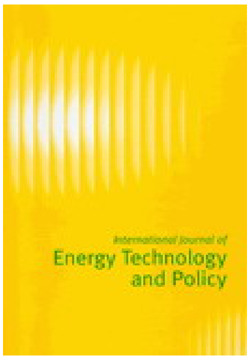
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Energy Technology and Policy

ISSN online: 1741-508X - ISSN print: 1472-8923
https://www.inderscience.com/ijetp

A novel residential electricity load prediction algorithm based on hybrid seasonal decomposition and deep learning models

Shan Gao, Xinran Zhang, Lihong Gao, Yancong Zhou

DOI: 10.1504/IJETP.2025.10071692

Article History:

Received:	29 January 2025
Last revised:	01 April 2025
Accepted:	08 May 2025
Published online:	24 June 2025

A novel residential electricity load prediction algorithm based on hybrid seasonal decomposition and deep learning models

Shan Gao*, Xinran Zhang, Lihong Gao and Yancong Zhou

School of Information and Engineering,
Tianjin University of Commerce,
Tianjin, 300134, China
Email: gaoshan@tjcu.edu.cn
Email: yvonne6190@163.com
Email: g_lihong@126.com
Email: zycong78@126.com
*Corresponding author

Abstract: Residential electricity load prediction is of great significance for power system planning. With the increasing complexity and uncertainty of the power grid, traditional prediction models still have insufficient accuracy and neglect seasonal changes. In this paper, a data-driven multi-scale hybrid prediction model for residential electricity load is proposed, which integrates a convolutional neural network (CNN), long short-term memory (LSTM), and attention mechanism. The seasonal decomposition was applied to extract seasonal patterns of the electricity consumption data. The hybrid model integrates the parallel processing capability of CNN and the long time-series modelling capability of LSTM to capture the spatial-temporal characteristics of electricity load accurately. The attention mechanism is employed to calculate the critical weight to enhance the prediction accuracy dynamically. Finally, detailed comparison experiments show that the proposed hybrid model outperformed state-of-the-art algorithms. The MAPE of the hourly and daily prediction results of the proposed model are 2.36% and 0.76%, respectively.

Keywords: electricity consumption prediction; deep learning; convolutional neural network; CNN; long short-term memory; LSTM; attention mechanism.

Reference to this paper should be made as follows: Gao, S., Zhang, X., Gao, L. and Zhou, Y. (2025) 'A novel residential electricity load prediction algorithm based on hybrid seasonal decomposition and deep learning models', *Int. J. Energy Technology and Policy*, Vol. 20, No. 5, pp.1–23.

Biographical notes: Shan Gao received her PhD from Tianjin University in 2012. She is currently a Lecturer at Tianjin University of Commerce. Her research interests include intelligent optimisation algorithm, decision reasoning and knowledge model construction.

Xinran Zhang received her BEng from Tianjin University of Commerce in 2024. Her research interests include intelligent optimisation algorithm et al.

Lihong Gao received her BEng from Tianjin University of Commerce in 2024. Her research interests include intelligent optimisation algorithm et al.

Yancong Zhou received her PhD from Hebei University of Technology. She is currently a Professor at Tianjin University of Commerce. Her research interests include intelligent optimisation algorithm and knowledge model construction.

This paper was originally accepted for a special issue on 'OA Multiscale Energy Systems for Renewable Energy Storage' guest edited by Prof. Zheng Xu, Prof. Jemal H. Abawajy and Prof. Vijayan Sugumaran.

1 Introduction

Electricity is the basis of modern industrial and commercial activities and plays an irreplaceable role in efficient economic operation. Accompanied by the massive access of clean energy to the power system, the complexity and degree of uncertainty of the power grid are increasing. Accurate prediction of residential electricity load is crucial to improve the reliability and efficiency of the smart grid.

Electricity load prediction plays an essential role in the supply-demand collaboration of smart grids, and it has also been of focus for both academia and industry (Ahmad et al., 2022). The existing prediction models can be categorised into three main categories from the prediction algorithms: statistical theory, classical machine learning, and deep learning (DL). We proposed a multi-scale short-term prediction model for electricity load based on a hybrid of SD, convolutional neural network (CNN), long short-term memory (LSTM), and attention mechanism (SD-CNN-LSTM-Attention). This model reveals the electricity consumption behaviours of users under different conditions with respect to the results of the analysis. In summary, the model can achieve accurate multi-scale prediction at both the hourly and daily time scale, which provides an essential reference for power planning on both the supply and demand sides. The proposed algorithm has good prediction performance and generalisation ability. In addition, the proposed algorithm is an optimised combination of several traditional algorithms, which also have good interpretability. It provides a valuable perspective and methodological reference for residential electricity consumption forecasting research.

The main contributions of this study are as follows:

1 SD is applied to decompose the electricity load to reveal its evolving pattern under different frequency characteristics. Meanwhile, it adapts to the requirements of different time-scale data analysis and prediction accurately.

2 An electricity load prediction model based on SD-CNN-LSTM-Attention is proposed, which can effectively capture the long-term dependence and local features in the power consumption and strengthen the attention to crucial information.

3 The hybrid model is compared with the state-of-the-art algorithms for hourly and daily prediction of residential electricity load. The experimental results indicate that the proposed model outperformed the traditional models with a minimum MAPE of 2.36% and 0.76%, respectively.

The rest of this paper is organised as follows: Section 2 is related works. Section 3 is the methodology, which analyses the basic principle of SD-CNN-LSTM-Attention prediction algorithm in detail. Section 4 presents and analyses the experimental results. The conclusions are given in Section 5.

2 Related works

Traditional statistical methods construct forecasting models by analysing the mathematical relationship implied in the changing trend of electricity load. For instance, Rekhade and Sakhare (2021) employed four regression modeling methods to predict electricity consumption, offering simplicity and interpretability with reduced overfitting. Yang et al. (2018) proposed an integrated probability density prediction method using Gaussian process quartile regression (GPQR) to deal with the uncertainty of electricity load, providing robust modelling of uncertainty and nonlinear relationships. Although these regression prediction methods have fast prediction speeds, the model performance will be reduced when there are significant errors and defects in the historical data. Pan and Jia (2025) used univariate linear regression, ARIMA model, Fourier analysis method, univariate linear regression prediction method and binary linear regression prediction to predict the long-term trend, seasonal variation, periodic variation, and irregular variation of power demand in industrial parks. They also introduced the GARCH model to test the error sequence, which improved the accuracy of the prediction model. The grey prediction method is another statistical method that is usually used to deal with the problem of predicting systems with incomplete and imprecise information. Zhao and Guo (2016) used the grey model of hybrid optimisation to improve the accuracy of annual electricity load forecasting significantly. However, the statistical methods have significant limitations for nonlinear electrical load time series with strong randomness. The second category is prediction models based on classical machine learning algorithms. This kind of model realises the analysis and prediction of electricity consumption by learning from historical data. Jiang et al. (2020) used the support vector regression (SVR) model to mine the nonlinear relationship of electricity load. This approach improves model robustness and adaptability but comes with high computational costs and complexity. Tang et al. (2021) proposed an electricity consumption prediction model based on binary nonlinear fitted regression (BNFR) and SVR, enhancing accuracy and adaptability to complex consumption patterns. However, overfitting, high computational cost and complex feature engineering are still problems for researchers (Al-Alimi et al., 2023). Therefore, DL models can extract complex nonlinear features of data for time series prediction in recent years (Xu et al., 2024). RNN is a kind of neural network that can process sequence data and maintain the information state during the evolution of the sequence. Heydari et al. (2019) predicted electricity consumption in Russia based on RNN. Fekri et al. (2021) proposed an innovative online adaptive RNN. Online features are achieved by capturing temporal dependencies while updating RNN weights based on new data. Compared with RNN, LSTM introduces three ‘gate’ structures and a memory cell, which solves the problem of gradient explosion and gradient disappearance when RNN processes long time series data (Yang et al., 2023). Bareth et al. (2024) applied LSTM to estimate the monthly electricity demand as well as the daily load demand. Compared with the artificial neural network (ANN) model, the proposed LSTM model is more accurate. Wen et al. (2019) established a deep recurrent neural network with a long short-term memory (DRNN-LSTM) model to predict the aggregated electricity load in a community microgrid. While the approach enhances forecasting accuracy and adaptability to renewable energy fluctuations, it also introduces high computational complexity and strong data dependency.

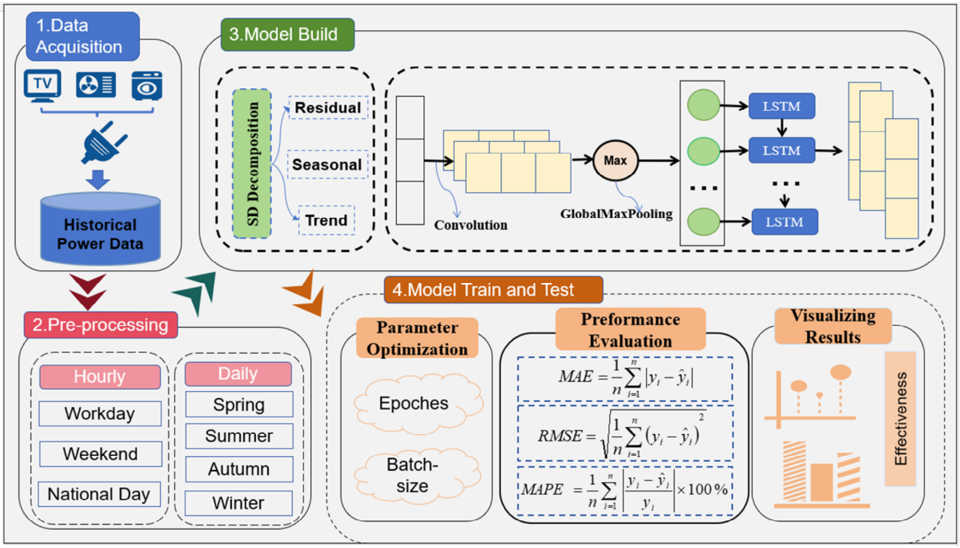
According to the prediction horizon, the existing methods for predicting electricity consumption are divided into ultra-short-term load forecasting (USTLF), short-term load forecasting (STLF), medium-term load forecasting (MTLF) and long-term forecasting (LTLF) (Butt et al., 2020). USTLF is affected by the substantial volatility of electricity load, which is difficult to predict accurately and robustly (Tan et al., 2020). Short-term forecasting offers significant savings potential for the economical and safe operation of power systems. Medium-term forecasts involve the scheduling of fuel supply and maintenance operations, while long-term forecasts facilitate the formulation of overall energy planning layouts and related strategic decisions.

Due to the fluctuation and non-stationarity of the electricity load, a single model cannot fully capture the intrinsically implied complex features. In order to improve the electricity consumption prediction performance, many researchers have focused on the development and training of hybrid models. Dong et al. (2016) developed an innovative hybrid modelling approach by combining data-driven techniques with physically based forward models. Li et al. (2021) proposed an electricity load prediction method based on multiple linear regression (MLR) and LSTM. The EMD is applied to decompose the original data to reduce the complexity of the electricity load. Lv et al. (2022) proposed a hybrid model based on variational modal decomposition (VMD) and long- and short-term memory, as well as the elimination of seasonal factors and error correction. Four real-world electricity load datasets from Singapore and the USA are used to verify the effectiveness and practicality of the proposed hybrid model. Zhang et al. (2018) proposed a hybrid model based on improved empirical mode decomposition (IEMD), autoregressive integrated moving average (ARIMA), and wavelet neural network (WNN). The simulation results show that the proposed model performs well in electricity load forecasting. Besides, the hybrid models also include the combination of statistical models and data-driven models (Dudek et al., 2022), the hybrid of machine learning and DL models (Bashir et al., 2022; Wang et al., 2021), the mixture of multiple DL models (Eskandari et al., 2021), etc. However, these existing models tend to neglect factors such as holidays, which makes it challenging to achieve high-precision prediction of electricity load over multiple time scales.

3 Methodology

The residential electricity load prediction model based on SD-CNN-LSTM-Attention is proposed in this paper, and the architecture is illustrated in Figure 1. It can be seen that the hybrid model contains the key components, such as the SD, CNN, LSTM, and attention mechanism. The process of the SD-CNN-LSTM-Attention prediction model mainly includes five parts: data acquisition and analysis, seasonal decomposition, model construction, parameter optimisation, and visualisation of prediction results. Firstly, data analysis and processing are performed to make the original data more convenient for prediction. Due to the periodic variation of electricity load, SD is performed on the original electricity load data set. Then, a hybrid model based on CNN-LSTM-Attention is constructed for electricity load prediction, and grid search is used to find the optimal parameter combination of the model. Finally, the prediction results were visualised to intuitively demonstrate the superiority of hourly and daily electricity load prediction.

Figure 1 The flowchart of residential electricity load prediction based on SD-CNN-LSTM-attention (see online version for colours)



3.1 SD

SD is often used to separate the original time series data into different sub-items so that we can understand the abstract structure and reveal hidden patterns and trends in the data. The historical electricity load of residents has prominent periodic characteristics. In this paper, SD is selected to effectively extract the periodic change trend in the electricity load data. SD divides the raw electricity load into trend, seasonal, and residual components. The decomposition process is as follows.

$$\boldsymbol{Y}_t = \boldsymbol{T}_t + \boldsymbol{S}_t + \boldsymbol{R}_t \quad (t = 1, 2, \ldots, n) \tag{1}$$

where $\boldsymbol{Y}_t$ is the raw electricity load data at the t^{th} moment, $\boldsymbol{T}_t$, $\boldsymbol{S}_t$ and $\boldsymbol{R}_t$ are the trend, seasonal and residual components at the corresponding moment respectively.

In additive models, the magnitude of seasonal fluctuations is fixed. The electricity load of residential buildings has stable seasonal fluctuations. The corresponding operations are as follows:

1 Trend analysis: The trend is obtained by smoothing the data with a convolutional filter. Calculating the long-term trends can contribute to revealing long-term growth or decrease trends in the data.

$$\boldsymbol{T}_t = \frac{1}{n} \sum_{i=-\frac{n-1}{2}}^{\frac{n-1}{2}} \boldsymbol{Y}_{t+i} \tag{2}$$

where $\boldsymbol{T}_t$ represents the electricity load trend component corresponding to the t^{th} time step; n indicates the average moving window size. $\boldsymbol{Y}_{t+i}$ denotes the raw data corresponding to the $(t+i)^{\text{th}}$ time step.

2 Seasonal factor extraction: Remove the trend from the original data according to the additive model. Calculate the seasonal mean value after removing the trend as the seasonal component.

$$\boldsymbol{D_t} = \boldsymbol{Y_t} - \boldsymbol{T_t} \tag{3}$$

$$\boldsymbol{S_t} = \frac{1}{m}\sum_{k=1}^{m} \boldsymbol{D_{t+k \times p}} \tag{4}$$

where $\boldsymbol{D_t}$ is the electricity load data after detrending corresponding to the t^{th} time step; $\boldsymbol{Y_t}$, $\boldsymbol{T_t}$ and $\boldsymbol{S_t}$ is the original, trend, and seasonal component of the t^{th} time step; m denotes the number of cycles; p is the length of the cycle; and is the detrended electricity load of the $(t + k \times p)^{\text{th}}$ time step.

3 Stochastic volatility analysis: Subtracting the seasonal component from the detrending data yields a residual component containing noise and other unpredictable factors.

$$\boldsymbol{R_t} = \boldsymbol{D_t} - \boldsymbol{S_t} \tag{5}$$

where $\boldsymbol{R_t}$, $\boldsymbol{D_t}$ and $\boldsymbol{S_t}$ are the residual electricity load component corresponding to time step t, the electricity load after detrending, and the seasonal electricity load component, respectively.

3.2 CNN

CNN automatically learns and extracts significant features from the original data through convolution operation, which makes up for the shortcomings of LSTM in feature extraction. The convolution and pooling operations of CNN make the model adaptable to the scale and shape of the input data. It means that CNN can adaptively process electricity load data of different lengths without complex pre-processing or scaling operations.

The core principle of CNN is to extract the features of the data through convolutional and pooling operations, and finally to perform regression or classification through the fully connected layer, and the principle is shown in the first part of Figure 2.

The main mathematical principles and formulas corresponding to CNN are as follows:

1 Convolutional operation extracts features of the input data. For the i^{th} convolution kernel, its output $z_i(t)$ at the t^{th} time step is given as:

$$z_i(t) = \mathrm{Re}LU\left(\sum_{k=-\frac{n}{2}}^{\frac{n}{2}} w_i(k) \cdot x(t+k) + b_i\right) \tag{6}$$

where n is the size of the convolution kernel, $w_i(k)$ is the weight of the i^{th} convolution kernel at position k, $x(t + k)$ is the value of the input data at position $t + k$, and b_i is the bias term of the i^{th} convolution kernel.

2 Pooling operation reduces the size of the feature map and keeps the most important features, commonly used way is maximum pooling.

$$p_i = \max_t \{z_i(t)\} \tag{7}$$

where p_i is the maximum value of the i^{th} convolution kernel in all time steps, which is the value of the i^{th} feature.

3 The fully connected layer flattens the output of the convolutional layer or pooling layer into a vector and performs a linear transformation through the weight matrix and the bias vector.

$$y = \sum_i w_i \cdot p_i + b \tag{8}$$

where w_i is the weight of the i^{th} neuron in the output layer, that is, the weight of the i^{th} feature; b is the bias term of the output layer.

3.3 LSTM

There are two primary components of LSTM: cell states and gate mechanisms. Among the gate mechanisms are the input gate, output gate, and forgetting gate. These two components are primarily used to filter what needs to be forgotten and remembered in the input data. The second part of Figure 2 shows the flowchart of the internal state of the LSTM hidden cell corresponding to the t time step. Where the symbols ⊙ and ⊕ denote multiplication and addition in the model, and the arrows denote the direction of the flow of information.

The leading information flow of an LSTM cell can be described mathematically.

1 Forget gate: It determines what information should be forgotten.

$$f_t = \sigma\left(\boldsymbol{W_f} \times x_t + \boldsymbol{U_f} \times h_{t-1} + b_f\right) \tag{9}$$

where $\boldsymbol{f_t}$ represents the activation vector of the forget gate. σ represents the activation function *sigmoid*. x_t denotes the input vector. h_{t-1} represents the hidden state vector of the LSTM unit. $\boldsymbol{W_f}$ and $\boldsymbol{U_f}$ represent the weight matrix of the input vector and the hidden state vector, respectively. b_f is the bias vector parameter of the forget gate.

2 Input gate and candidate cell: The input gate controls the amount of information about the current input x_t and the hidden state h_{t-1} at the last moment. The candidate units contain potentially new information i_t.

$$i_t = \sigma\left(\boldsymbol{W_i} \times x_t + \boldsymbol{U_i} \times h_{t-1} + b_i\right) \tag{10}$$

$$\tilde{c}_t = \sigma\left(\boldsymbol{W_c} \times x_t + \boldsymbol{U_c} \times h_{t-1} + b_c\right) \tag{11}$$

where $\tilde{c}_t$ denote the activation vector of the input gate and candidate unit. $\boldsymbol{W_i}$, $\boldsymbol{U_i}$ represents the input vector corresponding to the input gate and the weight matrix of the hidden state vector; $\boldsymbol{U_c}$ denotes a weight matrix of input vectors and hidden state vectors of the candidate unit. b_i and b_c denote the bias vector parameters of the input gates and candidate cells.

3 Memory cell c_t update: Combining the decisions of the forgetting gate and the input gate, the long-term state of the cell is updated.

$$c_t = f_t \times c_{t-1} + i_t \times \widetilde{c_t} \tag{12}$$

where c_t is the cell state vector; c_{t-1} is the cell state at the previous moment.

4 Output gate and final hidden state: The output gate determines which information is output from the unit state to the hidden state, where the hidden state is used both for the output of the current time step and as part of the input of the next time step.

$$o_t = \sigma\left(\boldsymbol{W_o} \times x_t + \boldsymbol{U_o} \times h_{t-1} + b_o\right) \tag{13}$$

$$h_t = o_t \times \tanh\left(c_t\right) \tag{14}$$

where o_t is the output vector at the time step t; $\boldsymbol{W_o}$ and $\boldsymbol{U_o}$ denote the weight matrix of the input vector and the hidden state vector corresponding to the output gate; b_o denotes the bias vector parameter of the output gate; h_t is the hidden state vector at time step t; tanh is the activation function.

In summary, LSTM can effectively solve the problem of gradient disappearance by introducing a 'gate' mechanism to control the flow of information.

3.4 Attention

The dot product attention mechanism automatically assigns weights to different parts of the electricity load input sequence by directly capturing the correlation between query vectors and key vectors, allowing the model to focus on features that are critical for the prediction results. In this paper, we add the attention layer as an extra layer after the CNN-LSTM model and apply the attention mechanism at each time step. The principle of the corresponding attention mechanism is shown in the third part of Figure 2. The specific operation is as follows:

1 Calculate similarity score. The similarity score is obtained by taking the dot product of the transpose of the query vector and the key vectors.

$$\text{score}(\boldsymbol{Q}, \boldsymbol{K}) = \boldsymbol{Q} \cdot \boldsymbol{K}^{\boldsymbol{T}} \tag{15}$$

where **score**($\boldsymbol{Q}$, $\boldsymbol{K}$) is the similarity score matrix; $\boldsymbol{Q}$ is the query vector matrix; $\boldsymbol{K}^{\boldsymbol{T}}$ is the transpose matrix of the key vector.

2 Combined with a scaling factor: The score is scaled in order to avoid the dot product result being too large.

$$\text{scaled\_score}(\boldsymbol{Q}, \boldsymbol{K}) = \frac{\textbf{score}(\boldsymbol{Q}, \boldsymbol{K})}{\sqrt{d_k}} \tag{16}$$

where **scaled_score**($\boldsymbol{Q}$, $\boldsymbol{K}$) is the scaled score result; $\sqrt{d_k}$ is the dimension of the key vector.

3 Calculate the attention weight: The attention score is converted into a probability distribution by performing the SoftMax operation on it in order to represent the attention weight of each LSTM hidden unit.

$$\alpha = soft\max\left(\text{scaled\_score}(\boldsymbol{Q}, \boldsymbol{K})\right) \tag{17}$$

where α is the attention weight matrix corresponding to the query vector $\boldsymbol{Q}$.

4 Weighted sum: The attention output is obtained by the weighted sum of the value vectors using the attention weights. This output will contain the most relevant information to the current task.

$$\text{Attention}(\boldsymbol{Q}, \boldsymbol{K}, \boldsymbol{V}) = \boldsymbol{\alpha} \cdot \boldsymbol{V} \tag{18}$$

where Attention($\boldsymbol{Q}$, $\boldsymbol{K}$, $\boldsymbol{V}$) is the attention output; V denotes the value vector.

3.5 The principle of the proposed hybrid prediction algorithm based on SD-CNN-LSTM-Attention

A model based on CNN-LSTM-Attention is constructed for electricity load prediction, and the calculation of the proposed algorithm is shown in equation (19).

$$\hat{\boldsymbol{Y}} = f_{SD-CNN-LSTM-Attention}(\boldsymbol{Y}, \boldsymbol{D}) \tag{19}$$

where $\hat{\boldsymbol{Y}}$ denotes the residential electricity consumption in the following m hours to be predicted;

$f_{SD\text{–}CNN\text{–}LSTM\text{–}Attention}$ is the SD-CNN-LSTM-Attention model; ($\boldsymbol{Y}$, $\boldsymbol{D}$) – residential electricity consumption ($kw \cdot h$), time (*hour* or *day*), and all contain data from the past j time steps, corresponding to $(y_1, y_2, \cdots, y_j)$, $(d_1, d_2, \cdots, d_j)$ respectively. The flowchart of the CNN-LSTM-Attention prediction algorithm is demonstrated in Figure 3.

3.6 Evaluation criteria

In this study, four metrics are selected to evaluate the performance of the proposed model: mean absolute percentage error (MAPE), mean square error (MSE), root mean square error (RMSE), and mean absolute error (MAE).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \widehat{y_i}\right| \tag{20}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y_i}\right)^2 \tag{21}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y_i}\right)^2} \tag{22}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y_i}}{y_i}\right| * 100\% \tag{23}$$

where y_i is the actual value, $\widehat{y_i}$ is the predicted value, and n is the number of samples.

Figure 2 Corresponding schematic diagram of CNN, LSTM, and Attention (see online version for colours)

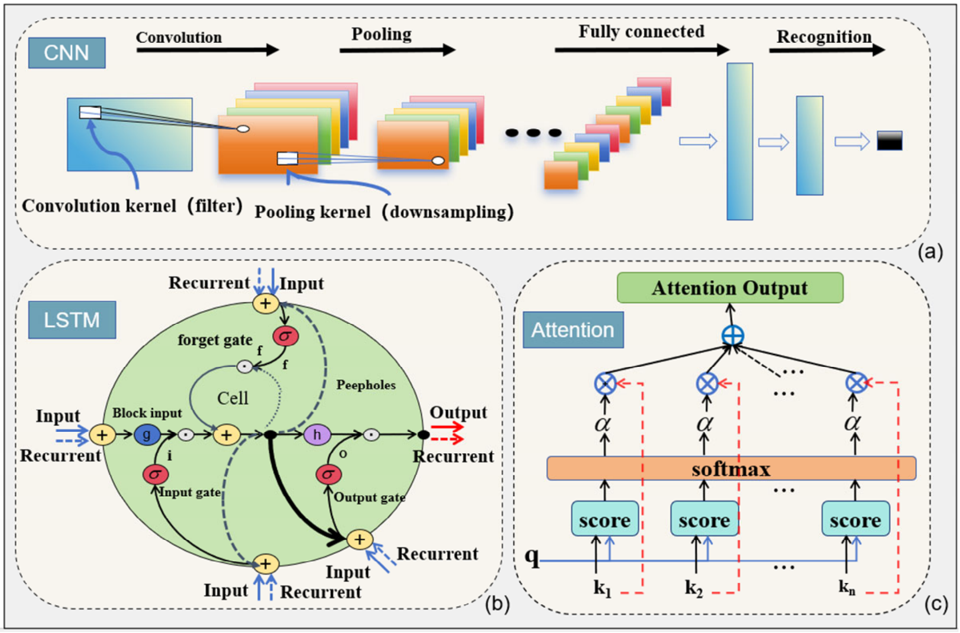
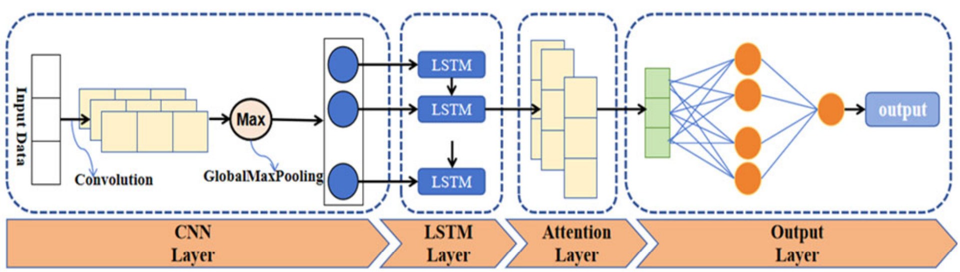


Figure 3 The flowchart of CNN-LSTM-Attention algorithm (see online version for colours)



4 Experimental results and analysis

4.1 Data sources and feature selection

The residential electrical load dataset for this study is shared in the *UCI Machine Learning repository*. Three households were selected for evaluation and validation. The electricity consumption data of the three households are aggregated and integrated hourly and daily.

Residents' electricity load is affected by multi-dimensional factors such as indoor and outdoor temperature and holidays. Different electricity consumption habits and control strategies of power supply companies also directly affect the actual electricity load of residents. Under the unified supervision of the power supply management platform, these factors are implied in the historical electricity load data characteristics. Based on the above premise, this study mainly considers time and historical electricity consumption as influencing factors.

4.2 Data processing and analysis

1 Data pre-processing

Before training the model, the data is preprocessed with outlier detection and missing data filling to ensure data integrity. The outlier is detected based on the traditional 3-sigma rule, and the missing data is filled with Spline interpolation. Subsequently, the data is normalised using Min-Max Scaling to scale the data to the range [0–1].

$$x_{scaled} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{24}$$

2 SD

In order to obtain the best model accuracy, the original electricity load data is decomposed into three parts: seasonal, trend, and residual components based on SD. The output of the seasonal decomposition diagram of electricity load is shown in Figure 4. Here, the left and right are the hourly and daily electricity load decomposition results, respectively.

4.3 The optimisation of hyperparameters

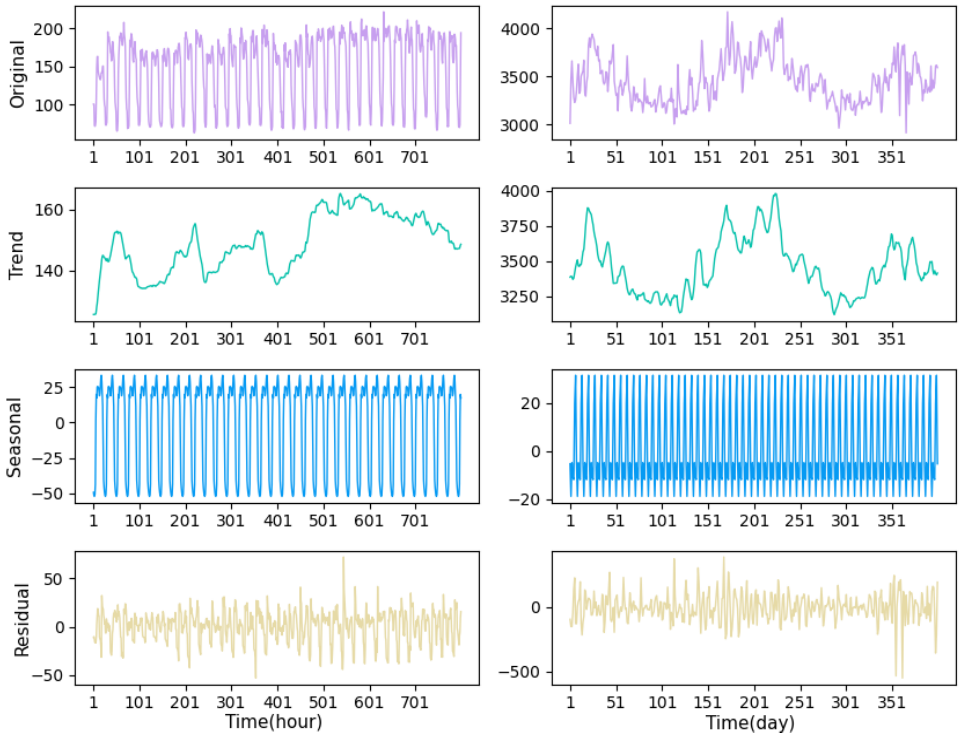
Hyperparameters play an important role in the performance of the proposed model, and the grid search method is employed in this paper to find the optimal hyperparameters. By defining a space of all possible hyperparameters and then evaluating the performance of each combination successively, the best performing one is ultimately selected. The approach is simple, intuitive and easy to implement, and can find the global optimal solution.

4.4 Prediction results

4.1.1 Hourly electricity load prediction

Figure 5 shows the electricity load prediction curve and performance comparison with different algorithms for three residential houses. It can be seen from Figure 5 that residents usually use more electrical appliances, such as lighting and cooking appliances, after getting up in the morning, and the electricity consumption increases accordingly. Therefore, the electricity load has a clear upward trend from 5:00 to 7:00. During the day, residents go out to work, and the electricity load may decrease. Therefore, the electricity load is in a relatively stable state from 9:00 to 17:00, and main appliances such as refrigerators maintain essential electricity consumption. From 17:00 to 19:00, electricity consumption increases again as residents return home from work. At dinner time and evening leisure activities, we will use more electrical appliances, such as a TV, computer, etc.

Figure 4 The hourly and daily electricity load plots of SD of household A (see online version for colours)



The residential electricity consumption pattern is relatively regular, with noticeable peaks in the morning and evening: 7:00 a.m.–9:00 a.m. and 6:00 p.m.–10:00 p.m. The troughs are usually in the middle of the night and during the afternoon work hours, with a smoother in the middle part. From Figure 5, we can see that the SD-CNN-LSTM-Attention algorithm fits better in all three time periods (morning, daytime, and evening), and the curves are in good agreement with the actual values. Especially in the peak periods of morning and evening, SD-CNN-LSTM-Attention can capture the fluctuating changes in electricity load better than simple models such as SVR. As shown in Table 1, the MAE, RMSE, and MAPE of the proposed algorithm based on SD-CNN-LSTM-Attention for predicting the electricity load of three households are minimal compared to other algorithms. In residents A and C, the SVR performs the worst in all the metrics and has the highest error. Although both CNN-LSTM and Autoformer offer strong capabilities in time-series prediction, the SD-CNN-LSTM-Attention model consistently performs better, especially during critical periods of morning and evening peak electricity consumption, where capturing sudden fluctuations in load demand is essential.

The electricity load prediction curves and performance comparisons for different models applied to Household A across three periods – working days, weekends, and national day – are illustrated in Figure 6. As shown, the electricity load on weekends and National Day is generally higher than on working days, with similar trends observed across the forecast curves for all three periods.

Table 1 Performance performances of the three residents predicted by different models

Model	*Residence A*			*Residence B*			*Residence C*		
	MAPE (%)	*MAE* (kW·h)	*RMSE* (kW·h)	*MAPE* (%)	*MAE* (kW·h)	*RMSE* (kW·h)	*MAPE* (%)	*MAE* (kW·h)	*RMSE* (kW·h)
SVR	9.70	10.36	11.40	7.93	6.65	8.42	9.02	9.60	11.64
CNN	4.00	5.38	7.60	5.72	5.92	7.38	6.90	8.22	11.07
GRU	4.09	5.49	7.79	5.05	5.35	7.28	5.96	6.96	9.66
LSTM	5.45	6.89	8.82	5.13	5.43	7.33	5.37	6.39	9.64
CNN-LSTM	4.06	5.20	7.23	4.86	5.34	6.87	7.50	9.12	12.07
Autoformer	4.87	6.35	8.50	7.53	6.93	9.24	10.91	11.47	13.42
Proposed	*3.56*	*4.78*	*6.81*	*4.41*	*4.86*	*6.58*	*4.91*	*5.83*	*8.66*

Table 2 Different period performances of the electricity consumption predicted by different models

Model	*Weekdays*			*Weekend*			*National day*		
	MAPE (%)	*MAE (kW·h)*	*RMSE (kW·h)*	*MAPE (%)*	*MAE (kW·h)*	*RMSE (kW·h)*	*MAPE (%)*	*MAE (kW·h)*	*RMSE (kW·h)*
SVR	9.70	10.36	11.40	7.29	8.25	9.71	6.22	7.55	9.20
CNN	4.00	5.38	7.60	2.93	3.85	4.94	4.88	5.67	7.54
GRU	4.09	5.49	7.79	3.41	4.60	5.83	5.17	5.94	8.34
LSTM	5.45	6.89	8.82	4.47	5.51	6.88	5.29	5.97	7.26
CNN-LSTM	4.06	5.20	7.23	3.85	5.18	6.11	4.81	5.36	6.65
Autoformer	4.87	6.35	8.50	5.65	7.63	9.55	6.70	7.86	9.41
Proposed	*3.56*	*4.78*	*6.81*	*2.36*	*3.19*	*4.02*	*3.92*	*4.55*	*6.54*

From Table 2, it is evident that the SD-CNN-LSTM-Attention algorithm outperforms other models, achieving the lowest MAE, RMSE, and MAPE across all three periods. In contrast, the Autoformer model performs poorly compared to the other DL models. Additionally, when compared to the regression-based model SVR, the proposed SD-CNN-LSTM-Attention model improves residential electricity consumption prediction accuracy by 6.14%, 4.93%, and 2.3% for working days, weekends, and National Day, respectively.

Notably, the best prediction performance occurs during the weekend, where the proposed model achieves a minimum MAPE of 2.36%, indicating that the model adapts well to periods of higher variability in household energy usage. This suggests that the proposed model is highly effective at capturing the unique consumption patterns seen during non-working periods. The overall improvements demonstrate the robustness of the model in different real-world scenarios, providing valuable insights for energy management and consumption optimisation in residential environments.

Figure 5 Prediction curves and performance comparison of weekday electricity load, (a) Residence A (b) Residence B (c) Residence C (see online version for colours)

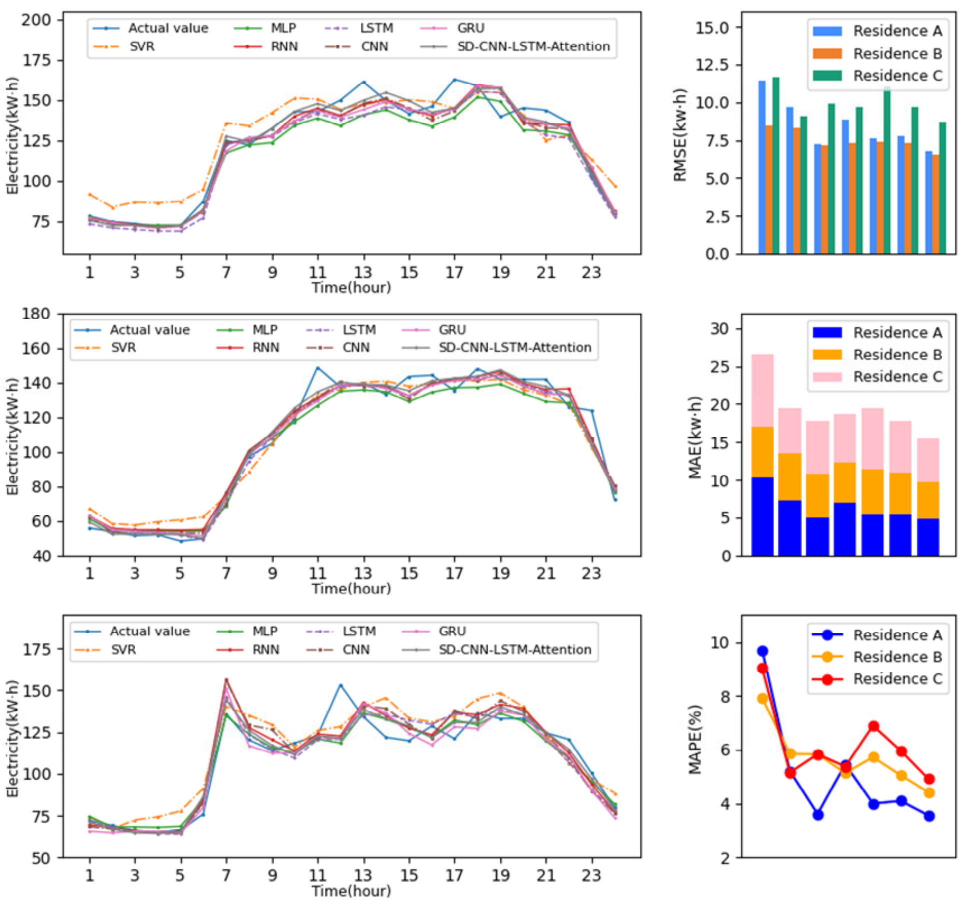
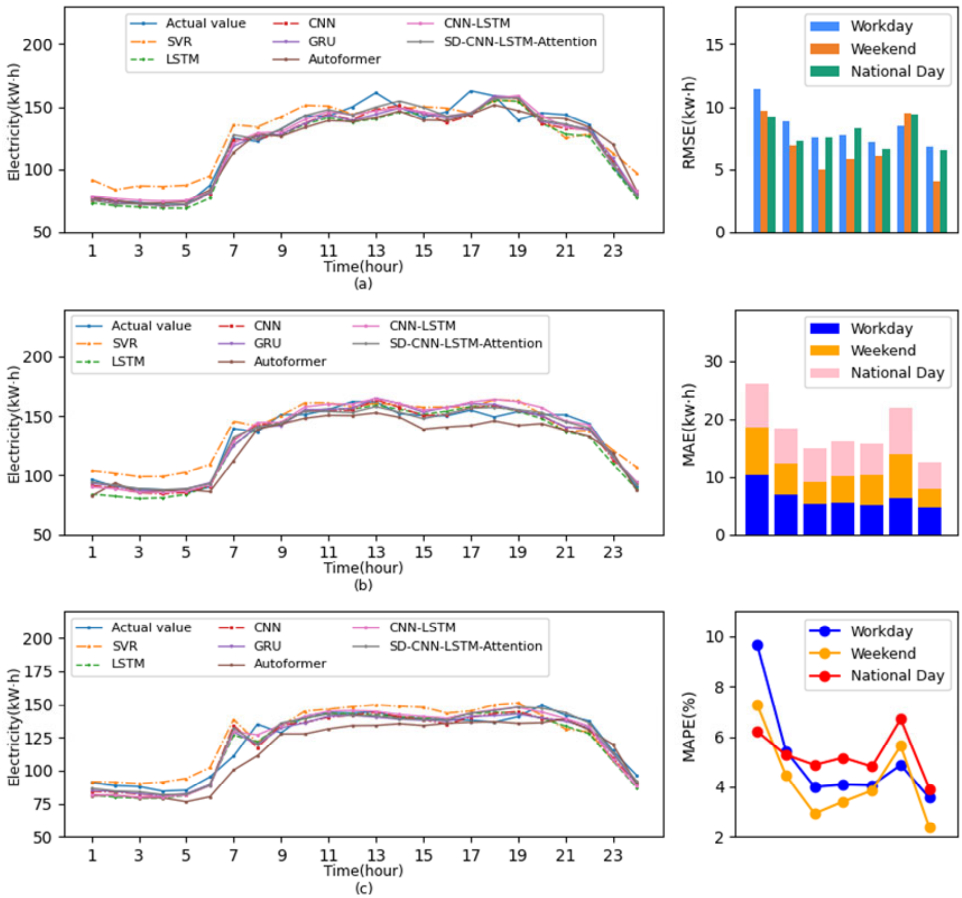


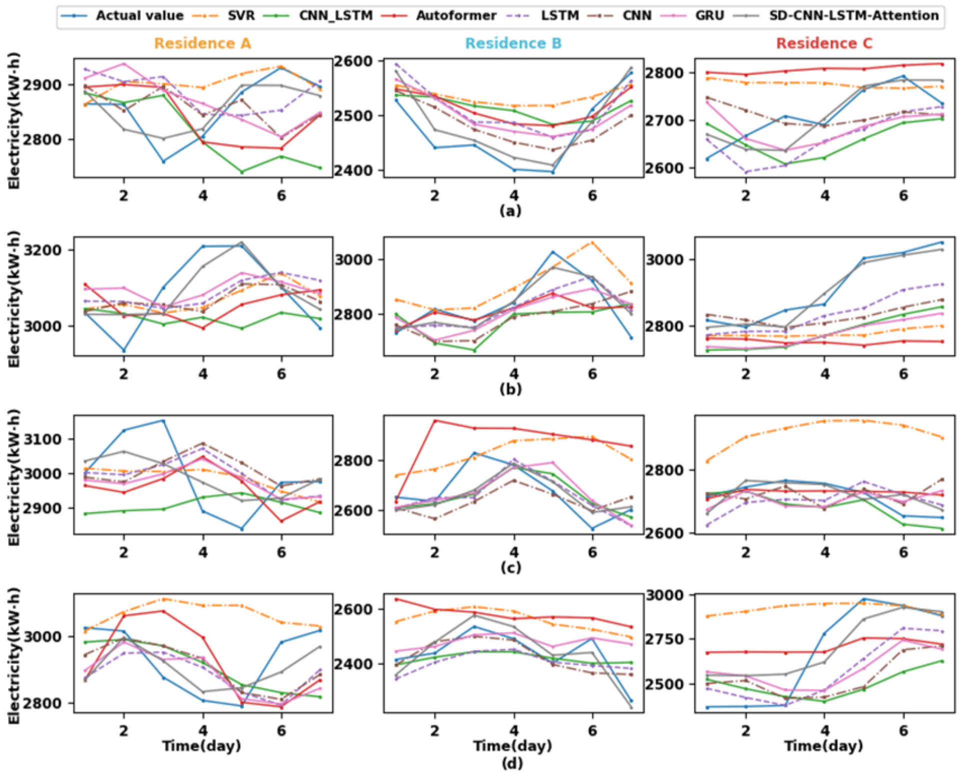
Figure 6 Electricity load forecasting curve and performance comparison of household A, (a) workday (b) weekend (c) national day (see online version for colours)



4.1.2 Daily electricity load prediction

The daily electricity load dataset is divided into four seasons, spring, summer, autumn and winter, and the prediction results of the three residents are shown in Figure 7. It can be seen that the electricity consumption of the three households fluctuates less in spring. The decline in electricity consumption in the early stage may be due to the gradual rise of temperature in spring and the reduction of heating demand consumption. It could also be that the increase in daylight hours in spring reduces the need for indoor lighting. The subsequent rise in electricity consumption may be due to the fact that residents begin to use air conditioning and refrigeration or other electrical appliances as the temperature rises; it may also be due to the increase in evening activities as the hours of daylight increase, which leads to a subsequent increase in electricity consumption. As can be seen from Table 3, the proposed SD-CNN-LSTM-Attention prediction model performs the best in all three households with the lowest MAE, MAPE and RMSE, and MAPE reaches a minimum of 0.9% for resident B, significantly lower than CNN-LSTM and Autoformer.

Figure 7 Electricity load prediction curves and performance comparison of three residents in four seasons, (a) spring (b) summer (c) autumn (d) winter (see online version for colours)



During the summer months, all three residents show a significant increase in electricity consumption. The change of resident A is relatively stable, but there is an obvious peak in the middle period, which may correspond to a large increase in air conditioning usage caused by extremely high-temperature weather. Moreover, there is usually a summer vacation in summer. Under the effect of holidays, family members spend more time at home, and the frequency of electrical appliances increases, resulting in increased electricity load. As can be seen from Table 2, SD-CNN-LSTM-Attention shows significant advantages in all three households, and it is especially prominent in residents B and C.

In autumn, the electricity load of the three residents shows different trends. The electricity load of resident A rises and falls with large fluctuations. There is a clear peak in residents B and C. The reason for this large change may be that air conditioners are used less in autumn when the temperature gradually decreases, but the heating demand has not yet fully appeared. From Table 2, it can be seen that the error of the proposed SD-CNN-LSTM-Attention model is significantly lower than that of other algorithms in all residents, with higher stability and accuracy.

Table 3 The prediction performance of different models in three residents in four different seasons

Season	*Model*	*Resident A*			*Resident B*			*Resident C*		
		MAPE	*MAE*	*RMSE*	*MAPE*	*MAE*	*RMSE*	*MAPE*	*MAE*	*RMSE*
Spring	SVR	1.6	44.9	66.5	2.8	68.8	80.1	2.7	72.0	89.2
	CNN	2.0	59.1	76.2	1.9	48.6	52.7	1.9	51.8	65.3
	GRU	2.6	76.3	83.4	2.2	56.0	59.0	2.2	59.5	69.8
	LSTM	2.1	61.5	74.9	2.3	56.5	62.0	2.1	59.5	67.2
	CNN-LSTM	2.2	64.2	70.1	2.1	52.8	55.5	2.6	70.6	77.2
	Autoformer	2.0	58.0	67.5	2.7	68.0	72.2	3.5	95.9	107.8
	Proposed	*0.9*	*26.6*	*29.4*	*0.9*	*22.9*	*27.2*	*1.2*	*32.7*	*40.0*
Summer	SVR	2.7	85.3	98.1	3.2	90.1	110.0	4.5	134.4	160.9
	CNN	2.4	74.6	94.5	3.7	107.4	124.4	3.1	93.0	114.6
	GRU	2.7	82.9	95.0	2.7	79.0	94.3	4.6	135.7	148.2
	LSTM	2.8	88.4	99.0	1.8	53.8	70.8	2.5	76.3	89.8
	CNN-LSTM	2.9	90.8	103.8	2.2	65.5	81.9	4.5	132.5	142.1
	Autoformer	4.3	136.5	167.7	2.7	76.9	96.5	5.3	158.9	188.1
	Proposed	*1.3*	*40.1*	*52.0*	*1.3*	*36.6*	*45.6*	*0.7*	*21.9*	*25.9*

Table 3 The prediction performance of different models in three residents in four different seasons (continued)

Season	*Model*	*Resident A*			*Resident B*			*Resident C*		
		MAPE	*MAE*	*RMSE*	*MAPE*	*MAE*	*RMSE*	*MAPE*	*MAE*	*RMSE*
Autumn	SVR	3.0	90.5	104.7	6.1	161.4	193.9	7.5	203.1	210.8
	CNN	3.2	97.7	126.1	2.7	73.4	91.5	1.6	45.6	59.9
	GRU	3.4	102.6	117.4	2.7	73.4	91.9	1.8	49.1	58.1
	LSTM	3.3	98.7	116.6	2.4	66.7	85.4	2.1	57.0	59.5
	CNN-LSTM	3.5	106.9	119.2	2.3	62.1	81.7	1.3	37.2	45.3
	Autoformer	4.4	136.8	163.5	7.8	207.0	236.9	1.2	32.1	42.8
	Proposed	*2.0*	*62.5*	*71.8*	*1.8*	*48.7*	*67.8*	*1.0*	*28.7*	*36.3*
Winter	SVR	5.2	138.6	184.6	10.6	124.4	137.7	4.8	257.4	355.6
	CNN	1.9	95.8	107.8	8.1	46.8	63.3	3.2	226.4	267.6
	GRU	2.2	104.2	121.5	8.1	53.5	83.6	3.5	220.9	240.0
	LSTM	2.6	106.3	116.5	5.2	64.6	74.7	3.6	145.7	189.2
	CNN-LSTM	3.5	105.1	113.9	2.5	60.2	74.7	9.2	259.4	302.6
	Autoformer	3.6	106.8	110.6	6.0	145.2	164.0	8.7	226.0	237.9
	Proposed	*1.6*	*63.8*	*77.7*	*1.6*	*40.3*	*41.7*	*2.1*	*119.0*	*137.1*

In winter, the electricity load of resident A fluctuates greatly, falling in the middle period and rising in the later period. The electricity load of resident B first rises and then falls. There is a clear rise and fall in resident C. This change may be due to the frequent use of heating equipment (electric heaters, air conditioners) in winter, when the temperature is lower. In addition, the long nights and short days in winter increase the demand for lighting, and the frequency of use of electrical appliances such as washing machines and water heaters increases in cold weather. Although it is difficult to predict accurately in winter, Table 3 shows that SD-CNN-LSTM-Attention outperforms other algorithms in all residents. Autoformer performs well in long-term trend prediction but lags in handling the sharper load variations during peak hours.

5 Conclusions

In this paper, we propose a multi-scale electricity load prediction model based on SD-CNN-LSTM-Attention. The model accurately captures critical information in historical data. The proposed model has a smaller MAPE and smaller error distribution range in all cases. From the perspective of prediction results, the proposed model has higher prediction accuracy for the three residents. Specifically, in the workday prediction of the three residents, our model achieves a MAPE of 3.56% in residence A, 4.78% in residence B and 6.81% in residence C, which significantly outperforms the performance of other traditional models such as CNN, GRU and LSTM. In the weekend prediction, the MAPE of the model is 2.36%, 3.19% and 4.02% in residential A, B and C, respectively, which also shows high accuracy. The variation in electricity consumption of residential buildings is linearly related to the activity of the day. Compared with working days, the change in electricity consumption on weekends and holidays is backward with the delay of the start time of residential activities. The model exhibits superior robustness. According to the seasonal characteristics, electricity consumption is closely related to the electricity consumption behaviour caused by climate change. In the daily forecasting results of the four seasons, the fluctuation of electricity load data in winter is increased due to the large temperature change. By considering the prediction results of seasonal characteristics, it provides a basic guarantee for the subsequent analysis of the difference in electricity load between the four seasons.

Overall, the proposed method achieves accurate electricity consumption estimation of residents in different situations, and the model is more suitable for different data sets. The seasonal pattern of electricity load in practical applications may be affected by many factors, such as energy policy, economic development, etc., resulting in fluctuations or changes in the seasonal pattern. In the following work, the influencing factors of electricity load will be further studied to improve the accuracy of the mode.

Data availability

Sequence data of this study have been deposited in the following WebURL link: https://archive.ics.uci.edu/dataset/321/electricityloaddiagrams20112014.

Declarations

All authors declare that they have no conflicts of interest.

References

Ahmad, N., Ghadi, Y., Adnan, M., et al. (2022) 'Load forecasting techniques for power system: research challenges and survey', *IEEE Access*, Vol. 10, pp.71054–71090, https://doi.org/10.1109/ACCESS.2022. 3187839.

Al-Alimi D., Alrassas A.M., Al-Qaness M.A.A., et al. (2023) 'TLIA: time-series forecasting model using long short-term memory integrated with artificial neural networks for volatile energy markets', *Applied Energy*, Vol. 343, p.14, https://doi.org/10.1016/j.apenergy.2023.121230.

Bareth, R., Yadav, A., Gupta, S., et al. (2024) 'Daily average load demand forecasting using LSTM model based on historical load trends', *IET Generation Transmission & Distribution*, Vol. 18, No. 5, pp.952–962, https://doi.org/10.1049/gtd2.13132.

Bashir, T., Chen, H.Y., Tahir, M.F., et al. (2022) 'Short term electricity load forecasting using hybrid prophet-LSTM model optimized by BPNN', *Energy Reports*, Vol. 8, pp.1678–1686, https://doi.org/10.1016/j.egyr.2021.12.067.

Butt, F.M., Hussain, L., Mahmood, A., et al. (2020) 'Artificial Intelligence based accurately load forecasting system to forecast short and medium-term load demands', *Mathematical Biosciences and Engineering*, Vol. 18, No. 1, pp.400–425, https://doi.org/10.3934 /mbe.2021022.

Dong, B., Li, Z.X., Rahman, S.M.M., et al. (2016) 'A hybrid model approach for forecasting future residential electricity consumption', *Energy and Buildings*, Vol. 117, pp.341–351, https://doi.org/10.1016/j.enbuild.2015.09.033.

Dudek, G., Pelka, P. and Smyl, S. (2022) 'A hybrid residual dilated LSTM and exponential smoothing model for midterm electric load forecasting', *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 33, No. 7, pp.2879–2891, https://doi.org/10.1109 /TNNLS.2020.3046629.

Eskandari, H., Imani, M. and Moghaddam, M.P. (2021) 'Convolutional and recurrent neural network based model for short-term load forecasting', *Electric Power Systems Research*, Vol. 195, p.14, https://doi.org/10.1016/j.epsr.2021.107173.

Fekri, M.N., Patel, H., Grolinger, K., et al. (2021) 'Deep learning for load forecasting with smart meter data: online adaptive recurrent neural network', *Applied Energy*, Vol. 282, p.17, https://doi.org/10.1016/ j.apenergy.2020.116177.

Heydari, A., Garcia, D.A., Keynia, F., et al. (2019) 'Renewable energies generation and carbon dioxide emission forecasting in microgrids and national grids using GRNN-GWO methodology', *Energy Procedia*, Vol. 159, pp.154–159, https://doi.org/10.1016/j.egypro .2018.12.044.

Jiang, P., Li, R.R., Liu, N.N., et al. (2020) 'A novel composite electricity demand forecasting framework by data processing and optimized support vector machine', *Applied Energy*, Vol. 260, https://doi.org/10.1016/j.apenergy.2019.114243.

Li, J., Deng, D.Y., Zhao, J.B., et al. (2021) 'A novel hybrid short-term load forecasting method of smart grid using MLR and LSTM neural network', *IEEE Transactions on Industrial Informatics*, Vol. 17, No. 4, pp.2443–2452, https://doi.org/10.1109/TII.2020.3000184.

Lv, L.L., Wu, Z.Y., Zhang, J.H., et al. (2022) 'A VMD and LSTM Based Hybrid Model Of Load Forecasting For Power Grid Security', *IEEE Transactions on Industrial Informatics*, Vol. 18, No. 9, pp.6474–6482, https://doi.org/10.1109/TII.2021.3130237.

Pan, Y.R. and Jia, C.Y. (2025) 'A time series-based method for predicting electricity demand in industrial parks', *International Journal of Energy Technology and Policy*, Vol. 20, Nos. 1/2, pp.95–109, https://doi.org/10.1504/ IJETP.2025.144301.

Rekhade, R. and Sakhare, D.K. (2021) 'Forecasting sector-wise electricity consumption for India using various regression models', *Current Science*, Vol. 121, No. 3, pp.365–371, https://doi.org/10.18520/cs/v121/i3/365-371.

Tan, M., Yuan, S.P., Li, S.H., et al. (2020) 'Ultra-short-term industrial power demand forecasting using LSTM based hybrid ensemble learning', *IEEE Transactions on Power Systems*, Vol. 35, No. 4, pp.2937–2948, https://doi.org/10.1109/TPWRS.2019.2963109.

Tang, Z.H., Yin, H., Yang, C.Y., et al. (2021) 'Predicting the electricity consumption of urban rail transit based on binary nonlinear fitting regression and support vector regression', *Sustainable Cities and Society*, Vol. 66, p.14, https://doi.org/10.1016/j.scs.2020.102690.

Wang, Y.Y., Chen, J., Chen, X.Q., et al. (2021) 'Short-term load forecasting for industrial customers based on TCN-LightGBM', *IEEE Transactions on Power Systems*, Vol. 36, No. 3, pp.1984–1997, https://doi.org/10.1109/TPWRS.2020.3028133.

Wen, L.L., Zhou, K.L., Yang, S.L., et al. (2019) 'Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting', *Energy*, Vol. 171, pp.1053–1065, http://doi.org/10.1016/j.energy.2019.01.075.

Xu, H.F., Hu, F.H., Liang, X.H., et al. (2024) 'A framework for electricity load forecasting based on attention mechanism time series depthwise separable convolutional neural network', *Energy*, Vol. 299, p.14, https://doi.org/10.1016/j.energy.2024.131258.

Yang Y.D., Li S.F., Li W.Q., et al. (2018) 'Power load probability density forecasting using Gaussian process quantile regression', *Applied Energy*, Vol. 213, pp.499–509, https://doi.org/10.1016/j.apenergy.2017.11.035.

Yang, M., Liu, D.X., Su, X., et al. (2023) 'Ultra-short-term load prediction of integrated energy system based on load similar fluctuation set classification', *Frontiers in Energy Research*, Vol. 10, p.14, https://doi.org/10.3389/fenrg.2022.1037874.

Zhang, J.L., Wei, Y.M., Li, D.Z., et al. (2018) 'Short term electricity load forecasting using a hybrid model', *Energy*, Vol. 158, pp.774–781, https://doi.org/10.1016/j.energy.2018.06.012.

Zhao, H.R. and Guo, S. (2016) 'An optimized grey model for annual power load forecasting', *Energy*, Vol. 107, pp.272–286, https://doi.org/10.1016/j.energy.2016.04.009.

Nomenclature

CNN	Convolutional neural network
LSTM	Long short-term memory
Attention	Attention mechanism
R2	Coefficient of determination
MAPE	Mean absolute percentage error
MLR	Multiple linear regression
GPQR	Gaussian process quartile regression
SVR	Support vector regression
BNFR	Binary nonlinear fitted regression
DL	Deep learning
RNN	Recurrent neural network
ANN	Artificial neural network
DRNN-LSTM	Deep recurrent neural network with long short-term memory
USTLF	Ultra-short-term load forecasting
STLF	Short-term load forecasting
MTLF	Medium-term load forecasting
LTLF	Long-term forecasting
EMD	Empirical mode decomposition
BiLSTM	Bidirectional LSTM
ARIMA	Autoregressive integrated moving average
WNN	Wavelet neural network
SD-CNN-LSTM-Attention	Seasonal decomposition-convolutional neural network-long and short-term memory network-attention mechanism
MAPE	Mean absolute percentage error
MSE	Mean square error
RMSE	Root mean square error
MAE	Mean absolute error