Air Conditioning Load Prediction Based on HRD-TCN Model

Ning He[1], Liqiang Liu[1], Danlei Chu[2], Cheng Qian[*1] and Shang Li[1]

[1] Xi'an University of Architecture and Technology, Xi'an, China
[2] Optimal Process Control Techonlogies Co., Ltd, Xiamen, China
[*] E-mail: qiancheng@xauat.edu.cn

Abstract. Accurate air conditioning load prediction is of great importance for building energy management. This paper presents a hybrid data decomposition method (HRD) combined with temporal convolutional network (TCN) for air conditioning load prediction. First, the original data are decomposed using HRD to extract its linear features and frequency domain features. Then, the extracted significant features are input to TCN for predicting the next day's air conditioning load. Finally, the real values of the next day's air conditioning load are added to the training set, which are decomposed and predicted again to form a rolling mechanism. The experimental results show that the proposed method has excellent prediction performance.

Keywords: air conditioning load prediction, hybrid data decomposition method, temporal convolutional network, rolling mechanism

1. INTRODUCTION

In recent years, the issue of energy consumption has received widespread attention from academia and industry due to environmental degradation. Air conditioning energy consumption accounts for about 15% of social energy consumption [1, 2], and optimal control of air conditioning is an effective means to save energy. Predicting the air conditioning load in advance helps to make appropriate and suitable optimization strategies for air conditioning. In the context of big data, data-driven methods are now the mainstream approach to air conditioning load prediction. Common prediction methods include back propagation neural network (BP), support vector regression (SVR), long short-term memory networks (LSTM), etc [3].

BP and SVR are traditional neural network models that are widely used in load prediction. However, the single model performance is limited. Yu et al. [4] used particle swarm optimization (PSO) to optimize the initial weights and thresholds of BP to improve the performance of the model. Cheng et al. [5] used improved gray wolf optimizer to optimize the parameters of SVR, saving the tuning time and improving the prediction accuracy. LSTM is considered as a classical method to deal with time series problems. Zhou et al. [6] used LSTM to predict the cold load of Guangzhou library, and the results showed that its prediction accuracy was better than BP.

The accuracy of the prediction model is highly dependent on the quality of the data. Air conditioning is disturbed by many kinds of noise during operation, and the data should be processed for noise reduction before predicting the load. Signal decomposition is an effective means of noise reduction, and Zhang et al. [7] used discrete wavelet transform (DWT) and empirical mode decomposition (EMD) to decompose the data to improve the data quality and thus the prediction accuracy. Xuan et al. [8] used Chaos-SVR and wavelet decomposition SVR to predict the cooling load, and the experimental results showed that the accuracy of the hybrid model was better than that of the single model and the running time was shorter. Zhang et al. [9] proposed a hybrid model of seasonal decomposition and ARIMA, which effectively extracted the trend of air conditioning load and reduced the influence of random components on the model accuracy. Gao et al. [10] used CEEMDAN to decompose the data and then corrected the prediction error of PSO-BP, effectively improving the prediction accuracy and model stability. Jia et al. [11] used locally weighted regression de-noise the data to improve the data quality, and then obtained a high-precision prediction model. Although the above methods have achieved certain results, there are still some drawbacks, including limited model performance, unstable accuracy, high accuracy affected by data, and failure to fully extract important features of data.

To address the above existing problems, this paper proposes a hybrid data decomposition method combined with temporal convolutional network (HRD-TCN) for air conditioning load prediction hybrid model in order to improve the prediction accuracy and increase the robustness of the model. Firstly, HRD is used to decompose the original air conditioning load data, which can fully extract the linear features and frequency domain features of the original data. Then, TCN, which has excellent ability to process time series, uses important features to predict future values. Finally, the true values corresponding to the predicted values are added to the training set, decomposed and predicted

again to form a rolling mechanism. This mechanism only uses historical known information, does not disclose future information, and conforms to the logic of realistic prediction.

2. Methodology

2.1. The proposed method framework

The irregularity, nonlinearity, coupling and high fluctuation of air conditioning system load increase the difficulty of load prediction. In order to improve the accuracy and robustness of air conditioning load prediction, an effective rolling prediction model based on hybrid data decomposition combined with TCN is proposed. In order to make full use of the known information of air conditioning load variables and improve the authenticity of prediction results, a rolling mechanism is introduced in data decomposition. The proposed HRD-TCN model implementation process is summarized as follows:

(a) First, HRD model is used to decompose the original air conditioning load data.

(b) TCN prediction. The extracted features are input into TCN for air conditioning load prediction.

(c) Rolling decomposition and prediction. After the prediction is completed, the measured real values of corresponding points are added to the training set for re-decomposition and prediction, forming a rolling decomposition and prediction mechanism. This method makes full use of the known information, improves the prediction accuracy, authenticity and reliability, and conforms to the practical application prediction logic.

2.2. Correlation analysis

The air-conditioning load has the characteristics of non-linear and high fluctuation due to the interference of personnel flow, meteorology, equipment operation and other factors. Although the change of air conditioning load has the characteristics of time series, it is difficult to capture its change law only considering the historical data, and other influencing factors need to be taken into account. However, the variables with low correlation have little impact on the air conditioning load. Taking these variables as input variables will not only increase the operation cost of the model, but also may reduce the prediction accuracy. Therefore, Pearson correlation analysis was used to screen the variables. The formula is as follows:

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \quad (1)$$

where, x, y represents two kinds of variables, and $-$ represents the average value of variables.

2.3. Hybrid data decomposition

Since many time series are irregular and nonlinear, a single model cannot fully extract their inherent features, resulting in low prediction accuracy. Many studies show that the hybrid model solves this problem well. The hybrid model analyzes data from different angles and integrates the characteristics of each model to improve the prediction accuracy and reduce the problem of poor adaptability of the model due to the use of wrong methods.

2.3.1 Linear-nonlinear decomposition

Büyükşahin et al. [12] put forward the assumption that time series are composed of linear series and nonlinear series. Using this decomposition idea (LN), the moving average filter (MA) is used to decompose the air conditioning load data into linear and nonlinear parts, as shown in the formulas:

$$C_t = L_t + N_t \quad (2)$$

$$L_t = \frac{1}{m}\sum_{i=t-m+1}^{t} C_i \quad (3)$$

$$N_t = C_t - L_t \quad (4)$$

where, m is the moving window of MA, L_t is a relatively stable linear sequence, and C_t is a nonlinear sequence with high volatility.

After obtaining the linear component, TCN is used to predict the linear component.

$$L_t = TCN(L_{t-1}, L_{t-2}, \cdots, L_{t-m}, x) \quad (5)$$

where, L_{t-m} is historical data and x is other influencing factors, such as meteorological data.

For the non-linear sequence, (6) is used for modeling. The non-linear modeling is used to establish the relationship between these characteristics and the future air conditioning load and to make the final prediction.

$$\begin{aligned} C_t = TCN(&C_{t-1}, C_{t-2}, \cdots, \\ &C_{t-p}, L_t, N_{t-1}, N_{t-2}, \cdots, N_{t-q}, x) \end{aligned} \quad (6)$$

2.3.2 VMD

As a signal processing method with rigorous mathematical derivation, variational mode decomposition (VMD) is widely used. It can adaptively decompose complex, non-stationary and nonlinear air conditioning load signals into a series of modes with different frequency scales.

VMD seeks k sequence components with specific sparsity through constrained variational model to achieve the goal of minimizing the estimated bandwidth of each component.

$$\min_{\{u_m\},\{\omega_m\}} \left\{ \sum_{m=1}^{k} \left\| \left[\left(\partial(t) + \frac{j}{\pi t} \right) * u_m(t) \right] e^{jw_m t} \right\|_2^2 \right\} \quad s.t. \sum_{m=1}^{k} u_m(t) = f(t) \tag{7}$$

where, $f(t)$ is the original signal, t is the time, m is the number of modes, u_m is the m-th mode, $\delta(t)$ is Dirac distribution, ω_m is the center frequency, and * is convolution operation.

Solving optimization problems with augmented Lagrange function.

$$L(\{u_m\},\{\omega_m\},\{\lambda\}) = \alpha \sum_{m=1}^{k} \left\| \partial_t [(\delta(t) + \frac{j}{\pi t}) u_m(t)] e^{-jw_m t} \right\|_2^2 + \left\| f(t) - \sum_{m=1}^{k} u_m(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{m=1}^{k} u_m(t) \right\rangle \tag{8}$$

where, α is the quadratic penalty factor and λ is the Lagrange multiplication operator.

Eq. (8) is solved by alternating direction multiplier method (ADMM), and the optimal solution is obtained. The variable and is updated by (9) - (11) as follows:

$$\hat{u}_m^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{p \neq m}^{k} \hat{u}_p^n(\omega) + 0.5\hat{\lambda}(\omega)}{1 + 2\alpha(\omega - \omega_m^n)^2} \tag{9}$$

$$\omega_m^{n+1} = \frac{\int_0^\infty \omega \left| \hat{u}_m^{n+1}(\omega) \right|^2 d\omega}{\int_0^\infty \left| \hat{u}_m^{n+1}(\omega) \right|^2 d\omega} \tag{10}$$

$$\lambda^{n+1}(\omega) = \lambda^n(\omega) + \tau[x(\omega) - \sum_{m=1}^{k} u_m^{n+1}(\omega)] \tag{11}$$

where τ is step and $\tau > 0$, and ^ stands for Fourier transform operation symbol.

2.3.3 Hybrid rolling decomposition

In order to fully capture the inherent variation characteristics from the original air conditioning load data, HRD is used to decompose the original sequence, which can well extract the linear and nonlinear frequency domain features of the air conditioning load series. Firstly, the original series is decomposed into linear subsequence and residual subsequence (nonlinear series) by using LN method, and then the linear stationary components are modeled by TCN to obtain their predictive values. After that, the residual sequence is decomposed by VMD to extract its frequency domain features. Finally, taking the historical load value, linear prediction value, nonlinear decomposition value and other influencing factors as input variables, a nonlinear model is established to predict the air conditioning load, as shown in (12).

$$C_t = TCN(C_{t-1}, C_{t-2}, \cdots, C_{t-p}, L_t, u_1, u_2, \cdots, u_n, x) \tag{12}$$

where, $u_i = \{u_i(t-1), u_i(t-2), \cdots, u_i(t-p)\}$.

The specific process of HRD can be summarized as follows:

(a) LN decomposition. Use MA to decompose the original air conditioning load data into linear series and nonlinear series, as shown in (3) and (4).

(b) Linear prediction. TCN is used to predict linear subsequence, as shown in (5).

(c) Nonlinear decomposition. The nonlinear residual subsequence is decomposed into n subsequences $\{u_1, u_2, \cdots, u_n\}$ with different modes by VMD.

(d) Nonlinear modeling. Take the historical observation value of the original air conditioning load series, the predicted value of the linear subsequence, the historical observation value of the non-linear VMD decomposition sub series and other influencing factors as input characteristics to predict the future value of the air conditioning load, as shown in (12).

(e) Rolling mechanism. In order to make full use of the known information of the training set, a rolling decomposition mechanism is introduced. Some existing decomposition models decompose the original sequence first, and then divide the training set and the test set. This method leaks the unknown information of the test set to the prediction model, resulting in false high-precision prediction. However, in the actual working conditions, the test data are unknown, so the method of unified decomposition of training set and test set data in advance is not applicable. In order to meet the actual application scenarios, a rolling decomposition and prediction model is proposed to ensure the feasibility and accuracy of the prediction. In the rolling decomposition and prediction mechanism, only known historical data should be decomposed. For single-step prediction, only one point is predicted at a time. When the next point is predicted, the current test point is regarded as known, and the currently known data is incorporated into the training set for decomposition and prediction, and this process is repeated to form rolling. It should be noted that the actual observed values of the first test data points are used instead of the predicted values during the cycle, because when the prediction model has predicted the corresponding test set data points and compared them with the actual values, they are considered as known data. And the same is true for multistep prediction. As shown in Fig. 1, the original air conditioning load was decomposed by HTD and predicted by TCN in the first stage. After the prediction is completed, the actual observed value of the test data corresponding to the predicted value is carried into the second stage as known

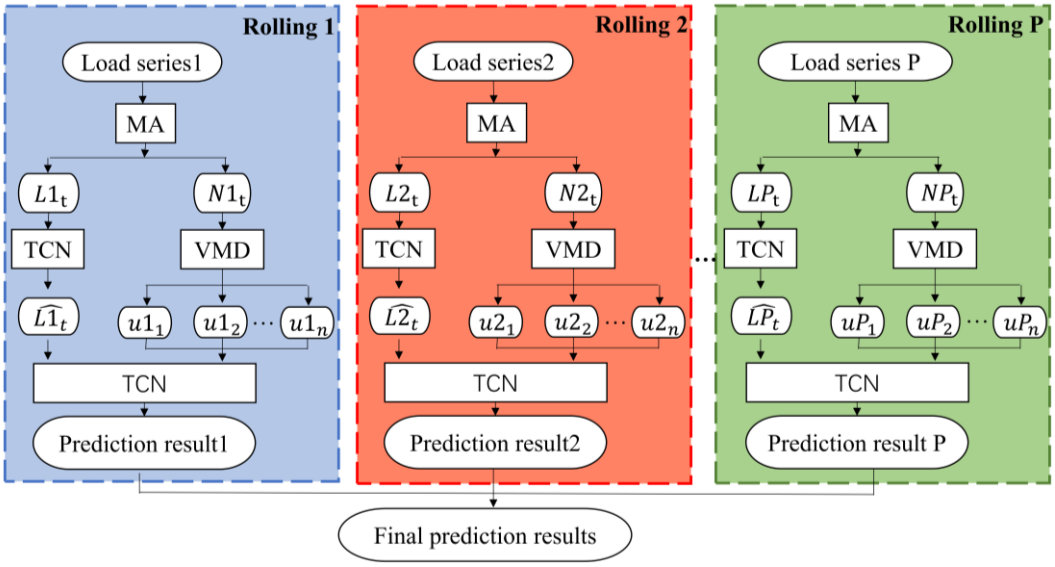


Fig. 1 Rolling decomposition and prediction mechanism

data. Then decompose the new training data and predict the next point, repeat the above steps, and roll P times to predict P points.

Note that nonlinear modeling uses the TCN model, which will be covered in detail later.

2.4. TCN

The TCN based on dilated convolution and causal convolution can effectively avoid the gradient vanishing problem, and is widely used in time series prediction.

The time series requires that only historical data can be used without disclosing future information, that is, $x_{t-1}, x_{t-2}, ..., x_{t-m}$ can be used to predict $y_t, y_{t+1}, ..., y_{t+n}$, which is expressed by (13)

$$y = f_{tcn}(x) \tag{13}$$

TCN strictly abides by this principle and uses the known information in the past to predict the future through causal convolution.

Under the premise of causal convolution, TCN uses dilated convolution to increase the field of view, which is conducive to better extracting historical information and capturing the overall features of data, thus reducing the influence of local features. TCN is a good solution to the problem that traditional convolutional networks require more network layers and larger convolutional kernels to process extremely informative time series, and the perceptual field of view increases exponentially with the expansion factor, reducing the operating cost. The schematic diagram is shown in Fig.2.

$$D(x) = (x *_c f)(x) = \sum_{i=0}^{k-1} f(i) \cdot x_{x-c \cdot i} \tag{14}$$

where, $*_C$ is the dilation convolution operator, $f(i)$ represents the i-th convolution kernel, c is the dilation rate, and k is the convolution kernel size.

As shown in the Fig. 2, the size of the convolution kernel k and the dilation convolution factor dc affect the field of view dc of the TCN. dc is calculated using the (15).

$$\mathrm{dc} = k^{v} \tag{15}$$

where, k is chosen as 2 and v is 1,2,3,4 in this paper.

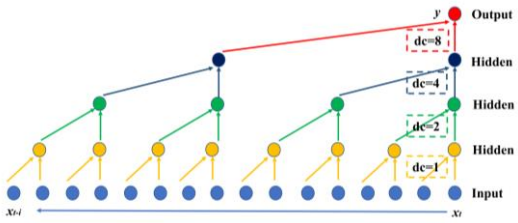


Fig. 2 Dilated convolution schematic

Further, the TCN uses residual connectivity, which can perform feature extraction in a cross-layer manner, effectively avoiding the problem of losing important information to reduce the accuracy of the model when training deeper networks. Its structure is shown in Fig.3

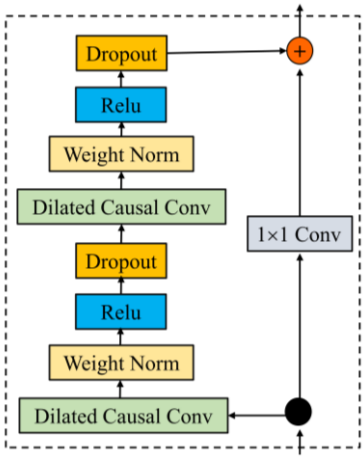


Fig. 3 TCN Residual Block

3. Experiment

A central air conditioning system in a city was used as a study case and its actual operational data from October 5, 2016 to December 12, 2016 was selected for the experiment. October 5, 2016 to December 9, 2016 is the dataset as the training set and the dataset from December 10 to 12, 2016 as the test set.

3.1 Correlation analysis

RH : Outdoor humidity
DT : Outdoor dry bulb temperature
WT : Outdoor wet bulb temperature
P : Air conditioning load

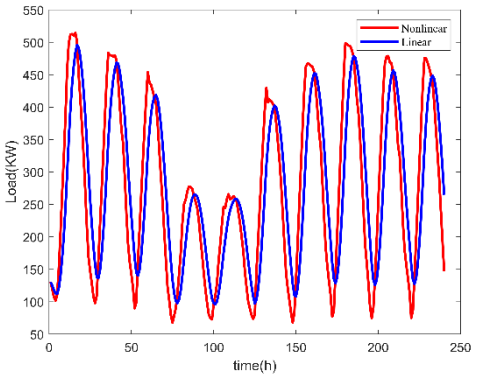
In this paper, we choose the previous day's historical load and meteorological factors as input variables, and do correlation analysis on meteorological factors, and those with high correlation were selected as the input variables. The results of Pearson correlation analysis are shown in Table 1. The correlations of *RH* , *DT* and *WT* were all greater than 0.4, indicating moderate correlations.

Table. 1 Correction of input variables with *P* at moment *t*

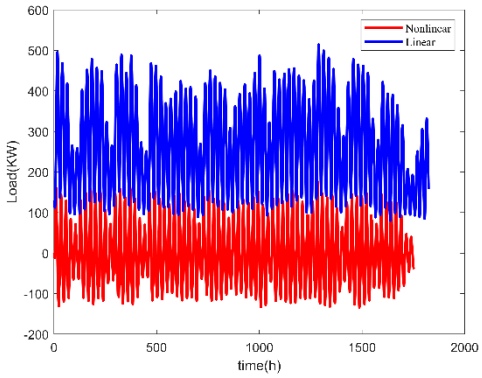
The input variable	correlation
RH	0.46
DT	0.50
WT	0.45

3.2 HRD Data decomposition

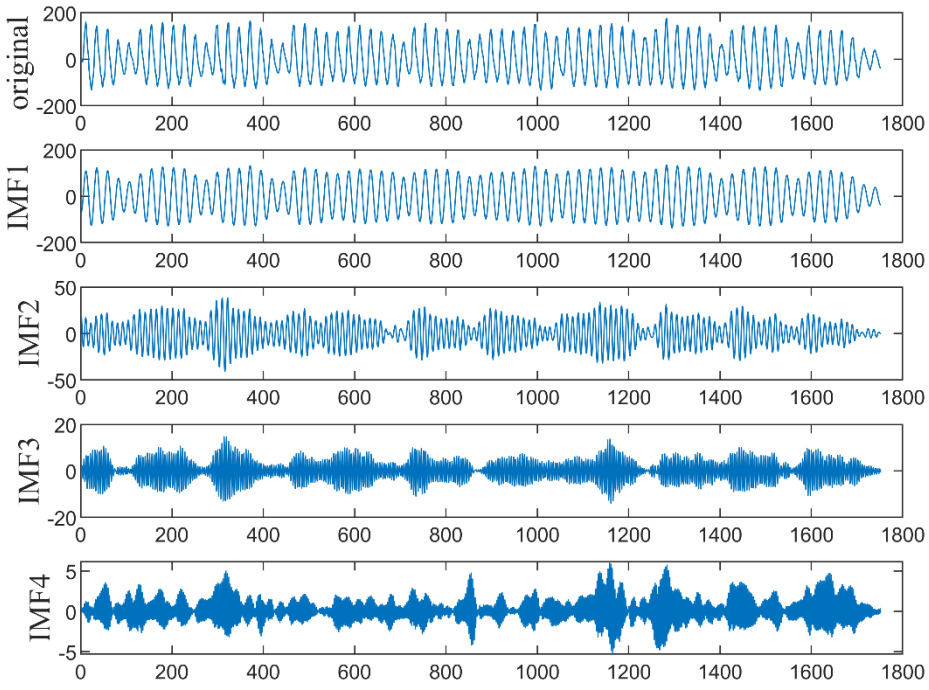
HRD decomposition method combines the advantages of LN method and VMD method. It not only considers the original linear characteristics of data, but also performs frequency domain analysis on data. As can be seen from Fig. 4(a), the data decomposed by LN is smoother with more significant features, and the linear and nonlinear components are shown in Fig. 4(b). The nonlinear components were decomposed by VMD, and the important features were further extracted to improve the prediction accuracy. The decomposition effect of VMD is shown in Fig. 4(c).



(a) LN decomposition



(b) linear and nonlinear components



(c) VMD decomposition

Fig. 4 HRD decomposition

3.3 HRD - TCN prediction

The extracted important features are input into TCN as input variables to predict air conditioning load, as shown in (12).

To verify the effectiveness of the proposed method, it was compared with TCN, VMD-TCN and CNN-LSTM, and its prediction effect was shown in Fig. 5. It can be seen from the Fig. 5 that the prediction effect of the proposed method is the best, and the predicted value is closest to the actual load value of air conditioning.

To accurately evaluate the performance of the proposed method, RMSE, CVRMSE, MAE and MAPE were used as evaluation indexes, and the results are shown in Table. 2.

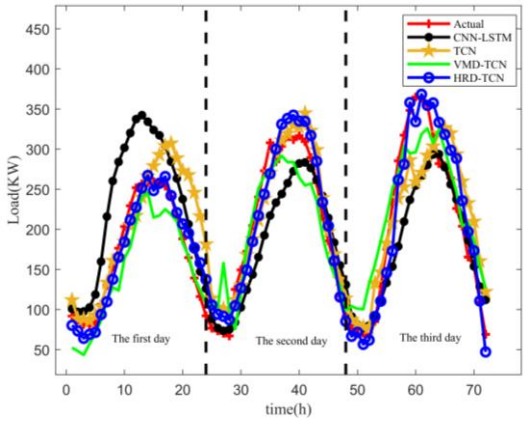


Fig. 5 prediction result

Table. 2 Prediction accuracy

Models	MAE (KW)	RMSE (KW)	MAPE (%)	CVRMSE (%)
CNN-LSTM	44.46	55.21	23.54	28.31
TCN	34.82	44.78	21.26	22.96
VMD-TCN	29.02	34.55	21.20	17.71
HRD-TCN	20.77	24.76	12.73	12.69

4.conclusion

In this paper, a novel hybrid prediction model based on HRD-TCN for air conditioning load is proposed. HRD is used to decompose the original air-conditioning load data and extract its internal important features as input variables into TCN for load prediction. At the same time, HRD introduces rolling mechanism to make full use of the known information, which makes the prediction process conform to the actual logic and the results are more real. Experimental results show that the proposed method meets ASHARE's standard (CVRMSE less than 30%) and has excellent performance in time series problems.

In the future work, we will further study the parameters selection of HRD to improve the prediction accuracy.

Acknowledgements

This work was supported by NSFC (61903291), and China Postdoctoral Science Foundation (2019M660257).

References

[1] Xu G, Wang W. China's energy consumption in construction and building sectors: An outlook to 2100[J]. Energy, 2020, 195: 117045.

[2] Kumar S, Pal S K, Singh R P. A novel method based on extreme learning machine to predict heating and cooling load through design and structural attributes[J]. Energy and Buildings, 2018, 176: 275-286.

[3] Kadir Amasyali, Nora M. El-Gohary, "A review of data-driven building energy consumption prediction studies, " Renewable and Sustainable Energy Reviews,vol. 81,no. 1, pp. 1192-1205, 2018.

[4] Yu J Q, Jing W Q, Zhao A J, et al. Application of improved PSO-BP neural network in cold load forecasting of mall air-conditioning[J]. Journal of Control Science and Engineering, 2019, 2019.

[5] Cheng R, Yu J, Zhang M, et al. Short-term hybrid forecasting model of ice storage air-conditioning based on improved SVR[J]. Journal of Building Engineering, 2022, 50: 104194.

[6] Zhou C, Fang Z, Xu X, et al. Using long short-term memory networks to predict energy consumption of air-conditioning systems[J]. Sustainable Cities and Society, 2020, 55: 102000.

[7] Zhang L, Alahmad M, Wen J. Comparison of time-frequency-analysis techniques applied in building energy data noise cancellation for building load forecasting: A real-building case study[J]. Energy and Buildings, 2021, 231: 110592.

[8] Xuan Z, Xuehui Z, Liequan L, et al. Forecasting performance comparison of two hybrid machine learning models for cooling load of a large-scale commercial building[J]. Journal of Building Engineering, 2019, 21: 64-73.

[9] Zhang L, Shi P, Lai X, et al. Air-conditioning load forecasting based on seasonal decomposition and ARIMA model[C]//2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). IEEE, 2021: 664-669.

[10] Gao Y, Hang Y, Yang M. A cooling load prediction method using improved CEEMDAN and Markov Chains correction[J]. Journal of Building Engineering, 2021, 42: 103041.

[11] Jia X, Sang Y, Li Y, et al. Short-term forecasting for supercharged boiler safety performance based on advanced data-driven modelling framework[J]. Energy, 2022, 239: 122449.

[12] Büyükşahin Ü Ç, Ertekin Ş. Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition[J]. Neurocomputing, 2019, 361: 151-163.