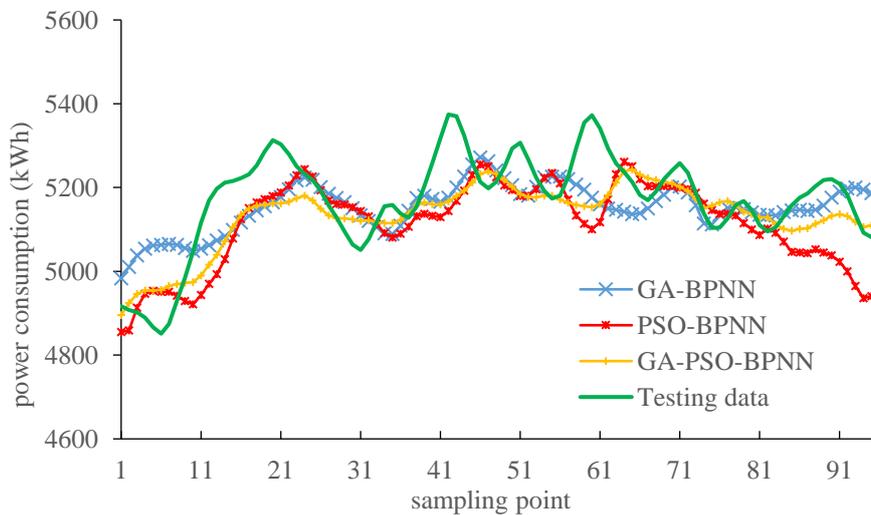


Fig. 6. The preprocessed original data

The first 59 days of the processed data are used for training, and the data on the 60th day are used to verify the forecasting performance. The parameters for GA-PSO-BPNN are set as follows: upper limit of position, i.e., $\max(P_{load})$, lower limit of position, i.e., $\min(P_{load})$, the neural number of input layer, i.e., 7, the neural number of hidden Layer, i.e., 10, the neural number of output layer, i.e., 1. And the parameters of GA-BPNN and PSO-BPNN are set as the same as the GA-PSO-BPNN. The forecasting results are shown in Fig. 7. Fig. 7(a) is a comparison of the forecasting performance between three different models, and Fig. 7(b) is the relative error for them. The forecasting result shows that the GA-PSO-BPNN based forecasting model achieve the highest accuracy and the smallest error when compared with the other two employed models in Case 2.



(a) Forecasting result

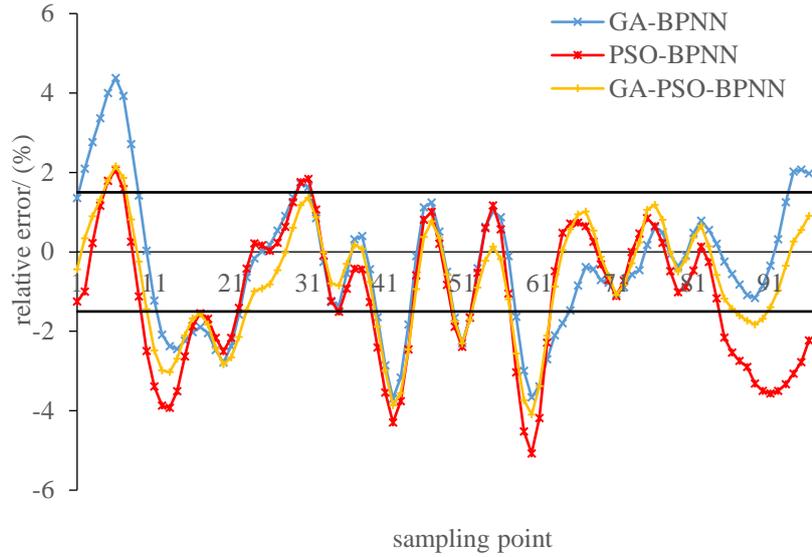


Fig.7. Forecasting results comparison of the three models for Case 2

4. Forecasting performance discussion

The comparative analysis of the forecasting performances of the GA-PSO-BPNN, GA-BPNN and PSO-BPNN are presented in this section. All the three hybrid forecasting models develop short term load forecasting under the same two case studies.

The half-hourly forecasting error of GA-PSO-BPNN, GA-BPNN and PSO-BPNN in case 1 and case 2 are shown in Fig. 8 (b), Fig. 9 (b) respectively. Setting a benchmark of $[-1.5\%, 1.5\%]$ in forecasting error makes it easy to find out the best consistent performer among the employed hybrid forecasting models. The discretized time points where the forecasting error lies within this benchmark are specifically shown in Fig. 8 (b), Fig. 9 (b). Considering the mentioned benchmark in forecasting error, the proposed GA-PSO-BPNN model provides permissible forecasting error in the two cases with respect to time points (89 and 64 in Case 1 and Case 2 respectively). On the contrary, in the PSO-BPNN model, the number of time points which lie within the forecasting

error [-1.5%, 1.5%] are 81 and 53 in Case 1 and Case 2 respectively. In the case of GA-BPNN model, the number of time points, which lies within the forecasting error [-1.5%, 1.5%] are only 22 and 58 in Case 1 and Case 2 respectively. This clearly reveals that the GA-PSO-BPNN model has the best consistent performance among all the employed models.

The generated forecasting error of GA-BPNN, PSO-BPNN and proposed GA-PSO-BPNN models for the two cases are presented in Table 3. The MAPE of the GA-PSO-BPNN model in Case 1 is 2 times less than that of GA-BPNN, and 0.2% lower than that of PSO-BPNN. In Case 2, the MAPE of the GA-PSO-BPNN model is reduced by 45.6% compared with GA-BPNN, and by 12.8% compared with PSO-BPNN. The verification results using industrial data show that the proposed GA-PSO-BPNN model achieve a higher accuracy than the compared models in all the cases.

Table 3. The forecasting performance analysis

Cases	GA-BPNN			PSO-BPNN			GA-PSO-BPNN		
	MAFRE (%)	MIFRE (%)	MAPE (%)	MAFRE (%)	MIFRE (%)	MAPE (%)	MAFRE (%)	MIFRE (%)	MAPE (%)
Case 1	-4.55	-0.09	1.88	-4.05	0.03	0.95	-1.86	0.04	0.77
Case 2	3.36	-0.08	1.82	-3.56	0.00	1.41	-3.55	-0.01	1.25

Note: MAFRE represents Maximum Forecasting Relative Error, and MIFRE represents Minimum Forecasting Relative Error.

In order to show the advantages of the proposed model in the state-of-the-art studies, this study compares the proposed method with respect to the accuracy and the characteristics of the predicted objects in different fields. The results are shown in Table 4. Table 4 shows that compared with recent researches, the method proposed in this

paper shows a wider range of applications and higher accuracy.

Table 4. Summary of studies on load forecasting for different application scenarios via AI-based algorithms

Application scenario	Core algorithms	The electric load characters	MAPE (%)	Period	References
Processing industry	GA-PSO-BPNN	Unstable and aperiodicity	0.77		This study
	CS-WNN	Stable and periodicity	1.39		[45]
	GPSO-ANN	Stable and periodicity	>1		[46]
Power grids	DMD	Unstable and periodicity	1.11		[50]
	SDPSO-ELM	Unstable and aperiodicity	2.18		[40]
	EMD-FOA-GRNN	Stable and aperiodicity	0.81		[49]
	GOA-SVM	Stable and periodicity	1.39		[18]
	SSA-BFGS-CS-WNN	Stable and aperiodicity	1.19		[45]
Wind	CSV-MOSBO-ENN	Stable and aperiodicity	1.57	Short term	[52]
	MWDO-CEEMD-BP	Unstable and aperiodicity	3.11		[53]
Building	SWEMD-ELM	Unstable and periodicity	2.14		[48]
	CS-WNN	Unstable and aperiodicity	8.0		[45]
Electricity price	ICEEMDAN-VMD-MOGWO-ENN	Unstable and aperiodicity	3.4		[54]
Asphalt pavements	SOS-LSSVR	-	12.9		[47]
Traffic	EMD-SAE	Unstable and periodicity	8.9		[51]
Energy	ACS	Stable and aperiodicity	0.74~1.46	Long term	[26]
	GEP-BSA	Stable and aperiodicity	2.89		[21]
	GEP-MLP	Stable and aperiodicity	3~23.2		[25]

5. Economic benefits analysis

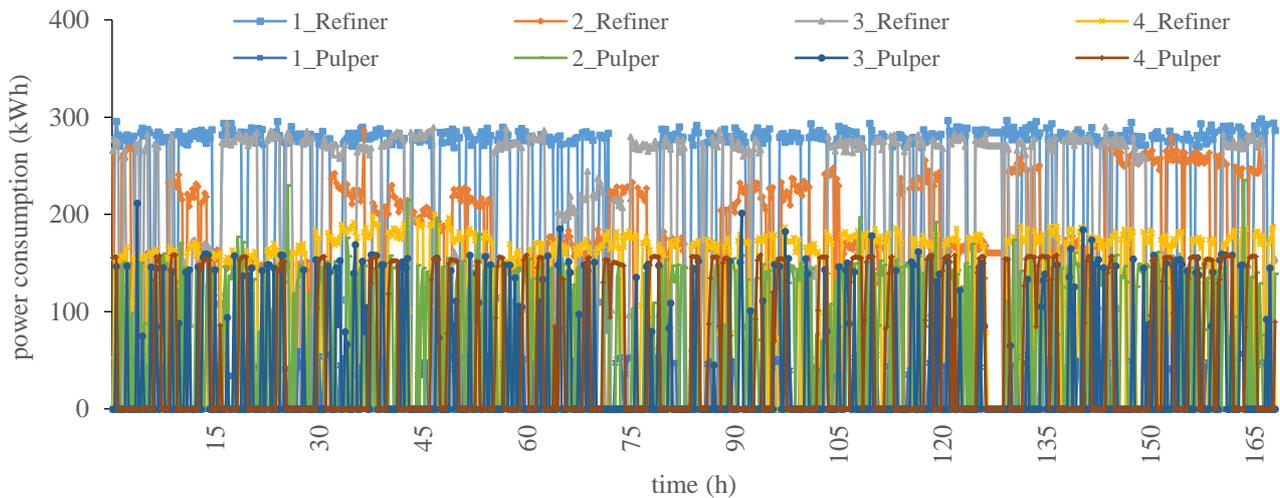
In China, to calculate the electricity cost for industrial electricity consumption is according to peak and off-peak electricity consumption. Taking Hubei Province as an instance, the electricity price for peak period (10:00-11:59; 18:00-21:59) is 1.0906 RMB/kWh, day period (08:00-09:59; 12:00-17:59; 22:00-23:59) is 0.9086 RMB / kWh

and night period (00:00-07:59) is 0.4246 RMB / kWh. In comparison with the electricity price for peak hours, the electricity price for the day and night was decreased by 17% and 61%. Reasonable use of the electricity according to the electricity price of peak period and off-peak period (for example, mitigate the peak power demand and increase the off-peak power demand) could effectively reduce the production cost and alleviate the pressure of electricity consumption of the electricity grid at peak hours.

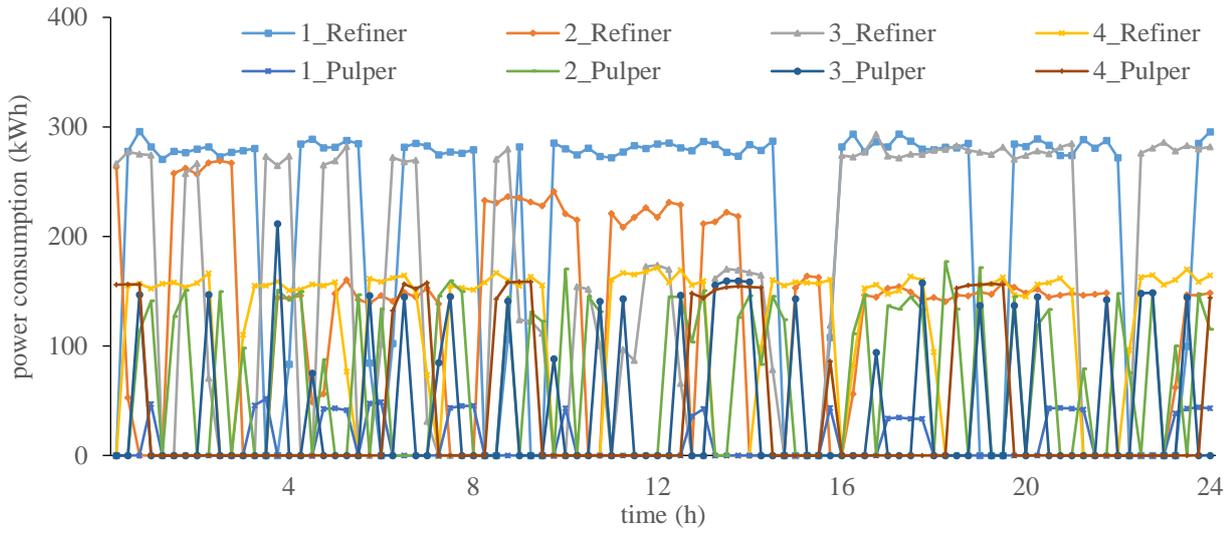
The papermaking process has a large amount of intermittent electric equipment. In order to shift peak load to off-peak period, the operating hours of intermittent equipment and production planning need to be adjusted. Based on the electric load forecasting, the operating hours of the intermittent equipment can be rescheduled. This study takes Case 2 as the case study to calculate the economic benefits from the electric load forecasting.

In Case 2, the annual production of the paper mill is 100,000 tons. The electricity consumption of this paper mill for one week is taken as the instance for the economic analysis. In the papermaking process, pulp beating and pulping processes are the batch processes. Fig. 8 (a) shows the trend of effective electric power of pulp beating and pulping process for one week, and Fig. 8 (b) shows the trend of effective electric power for the first day of the week. Fig. 9 and Fig. 10 show the electricity consumption of the refiner and pulper in the peak and off-peak periods in this week respectively. Combining these four figures, it can be found that the ratio of electricity consumption for different refiners and pulpers is close to 1:1:1. The electricity consumption of the refiner is more than twice as much as that of the pulpers. According to Fig. 10, the refiners and pulpers

have the longest operating hours in the day period, the peak period is second, and the night period is the shortest. However, the total operating hours of the pulper is much smaller than that of the refiner. Since the capacity the pulping tank in this paper mill is 100 m^3 , the required pulp could be made and stored in the night period. According to the forecasting results, the operating hours of the refiners could be adjusted to decrease the energy consumption during the peak period. The electric load forecasting results based on the optimized production schedule is shown in Fig. 11. The electricity cost of optimized production schedule based on the electric load forecasting results can save 400,000 RMB per year.



(a) Effective electric power for beating and pulping in one week



(b) Effective electric power for beating and pulping in the first day

Fig. 8. The chart of Effective electric power for beating and pulping

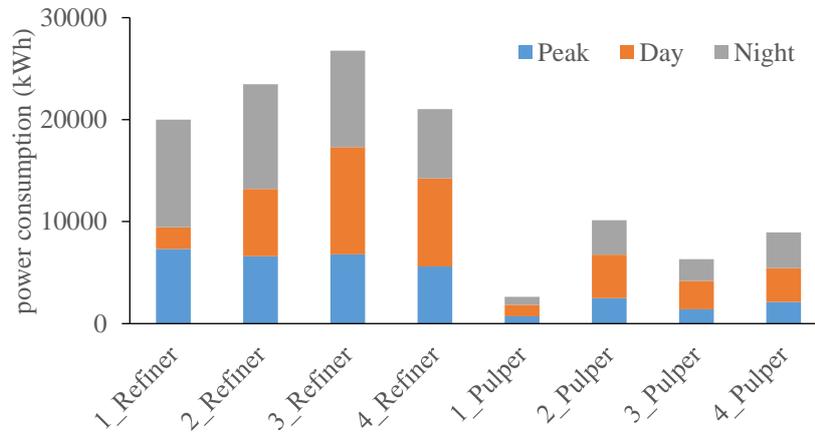


Fig.9. Total effective electric power of beating and pulping in peak and off-peak periods for a

week

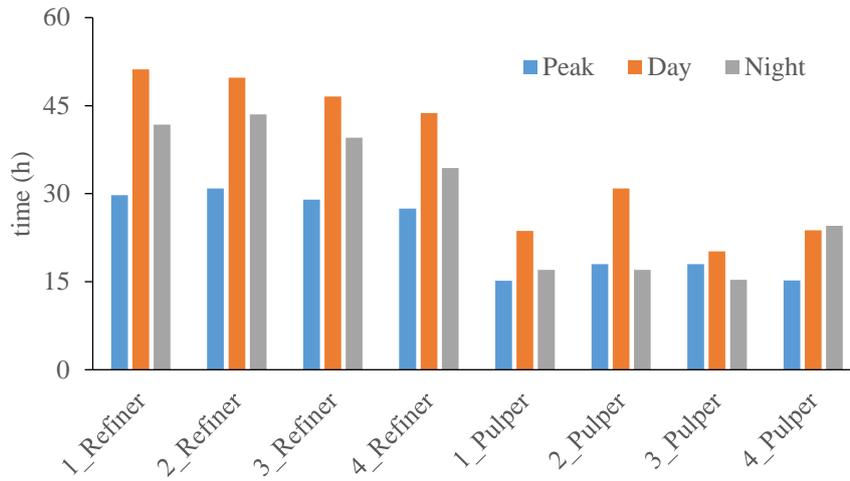


Fig.10. Total hours of beating and pulping in peak and off-peak periods

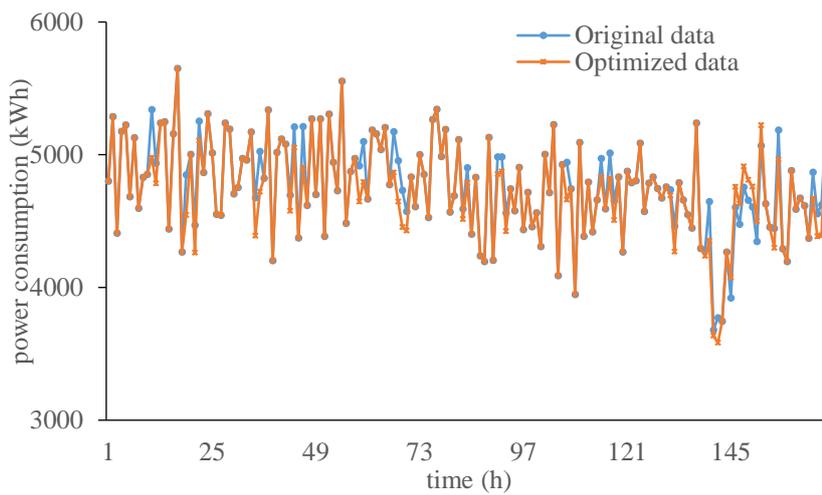


Fig.11. The electric load trend before and after optimization

6. Conclusions

Short-term electric load forecasting has been applied in many fields. Forecasting electric load could help to optimize electricity consumption, and reduce the production cost of the production process.

This study proposes a short term electric load forecasting model based on the

hybrid GA-PSO-BPNN algorithm for papermaking industry. The model is structured based on the major highly correlated electric equipment and the environmental factors. The industrial real-time data from two different papermaking enterprises are taken as the study cases for training and testing the model.

Compared with two employed contrast forecasting models (GA-BPNN and PSO-BPNN), the forecasting performance shows that the proposed GA-PSO-BPNN model has the highest accuracy. It is also revealed that the GA-PSO-BPNN model is superior to the other two hybrid forecasting models for future application in the papermaking process since its MAPE is only 0.77%. The MAPE of the GA-PSO-BPNN model could be 2 times less than that of GA-BPNN and could be reduced by 12.8% compared with PSO-BPNN. Compared with some state-of-the-art studies, the proposed model also shows good reliability and high accuracy of the forecasting performance. The electric load forecasting results could help the papermaking enterprises to optimize the production schedule. As a case study, a medium scale tissue paper enterprise with 100,000 tons annual production capacity could save 400,000 RMB per year on the cost of electricity.

Acknowledgements

This work is supported by the Fund of State Key Laboratory of Pulp and Paper Engineering (No. 201830), the Fundamental Research Funds for the Central Universities (No. 2017BQ023), the Science and Technology Project of Guangdong Province (2015B010110004 & 2015A010104004), and the Nature Science Funds of

Appendix

A. Interpolation method selection

After removing the low-quality data and outliers, the remaining data have a problem of data missing in the time series. It is necessary to add reasonable data to re-fill the data set. Because the raw data set used in this work is too large, here a 4 day data sample from the total data set is intercepted as an illustration, as shown in Fig. A.1. In this study, four different interpolation methods are used to fill the missing data. They are: nearest neighbor interpolation method, linear interpolation method, cubic spline interpolation method, and Hermite interpolation method.

(1) Nearest neighbor interpolation

It is a method of assigning the gray value of the nearest pixel point of the original pixel point to the original pixel point in the changed image. That is, the point closest to the position where interpolation is required is used as the point to interpolate.

(2) Linear interpolation

Suppose that the removing data exists on a certain straight line, the function is $y=ax+b$, and the position x of the missing data could be used to find the needed value of y .

(3) Cubic spline interpolation

Suppose $S(x)$ is a piecewise function, and it is a cubic polynomial between each interval $[x_j, x_{j+1}]$, where $a = x_0 < \dots < x_n = b$ is a given set of nodes, $S(x)$ is the cubic spline

function on nodes x_0, x_1, \dots, x_n . If the function value $Y_j=f(X_j)$ is given on the node $x_j, (j=0, 1, \dots, n)$, and $S(x_j)=y_j$ is established, $(j=0, 1, \dots, n)$, then $S(x)$ is defined as a cubic spline interpolation function, and the removing data in the interval $[a,b]$ can be interpolated using the function $S(x)$.

(4) Hermite interpolation

It satisfies that the interpolation on the node is equal to the value of the given function, and the derivative value on the node is also equal to the given derivative value at the same time. In the case of higher derivatives, the Hermitian interpolation method is more complicated. In the real word situation, it is often the case that the function value and the first derivative are given. In this case, the expressions of the Hermitian interpolation of n nodes x_1, x_2, \dots, x_n are as follows^[55]:

$$H(x) = \sum_{i=1}^n h_i \times [(x_i - x) \times (2 \times a_i \times y_i') + y_i] \quad (A.1)$$

where $y_i=y(x_i)$, y_i' is the derivative of y_i , $h_i = \prod_{\substack{j=1 \\ j \neq i}}^n \left(\frac{x-x_j}{x_i-x_j}\right)^2, a_i = \sum_{\substack{j=1 \\ j \neq i}}^n \frac{1}{x_i-x_j}$.

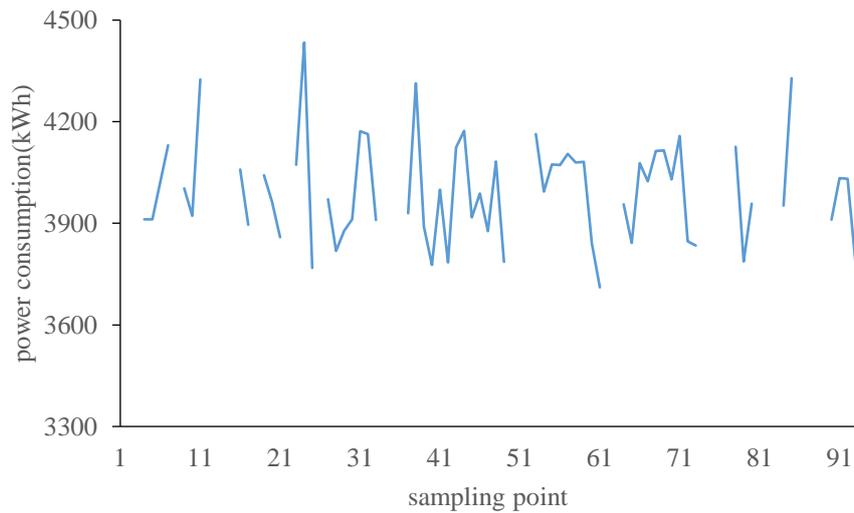


Fig. A.1. The relationship between energy consumption and time

The comparison of the four interpolation methods is shown in Figure A.2. The

result shows that except for the poor effect of cubic spline interpolation, the other three interpolation methods work well, so the nearest neighbor interpolation which is more convenient to calculate is selected to fill the removing data in this study.

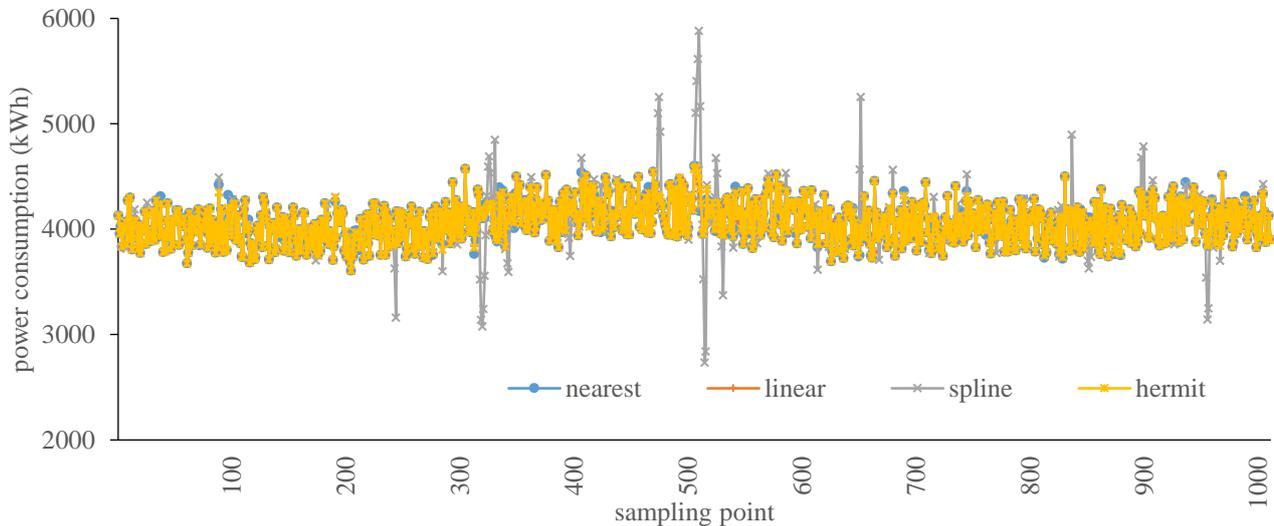


Fig. A.2. The comparison chart of four interpolation methods

B. Filter method selection

In this study, two common filtering methods, Kalman filter and moving average filter, are selected. The two methods are described in detail as follows:

(1) Moving average filter

The sample values obtained in succession could be considered as a queue. The length of the queue is n . Putting the new data from every data acquisition into the end of the queue, and throwing away the data of the original team leader (following the principle of “first in, first out” (FIFO)). The n data in the queue is subjected to an arithmetic averaging operation to get a new filtering result. The biggest advantage of the moving average filtering algorithm is that it has good real-time performance and improves the calculation speed of the system. It is defined as follows:

$$f(t) = \frac{\sum_{i=t}^{t+n} y(i)}{n} \quad (\text{A.2})$$

where $f(t)$ is the filter result at time t , $y(i)$ is the sampling data at time i , and n is the filter step size.

(2) Kalman filter

Kalman filtering is an algorithm that uses the linear system state equation to estimate the state of the system through the input and output observed data of the system. Due to the observed data includes the effects of noise and interference in the system, the optimal estimate could be seen as the process of filtering. The formula for the whole process is as follows [56]:

$$X(k|k-1) = A \times X(k-1|k-1) + B \times U(k) \quad (\text{B.1})$$

$$F(k|k-1) = A \times F(k-1|k-1) \times A' + Q \quad (\text{B.2})$$

$$X(k|k) = X(k|k-1) + K_g(k) * (Z(k) - H \times X(k|k-1)) \quad (\text{B.3})$$

$$K_g(k) = \frac{P(k|k-1) \times H'}{(H \times F(k|k-1) \times H')} \quad (\text{B.4})$$

$$F(k|k) = (I - K_g(k) \times H) \times F(k|k-1) \quad (\text{B.5})$$

where $X(k|k-1)$ is the previous state forecasting result, and $X(k-1|k-1)$ is the previous state optimization result, $U(k)$ is the control quantity of the current state. If there is no control quantity, its value will be 0, F represents the covariance, A represents a state transition matrix, Q is the covariance of the system process, $Z(k)$ is the average of the observed values, K_g is the Kalman gain.

After filling data, the data set need to be filtered. The result is shown in Fig.B.1. Due to the result of Kalman filter still contains low-quality data generated by noise or interference, the data obtained by moving average filter is smoother and more stable.

Therefore, this study uses the moving average method to filter the data.



Fig. B.1. The comparison chart of two filters

References

- [1] National Bureau of Statistics of the People's Republic of China. 2017 China Energy Statistical Yearbook. China Statistical Press: Beijing (China), 2018. Chinese.
- [2] George-Ufot G, Qu Y, Orji IJ. Sustainable lifestyle factors influencing industries' electric consumption patterns using Fuzzy logic and DEMATEL: The Nigerian perspective. *J Clean Prod.* 2017;162:624-34.
- [3] Man Y, Han Y, Li J, Hong M. Review of Energy Consumption Research for Papermaking Industry Based on Life Cycle Analysis. *Chinese J Chem Eng.* 2018. DOI: 10.1016/j.cjche.2018.08.017.
- [4] Man Y, Hong M, Li J, Yang S, Qian Y, Liu H. Paper mills integrated gasification combined cycle process with high energy efficiency for cleaner production. *J Clean Prod* 2017;156:244–52.

- [5] Yang Y, Wang M, Liu Y, Zhang L. Peak-off-peak load shifting: Are public willing to accept the peak and off-peak time of use electricity price? *J Clean Prod.* 2018;199:1066–71.
- [6] Kuster C, Rezgui Y, Mourshed M. Electrical load forecasting models: A critical systematic review. *Sustain Cities Soc.* 2017;35:257-70.
- [7] Ghasemi A, Shayeghi H, Moradzadeh M, Nooshyar M. A novel hybrid algorithm for electricity price and load forecasting in smart grids with demand-side management. *Appl Energy.* 2016;177:40-59.
- [8] Meng M, Wang L, Shang W. Decomposition and forecasting analysis of China's household electricity consumption using three-dimensional decomposition and hybrid trend extrapolation models. *Energy* 2018;165:143–52.
- [9] Dudek G. Pattern-based local linear regression models for short-term load forecasting. *Electr Power Syst Res.* 2016;130:139-47.
- [10] Lee CM, Ko CN. Short-term load forecasting using lifting scheme and ARIMA models. *Expert Syst Appl.* 2011;38:5902-11.
- [11] Zhang J, Wei YM, Li D, Tan Z, Zhou J. Short term electricity load forecasting using a hybrid model. *Energy.* 2018;158:774–81.
- [12] Jetcheva JG, Majidpour M, Chen WP. Neural network model ensembles for building-level electricity load forecasts. *Energy Build.* 2014;84:214-23.
- [13] Ding S, Hipel KW, Dang Y. Forecasting China's electricity consumption using a new grey prediction model. *Energy.* 2018;149:314-28.
- [14] Egrioglu E, Aladag CH, Yolcu U, Basaran MA, Uslu VR. A new hybrid approach

- based on SARIMA and partial high order bivariate fuzzy time series forecasting model. *Expert Syst Appl.* 2009;36(4):7424-34.
- [15] Tarsitano A, Amerise IL. Short-term load forecasting using a two-stage sarimax model. *Energy.* 2017;133:108-14.
- [16] Chang GW, Lu HJ, Chang YR, Lee YD. An improved neural network-based approach for short-term wind speed and power forecast. *Renew Energy.* 2017;105:301-11.
- [17] Cervone G, Clemente-Harding L, Alessandrini S, Delle Monache L. Short-term photovoltaic power forecasting using Artificial Neural Networks and an Analog Ensemble. *Renew Energy.* 2017;108:274-86.
- [18] Barman M, Dev Choudhury NB, Sutradhar S. A regional hybrid GOA-SVM model based on similar day approach for short-term load forecasting in Assam, India. *Energy.* 2018;145:710-20.
- [19] Safari A, Davallou M. Oil price forecasting using a hybrid model. *Energy.* 2018;148:49-58.
- [20] Oliveira EMD, Cyrino Oliveira FL. Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. *Energy.* 2018;144:776-88.
- [21] Kaboli SHA, Fallahpour A, Selvaraj J, Rahim NA. Long-term electrical energy consumption formulating and forecasting via optimized gene expression programming. *Energy* 2017;126:144-64.
- [22] Wang S, Zhang N, Wu L, Wang Y. Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method.

Renew Energy 2016;94:629-36.

- [23] Ardakani FJ, Ardehali MM. Long-term electrical energy consumption forecasting for developing and developed economies based on different optimized models and historical data types. *Energy* 2014;65:452-61.
- [24] Aghay Kaboli SH, Selvaraj J, Rahim NA. Rain-fall optimization algorithm: A population based algorithm for solving constrained optimization problems. *J Comput Sci* 2017;19:31-42.
- [25] Kaboli SHA, Fallahpour A, Kazemi N, Selvaraj J, Rahim NA. An Expression-Driven Approach for Long-Term Electric Power Consumption Forecasting. *Am J Data Min Knowl Discov* 2016;24:16-28.
- [26] Kaboli SHA, Selvaraj J, Rahim NA. Long-term electric energy consumption forecasting via artificial cooperative search algorithm. *Energy* 2016;115:857-71.
- [27] Modiri-Delshad M, Aghay Kaboli SH, Taslimi-Renani E, Rahim NA. Backtracking search algorithm for solving economic dispatch problems with valve-point effects and multiple fuel options. *Energy* 2016;116:637-49.
- [28] Kiran MS, Özceylan E, Gündüz M, Paksoy T. A novel hybrid approach based on Particle Swarm Optimization and Ant Colony Algorithm to forecast energy demand of Turkey. *Energy Convers Manag* 2012;53:75-83.
- [29] Yaşar C, Özyön S. A new hybrid approach for nonconvex economic dispatch problem with valve-point effect. *Energy* 2011;36:5838-45.
- [30] Ghorbani N, Kasaeian A, Toopshekan A, Bahrami L, Maghami A. Optimizing a Hybrid Wind-PV-Battery System Using GA-PSO and MOPSO for Reducing Cost

and Increasing Reliability. *Energy* 2017;154:581-91.

- [31] Xu J, Xiao Q, Fei Y, Wang S, Huang J. Accurate estimation of mixing time in a direct contact boiling heat transfer process using statistical methods. *Int Commun Heat Mass Transf.* 2016;75:162-8.
- [32] Ahmad T, Chen H. Short and medium-term forecasting of cooling and heating load demand in building environment with data-mining based approaches. *Energy Build.* 2018;166:460-76.
- [33] García S, Luengo J, Herrera F. Tutorial on practical tips of the most influential data preprocessing algorithms in data mining. *Knowledge-Based Syst.* 2016;98:1-29.
- [34] Xu J, Gu Y, Chen D, Li Q. Data mining based plant-level load dispatching strategy for the coal-fired power plant coal-saving: A case study. *Appl Therm Eng* 2017;119:553-9.
- [35] Cao D, Modiri A, Sureka G, Kiasaleh K. DSP implementation of the particle swarm and genetic algorithms for real-time design of thinned array antennas. *IEEE Antennas Wirel Propag Lett* 2012;11:1170-73.
- [36] Patwal RS, Narang N, Garg H. A novel TVAC-PSO based mutation strategies algorithm for generation scheduling of pumped storage hydrothermal system incorporating solar units. *Energy* 2018;142:822-37.
- [37] Awasthi A, Venkitusamy K, Padmanaban S, Selvamuthukumaran R, Blaabjerg F, Singh AK. Optimal planning of electric vehicle charging station at the distribution system using hybrid optimization algorithm. *Energy* 2017;133:70-8.

- [38] Mu Y, Liu X, Wang L. A Pearson's correlation coefficient based decision tree and its parallel implementation. *Inf Sci (Ny)* 2018;353:339-45.
- [39] Sun W, Xu Y. Financial security evaluation of the electric power industry in China based on a back propagation neural network optimized by genetic algorithm. *Energy* 2016;101:366-79.
- [40] Zeng N, Zhang H, Liu W, Liang J, Alsaadi FE. A switching delayed PSO optimized extreme learning machine for short-term load forecasting. *Neurocomputing* 2017;240:175-82.
- [41] Gong X, Plets D, Tanghe E, De Pessemier T, Martens L, Joseph W. An efficient genetic algorithm for large-scale planning of dense and robust industrial wireless networks. *Expert Syst Appl* 2018;96:311-29.
- [42] Lipowski A, Lipowska D. Roulette-wheel selection via stochastic acceptance. *Phys A Stat Mech Its Appl* 2012;391:2193-6.
- [43] Gou J, Fan ZW, Wang C, Guo WP, Lai XM, Chen MZ. A minimum-of-maximum relative error support vector machine for simultaneous reverse prediction of concrete components. *Comput Struct* 2016;172:59-70.
- [44] Amber KP, Ahmad R, Aslam MW, Kousar A, Usman M, Khan MS. Intelligent techniques for forecasting electricity consumption of buildings. *Energy* 2018;157:886-93.
- [45] Xiao L, Shao W, Yu M, Ma J, Jin C. Research and application of a hybrid wavelet neural network model with the improved cuckoo search algorithm for electrical power system forecasting. *Appl Energy* 2017;198:203-22.

- [46] Raza MQ, Nadarajah M, Hung DQ, Baharudin Z. An intelligent hybrid short-term load forecasting model for smart power grids. *Sustain Cities Soc* 2017;31:264-75.
- [47] Cheng MY, Prayogo D, Wu YW. Prediction of permanent deformation in asphalt pavements using a novel symbiotic organisms search–least squares support vector regression. *Neural Comput Appl* 2018. DIO: 10.1007/s00521-018-3426-0.
- [48] Liu Y, Wang W, Ghadimi N. Electricity load forecasting by an improved forecast engine for building level consumers. *Energy* 2017;139:18-30.
- [49] Liang Y, Niu D, Hong W. Short term load forecasting based on feature extraction and improved general regression neural network model. *Energy* 2019;166:653-63
- [50] Mohan N, Soman K.P., Kumar S. S, A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model. *Appl Energy* 2018;232:229-44.
- [51] Yang H, Chen Y P. Hybrid deep learning and empirical mode decomposition model for time series applications. *Expert Syst Appl* 2019;120:128-38.
- [52] Tian C, Hao Y, Hu J. A novel wind speed forecasting system based on hybrid data preprocessing and multi-objective optimization. *Appl Energy* 2018;231:201-19.
- [53] Yang Z, Wang J. A hybrid forecasting approach applied in wind speed forecasting based on a data processing strategy and an optimized artificial intelligence algorithm. *Energy* 2018;160:87-100.

- [54] Yang W, Wang J, Niu T, Du P. A hybrid forecasting system based on a dual decomposition strategy and multi-objective optimization for electricity price forecasting. *Appl Energy* 2019;235:1205-25.
- [55] Annaby MA, Hassan HA. Trigonometric sums by Hermite interpolations. *Appl Math Comput* 2018;330:213-24.
- [56] Garcia RV, Pardal PCPM, Kuga HK, Zanardi MC. Nonlinear filtering for sequential spacecraft attitude estimation with real data: Cubature Kalman Filter, Unscented Kalman Filter and Extended Kalman Filter. *Adv Space Res* 2018. DOI: 10.1016/j.asr.2018.10.003.