Short-term Load Prediction of Integrated Energy System with Wavelet Neural Network Model Based on Improved Particle Swarm Optimization and Chaos Optimization Algorithm

Leijiao Ge, *Senior Member*, *IEEE*, Yuanliang Li, Jun Yan, *Member*, *IEEE*, Yuqian Wang, and Na Zhang

***Abstract*—To improve energy efficiency and protect the environment, the integrated energy system (IES) becomes a significant direction of energy structure adjustment. This paper innovatively proposes a wavelet neural network (WNN) model optimized by the improved particle swarm optimization (IPSO) and chaos optimization algorithm (COA) for short-term load prediction of IES. The proposed model overcomes the disadvantages of the slow convergence and the tendency to fall into the local optimum in traditional WNN models. First, the Pearson correlation coefficient is employed to select the key influencing factors of load prediction. Then, the traditional particle swarm optimization (PSO) is improved by the dynamic particle inertia weight. To jump out of the local optimum, the COA is employed to search for individual optimal particles in IPSO. In the iteration, the parameters of WNN are continually optimized by IPSO-COA. Meanwhile, the feedback link is added to the proposed model, where the output error is adopted to modify the prediction results. Finally, the proposed model is employed for load prediction. The experimental simulation verifies that the proposed model significantly improves the prediction accuracy and operation efficiency compared with the artificial neural network (ANN), WNN, and PSO-WNN.**

***Index Terms*—Integrated energy system (IES), load prediction, chaos optimization algorithm (COA), improved particle swarm optimization (IPSO), Pearson correlation coefficient, wavelet neural network (WNN).**

Manuscript received: September 13, 2020; accepted: April 8, 2021. Date of CrossCheck: April 8, 2021. Date of online publication: June 7, 2021.

This research was supported in part by the National Key Research and Development Program of China (No. 2018YFB1500800), the National Natural Science Foundation of China (No.51807134), and the State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology (No. EERI_KF20200014).

L. Ge (corresponding author) and Y. Li are with the School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China (e-mail: legendglj99@tju.edu.cn; tjlyliang@foxmail.com).

J. Yan is with the Concordia Institute for Information Systems Engineering, Concordia University, Montréal, QC H3G 1M8, Canada (e-mail: jun.yan@concordia.ca).

Y. Wang is with Sany Heavy Machinery Co., Ltd., Suzhou 215300, China (e-mail: 18361397336@163.com).

N. Zhang is with the Department of Electrical Engineering, Inner Mongolia University of Technology, Hohhot 010321, China (e-mail: zhangna2010337@163.com).

DOI: 10.35833/MPCE.2020.000647

I. Introduction

THE integrated energy system (IES) is an integrated energy supply platform based on the multi-energy complementation and energy cascade utilization [1], which realizes the conversion of various energy sources such as combined cooling, heating and power (CCHP) [2]. With the rapid development of the Energy Internet, the coupling of different types of energy becomes significant, which calls for more advanced load prediction with higher accuracy and efficiency [3], [4].

Energy load can be broadly divided into three categories: electrical load, thermal load, and cooling load. The cooling load keeps indoor temperature and humidity at specified levels, including the lighting heat dissipation, human body heat dissipation, and other ways of heat dissipation. The thermal load refers to the amount of heat released in a time domain when the fuel burns. The sum of the electricity taken from the power system by the user's electrical equipment represents the electrical load, whose related characteristics can be observed through the load curves.

Early short-term load prediction (STLP) did not fully consider the impact of the environment, resulting in low load prediction accuracy on weekends [5]. Statistical models have been widely applied later in STLP, including the autoregressive (AR) model [6], exponential smoothing (ES) model [7], and autoregressive integrated moving average (ARIMA) model [8]. Among others, the AR model is the most popular due to its simplicity with effectiveness [6].

Most statistical methods are unsuitable in IES due to the complicated and non-linear relationships between multiple IES loads and related influencing factors. With the development of artificial intelligence technology, STLP is moving in an intelligent direction. To predict and analyze IES performance, extensive artificial intelligence prediction models such as Kalman filter [9], support vector machine (SVM) [10], and deep neural network (DNN) [11], [12] have been proposed for STLP. Reference [9] proposed the STLP conditioning applying the Kalman filter algorithm, which defined the load prediction as an observation model and a state-space model in time domain. Reference [10] presented a

new STLP method based on SVM to improve the prediction accuracy and timeliness. The neural network has a more robust capability of non-linear approximation and self-improving than statistical methods. Based on the time-series prediction of long short-term memory DNN, [11] proposed an energy prediction strategy for the multi-energy information interaction in regional IESs from the perspective of vertical interaction and horizontal interaction. Considering the influences among three types of load, [12] proposed a DNN model based on a multi-layer restricted Boltzmann machine (RBM), which achieved high accuracy at the costs of large training data size and slow prediction speed.

In [13], the wavelet neural network (WNN) was introduced to predict the cooling, thermal, and electrical loads of the IES with competitive performances. Compared with other methods mentioned above, WNN has stronger learning and feature extraction ability over a relatively simple structure, which is suitable for fast load prediction of IES with many influencing factors. As the traditional WNN prediction method is sensitive to the initial connection weights and wavelet parameters, a poor selection of parameters will cause a slow convergence rate or even non-convergence. The gradient descent method cannot obtain satisfactory convergence performance in WNN [14]. Besides, IES is a non-linear system with complex influencing factors and choosing the appropriate influencing factors is challenging.

In this paper, we propose a new WNN model optimized by the improved particle swarm optimization (IPSO) and chaos optimization algorithm (COA) for the STLP of IES. First, representative influencing factors are selected by the Pearson correlation coefficient, and redundant influencing factors can be discarded. As the particle swarm optimization (PSO) often falls into the local optimum and results in premature convergence, the dynamic inertia weight is presented to control the global search and convergence to the global best solution in the IPSO. To further improve the optimization effectiveness, the COA is employed to search for the personal optimal particle in chaos. The IPSO-COA is used to optimize the connection weights and wavelet parameters in the WNN. Besides, a feedback link is added to the prediction model, where the output error is adopted to modify the prediction results. Ultimately, the optimized WNN is used to predict the IES load. The experimental simulation and analysis verify that the IPSO-WNN has a satisfactory improvement in the prediction accuracy and operation efficiency compared with the ANN, WNN, and PSO-WNN.

The rest of this paper is organized as follows. In Section II, the problem formulation is presented. Section III employs the Pearson correlation coefficient to select the key influencing factors of load prediction. Section IV introduces the principle and process of the proposed model in detail. Section V verifies the feasibility of the proposed model via simulations. Section VI concludes the paper.

II. Problem Formulation

The structure of the load prediction model of IES is shown in Fig. 1. The IES loads are composed of three types, i.e., the electrical load, thermal load, and cooling load, whose predicted values are denoted as $\hat{y}_1 + \Delta y_{1t}$, $\hat{y}_2 + \Delta y_{2t}$, and $\hat{y}_3 + \Delta y_{3t}$, respectively. The input data in the load prediction include the historical load data, meteorological data, and daily parameters of IES. The key influencing factors of the three types of load are determined by the Pearson correlation coefficient ρ_{XY}. To reduce the prediction dimensions and forecast the three types of load with high accuracy and efficiency, influencing factors with a Pearson correlation coefficient greater than the threshold are regarded as representative influencing factors and used as the input for prediction.

The training data are input into the proposed model for the predictive outcome training and error correction, and the application data are input to predict each type of load. To optimize the performance of WNN, the particle iteration velocity and position in the IPSO are adopted to replace the connection weights and wavelet parameters in the WNN, and the COA is conducted on the global optimal particle in each iteration. The output of the proposed model includes electrical load, thermal load, and cooling load, which is recorded as $\hat{y}_i$ ($i = 1, 2, 3$). Finally, the STLP result of IES $\hat{y}_i + \Delta y_{it}$ is obtained after feedforward error correction.

III. Pearson Correlation Coefficient

Unlike STLP in traditional power systems, the IES has abundant energy conversion equipment, and the three types of load in the IES have high coupling and strong correlation. Reference [15] suggested that the load characteristics and the relationship of different loads should be considered in the STLP of the IES. STLP of IES is affected not only by the interaction of loads but also by many other influencing factors such as the operation environment and day types. Reference [16] presented that proper quantification and selection of influencing factors are the premises of accurate prediction. Reference [17] selected five influencing factors related to ignition temperature and activation energy by Pearson correlation coefficient. It developed a three-layer back-propagation neural network model with the selected five influencing factors to forecast the ignition characteristics of coal blends [17]. Therefore, the Pearson correlation coefficient is introduced to quantify the correlation between multiple influencing factors and predicted objects (IES loads) [18], which is calculated by:

$$\rho_{AB} = \frac{\sum_{i=1}^{I}\left(A_i - \overline{A}\right)\left(B_i - \overline{B}\right)}{\sqrt{\sum_{i=1}^{I}\left(A_i - \overline{A}\right)^2}\sqrt{\sum_{i=1}^{I}\left(B_i - \overline{B}\right)^2}} \tag{1}$$

where $\overline{A}$ and $\overline{B}$ are the average values of variables A_i and B_i, respectively; ρ_{AB} is the Pearson correlation coefficient; and I is the number of statistics.

The Pearson correlation coefficient ranges from -1 to 1, as shown in Fig. 2. The higher the absolute value of the Pearson correlation coefficient, the greater the correlation between influencing factors. If two variables are positively correlated, the range of the Pearson correlation coefficient is (0, 1]; on the contrary, the range of the Pearson correlation coefficient is $[-1, 0)$. When the Pearson correlation coefficient is 0, the two variables are completely independent of each other.

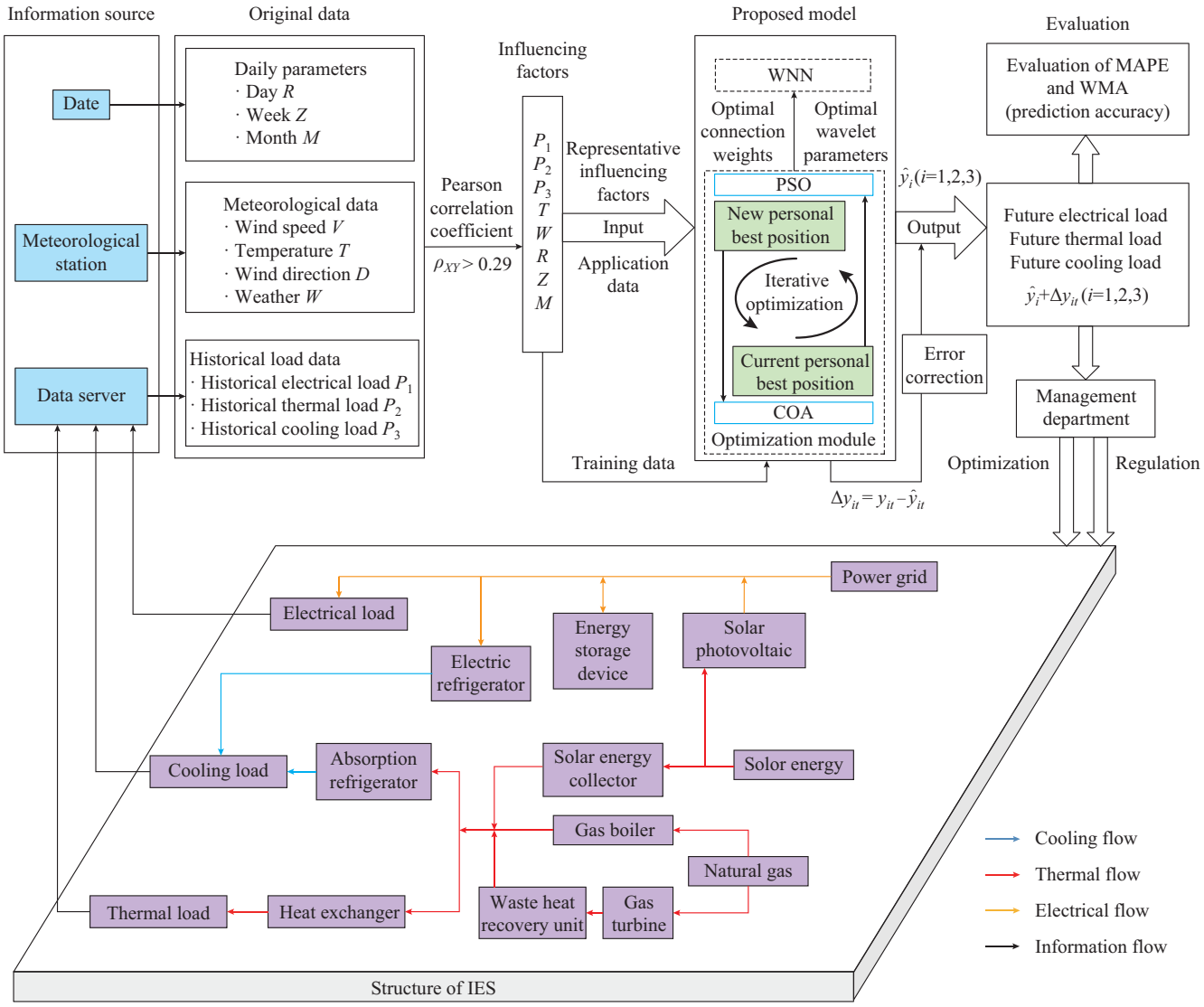

Fig. 1. Structure of load prediction model of IES.

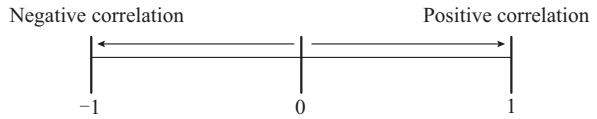

Fig. 2. Interval of Pearson correlation coefficient.

To improve the accuracy and efficiency of STLP of IES, influencing factors of STLP of IES can be selected based on the Pearson correlation analysis. The main influencing factors include historical electrical load P_1, historical thermal load P_2, historical cooling load P_3, wind speed V, temperature T, wind direction D, weather W, day R, week Z, and month M. The predicted objects are the electrical load, thermal load, and cooling load. Too many selected influencing factors may complicate the network structure and decrease the operation efficiency. Therefore, the influencing factors with a correlation coefficient below a certain threshold should be ignored in the subsequent prediction process. The determination of the appropriate threshold can be set through multiple simulation experiments.

IV. Principle and Process of Proposed Model

This section will briefly introduce the principle and process of the proposed model. The flowchart of the proposed WNN model based on IPSO and COA is shown in Fig. 3.

A. WNN

There are many potential influencing factors in the STLP of IES, and each factor is interrelated. It is challenging to extract data features directly. Due to the scalability and transformability of the wavelet function, WNN makes the data feature more obvious and reduces the extraction difficulty. Besides, WNN takes wavelet space as the feature space for pattern recognition and extracts features from the signal by weighting the inner product of the wavelet base and signal vector. The procedure can effectively learn input and output characteristics of the system without much prior knowledge such as structure and characteristics [19]. Compared with generic neural networks, WNN has a better generalization ability, convergence rate, and robustness to complex non-linear systems [20].

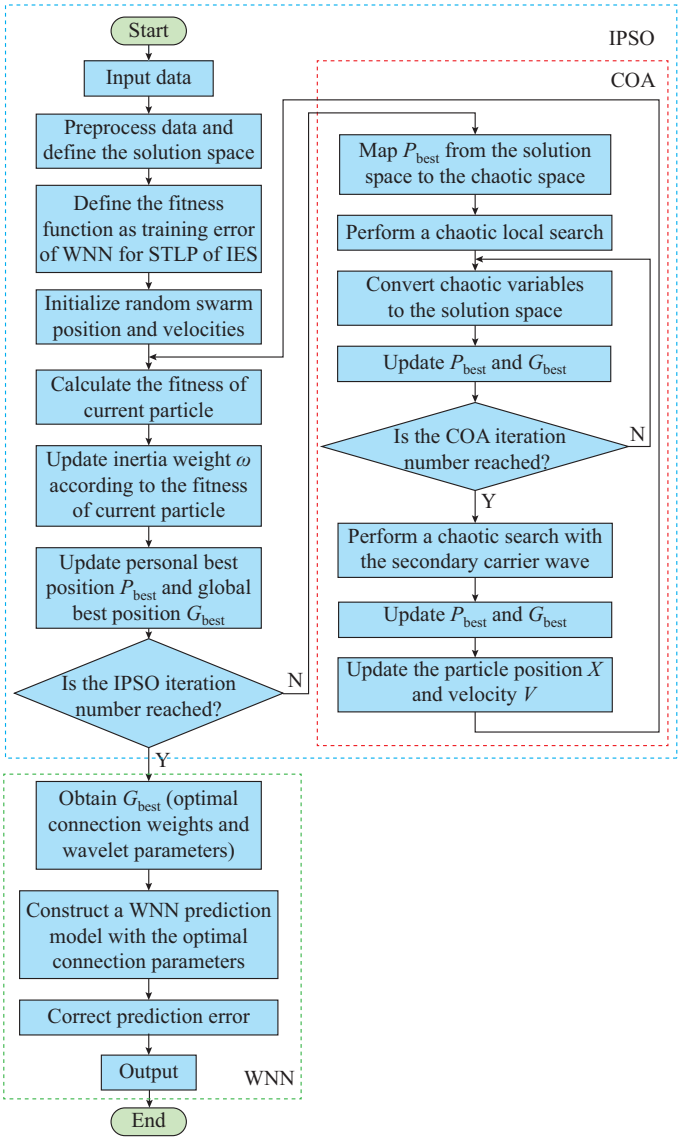

Fig. 3. Flowchart of proposed WNN model based on IPSO and COA.

The structure in Fig. 4 shows that the WNN consists of an input layer, a hidden layer (wavelet layer), and an output layer [21]. The WNN combines artificial neural network (ANN) [22] and wavelet analysis. The wavelet function is adopted as the activation function in the hidden layer of the WNN. The input in the input layer are $[x_1, x_2, ..., x_M]$, which are transmitted to the hidden neurons $[F_1, F_2, ..., F_n]$ in the hidden layer through the connection weight w_{ij}. The predicted loads in the output layer are $[y_1, y_2, ..., y_N]$.

In Fig. 4, $w_{jk}(j=1,2,\ldots,n; k=1,2,\ldots,N)$ is the connection weight between the hidden layer and the output layer. The output variable (predicted load) of the WNN $y_k(k=1,2,\ldots,N)$ is defined as:

$$y_k = \sum_{j=1}^{n} w_{jk} F_j(x_1, x_2, \ldots, x_M) \tag{2}$$

$$F_j(x_1, x_2, \ldots, x_M) = \tau\left(\frac{X_j - b_j}{a_j}\right) \quad \forall j = 1, 2, \ldots, n \tag{3}$$

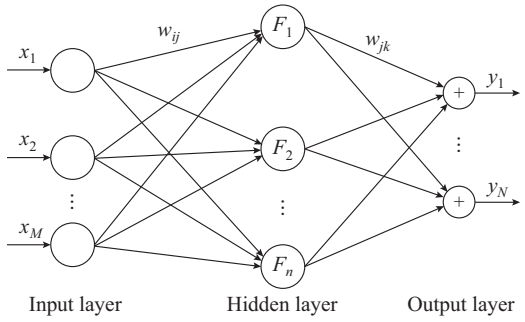

Fig. 4. Structure of WNN.

$$X_j = \sum_{i=1}^{M} w_{ij} x_i \tag{4}$$

where a_j is the dilation factor; b_j is the translation factor of the wavelet function; X_j is the input of the hidden layer; and $\tau(\cdot)$ is the Morlet wavelet function, which is a complex wavelet modulated by the Gaussian function and has the advantages of time-frequency regularity of Gaussian function. Compared with real wavelets, the Morlet wavelet can reflect the size of different time scales. Besides, the strength of the time-scale signal is determined by the modulus of the complex wavelet transform coefficient. The real part describes the intensity and phase of time-scale signals with different characteristics. $\tau(x)$ is defined as:

$$\tau(x) = \mathrm{e}^{-0.5x^2} \cos(1.75x) \tag{5}$$

However, the proposed model is sensitive to the initial connection weights w_{ij} and w_{jk} and wavelet parameters a_j and b_j. A poor selection of parameters will cause a slow convergence rate or even non-convergence.

B. IPSO

The gradient descent algorithm is adopted by the traditional WNN to optimize the parameters such as connection weights and wavelet parameters. However, the transmission information of neural networks with a complex structure is sparse, which may lead to low prediction effectiveness. As a mature optimization method, the PSO algorithm has the characteristics of fast optimization speed and satisfactory optimization effect [23], which is suitable for optimizing the connection weights and wavelet parameters in the WNN.

Each individual in the swarm represents a particle. All the particles follow the same principle: while continuously acquiring the fitness f (the effectiveness of the given solution) of its current position, they accelerate toward the personal best position P_{best} and the global best position G_{best}. Besides, the particle position represents the solution to the problem. The inertia weight ω is significant in the PSO algorithm and determines to what extent the particle remains along its original course unaffected by the pull of G_{best} and P_{best} [24]. A small inertia weight indicates excellent local searchability, while a large inertia weight leads to excellent global searchability. Therefore, the optimization effectiveness and convergence speed can be improved by selecting the appropriate inertia weight ω (from 0 to 1) in the IPSO algorithm. The algorithm implementation process is as follows.

Step 1: define the solution space. The input data come

from the influencing factors selected in Section III. To optimize the WNN model, the particle position is designated as connection weights w_{ij} and w_{jk} and wavelet parameters a_j and b_j. Besides, the position and velocity of the particles should be given upper and lower limits according to the actual situation.

Step 2: define the fitness function. The predicted value is calculated by the WNN with optimal connection weights and wavelet parameters. The mean absolute percentage error (MAPE) is selected as the fitness function to quantify the training error and the goodness of a given solution, which can be expressed as:

$$f_c = M_{\text{APE}} = \frac{1}{N_{pre}} \sum_{t=1}^{N_{pre}} \left| \frac{y_t - y_t'}{y_t} \right| \times 100\% \tag{6}$$

where f_c is the fitness of the current particle; y_t is the actual value; y_t' is the predicted value; M_{APE} is the value of MAPE; and N_{pre} is the number of predicted samples.

Step 3: initialize the random swarm location and velocities. To start searching for the optimal position, each particle begins at a random position, whose velocity is random in direction and magnitude.

Step 4: calculate f_c. If f_c is less than the fitness value at the respective P_{best} for the particle ($f(P_{\text{best}})$) or the fitness value at the global G_{best} ($f(G_{\text{best}})$), then P_{best} and G_{best} are replaced with the current position.

Step 5: update the inertia weight ω according to f_c. According to the optimization demand of different particle fitness, the dynamic weight selection strategy is established. By dynamically adjusting ω, fine particles tend to perform exploitation to refine local search results, while weak particles perform extensive modifications to explore space with massive steps. Therefore, the possibility of falling into a local optimum can be reduced. Assume that $\bar{f}$ and f_p are the average fitness values of the current particle and the global best particle, respectively.

1) When $f_c < \bar{f}$, the current particle is far from the G_{best}, thus the inertia weight ω should be larger, which is calculated as:

$$\omega = \omega_{avg} + \frac{\omega_{\max} - \omega_{\min}}{2} r_{and} \tag{7}$$

$$\omega_{avg} = \frac{\omega_{\max} + \omega_{\min}}{2} \tag{8}$$

where $\omega_{\max}$ and $\omega_{\min}$ are the maximum and minimum inertia weights, respectively; ω_{avg} is the average inertia weight; and r_{and} is a random value in [0,1].

2) When $f_c > f_p$, the current particle is closer to the G_{best}; thus, the inertia weight ω should be smaller, which is calculated as:

$$\omega = \omega_{avg} - \left| \frac{f_c - f_p}{f\left(G_{\text{best}}\right) - f_p} \right| (\omega_{avg} - \omega_{\min}) \tag{9}$$

3) When $\bar{f} \le f_c \le f_p$, the non-linear decrease of the inertia weight ω can be calculated as:

$$\omega = \omega_{\max} - \frac{(\omega_{\max} - \omega_{\min}) i_{ter}}{M_{gen}} \tag{10}$$

where i_{ter} is the current iteration number of IPSO; and M_{gen} is the total number of IPSO iterations.

Step 6: update the particle position and velocity. The updated formula of the particle position can be expressed as:

$$X_i(k+1) = X_i(k) + V_i(k+1) \tag{11}$$

where $X_i(k)$ is the position of particle i in the iteration k; and $V_i(k)$ is the velocity of particle i in the iteration k, which is the core element of the entire optimization. The particle velocity is accelerated based on the relative positions of P_{best} and G_{best}. The goal of $V_i(k+1)$ is to guide particles to the position with the best fitness, which is defined as:

$$V_i(k+1) = \omega V_i(k) + c_1 r_{and1} (P_{\text{best}} - X_i) + c_2 r_{and2} (G_{\text{best}} - X_i) \tag{12}$$

where r_{and1} and r_{and2} are the random numbers in the range of [−1, 1]; and c_1 and c_2 are the learning factors.

Step 7: repeat *Steps 4* to *6*.

Return to *Step 4* until i_{ter} reaches M_{gen}. G_{best} is regarded as the optimal connection weight and wavelet parameters.

C. COA

The PSO has several shortcomings, e.g., being trapped in the local optimum and premature convergence, resulting in sizable STLP error [25]. Due to the easy implementation and remarkable ability to avoid being trapped into the local optimum, chaos has been applied to many scenarios. To apply chaos for optimization, COA is incorporated into the iterative process in the IPSO to enrich the searching behavior and jump out of the local optimum. The COA includes two main steps: mapping from the chaotic space to the solution space and searching optimal regions using chaotic dynamics [26], [27].

Considering the computational complexity, only the P_{best} of the current personal optimal particle is searched by chaos in each iteration. If the new particle produced by the chaotic search is better than P_{best}, then P_{best} is replaced with the new particle. The chaotic search process is as follows.

Step 1: map from the solution space to the chaotic space. $\boldsymbol{P}_{\text{best}} = [p_1, p_2, \ldots, p_D]$ represents the vector of the personal optimal positions. Each dimension p_i in the $\boldsymbol{P}_{\text{best}}$ is mapped to the chaotic variable δ_d^k, and $\delta_d^k \in [0,1]$. The mapping formula is presented as:

$$\delta_d^k = \frac{p_{\text{best},d} - p_{\max,d}}{p_{\max,d} - p_{\min,d}} \quad \forall d = 1, 2, \ldots, D \tag{13}$$

where D is the number of dimensions in $\boldsymbol{P}_{\text{best}}$; $P_{\text{best},d}$ is the d^{th} variable of $\boldsymbol{P}_{\text{best}}$; and $P_{\min,d}$ and $P_{\max,d}$ are the lower and upper limits of the d^{th} variable, respectively.

Step 2: perform a chaotic local search. The iterative logistic equation is employed to deal with δ_d^k, as shown in (14).

$$\delta_d^{k+1} = \mu \delta_d^k (1 - \delta_d^k) \tag{14}$$

where μ is the control parameter. When $\mu = 4$, the iterative logistic equation exhibits chaotic dynamics, exhibiting sensitive dependence on initial conditions. A minute difference in the initial value of the chaotic variable would make a considerable difference in its long-time behavior. The track of chaotic variables can travel ergodically over the entire search space. In general, the above chaotic variable has extraordinary characteristics such as ergodicity, pseudo-randomness,

and irregularity.

Step 3: convert the chaotic variables to the solution space. The iterative chaotic variable δ_d^{k+1} is inversely mapped to the original solution space, as shown in (15).

$$x_d^{k+1} = (p_{\max,d} - p_{\min,d})\delta_d^{k+1} + p_{\max,d} \tag{15}$$

Step 4: calculate the fitness of the current particle x_d^{k+1}. Based on the fitness function proposed in the IPSO algorithm, if the current particle fitness is less than the value at the respective P_{best} for the particle or the global G_{best}, then the appropriate positions are replaced with the current position.

Step 5: repeat *Steps 1* to *4*. If the iteration number of COA C_{itr} has not been reached, go back to *Step 1* and conduct a chaotic local search again; otherwise, proceed to *Step 6* for the secondary carrier wave.

Step 6: generate the initial secondary carrier wave. For more precise optimization, the second carrier wave can be carried out. Re-assign D initial values from 0 to 1 of δ_d^k in (13) with slight differences.

Step 7: perform a chaotic local search using the chaotic variables of the secondary carrier. According to (14), generate δ_d^{k+1} through chaotic local search. Then, perform precise optimization as:

$$P_{\text{best}}^{k+1} = P_{\text{best}}^{k} + \alpha\delta_d^{k+1} \tag{16}$$

where α is the adjustment coefficient (less than 1).

Step 8: calculate the fitness of the new personal best position P_{best}^{k+1}. If the fitness value of P_{best}^{k+1} ($f(P_{\text{best}}^{k+1})$) is less than $f(P_{\text{best}})$ or $f(G_{\text{best}})$, then the P_{best} and G_{best} are replaced with the new personal best position.

Step 9: repeat *Steps 7* and *8*.

Go back to *Step 7* and conduct a chaotic local search until it reaches the iteration number C_{itr}.

V. Experimental Simulation and Analysis

To verify the effectiveness of the proposed model, an IES of the Binhai district in Tianjin, China, is selected for daily load prediction. The historical load, temperature, weather, and date data from January 2017 to December 2017 are extracted for network training and error correction. The hyper-parameter validating data are from January 2016 to December 2016 for the hyper-parameter setting. The testing data are from January 2018 to December 2018 for load prediction and effectiveness verification. Simultaneously, the ANN, WNN, and PSO-WNN are introduced to compare with the proposed model. The time resolution of the thermal load and the electrical load is 15 min, and the time resolution of the cooling load is 30 min. Therefore, a day can be divided into 94 or 48 intervals, represented by time t. The electrical load and thermal load can realize numerical prediction after 15 min, and the cooling load can realize numerical prediction after 30 min.

A. Data Processing

To guarantee the integrity of data sequence, the missing value is replaced by the average of the data before and after the moment. The data normalization scheme is as follows:

$$H' = \frac{H - H_{\min}}{H_{\max} - H_{\min}} \tag{17}$$

where $H_{\max}$ and $H_{\min}$ are the maximum and minimum values in the data sequence, respectively; H' is the normalized data; and H is the data before processing.

In this paper, the input data have been normalized by the MinMaxScaler. The threshold of the Pearson correlation coefficient from 0.1 to 0.7 is tested for STLP. The simulation results prove that the prediction accuracy of the proposed model is the highest when the threshold is 0.29. The Pearson correlation coefficients of multiple influencing factors and the predicted objects are shown in Fig. 5, and the input variables of various load predictions are as follows.

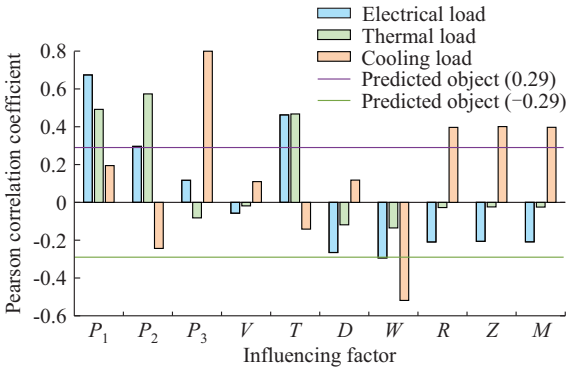

Fig. 5. Pearson correlation coefficients of influencing factors and predicted objects.

1) Supposing the electrical load at time t is to be predicted, the input variables are: the temperature of the day T, the weather of the day W, the electrical load at 15 min ago $P_1(t-1)$, the electrical load at 30 min ago $P_1(t-2)$, the electrical load at 45 min ago $P_1(t-3)$, the average electrical load of the previous day $P_{1,\text{ave}}$, the electrical load from the same time t of the previous day $P_1'(t)$, the thermal load of the day at 15 min ago $P_2(t-1)$, the thermal load of the day at 30 min ago $P_2(t-2)$, the thermal load of the day at 45 min ago $P_2(t-3)$, the average thermal load of the previous day $P_{2,\text{ave}}$, and the thermal load from the same time t of the previous day $P_2'(t)$. When electrical load prediction is conducted, the dimensions of the input layer are 12.

2) Supposing the thermal load at time t is to be predicted, the input variables are: T, $P_1(t-1)$, $P_1(t-2)$, $P_1(t-3)$, $P_{1,\text{ave}}$, $P_1'(t)$, $P_2(t-1)$, $P_2(t-2)$, $P_2(t-3)$, $P_{2,\text{ave}}$, and $P_2'(t)$. When thermal load prediction is conducted, the dimensions of the input layer are 11.

3) Supposing the cooling load at time t is to be predicted, the input variables are: W, R, Z, M, the cooling load at 15 min ago $P_3(t-1)$, the cooling load at 30 min ago $P_3(t-2)$, the cooling load at 45 min ago $P_3(t-3)$, the average cooling load of the previous day $P_{3,\text{ave}}$, and the cooling load from the same time t of the previous day $P_3'(t)$. When cooling load prediction is conducted, the dimensions of the input layer are 9.

B. Evaluating Indicators

The MAPE and weighted mean accuracy (WMA) are se-

lected as the evaluation indexes. The MAPE can quantify the accuracy of the prediction method in predicting a certain type of IES load, and the WMA can comprehensively evaluate the accuracy of prediction methods in IES. The value of WMA is calculated by:

$$W_{\mathrm{MA}}=\alpha_e(1-M_{\mathrm{APE},e})+\alpha_t(1-M_{\mathrm{APE},t})+\alpha_c(1-M_{\mathrm{APE},c}) \quad (18)$$

where α_e, α_t, and α_c are the weights of electrical, thermal, and cooling loads, respectively; and $M_{\mathrm{APE},e}$, $M_{\mathrm{APE},t}$, and $M_{\mathrm{APE},c}$ are the MAPEs of electrical, thermal, and cooling loads, respectively.

The self-learning interval analytic hierarchical process in [28] is adopted to calculate the weights. N experienced engineers are invited to score the importance intervals of different loads to obtain the interval weight of each load by the interval analytic hierarchical process. In addition, taking the maximum entropy criterion as the optimization goal, the simulated annealing algorithm is employed to convert the interval weights into the constant weights. By completing the above process, α_e, α_h, and α_c are set to be 0.4, 0.3, and 0.3, respectively.

C. Hyper-parameter Setting

The value range of the particle position is set to be [−5, 5], and the maximum absolute velocity of the particle is 5. The control parameter μ is set to be 4. The value range of the inertia weight ω is set to be [0.4, 0.9], the initial value of the inertia weight ω is set to be 0.65, and the iteration number of IPSO is set to be 1000. To obtain the best prediction effectiveness, the appropriate hyper-parameter setting of the proposed model is significant. The data of three types of load from January 2016 to December 2016 are selected as the hyper-parameter validating data of the proposed model. The average WMA of prediction results is employed to quantify the effectiveness of the hyper-parameter setting.

1) Number of Hidden Layer Neurons n and Number of IPSO Particles N_{size}

To limit the number of parameter variables, the COA is temporarily eliminated in the proposed model, and the parameter settings of the IPSO-WNN are studied first. To avoid affecting the optimization speed of particles, the learning factors a and b are set to be 1. Therefore, the parameter settings of n and N_{size} are discussed.

Hidden layer neurons are a significant and sensitive part of the WNN structure. Too few neurons in the hidden layers can lead to under-fitting, while too many neurons may result in over-fitting. Besides, the particle number is a key factor affecting the efficiency and speed of the IPSO. A large number of particles will slow down the calculation speed, while too few particles will result in a local optimum. Therefore, the choice of n and N_{size} is a significant challenge. The WMA with different n and N_{size} is shown in Fig. 6. It can be observed that when $n=12$ and $N_{\mathrm{size}}=24$, the IPSO-WNN has the highest operational efficiency and prediction accuracy.

2) Learning Factors c_1 and c_2

After $n=12$ and $N_{\mathrm{size}}=24$ are determined, the parameter settings of c_1 and c_2 are discussed. As demonstrated in [23], increasing c_1 encourages the exploration of the solution space as each particle moves toward the respective P_{best}; and increasing c_2 encourages the exploitation of the supposed global optimum. The WMA with different c_1 and c_2 is shown in Fig. 7.

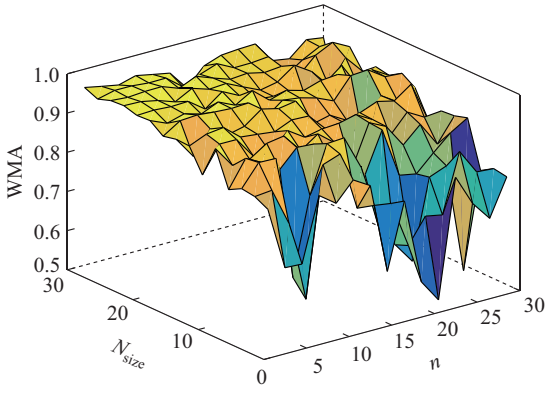

Fig. 6. WMA with different n and N_{size}.

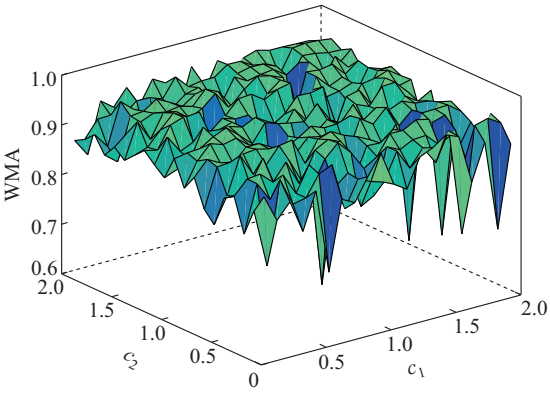

Fig. 7. WMA with different c_1 and c_2.

Due to the randomness in the optimization process of IPSO and WNN, the prediction under different parameters fluctuates significantly. The experiments found that when $c_1=1.5$ and $c_2=0.8$, the optimization effectiveness is the best.

3) Iteration Number of COA C_{itr} and Adjustment Coefficient α

Based on the determined IPSO-WNN parameters, the COA parameters are studied. The COA is incorporated into the iterative process in the IPSO to form the proposed model. When $n=12$, $N_{\mathrm{size}}=24$, $c_1=1.5$, and $c_2=0.8$, the parameter settings of C_{itr} and α are discussed. For a given P_{best}, the COA with different parameters is adopted to optimize P_{best} in an iteration of IPSO. The WMA with different C_{itr} and α is shown in Fig. 8.

Increasing C_{itr} will improve the optimization accuracy with increased costs of data processing. A large α will reduce the accuracy of the secondary carrier wave of COA, while a small α will reduce the optimization efficiency. To meet the prediction accuracy and efficiency requirements simultaneously, α and C_{itr} are selected as 0.0013 and 700, respectively.

D. Result Analysis

In this subsection, the prediction accuracies of the ANN,

WNN, PSO-WNN, and proposed model are compared.

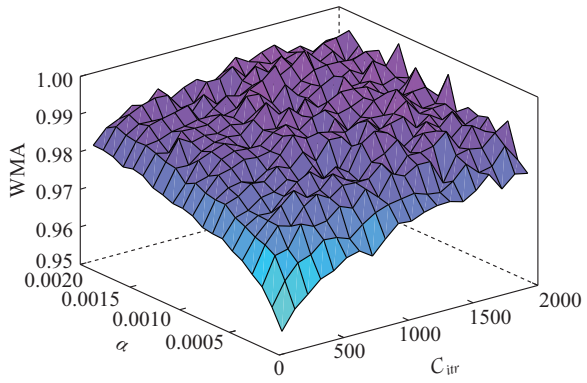

Fig. 8. WMA with different C_{itr} and α.

The data of three types of load from January 2018 to December 2018 are selected as the testing data of the proposed model. Since the load prediction accuracy of the same season is similar, the prediction results in four typical days (January 15, April 15, July 15, and October 15) are shown in Fig. 9, including the electrical, thermal, and cooling loads. January, April, July, and October are used to represent the winter, spring, summer, and autumn, respectively. The following conclusions can be drawn.

1) The prediction accuracy of three types of load in peak and valley periods is lower than those in other periods.

2) Compared with the thermal load and electrical load, the prediction accuracy of the cooling load is lower, the regularity is lower, and the fluctuation during the day is more frequent. The prediction errors of the cooling load are all within acceptable limits.

3) Due to the occasional large deviation of prediction, the accuracy of the ANN in predicting load is unstable. The WNN strengthens the non-linear optimization with a more accurate approximation of the results based on ANN. Although PSO-WNN further improves the prediction performance of WNN, there are still large deviations sometimes. In general, the prediction performance of the proposed model is relatively stable, which is better than that of the ANN, PSO-WNN, and WNN.

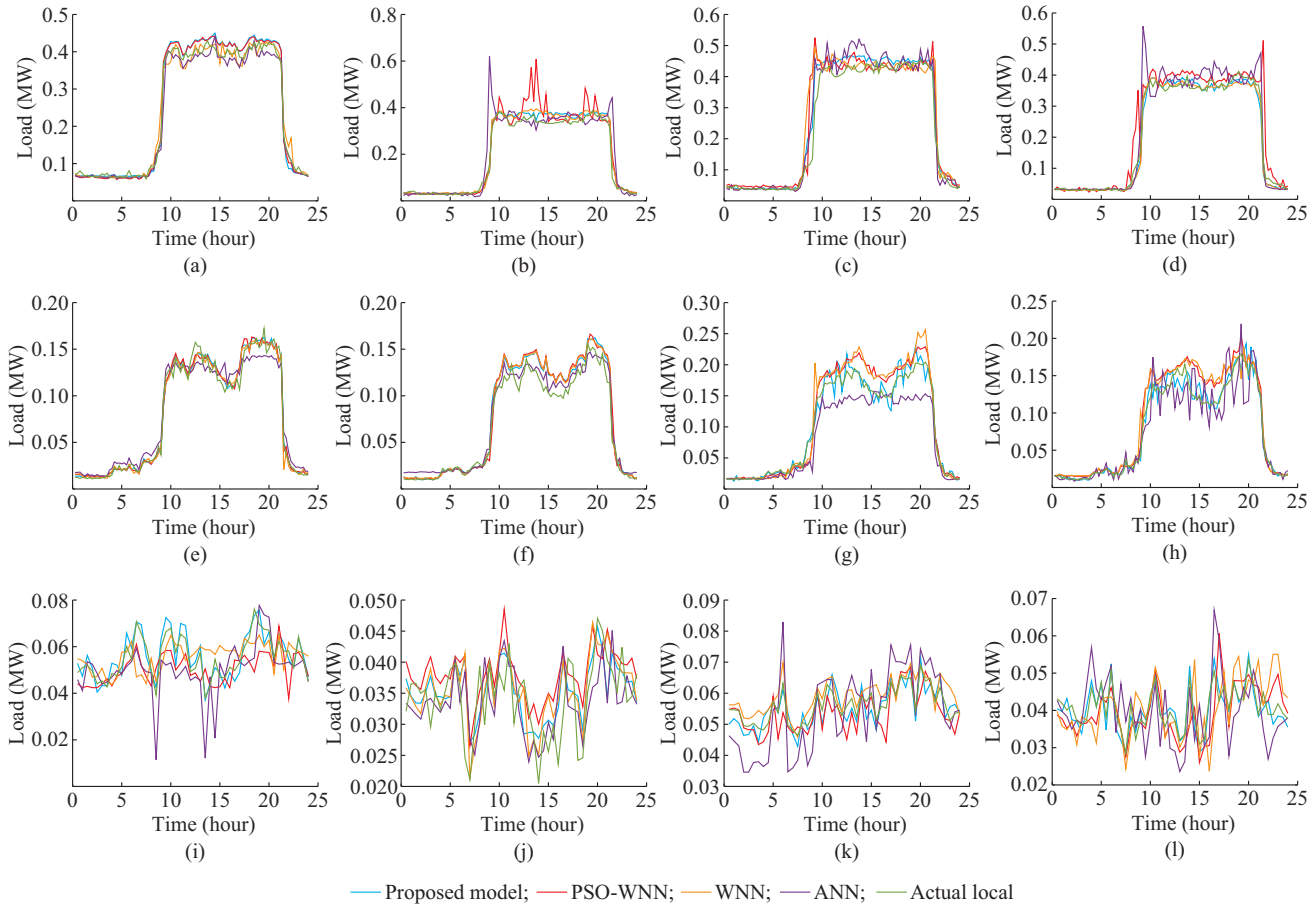

Fig. 9. Prediction results of different prediction models. (a) Electrical load prediction in January. (b) Electrical load prediction in April. (c) Electrical load prediction in July. (d) Electrical load prediction in October. (e) Thermal load prediction in January. (f) Thermal load prediction in April. (g) Thermal load prediction in July. (h) Thermal load prediction in October. (i) Cooling load prediction in January. (j) Cooling load prediction in April. (k) Cooling load prediction in July. (l) Cooling load prediction in July.

4) Due to temperature changes, the types of main electrical equipment have changed, and the loads in winter and summer have apparent regularity and periodicity. Compared with April (spring) and October (fall), the prediction accuracy in January (winter) and July (summer) increases by 2.80%.

The MAPE is used to compare and analyze the prediction

accuracy of different models, and its exact values are shown in Fig. 10 and Table I. The conclusions can be drawn as follows.

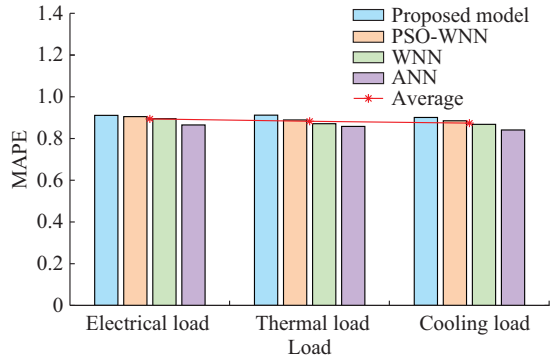

Fig. 10. MAPEs of different prediction models.

1) Due to the lack of wavelet analysis, the prediction accuracy of ANN is the lowest among all models. Compared with WNN, the proposed model and PSO-WNN increase the prediction accuracies by 14.2% and 7.1%, respectively. The introduction of new swarm intelligence algorithms can improve the prediction accuracy of WNN.

2) Compared with the other three models, the proposed model has the highest prediction accuracy of 90.8%. It is verified that the introduction of the COA and dynamic inertia weight selection strategy in IPSO can overcome the poor convergence and the tendency to fall into the local optimum. The accuracy of the STLP of IES has been effectively improved by the proposed model.

1) Efficiency Comparison

The operation efficiency of prediction models is also a key aspect that needs to be considered in actual engineering. In each iteration, the computational efficiency and the prediction time of different models can be effectively compared by observing the decline of fitness. The prediction data of electrical load in spring are selected as the testing data. Through 100 consecutive experiments, the average prediction time of different prediction models is shown in Table II. The declining trend of fitness of G_{best} during the iterations of IPSO is shown in Fig. 11.

TABLE I
MAPEs of Different Prediction Models

Load	Model	MAPE												Average
		January	February	March	April	May	June	July	August	September	October	November	December	
Electrical	Proposed	0.069	0.064	0.072	0.100	0.119	0.086	0.091	0.089	0.099	0.094	0.087	0.097	0.089
	PSO-WNN	0.073	0.070	0.066	0.125	0.121	0.119	0.123	0.081	0.102	0.139	0.112	0.011	0.095
	WNN	0.066	0.067	0.074	0.110	0.108	0.129	0.123	0.132	0.101	0.102	0.131	0.118	0.105
	ANN	0.070	0.084	0.121	0.208	0.107	0.182	0.143	0.091	0.138	0.143	0.178	0.154	0.135
Thermal	Proposed	0.075	0.089	0.077	0.110	0.089	0.092	0.077	0.078	0.094	0.079	0.083	0.108	0.088
	PSO-WNN	0.070	0.101	0.087	0.149	0.131	0.108	0.132	0.089	0.113	0.118	0.098	0.132	0.111
	WNN	0.057	0.101	0.087	0.183	0.172	0.082	0.160	0.152	0.118	0.125	0.148	0.156	0.129
	ANN	0.149	0.129	0.137	0.147	0.138	0.101	0.171	0.082	0.156	0.203	0.118	0.179	0.142
Cooling	Proposed	0.077	0.098	0.087	0.114	0.119	0.119	0.076	0.109	0.089	0.079	0.119	0.098	0.099
	PSO-WNN	0.121	0.129	0.110	0.167	0.127	0.086	0.088	0.118	0.109	0.106	0.118	0.096	0.115
	WNN	0.161	0.099	0.145	0.135	0.135	0.119	0.087	0.171	0.181	0.146	0.108	0.099	0.132
	ANN	0.143	0.157	0.162	0.154	0.123	0.176	0.201	0.211	0.172	0.165	0.119	0.129	0.159

TABLE II
Average Prediction Time of Different Prediction Models

Model	Prediction time (s)	Model	Prediction time (s)
ANN	229	PSO-WNN	220
WNN	275	Proposed	256

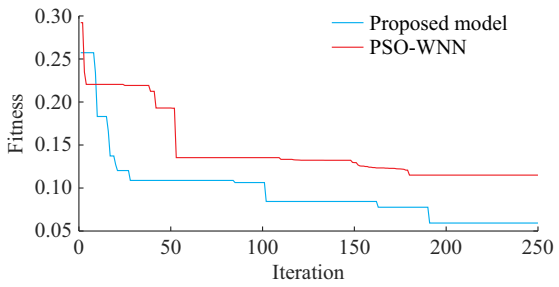

Fig. 11. Declining trend of fitness of G_{best} during iterations of IPSO.

The conclusions drawn from Table II and Fig. 11 are as follows.

1) When accuracy is similar, the prediction of the WNN is the slowest, and that of the PSO-WNN is the fastest. The introduction of the PSO algorithm improves the operation efficiency of the WNN.

2) Although the COA increases the complexity of the prediction model, it does not significantly reduce the prediction speed of the proposed model due to the dynamic inertia weight. Meanwhile, there is a 6.29% improvement in the prediction accuracy.

3) As shown in Fig. 11, the prediction speed of the proposed model is faster than that of the PSO-WNN. The update frequency of G_{best} in the proposed model is higher, achieving better optimization effectiveness with fewer iterations than that of the PSO-WNN. Besides, the calculation speed of the proposed model can be increased by reducing the iteration number of IPSO while ensuring good prediction accuracy.

VI. Conclusion

To overcome the poor convergence and the tendency to fall into local optimum of the traditional WNN for STLP of IES, the WNN model based on the IPSO and COA is put forward for the STLP of IES in this paper. The Pearson correlation coefficient is used to select appropriate influencing factors as the prediction input. The dynamic inertia weight strategy is incorporated into the PSO. The IPSO-COA is used to optimize the connection weights and wavelet parameters in the WNN, and the training errors are referenced for feedback correction. In the experimental simulation and analysis, the proposed model, PSO-WNN, WNN, and ANN are compared in the STLP of IES. The simulation results prove that the proposed model has higher prediction accuracy and operation efficiency and overcomes the shortcomings of the other prediction models.

References

[1] S. Ren, X. Dou, Z. Wang *et al.*, "Medium- and long-term integrated demand response of integrated energy system based on system dynamics," *Energies*, vol. 13, no. 3, pp. 1-24, Feb. 2020.

[2] B. Tang, G. Gao, X. Xia *et al.*, "Integrated energy system configuration optimization for multi-zone heat-supply network interaction," *Energies*, vol. 11, no. 11, pp. 1-18, Nov. 2018.

[3] J. He, Y. Li, H. Li *et al.*, "Application of game theory in integrated energy system systems: a review," *IEEE Access*, vol. 8, pp. 93380-93397, May 2020.

[4] Y. Wang, Y. Wang, Y. Huang *et al.*, "Optimal scheduling of the regional integrated energy system considering economy and environment," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 4, pp. 1939-1949, Oct. 2019.

[5] L. Ge, Y. Li, Y. Xian *et al.*, "A FA-GWO-GRNN method for short-term photovoltaic output prediction," in *Proceedings of 2020 IEEE PES General Meeting*, Montreal, Canada, Aug. 2020, pp. 1-5.

[6] H. Dong, Y. Gao, X. Meng *et al.*, "A multifactorial short-term load forecasting model combined with periodic and non-periodic features–a case study of Qingdao, China," *IEEE Access*, vol. 8, pp. 67416-67425, Apr. 2020.

[7] J. W. Taylor, "Short-term load forecasting with exponentially weighted methods," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 458-464, Feb. 2012.

[8] C. M. Lee and C. N. Ko, "Short-term load forecasting using lifting scheme and ARIMA models," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5902-5911, May 2011.

[9] H.-W. Jung, K.-B. Song, J.-D. Park *et al.*, "Very short-term electric load forecasting for real-time power system operation," *Journal of Electrical Engineering & Technology*, vol. 13, no. 4, pp. 1419-1424, Jul. 2018.

[10] Q. Liu, Z. Huang, and S. Li, "Research on power load forecasting based on support vector machine," *Journal of the Balkan Tribological Association*, vol. 22, no. 1, pp. 151-159, Jan. 2016.

[11] B. Wang, L. Zhang, H. Ma *et al.*, "Parallel LSTM-based regional integrated energy system multienergy source-load information interactive energy prediction," *Complexity*, vol. 2019, pp. 1-13, Nov. 2019.

[12] J. Yang, T. Chai, C. Luo *et al.*, "Intelligent demand forecasting of smelting process using data-driven and mechanism model," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 12, pp. 9745-9755, Dec. 2019.

[13] Y. Li, X. Yuan, J. Xu *et al.*, "Medium-term forecasting of cold, electric and gas load in multi-energy system based on VAR model," in *Proceedings of the 2018 13th IEEE Conference on Industrial Electronics and Applications*, Wuhan, China, May 2018, pp. 1676-1680.

[14] Z. Xiao, Z. Sun, L. Song *et al.*, "Simulated annealing-wavelet neural network for vibration fault diagnosis of hydro-turbine generating unit," *Journal of Optoelectronics and Advanced Materials*, vol. 17, no. 5-6, pp. 734-740, May 2015.

[15] R. Billinton and E. Khan, "A security based approach to composite power system reliability evaluation," *IEEE Transactions on Power Systems*, vol. 7, no. 1, pp. 65-72, Feb. 1992.

[16] R. Fang, "Life cycle cost assessment of wind power-hydrogen coupled integrated energy system," *International Journal of Hydrogen Energy*, vol. 44, no. 56, pp. 29399-29408, Nov. 2019.

[17] J. Cheng, X. Wang, T. Si *et al.*, "Ignition temperature and activation energy of power coal blends predicted with back-propagation neural network models," *Fuel*, vol. 173, pp. 230-238, Jun. 2016.

[18] H. Xu and Y. Deng, "Dependent evidence combination based on Shearman coefficient and Pearson coefficient," *IEEE Access*, vol. 6, pp. 11634-11640, Dec. 2017.

[19] W. Huang, S. K. Oh, and W. Pedrycz, "Hybrid fuzzy wavelet neural networks architecture based on polynomial neural networks and fuzzy set/relation inference-based wavelet neurons," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 8, pp. 3452-3462, Aug. 2018.

[20] F. Lin, R. Wai, and M. Chen, "Wavelet neural network control for linear ultrasonic motor drive via adaptive sliding-mode technique," *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, vol. 50, no. 6, pp. 686-698, Jun. 2003.

[21] P. Wolfs, K. Emami, Y. Lin *et al.*, "Load forecasting for diurnal management of community battery systems," *Journal of Modern Power Systems and Clean Energy*, vol. 6, no. 2, pp. 215-222, Mar. 2018.

[22] S. Admasie, S. B. A. Bukhari, T. Gush *et al.*, "Intelligent islanding detection of multi-distributed generation using artificial neural network based on intrinsic mode function feature," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 3, pp. 511-520, May 2020.

[23] J. Robinson and Y. Rahmat-Samii, "Particle swarm optimization in electromagnetics," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 2, pp. 397-407, Feb. 2004.

[24] F. Wu and L. Yang, "Simulation of the integrated energy system for isolated island," in *Proceedings of 2017 China International Electrical and Energy Conference (CIEEC 2017)*, Beijing, China, Oct. 2017, pp. 527-531.

[25] S. M. Arif, A. Hussain, T. T. Lie *et al.*, "Analytical hybrid particle swarm optimization algorithm for optimal siting and sizing of distributed generation in smart grid," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 6, pp. 1221-1230, Nov. 2020.

[26] B. Liu, L. Wang, Y. Jin *et al.*, "Improved particle swarm optimization combined with chaos," *Chaos, Solitons & Fractals*, vol. 25, no. 5, pp. 1261-1271, Sept. 2005.

[27] L. Ge, Y. Li, S. Li *et al.*, "Evaluation of the situational awareness effects for smart distribution networks under the novel design of indicator framework and hybrid weighting method," *Frontiers in Energy*, vol. 15, no. 1, pp. 143-158, Mar. 2021.

[28] L. Ge, Y. Li, X. Zhu *et al.*, "An evaluation system for HVDC protection systems by a novel indicator framework and a self-learning combination method," *IEEE Access*, vol. 8, pp. 152053-152070, Aug. 2020.

Leijiao Ge received the bachelor degree in electrical engineering and its automation from Beihua University, Jilin, China, in 2006, and the master degree in electrical engineering from Hebei University of Technology, Tianjin, China, in 2009, and the Ph.D. degree in electrical engineering from Tianjin University, Tianjin, China, in 2016. He is currently an Associate Professor in the School of Electrical and Information Engineering at Tianjin University. His main research interests include smart distribution networks, cloud computing, and big data.

Yuanliang Li received the B.E. degree in electrical engineering and automation from China University of Petroleum, Qingdao, China, in 2016. Currently, he is pursuing the M.S. degree in electrical engineering from Tianjin University, Tianjin, China. His main research interests include situational awareness of smart distribution networks and big data.

Jun Yan received the B.Eng. degree in information and communication engineering from Zhejiang University, Hangzhou, China, in 2011, and the M. Sc. and Ph.D. (with excellence in doctoral research) degrees in electrical engineering from the University of Rhode Island, Kingston, USA, in 2013 and 2017, respectively. Currently, he is an Assistant Professor at the Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Canada. He was a recipient of the IEEE ICC Best Paper Award (2014) and the IEEE WCCI Best Student Paper Award (2016). His research interests include computational intelligence and cyber-physical security with applications in smart grids, smart cities, and other smart critical infrastructures.

Yuqian Wang received the M.S degree in control engineering at Tianjin University, Tianjin, China. She is currently a member of the Sany Heavy Machinery Co., Ltd., Suzhou, China. Her research interest includes smart grid.

Na Zhang received Ph.D. degree in electrical engineering from Tianjin University, Tianjin, China, in 2014. She is currently an Associate Professor in Department of Electrical Engineering at Inner Mongolia University of Technology, Huhhot, China. Her research interests include distributed generation and smart distribution system.