Adaptive IES Load Forecasting Method Based on the Octopus Model

Na Zhang*, Xiao Pan, Yihe Wang, Mingli Zhang, Mengzeng Cheng and Wenying Shang

Economic and Technological Research Institute of State Grid Liaoning Electric Power Co., Ltd, Shenyang, China

Improving the accuracy and speed of integrated energy system load forecasting is a great significance for improving the real-time scheduling and optimized operation of the integrated energy system. In order to achieve rapid and accurate forecasting of the integrated energy system, this paper proposes an adaptive integrate energy system (IES) load forecasting method based on the octopus model. This method uses long short-term memory (LSTM), support vector machines (SVMs), restricted Boltzmann machines (RBMs), and Elman neural network as the octopus model quadrupeds. Through taking over differences in different data and training principles and utilizing the advantages of the octopus quadruped model, a special octopus-head and XGBoost algorithm were adopted to set the weight of the octopus' quadruped and prevent local minimum points in the model. We train the octopus model through RMSProp adaptive learning algorithm, constrain the learning rate, get the best parameters, and improve the model's adaptability to different types of data. In addition, for the incomplete comprehensive energy load data, the generative confrontation network is used to fill it. The simulation results show that compared with other prediction methods, the effectiveness and feasibility of the method proposed in this paper are verified.

Keywords: comprehensive energy, generative confrontation network, XGBoost algorithm, RMSProp adaptive learning, octopus model

INTRODUCTION

An integrate energy system (IES) (AlDahoul et al., 2021) refers to the reasonable distribution and utilization of electricity, gas, heat, wind, and other energies within a certain range or area, using professional technology and operating modes, to achieve energy efficiency, interaction, and complementarity. While ensuring that the energy consumption of individual users meets their needs, it also improves the energy consumption of the entire range and improves the overall energy utilization efficiency.

At present, the scale of our country's integrated energy system continues to expand, and the load forecast of the integrated energy system affects the operation and planning of the integrated energy system. Improving the speed and accuracy of load forecasting of the integrated energy system is of vital importance for realizing the real-time scheduling and optimized operation of the integrated energy system. In recent years, scholars at home and abroad have conducted some research on load forecasting; Bian et al. (2020) built gray correlation analysis under the LSTM (Dai and Zhao, 2020) neural network model, improved the traditional LSTM neural network's processing method of time series and nonlinear data, and improved the accuracy of short-term load forecasting, but did not consider the comprehensive energy relationship. Daniel et al. (2020) used the packet decomposition and cyclic neural network to decompose the electric, cold, and heat loads in frequency bands and determined the prediction method by judging the correlation between each frequency band. This

method can accurately predict the loads with strong and weak autocorrelation, but does not consider the impact on incomplete data onto the prediction accuracy. Dong et al. (2021) used Copula theory to analyze the correlation between the comprehensive energy load, established a time series, and used the K-means clustering algorithm to design a radial basis function neural network. The calculation method of the model is simple and easy to design, but the prediction accuracy of some outlier data points and isolated data points is low. Guo et al. (2021) considered that the training process of the traditional wavelet neural network (WNN) (Guo et al., 2020) is prone to the shortcoming of too fast convergence speed and proposed a WNN prediction method based on improved particle swarm optimization (IPSO). On the basis of traditional particle swarm optimization (PSO), chaos algorithm is added, and chaos algorithm is more random and more general and has deeper search ability to improve the overall prediction accuracy and prediction speed of the wavelet neural network. However, this method needs to continuously optimize the weights and parameters, and the model establishment is too complicated, so the speed slows down during the data training process. Jamal et al. (2020) and Khan et al. (2020) designed a multivariable phase space reconstruction Kalman filter method, which fully considered the coupling relationship between various energy sources, and used the five-step parameter trend method to dilute the influence of the old parameters on the current load forecasting. A larger training set was used to train the prediction models to improve the prediction accuracy. Because the loss function is set to a convex function, this does not guarantee that the global optimal solution can be achieved when the optimized nonconvex function is achieved. The final result will be greatly affected by the initial value of the parameter. It takes a very long time to calculate the loss function of all training data at the same time, so that the prediction accuracy may get a suboptimal solution. Matrenin et al. (2020) adopted the multitask learning method of deep structure (Pitchforth et al., 2021) to predict the complex energy load of the park type, using a combination of offline and online, but due to the fixed learning rate adopted by this network, it may cause network oscillations and make the speed of convergence slower and the prediction accuracy lower, and the optimal value is also not reached. Verma et al. (2021) solved the problem of overfitting and limited generalization ability of a single model, but they could not solve the application limitation of a single algorithm. The work of Wang et al. (2019) is composed of a variety of heterogeneous models, which overcome the shortcomings of the application limitation of a single algorithm, but the training set of each model is the same, and there are still problems of single model overfitting and limited generalization ability. Although, artificial intelligence, neural networks, support vector machines, and deep learning methods have made great progress of power system prediction. However, the abovementioned models have their specific application scopes, and they are less involved in the field of integrated energy.

On the basis of the abovementioned research, this paper proposes an adaptive IES load forecasting method based on the octopus model. First, this paper uses a generative

FIGURE 1 | Block diagram of the octopus model adaptive IES load forecasting method.
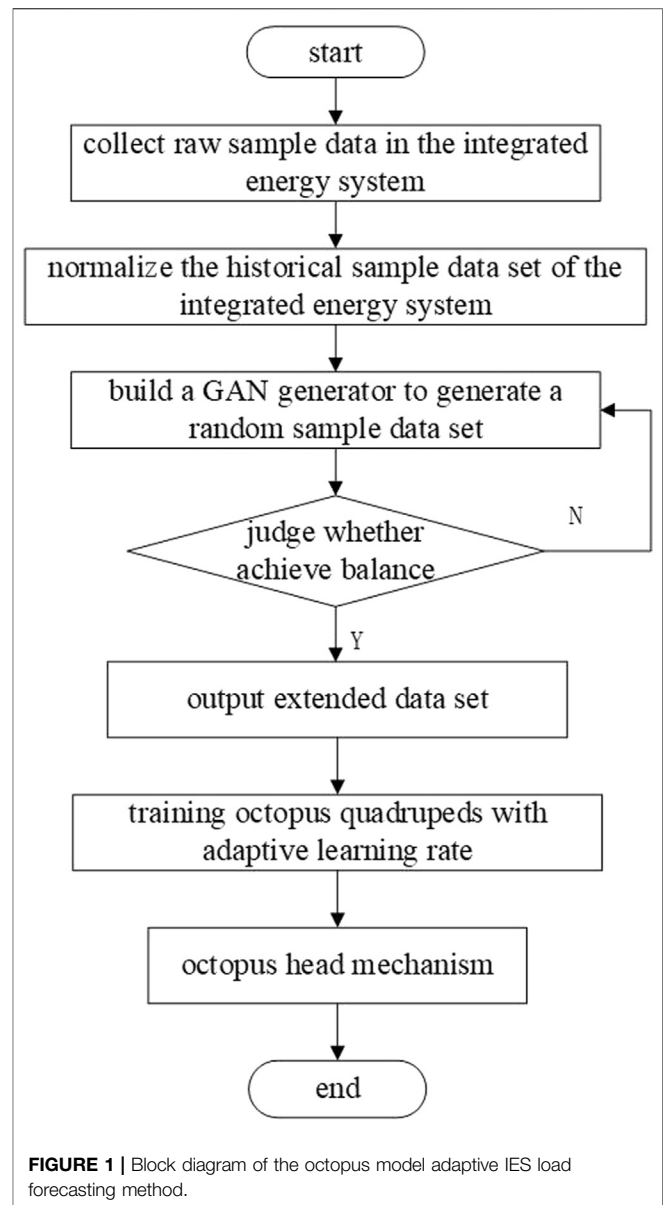
adversarial network (GAN) to supplement the incomplete data in the integrated energy system, thereby reducing the data-induced inadequate forecasting accuracy. The problem is that the head mechanism of the octopus model is used to change the weight of the octopus' foot, according to the prediction accuracy to prevent local minimum points and improve the prediction accuracy. Then, the RMSProp algorithm is used to train the octopus model, and the adaptive learning rate is used to obtain the optimal parameters, which not only improves the model's adaptability to different types of data but also improves the prediction speed. Finally, it is based on the operation data onto the integrated energy system of a residential district in Shenyang. A simulation analysis of the algorithm proposed to this paper is carried out. The prediction results show that, in the comprehensive energy system load forecasting, the octopus model using RMSProp algorithm, XGBoost algorithm, and

FIGURE 2 | Principle of the generative adversarial network.
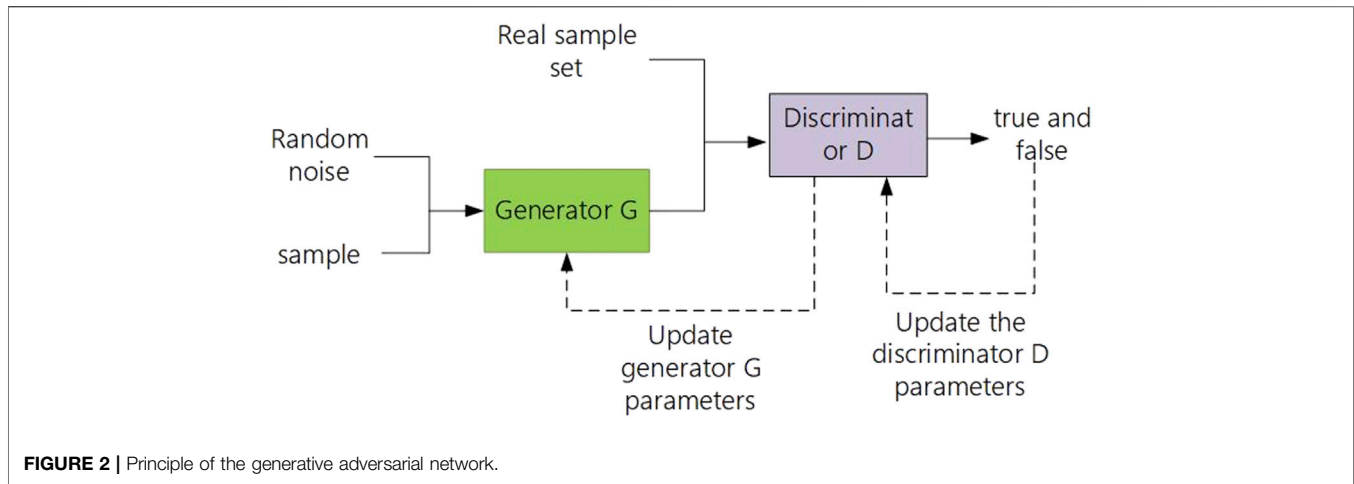
GAN has good prediction accuracy and prediction speed and can better solve the problem that traditional neural network prediction models tend to fall into local optimal solutions and shortcomings such as limited application of predictive models.

OCTOPUS MODEL ADAPTIVE INTEGRATE ENERGY SYSTEM PREDICTION METHOD

This paper proposes an adaptive long and short-term IES load forecasting method based on the octopus model. First, we collect the original sample data and normalize the sample data. Then, the processed sample dataset is expanded by GAN. When the generator and the discriminator reach the Nash equilibrium, the expanded sample dataset can be generated from the incomplete sample dataset. We divide the extended sample dataset into six subdatasets according to the time dimension and ensure that each subdataset does not overlap with each other. Finally, we build an octopus model. Three of the six subdatasets are used as the training set, one is used as the test set, and the remaining 2 validation sets are input into the octopus model quadrupeds, and the octopus quadruped prediction results are integrated by weighted average through the octopus-head mechanism, so as to predict the electrical load, air load, and thermal load. The load forecasting model used in this paper can reduce the time for selecting the network model in the early stage and, at the same time, improve the accuracy of load forecasting. The overall idea of the forecasting method is shown in **Figure 1**.

GENERATIVE ADVERSARIAL NETWORK

Generative adversarial networks (GANs) consist of two parts, the generator and the discriminator. The generator uses the processing of random noise to generate pseudosamples, and the pseudosample value is similar to the real sample value. The discriminator compares the real sample value of the fake sample value generated by the generator, distinguishes the difference between the fake sample and the real value, and

improves the recognition ability of its own network. The two realize the learning optimization processes through the game and finally reach a Nash equilibrium. The GAN schematic diagram is shown in **Figure 2**.

The generated data are generated by the original dataset training, so the generated data are similar to the original data. The data expanded by GAN will reduce the isolated points and discrete points of the original data. Using the data enhanced by GAN for load prediction will make the prediction accuracy higher. Since GAN already has good data generation capabilities, the use of extended data has little effect on the prediction results.

OCTOPUS FOUR-LEGGED MODEL

Long Short-Term Memory Principle

LSTM is a model that uses back propagation time to train a neural network. It is not a neuron, but a unit connected by layers. LSTM is a nonlinear prediction model, so it can build a larger and deeper recurrent neural network, effectively solving the problem of the disappearance of gradients in the prediction process, and is suitable for processing time series models. LSTM updates short-term memory through memory, forgets past memory information, and updates new information. But, when the sequence of continuous data becomes longer, the unfolding time step will be too long. The LSTM neural network has a long- and short-term memory structure, and LSTM is suitable for time series forecasting.

The memory unit memorizes the historical information about the sequence data together with the hidden state. The information about the memory unit is controlled by three gate units. The forget gate deletes the information about the memory unit according to h_{t-1} and x_t. The forget door is

$$f_t = \sigma\left(W_f \left[h_{t-1} \, x_t \right] \right) + b_f, \tag{1}$$

where $\sigma(\cdot)$ is the sigmoid activation function; $\boldsymbol{b}_f$ is the bias of the forgetting gate; and W_f is the weight of the forgetting gate.
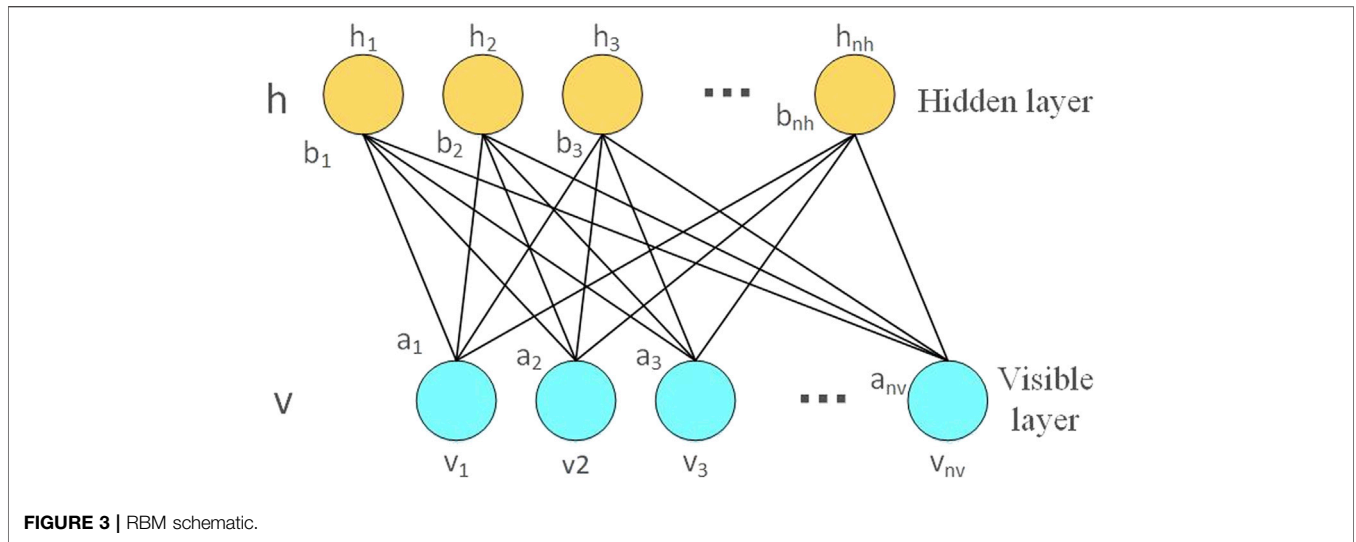
FIGURE 3 | RBM schematic.

The input gate adds information to the memory unit according to h_{t-1} and x_t, as shown in **Eqs 2** and **3**.

$$i_t = \sigma\left(W_i\left[\,h_{t-1}\ \ x_t\,\right]\right) + b_i, \tag{2}$$

$$\widehat{C}_t = \tanh\left(W_c[h_{t-1}\ x_t]\right) + b_c. \tag{3}$$

In the formula, i_t is the information that needs to be memorized; $\widehat{C}_t$ is the candidate memory unit, used to update the memory unit; W_i and W_c are the input gate weights; and b_i and b_c are the input gate biases.

After the calculation of the forget gate and the output gate is completed, the memory unit is updated using the following equation:

$$C_t = f_t \circ C_{t-1} + i_t \circ \widehat{C}_t, \tag{4}$$

where $\circ$ is the product of Hadamard.

The output gate determines h_t according to h_{t-1}, x_t, and C_t.

$$o_t = \sigma\left(W_0\left[\,h_{t-1}\ \ x_t\,\right]\right) + b_o, \tag{5}$$

$$h_t = o_t \circ \tanh C_t, \tag{6}$$

where W_o is the output gate weight; b_o is the output gate bias.

LSTM inherits the advantages of the recurrent neural network (RNN) well and has a long-term memory function. Compared with the prediction model constructed by the ordinary RNN, LSTM can solve the problems of gradient explosion and gradient disappearance. Therefore, LSTM has better performance in dealing with models that are highly correlated with time series.

Support Vector Machine Principle

SVM, as a relatively important learning method in machine learning, is based on statistical theory and supervised learning and can solve multivariate nonlinear problems well. The principle is to map the sample data one by one in a high-dimensional space, and this kind of mapping does not require a clear mapping function, so as to achieve a conversion from nonlinear to linear. Simply putting, it is to upgrade the data and linearize the data.

Compared with other linear models, SVM can better solve the dimension problem with the premise of the same computational complexity. Therefore, the SVM model requires small storage space and strong algorithm robustness. SVM is proposed by a binary classification problem. When making predictions, a prediction curve is made through linear regression. SVM predicts data through regression fitting, and it does not have memory function. Therefore, SVM can reduce the impact of discrete points and isolated points on the prediction results.

Suppose the training set samples are $\{(x_i, y_i),\ i = 1,2,3, \ldots ,l\}$, where x_i is the input column vector of the ith training sample, $x_i = [x_i^1, x_i^2, \ldots, x_i^d]^T$, and $y_i \in R$ is the corresponding output value.

Suppose the linear regression function established in the high-dimensional feature space is

$$f(x) = w\Phi(x) + b. \tag{7}$$

Among them, $\Phi(x)$ is a nonlinear mapping function.

The ε linear insensitive loss function is defined as

$$L\big(f(x), y, \varepsilon\big) = \begin{cases} 0 & |y - f(x)| \leqslant \varepsilon \\ |y - f(x) - \varepsilon| & |y - f(x)| > \varepsilon \end{cases}. \tag{8}$$

Among them, $f(x)$ is the predicted value returned by the regression function, and y is the corresponding true value.

Restricted Boltzmann Machine Principle

Boltzmann machine (RBM) is a model with a two-layer neural network, which is a probability distribution model based on energy.

The Boltzmann machine is divided into a hidden layer h and a visible layer v (that is, the input layer and the output layer). The Boltzmann machine can be regarded as a fully connected graph; that is, each neuron is fully connected with all neurons in this layer and neurons in other layers. The principle of the RBM is shown in **Figure 3**. The RBM layer is not connected and is connected to all neurons in other layers, where v_i is the visible layer neuron, h_j is the hidden layer neuron, a_i is the visible layer, b_j
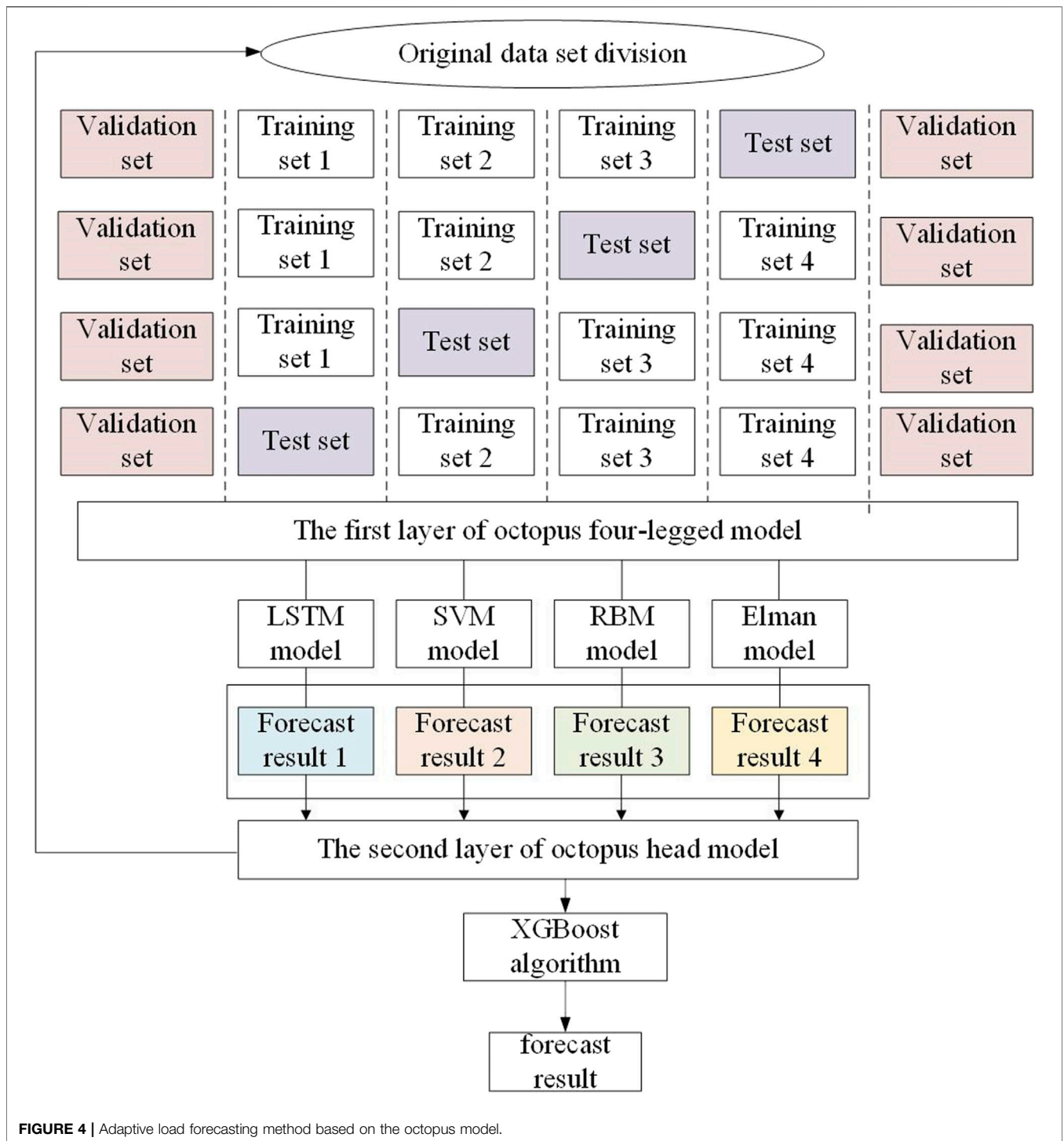
FIGURE 4 | Adaptive load forecasting method based on the octopus model.

is the hidden layer, and the W is the weight matrix. Its visible nodes are independent of other visible nodes, and there is no connection between hidden layer nodes. When the observation data are given in the visible layer, each node of the hidden layer is independent of each other. The RGM network is relatively simple and does not have a memory function. Important information entered early will be forgotten over time.

The hidden layer and visible layer of the RBM used in this paper are both binary, namely, $v_i \in \{0,1\}^n$, $h_j \in \{0,1\}^m$. The energy formula is shown as follows:

$$E(v, h) = -\sum_{i=1}^{n} \sum_{j=1}^{m} w_{ij} v_i h_j - \sum_{i=1}^{n} a_i v_i \sum_{j=1}^{m} b_j h_j. \qquad (9)$$
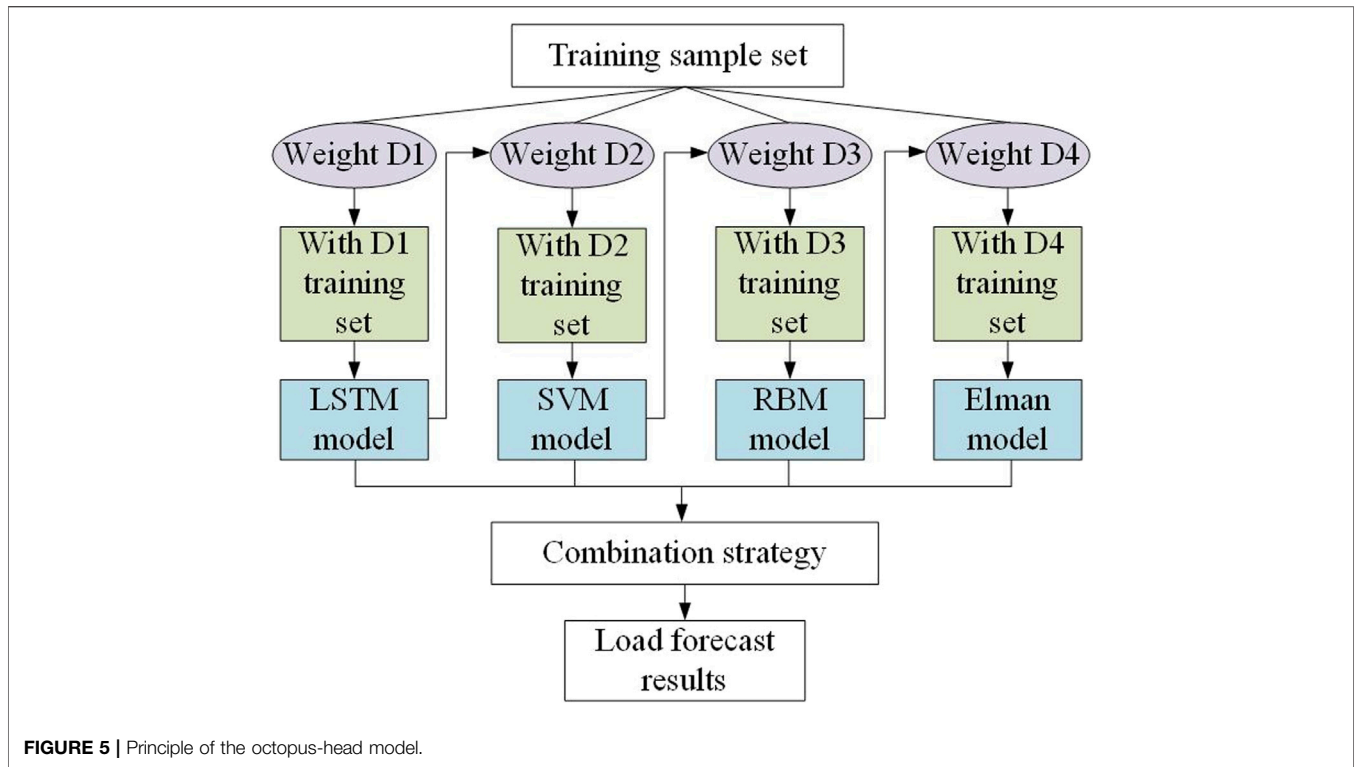
FIGURE 5 | Principle of the octopus-head model.

Among them, $w = [w_{ij}]^{n \times m}$, $a = (a_i)^n$, and $b = (b_j)^m$ constituted the parameter $\boldsymbol{\theta}$ of the RBM. m and n represent the number of visible units and winning units, v_i and h_j are the *i*th visible unit and the *j*th hidden unit, respectively, and w_{ij} is the weight between the visible unit and the hidden unit. a_i and b_j, respectively, represent the bias of the visible unit and the hidden unit.

The joint probability distribution *p* (v,h) of the visible unit and the hidden unit is defined as follows:

$$P(v, h) = \frac{1}{\sum_{v,h} e^{-E(v,h)}} e^{-E(v,h)}. \tag{10}$$

The corresponding two marginal probability distributions are

$$P(v) = \frac{1}{z} \sum_h e^{-E(r,h)}, \tag{11}$$

$$P(h) = \frac{1}{z} \sum_v e^{-E(v,h)}. \tag{12}$$

Elman Neural Network

The Elman neural network is a dynamic neural network. It adds a layer on the basis of the BP neural network, which is used as a delay factor. The Elman neural network can use, store, and feedback historical past time information. To a certain extent, the Elman neural network can perform load forecasting well. This paper uses a 4-layer Elman neural network. The connection to the output layer is similar to the feed forward network. The input layer unit only plays a role in signal transmission. The transfer function of the hidden layer unit adopts a nonlinear function. The receiving layer can be considered as a one-step delay operator.

Elman is a dynamic network, adapts to time-varying characteristics, has a short-term memory function, can internally feedback, store, and use the output information of the past moment, and is better than the BP network in terms of computing power and network stability.

In this paper, the Elman neural network of particle swarm optimization (PSO) is used. PSO is an efficient and rapid optimization method, suitable for solving continuous weights in Elman. The position and speed update formulas are

$$V_i^{k+1} = \omega \cdot V_i^k + c_1 \cdot rand_1 \cdot (P_{best} - X_i) + c_2 \cdot rand_2 \cdot (G_{best} - X_i), \tag{13}$$

$$X_i^{k+1} = X_i^k + V_i^{k+1}, \tag{14}$$

where V_i^k and X_i^k are the velocity and position of particle *i* at the *k*th iteration, respectively. $rand_1$ and $rand_2$ are randomly generated in (-1,1), c_1 and c_2 are the learning factors, and ω is the particle weight coefficient. G_{best} is the global optimum, and P_{best} is the individual optimum.

ADAPTIVE LEARNING

The traditional training method can reach the global optimal solution only when the loss function is convex. For the concave function, it cannot ensure that the training of the neural network will definitely reach the global optimum. On the contrary, the loss function can easily reach the local optimum. At the same time, the traditional training process takes a long time, and the training result is greatly affected by the initial value.

FIGURE 6 | Comprehensive energy long-term forecast results.
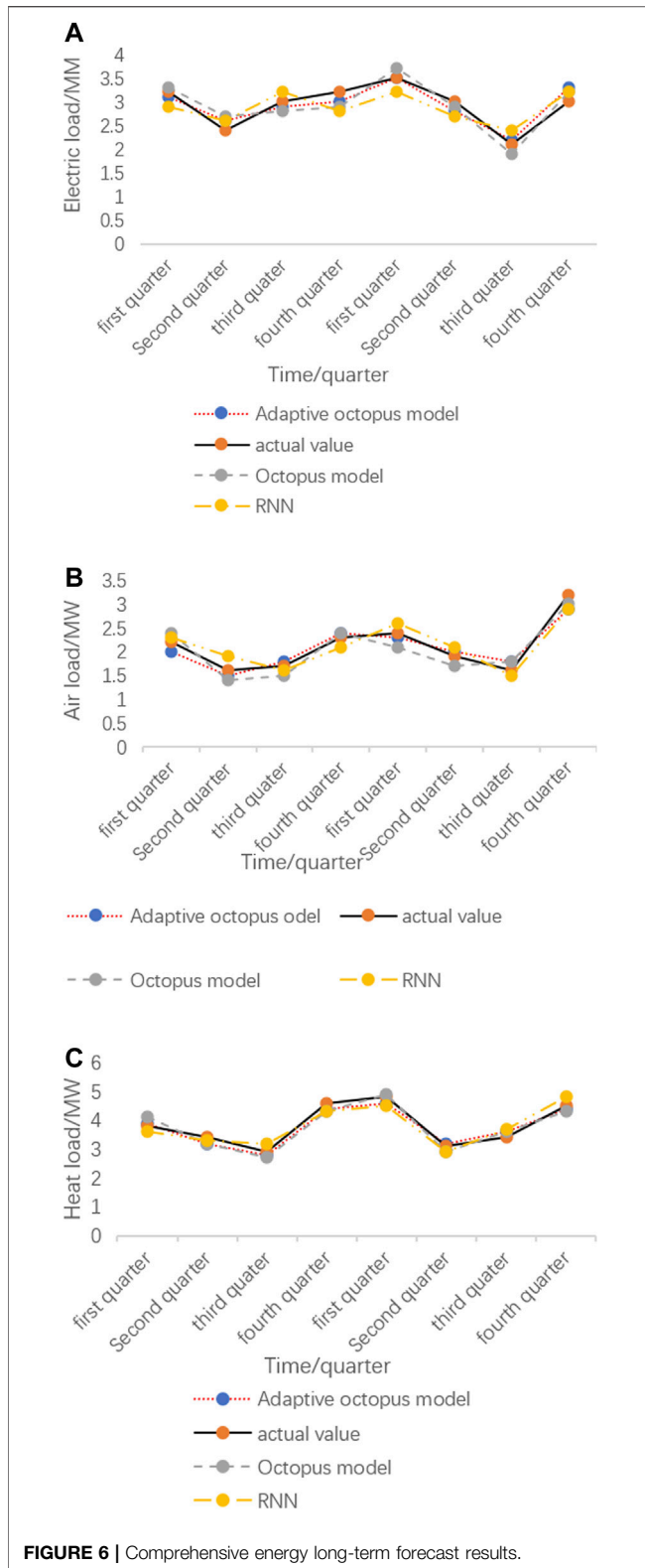
Aiming at the shortcomings of traditional training methods, the octopus model structure uses an adaptive learning rate algorithm to train four prediction models to

obtain the best parameters of the octopus model. Taking into account that different parameters require different learning rates to be adjusted, if the learning rate is too large, it will cause the training of some parameters to miss the global optimum and reach the local optimum. If the learning rate is too small, the convergence speed of the parameters will be slow, and the training process will take a long time. Therefore, it is particularly important to set different learning rates for different parameters in training of the model.

The RMSProp optimization algorithm can improve the prediction accuracy of the octopus model and make the octopus model more adaptive. Compared with the AdaGrad optimization algorithm, the RMSProp algorithm uses a new exponential decay algorithm, which reduces the impact on historical data. At the same time, RMSProp introduces a new parameter ρ, which can be expressed as the second derivative of the gradient value, which is used to control the decay rate of the historical gradient value. Therefore, RMSProp algorithm has better adaptability than AdaGrad. The main body of the algorithm executes the following loop steps and will not stop until the stop condition is reached.

We take out the small batch of data $\{x_1, x_2, \ldots, x_m\}$. The target corresponding to the data is denoted by y_i.

We calculate the gradient based on the small batch data according to the following formula:

$$g = \frac{1}{m} \vee_\omega \sum_i L\big(f(x_i; \omega), y_i\big). \tag{15}$$

We accumulate the square gradient and refresh r, and the process is as follows:

$$r = \rho r + (1 - \rho) g \odot g. \tag{16}$$

We calculate the parameter update amount as follows:

$$\Delta\omega = -\frac{\sigma}{\sqrt{\delta + r}} \odot g. \tag{17}$$

We update parameters according to $\Delta\omega$:

$$\omega = \omega + \Delta\omega. \tag{18}$$

The traditional stochastic gradient descent maintains a single learning rate to update all weights, and the learning rate does not change from the training process. The RMSProp algorithm uses different learning rates when optimizing deep neural networks, which is efficient and practical.

OCTOPUS MODEL

The octopus-head model in **Figure 4** uses XGBoost, which is an optimized integrated tree model, improved and extended for the gradient boosting tree model. The integrated model of the tree is as follows:

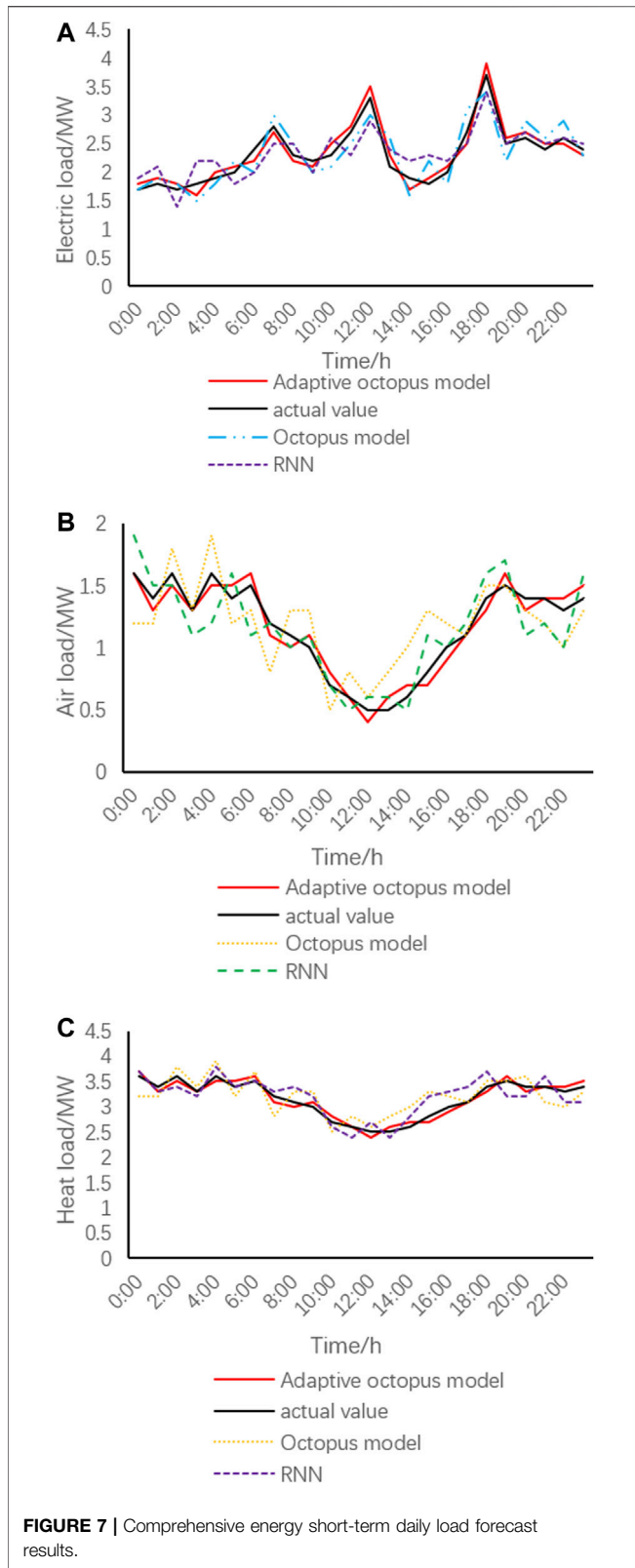$$\bar{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in F, \tag{19}$$

FIGURE 7 | Comprehensive energy short-term daily load forecast results.

where $\overline{y}_i$ is the predicted value of the ith sample; x_i is the feature vector of the ith data point, f_k is the structure q of the kth tree and the leaf weight w, K is the number of trees, and F is the collection space of trees.

The loss function of the model can be expressed as

$$L = \sum_{i=1}^{n} l(y_i, \overline{y}_i) + \sum_{k=1}^{K} \varepsilon(f_k). \qquad (20)$$

In the formula, the first half is the error between the predicted value and the true value, and the second half is the complexity of the tree.

In the process of minimizing the sequence, the loss function is reduced by adding the increment function $f_i(x_i)$. The objective function of the mth round is

$$L^{(m)} = \sum_{i=1}^{n} l(y_i, \overline{y}_i) + \sum_{k=1}^{K} \varepsilon(f_k)$$
$$= \sum_{i=1}^{n} l\left(y_i, \overline{y}_i^{(m-1)} + f_m(x_i)\right) + \varepsilon(f_m). \qquad (21)$$

For **Eq. 21**, the second-order Taylor expansion is used to approximate the objective function. This results in

$$L^{(m)} = \sum_{j=1}^{T} \left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \mu T, \quad (22)$$

$$g_i = l'\left(y_i, \widehat{y}_i^{(t-1)}\right),$$

$$h_i = l''\left(y_i, \widehat{y}_i^{(t-1)}\right).$$

We find the partial derivative of w to get

$$w_j = -\frac{G_j}{H_{j+\lambda}}. \qquad (23)$$

Substituting the weights into the objective function, we can get

$$L^{(m)} \cong -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j}{H_{j+\lambda}} \mu T. \qquad (24)$$

The octopus-head model mainly uses the weighted average method to train the octopus' quadrupeds with initial weights. The head adopts an integrated learning method to train the weight parameters according to the prediction accuracy of the four prediction models. The weights are changed according to the prediction accuracy of the octopus' quadrupeds, so that the weight of high prediction accuracy increases, and the weight of the other octopus' feet with relatively low prediction accuracy decreases, and then, the octopus' quadruped after weight adjustment continues to train and repeats this until the specified number of times is reached. Finally, these prediction accuracies will be integrated according to the weight, and the final prediction result will be obtained.

CASE ANALYSIS

Screening of Related Factors

We use the influencing factors and load data onto m days before the forecast date to predict the comprehensive energy load data onto the forecast day, and the model output is the comprehensive energy data onto the forecast day, as shown in **Eq. 20**.

$$\widehat{L}_d = \left\{\widehat{l}_{0,d}\big)\widehat{l}_{1,d}\big)\cdots\big)\widehat{l}_{t,d}\left(\cdots\right)\widehat{l}_{T,d}\right\}. \tag{25}$$

In the formula, T is the number of points to be predicted.

The comprehensive energy historical load data are shown in **Eq. 21**.

$$L_{t,d-w} = \left\{l_{t,d-1}\big)l_{t,d-2}\big)\cdots\big)l_{t,d-m}\right\}. \tag{26}$$

In the formula, $l_{t,d-m}$ is the load at time t on m days before the forecast.

The comprehensive energy load influencing factors are

$$F_{t,d} = \left\{w_{t)d}\left(e_{t)d}\right)\cdots\right)h_{t,d}\right\}. \tag{27}$$

Data Normalization

Formula **(23)** is used to normalize the comprehensive energy data and obtain real sample data of temperature, humidity, date, and economy.

$$x_{std} = \frac{x - x_{min}}{xmin_{max} + 1}. \tag{28}$$

In the formula, x represents each sample data of the integrated energy system, x_{max} represents the sample data with the largest absolute value in each sample data set, x_{min} represents the sample data with the smallest absolute value, and x_{std} represents the normalized value.

Result Analysis

This paper selects the comprehensive energy system operation data of a residential district in Shenyang from January 1, 2009, to January 1, 2020, and the comprehensive energy operation data of the district in the first week of August 2019 for analysis. The comprehensive energy data from 2009 to 2018 are used as the training set quarterly to predict the comprehensive energy operation data in 2019 and 2020; the data from the first 8 days of August 2019 are used as the training set to predict the electrical load, air load, and heat load data.

In order to verify the effect of adaptive load forecasting based on the octopus model, the long-term and short-term forecasting were carried out separately, and two cases were set up.

Case 1: considering the long-term forecast of the coupling of electricity, gas, and heat loads, the octopus model is used to analyze the comprehensive energy data from 2009 to 2018 to predict the electricity, gas, and heat loads in 2019 and 2020.

Case 2: considering the short-term forecast of the coupling of electricity, gas, and heat loads, the octopus model is used to predict one day's load data.

The long-term and short-term results of comprehensive energy in the two cases are shown in **Figure 5** and **Figure 6**, respectively.

It can be seen that whether it is a long-term forecast or a short-term forecast of electricity, gas, and heat load, the forecast curve has good tracking ability. Only when the real value fluctuates greatly, the predicted value will have a large error. From **Figure 7**, it can be seen that the electrical load fluctuates greatly throughout the year and the heat load changes significantly with the seasons.
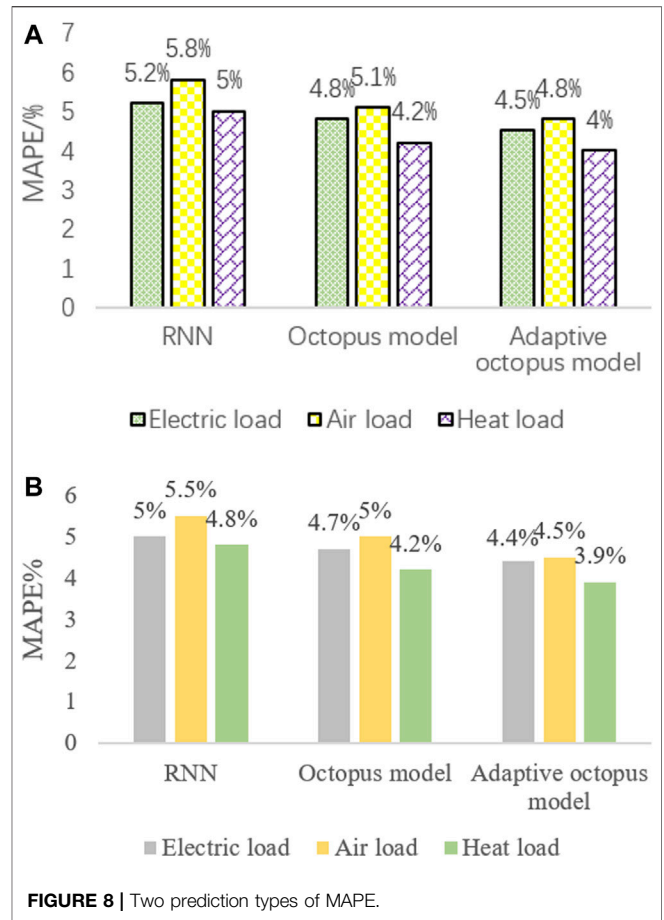

FIGURE 8 | Two prediction types of MAPE.

It can be seen from **Figure 5** that the electricity load fluctuates greatly throughout the day, with little electricity consumption during the day and large electricity consumption at night.

This article uses MAPE as the error evaluation standard, and the calculation formula of MAPE is

$$\delta_{MAPE} = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{\overline{y}_i - y_i}{y_i}\right| \times 100\%. \tag{29}$$

In the formula, y_t is the actual value; $y_t^{'}$ is the predicted value; and N is the number of predicted points.

Figure 8 shows the comparison between the MAPE value predicted by the octopus model and the MAPE value predicted by the ordinary RNN in two cases. The MAPE value of the adaptive IES load forecasting method is between 4 and 6%. Both of the MAPE values are not large.

It can be seen from **Figure 8** that the accuracy of short-term forecasts is higher than that of long-term forecasts. In the long-term and short-term forecasting, the adaptive octopus model has higher prediction accuracy and smaller error than the traditional load forecasting and ordinary octopus model and has a small MAPE value.

Compared with the results of gas and heat load forecasting, heat load forecasting has higher forecast accuracy, which is mainly due to the fact that heat load is greatly affected by temperature, humidity, economy, and date changes, especially

with a strong relationship between temperature and relevance. However, there are many uncontrollable factors of electricity and gas load, so the prediction accuracy is lower than that of heat load. From the perspective of long-term and short-term prediction results, the octopus model has better short-term prediction results. This is mainly due to the large training sample dataset for short-term load prediction and the small training error of the model.

CONCLUSION

Aiming at the background of integrated energy system, this paper proposes an adaptive integrated energy load forecasting method based on octopus model. This method not only uses the octopus model to effectively reduce the risk of model overfitting and prevents local minima but also improves the convergence speed of the model through the RMSProp algorithm and improves the prediction accuracy. It has high application value in the multielement load forecasting of the integrated energy system.

With the development of the energy Internet, the integrated energy system will receive more and more attention. The adaptive

IES load forecasting method based on the octopus model will have a higher development and application in the energy Internet system. The model's adaptive IES load forecast method has extremely high forecasting accuracy in ultra-short-term forecasting. In the future, it is hoped that, through further applications, it can predict abnormal conditions more accurately and improve the forecasting results.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

NZ conceived the idea for the manuscript and wrote the manuscript with input from XP, YW, MZ, MC, and WS. All authors have read and agreed to the published version of the manuscript.

REFERENCES

AlDahoul, N., Essam, Y., Kumar, P., Ahmed, A. N., Sherif, M., Sefelnasr, A., et al. (2021). Suspended Sediment Load Prediction Using Long Short-Term Memory Neural Network. *Sci. Rep.* 11 (1), 7826. doi:10.1038/s41598-021-87415-4

Bian, H., Zhong, Y., Sun, J., and Shi, F. (2020). Study on Power Consumption Load Forecast Based on K-Means Clustering and FCM-BP Model. *Energ. Rep.* 6 (S9), 693–700. doi:10.1016/j.egyr.2020.11.148

Dai, Y., and Zhao, P. (2020)., 279. A Hybrid Load Forecasting Model Based on Support Vector Machine with Intelligent Methods for Feature Selection and Parameter Optimization. *Appl. Energ.*, 115332. doi:10.1016/j.apenergy.2020.115332

Daniel, R, Teixeira, B., Pedro, F., Gomes, L., Omid, A., and Zita, V. (2020). Using Diverse Sensors in Load Forecasting in an Office Building to Support Energy Management[J]. *Energ. Rep.* 6 (S8). doi:10.1016/j.egyr.2020.11.100

Dong, Y., Dong, Z., Zhao, T., Li, Z., and Ding, Z. (2021). Short Term Load Forecasting with Markovian Switching Distributed Deep Belief Networks. *Int. J. Electr. Power Energ. Syst.* 130, 106942. doi:10.1016/j.ijepes.2021.106942

Guo, X., Gao, Y., Li, Y., Zheng, D., and Shan, D. (2021). Short-term Household Load Forecasting Based on Long- and Short-Term Time-Series Network. *Energ. Rep.* 7 (S1), 58–64. doi:10.1016/j.egyr.2021.02.023

Guo, X., Zhao, Q., Zheng, D., Ning, Y., and Gao, Y. (2020). A Short-Term Load Forecasting Model of Multi-Scale CNN-LSTM Hybrid Neural Network Considering the Real-Time Electricity price. *Energ. Rep.* 6 (S9), 1046–1053. doi:10.1016/j.egyr.2020.11.078

Jamal, F., Hamed, H-D., and Abbas, K. (2020).Multi-year Load Growth-Based Optimal Planning of Grid-Connected Microgrid Considering Long-Term Load Demand Forecasting: A Case Study of Tehran, Iran[J]. *Sustainable Energ. Tech. Assessments* 42. doi:10.1016/j.seta.2020.100827

Khan, Z. A. A., Ullah, A., Ullah, W., Rho, S., Lee, M., and Wook, S. W. B. (2020). Electrical Energy Prediction in Residential Buildings for Short-Term Horizons Using Hybrid Deep Learning Strategy. *Appl. Sci.* 10 (23), 8634. doi:10.3390/app10238634

Matrenin, P. V., Manusov, V. Z., Khalyasmaa, A. I., Antonenkov, D. V., Eroshenko, S. A., and Butusov, D. N. (2020). Improving Accuracy and Generalization Performance of Small-Size Recurrent Neural Networks Applied to Short-Term Load Forecasting. *Mathematics* 8 (12), 2169. doi:10.3390/math8122169

Pitchforth, D. J., Rogers, T. J., Tygesen, U. T., and Cross, E. J. (2021). Grey-box Models for Wave Loading Prediction. *Mech. Syst. Signal Process.* 159, 107741. doi:10.1016/j.ymssp.2021.107741

Verma, S., and Bala, A. (2021). Auto-scaling Techniques for IoT-Based Cloud Applications: a Review. *Cluster Comput.* doi:10.1007/s10586-021-03265-9

Wang, R., Sun, Q., Ma, D., and Liu, Z. (2019). The Small-Signal Stability Analysis of the Droop-Controlled Converter in Electromagnetic Timescale. *IEEE Trans. Sustain. Energ.* 10 (3), 1459–1469. doi:10.1109/tste.2019.2894633