frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 31 July 2023
DOI 10.3389/fenrg.2023.1227979

Check for updates

OPEN ACCESS

EDITED BY
Zhengmao Li,
Nanyang Technological University, Singapore

REVIEWED BY
Guibin Wang,
Shenzhen University, China
Jing Qiu,
The University of Sydney, Australia
Bing Luo,
Duke Kunshan University, China

*CORRESPONDENCE
Jiaqi Ruan,
jiaqiruan@link.cuhk.edu.cn

RECEIVED 24 May 2023
ACCEPTED 14 July 2023
PUBLISHED 31 July 2023

CITATION
Liao W, Ruan J, Xie Y, Wang Q, Li J, Wang R and Zhao J (2023), Deep learning time pattern attention mechanism-based short-term load forecasting method.
Front. Energy Res. 11:1227979.
doi: 10.3389/fenrg.2023.1227979

Deep learning time pattern attention mechanism-based short-term load forecasting method

Wei Liao[1], Jiaqi Ruan[2,3]*, Yinghua Xie[1], Qingwei Wang[1], Jing Li[1], Ruoyu Wang[1] and Junhua Zhao[2,3]

[1]Shenzhen Power Supply Co., Ltd., Shenzhen, China, [2]School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China, [3]Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen, China

Accurate load forecasting is crucial to improve the stability and cost-efficiency of smart grid operations. However, how to integrate multiple significant factors for enhancing load forecasting performance is insufficiently investigated in previous studies. To fill the gap, this study proposes a novel hybrid deep learning model for short-term load forecasting. First, the long short-term memory network is utilized to capture patterns from historical load data. Second, a time pattern attention (TPA) mechanism is incorporated to improve feature extraction and learning capabilities. By discerning valuable features and eliminating irrelevant ones, the TPA mechanism enhances the learning process. Third, fully-connected layers are employed to integrate external factors such as climatic conditions, economic indicators, and temporal aspects. This comprehensive approach facilitates a deeper understanding of the impact of these factors on load profiles, leading to the development of a highly accurate load forecasting model. Rigorous experimental evaluations demonstrate the superior performance of the proposed approach in comparison to existing state-of-the-art load forecasting methodologies.

KEYWORDS

load forecasting, deep learning, time pattern attention, smart grid, data driven

1 Introduction

With the increasing complexity of electrical infrastructure, the power industry has embraced the emergence of smart grids (Wang et al., 2018), which integrate conventional power system equipment with advanced intelligent digital communication devices, aiming to enhance system's performance, safety, and reliability (Yu et al., 2014). The integration of intelligent electronic devices into smart grids enables inter-device communication and real-time data sharing with control centers, resulting in the accumulation of significant amounts of data. Although initially unproductive, this data can be utilized for system assessments, ultimately improving the operational performance of smart grids.

As the energy internet emerges (Wang et al., 2019a) and energy demands escalate, there is a growing emphasis on energy conservation, leading to an amplified need for load forecasting, particularly in the commercial and industrial sectors. Load forecasting offers several advantages, including the mitigation of supply-demand imbalances and optimization of energy utilization benefits. With the increasing availability of big data, artificial intelligence techniques play a crucial role in load forecasting (Ruan et al., 2023b).

Short-term load forecasting, which commonly predicts loads within a range from hours to weeks, enables utilities and power plants to adjust generation in response to market demands. Research has shown that a mere 1% reduction in load forecasting errors can lead to an annual operating cost reduction of £10 million for a British power company (Gilanifar et al., 2019). Additionally, accurate load forecasts facilitate the implementation of dynamic pricing structures in the electricity market (Ruan et al., 2023b). However, due to the intricacies and uncertainties associated with power demands (Ruan et al., 2022a), load forecasting remains some significant challenges. To address it, recent advancements in data analysis techniques and data collection systems, such as smart meters (Li et al., 2021), have the potential to greatly enhance load forecasting accuracy. Specifically, machine learning-based load forecasting methods, including autoregressive integrated moving average (ARIMA), multiple linear regression (Yu et al., 2014), Gaussian process regression (Akorede et al., 2010), support vector regression (SVR) (Hossain et al., 2019), artificial neural networks (ANN) (Virote and Neves-Silva, 2012; Candanedo et al., 2017), and deep neural networks (DNN) (Menezes et al., 2014), have gained substantial attention.

Deep learning techniques have proven to be effective in developing highly accurate load forecasting models. For example, the literature (Wang et al., 2019a) proposed a deep belief network (DBN)-based model for short-term load forecasting, which is able to learn probability distribution so as to determine future load profiles. Other studies recommended the use of self-recurrent wavelet neural networks (SRWNN) for load forecasting in microgrids by introducing a Levenberg-Marquardt learning algorithm to improve the forecast accuracy for highly volatile and non-smooth time series of microgrid electricity load (Chitsaz et al., 2015), the employment of multi-layer perceptron (MLP) for non-residential building electric load forecasting with analyses of most relevant features (Massana et al., 2015), and the application of recurrent neural networks (RNN) for short-term load forecasting that can effectively handle time-series data (Wen et al., 2022). These models utilize historical data in digital formats to predict future electric load variations.

However, the extensive integration of renewable energy sources (Yang et al., 2021), the widespread adoption of electric vehicles (Hartvigsson et al., 2021; Yang et al., 2022), the large-scale deployment of energy storage systems (Zhang et al., 2021), and emerging cyber threats (Ruan et al., 2023a) have introduced greater uncertainty and disturbances in short-term load forecasting (Wang et al., 2019b). To address the limitations of existing models in capturing these dynamic changes, this paper proposes a deep learning-based approach that incorporates a time pattern attention (TPA) mechanism to construct a highly accurate load forecasting model. The contributions of this article can be summarized as follows.

- To our knowledge, it is the first study to propose an adaptive short-term load forecasting framework that can accommodate various critical features, thereby facilitating accurate forecasting results.
- A specific deep learning-based hybrid model is proposed. It incorporates the long short-term memory (LSTM) network and the TPA mechanism as well as various deep learning techniques that can effectively utilize historical load data and external factors (e.g., climate, economy, and date) to discern dynamic load trends for load forecasting.
- Comprehensive experiments are conducted by using Panama data to analyze of the proposed model and compare it with alternative state-of-the-art load forecasting models. The results demonstrate the superior performance of the proposed method.

The remainder of the paper is organized as follows. Section 2 introduces preliminaries of load forecasting, including its importance, features, and challenges. Section 3 elaborates on deep learning-based TPA mechanism and the proposed short-term load forecasting model as well as the overall framework. Section 4 demonstrates and discusses the case studies on the proposed load forecasting model. At last, section 5 summarizes the article.

2 Preliminaries of load forecasting

2.1 Importance of load forecasting

Load forecasting has consistently been a vital concern for the power industry (Li et al., 2023b), as forecasting data enables power generation and load management departments to bolster their performance and reliability. In addition to economic and environmental considerations, load forecasting serves the following essential functions.

1) Comprehending load profiles allows power companies to devise rational electricity demand plans for customers, make economically prudent decisions, and mitigate risks for the organization.
2) Load forecasting aids power generation enterprises in anticipating potential resource requirements, facilitating the storage of necessary resources, such as fuel, to guarantee an uninterrupted power supply.
3) It assists in projecting the evolution of electricity generation within society and determining the need for future power plants, thus guiding power companies in preparations for constructing additional generating units to accommodate escalating electricity demands.
4) It contributes to the analysis and planning of power system maintenance;
5) By reducing energy production shortages and surpluses, load forecasting helps power companies minimize economic and energy losses.

2.2 Load forecasting features

The outcomes of load forecasting techniques are influenced by various factors. To obtain accurate predictions, it is crucial to consider the relevant factors of the dataset and use them appropriately. Numerous variables may affect the load forecasting performance. Here are some factors related to load forecasting.

2.2.1 Time factors

Due to the deductive nature of electric load over time (Ruan et al., 2022b), the most critical aspect in forecasting is time. As the available data generated by various devices (such as smart meters, sensors, data servers, and other equipment) is time-series data, the importance of time in forecasting is paramount. Time has different attributes that can be used for prediction, such as "day of the week," "week of the month," "month of the season," and so on (Ruzic et al., 2003). The selection of time horizon in forecasting is also a key factor (Lusis et al., 2017). Employing a more extended time range allows for the utilization of additional historical data.

2.2.2 Climate factors

Climate stands as a paramount factor in load forecasting, as it substantially influences both the agricultural sector and household consumption behaviours. Specifically, the usage patterns of various electrical devices, contingent upon weather-related warmth or coldness, can give rise to distinct load profiles. Consequently, load forecasting models may incorporate weather data sourced from the nearest accessible meteorological station, encompassing variables such as temperature, precipitation, humidity, dew point temperature, solar radiation intensity, wind speed, wind chill index (WCI), temperature-humidity index (THI), and other meteorological parameters.

2.2.3 Other factors

Economic determinants, including market stability, electricity price fluctuations, load control, and industrial growth rates, profoundly influence system average load and peak demand (Li et al., 2023a). Moreover, the physical attributes of structural, housing, or surrounding areas exhibit distinct load characteristics. Load forecasting for edifices and other structures can generally be executed utilizing building attribute parameters, such as the number of rooms and floors, window-to-wall ratio, orientation, window-wall thermal efficiency, fresh air volume, and occupant density.

2.3 Challenges in load forecasting

For a long time, researchers have been dedicated to improving load forecasting techniques. However, when it comes to the specific modeling of load forecasting, there are still some obstacles.

First, weather is a key factor when performing load forecasting. Since the weather cannot be accurately estimated, it is impossible to accurately determine its impact on the load. Sudden weather changes can have significant effects on the expected load characteristics.

Second, the variety of meters utilized by consumers considerably influences load forecasting performance. Consumers employ an array of meters, encompassing smart and conventional meters, each with distinct measurement frequencies. As meter measurement frequencies and customer consumption behavior diverge, employing combined data for load forecasting may result in significant prediction errors.

Third, to further refine load forecasting, a series of other complex factors can be considered, but this adds to the difficulty of accommodating multiple variables, rendering the selection of an appropriate load forecasting model extremely challenging.

Fourth, since power systems may experience faults, power outages, and other intermittent events during dynamic operation, load forecasting models cannot account for such sudden occurrences, which also affect the load forecasting performance.

Fifth, consumer electricity demand is influenced by changes in economic market conditions or tariff changes. Although these economic factors significantly impact load forecasting outcomes, they are often overlooked by existing load forecasting methodologies.

3 Proposed short-term load forecasting model based on the time pattern attention mechanism

3.1 Long short-term memory network

In constructing a load forecasting model, the time dimension emerges as a critical factor influencing forecasting performance. Dynamic patterns can be discerned from historical load time series data. Consequently, employing neural networks adept at handling time series data can effectively extract inherent feature information and augment model accuracy. Long short-term memory (LSTM) networks, a unique variant of recurrent neural networks (RNNs), exhibit a natural advantage in processing sequential data (Hochreiter and Schmidhuber, 1997), as shown in Figure 1. The LSTM network manipulates the cell state through internal input gates, output gates, and forget gates, ultimately yielding their hidden state, as demonstrated in the ensuing equations:

$$\boldsymbol{i}_t = sigmoid\left(\boldsymbol{W}_{\boldsymbol{x}_i}\boldsymbol{x}_t + \boldsymbol{W}_{\boldsymbol{h}_i}\boldsymbol{h}_{t-1}\right) \tag{1}$$

$$\boldsymbol{o}_t = sigmoid\left(\boldsymbol{W}_{\boldsymbol{x}_o}\boldsymbol{x}_t + \boldsymbol{W}_{\boldsymbol{h}_o}\boldsymbol{h}_{t-1}\right) \tag{2}$$

$$\boldsymbol{f}_t = sigmoid\left(\boldsymbol{W}_{\boldsymbol{x}_f}\boldsymbol{x}_t + \boldsymbol{W}_{\boldsymbol{h}_f}\boldsymbol{h}_{t-1}\right) \tag{3}$$

$$\boldsymbol{c}_t = \boldsymbol{f}_t \odot \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \odot tanh\left(\boldsymbol{W}_{\boldsymbol{x}_c}\boldsymbol{x}_t + \boldsymbol{W}_{\boldsymbol{h}_c}\boldsymbol{h}_{t-1}\right) \tag{4}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \odot tanh\left(\boldsymbol{c}_t\right) \tag{5}$$

where $\boldsymbol{x}_t \in \mathbb{R}^n$ represents the input of the LSTM layer at time t; i_t, o_t, f_t, c_t, and $h_t \in R^m$ denote the input gate state, output gate state, forget gate state, cell state, and hidden layer state at time t, respectively; W_{x_i}, W_{x_o}, W_{x_f}, and $W_{x_c} \in R^{m\times n}$ are all learnable parameter matrices. The symbol $\odot$ denotes element-wise multiplication.

3.2 Time pattern attention mechanism

While the LSTM network has exhibited remarkable proficiency in managing time series data, the advent of the attention mechanism facilitates the extraction of pertinent information among features (Ruan et al., 2023c), thereby augmenting the model's learning capacity and accuracy. Consequently, this article incorporates a TPA mechanism, grounded in the LSTM network, to bolster the load forecasting model's ability to learn from historical load time series data.

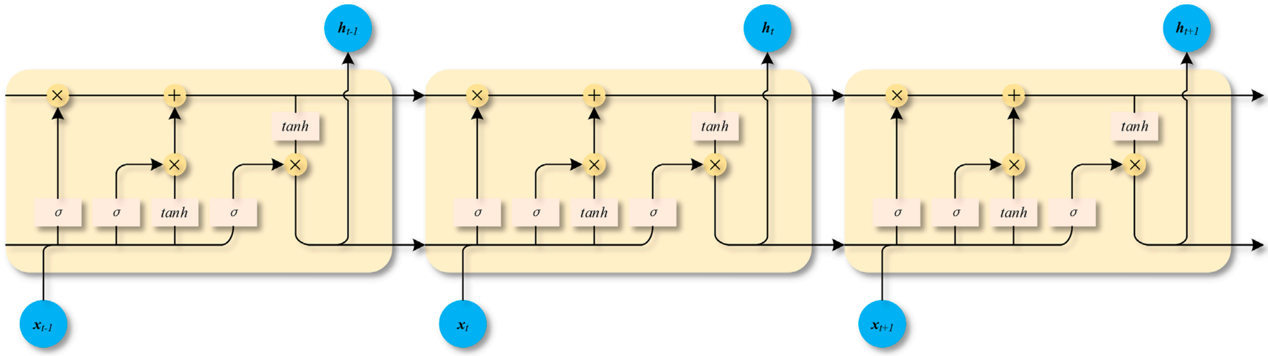

FIGURE 1
Diagram of the long short-term memory network.

First, a one-dimensional convolutional neural network (1-D CNN) layer is used to extract the feature learning capability of the LSTM network's hidden state. Let $\{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_t\} \in \mathbb{R}^{m \times t}$ represent the hidden states of the LSTM layer, where dimension m represents the number of features and dimension t represents the time steps. The hidden states in the past $t-1$ steps, i.e., $\boldsymbol{H} = \{h_1, \ldots, h_{t-1} \in \mathbb{R}^{m \times (t-1)}\}$, are processed by the one-dimensional convolution operation, as follows:

$$\boldsymbol{H}_{i,j}^{C} = \sum_{l=1}^{T} \boldsymbol{H}_{i,(t-T-1+l)} \times \boldsymbol{C}_{j,l} \tag{6}$$

where the convolution operation $\boldsymbol{H}^C \in R^{n \times k}$ is configured with k convolution kernels $\boldsymbol{C}_j \in \mathbb{R}^{1 \times T}$. The convolution kernels are applied along the row vectors of the hidden state matrix $\boldsymbol{H}$ to compute the convolution, extracting the temporal pattern matrix $\boldsymbol{H}^C$ within the visual field of the convolution kernels. $\boldsymbol{H}_{i,j}^{C}$ represents the result value of processing the i th row vector of $\boldsymbol{H}$ with the j th convolution kernel.

Subsequently, a scoring mechanism is employed to evaluate the relevance between the hidden state h_t and the row vectors of the convolutional temporal pattern matrix H^C, as follows,

$$s(\boldsymbol{H}_i^c, \boldsymbol{h}_t) = \boldsymbol{H}_i^c \boldsymbol{W}_a \boldsymbol{h}_t \tag{7}$$

where $H_i^C \in R^{1 \times k}$ represents the i th row vector of H^C; $W_a \in R^{k \times m}$ is the attention mapping matrix in the scoring mechanism.

By applying the Sigmoid activation function to the scoring mechanism, the attention coefficient α_i is obtained, which represents the relevance between h_t and H_i^C, making it easier to compare multivariate associations:

$$\alpha_i = sigmoid(s(\boldsymbol{H}_i^c, \boldsymbol{h}_t)) \tag{8}$$

Based on the obtained attention coefficients, performing attention-weighted summation and addition operations yields the output under the TPA mechanism:

$$\boldsymbol{h}_t' = \boldsymbol{W}_h \boldsymbol{h}_t + \boldsymbol{W}_v \left(\sum_{i=1}^{n} \alpha_i \boldsymbol{H}_i^C \right)^{\mathrm{T}} \tag{9}$$

where both $W_h \in \mathbb{R}^{m \times m}$ and $W_v \in R^{m \times k}$ are learnable parameter matrices for the TPA layer, and $h_t' \in \mathbb{R}^m$ represents the hidden state after being processed by the LSTM layer and the TPA layer.

3.3 Time pattern attention mechanism-based short-term load forecasting

An overall model for the TPA-LSTM-based short-term load forecasting considering multi-regional factors is described in Figure 2, encompassing three modules. Module 1 employs TPA-LSTM to learn from historical load data, initially establishing a load baseline. Module 2 assimilates various factors from distinct regions, such as climate and economy, utilizing fully connected layers (FCLs) to learn diverse climate conditions, including temperature, humidity, wind speed, precipitation, and economic factors like market stability, electricity price adjustments, load control, and industrial growth. Module 3 constitutes the date information learning module, which examines the influence of varying seasons and typical days on electric load and integrates the output of Module 2 through a concatenation operation to learn the impact of date information on the electric load across various regions. Ultimately, the three modules enter the fusion layer and yield the final load profile in the form of a fully connected layer.

The determination of hyperparameters can be accomplished through a combined implementation of random search and k-fold cross-validation methodologies. Initially, a predefined search space is established to encompass the range of potential values for each hyperparameter. From this search space, a series of random samples is generated to explore and discover the possibly optimal combination of hyperparameters. Subsequently, in order to enhance the robustness of the load forecasting model, the entire dataset is divided into k subsets. During each iteration of training, the model is trained k times using different subsets as the training set and one subset as the validation set. This process ensures that each subset serves as the validation set exactly once throughout the iterations. Following the completion of k iterations, the model's performance metrics are averaged over the training process. Finally, the set of hyperparameters that yields the best performance is selected for implementation in the load forecasting model.

The employment of the proposed load forecasting model is illustrated in Figure 3. The process begins with the careful

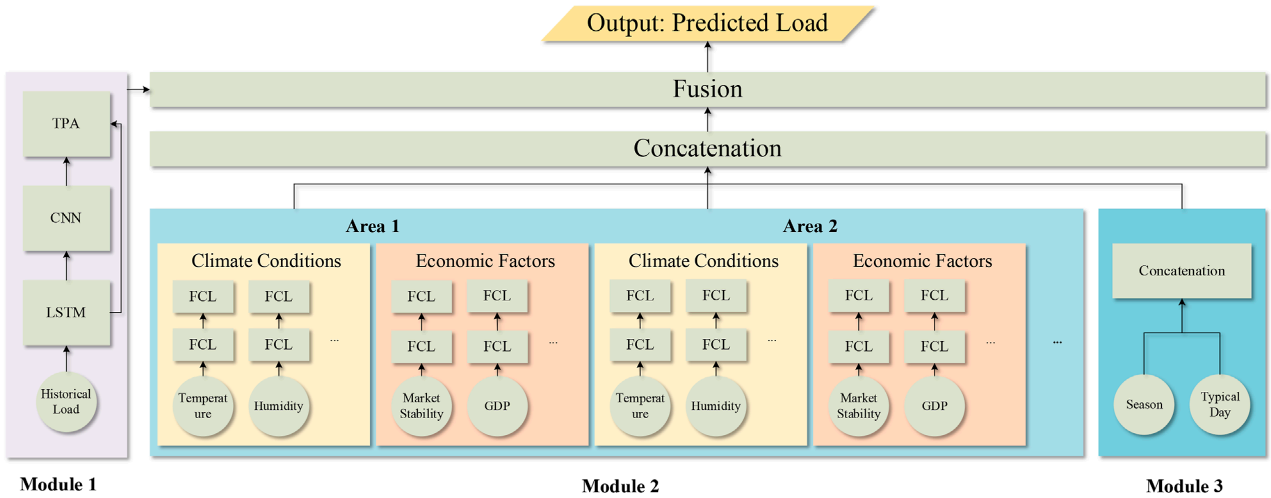

FIGURE 2
Framework of the short-term load forecasting model based on TPA-LSTM.

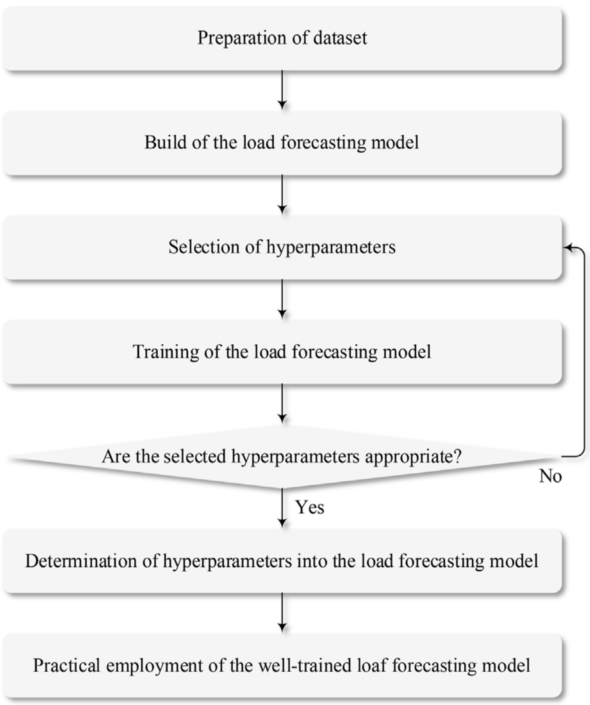

FIGURE 3
Illustration of employment of the proposed load forecasting model.

preparation of the dataset, followed by the construction of the load forecasting model as demonstrated in Figure 2. Subsequently, the hyperparameters of the load forecasting model are selected by employing a combination of random search and *k*-fold cross-validation techniques. The finalized hyperparameters are then implemented in the load forecasting model, which undergoes training for practical application.

4 Results and discussions

4.1 Set up

To evaluate the performance of the proposed load forecasting model, a historical load dataset encompassing all regions in Panama from 2015 to 2020 is employed for simulation. The data has a time granularity of 1 h and includes total load (MWh), temperature (°C), relative humidity (%), liquid precipitation (L/m2), wind speed (m/s), school day indicator (0/1), holiday indicator (0/1), and holiday index (integer) for three cities in Panama. The dataset is divided into training, validation, and test sets with a non-overlapping partition ratio of 8:1:1. The whole dataset starts from 3 January 2015, and ends by 27 June 2020. Accordingly, the sample sizes of the training, validation, and test sets are 1596, 199, 200, respectively.

In addition, three prevalent deep learning models serve as comparative models, as shown in Table 1. The MLP model learns climate information, economic factors, and date information to predict the load profile for the next 24 h. Based on the learning of external factors such as climate information and date information, the LSTM and GRU models learn from the historical load profile for the past week to predict the load profile for the next 24 h. Their model structure is similar to that shown in Figure 2, with the only difference being that they use LSTM and GRU instead of TPA-LSTM to process time-series data. The TPA-LSTM model predicts the load profile for the next 24 h by learning the impact of external factors on future loads while simultaneously extracting the features of historical load curves based on the temporal pattern attention mechanism.

4.2 Numerical results and discussions

During the model training process, standard deviation normalization transformation is applied to all data features to mitigate the influence of feature units on prediction outcomes. The

TABLE 1 Description of model scenarios.

Model	Brief description
MLP	The model inputs include climate information and date information, without considerations of previous load information, while the model output is a 24-h load profile. It removes Module 1 in Figure 2
LSTM	The model inputs include climate information, date information, and the load information from the previous week. The model output is a 24-h load profile. Module 1 in Figure 2 is replaced with the LSTM network to process the historical load information
GRU	The model inputs include climate information, date information, and the load information from the previous week. The model output is a 24-h load profile. Module 1 in Figure 2 is replaced with the GRU network to process the historical load information
TPA-LSTM	The proposed short-term load forecasting model based on the temporal pattern attention mechanism, with its network structure shown in Figure 2

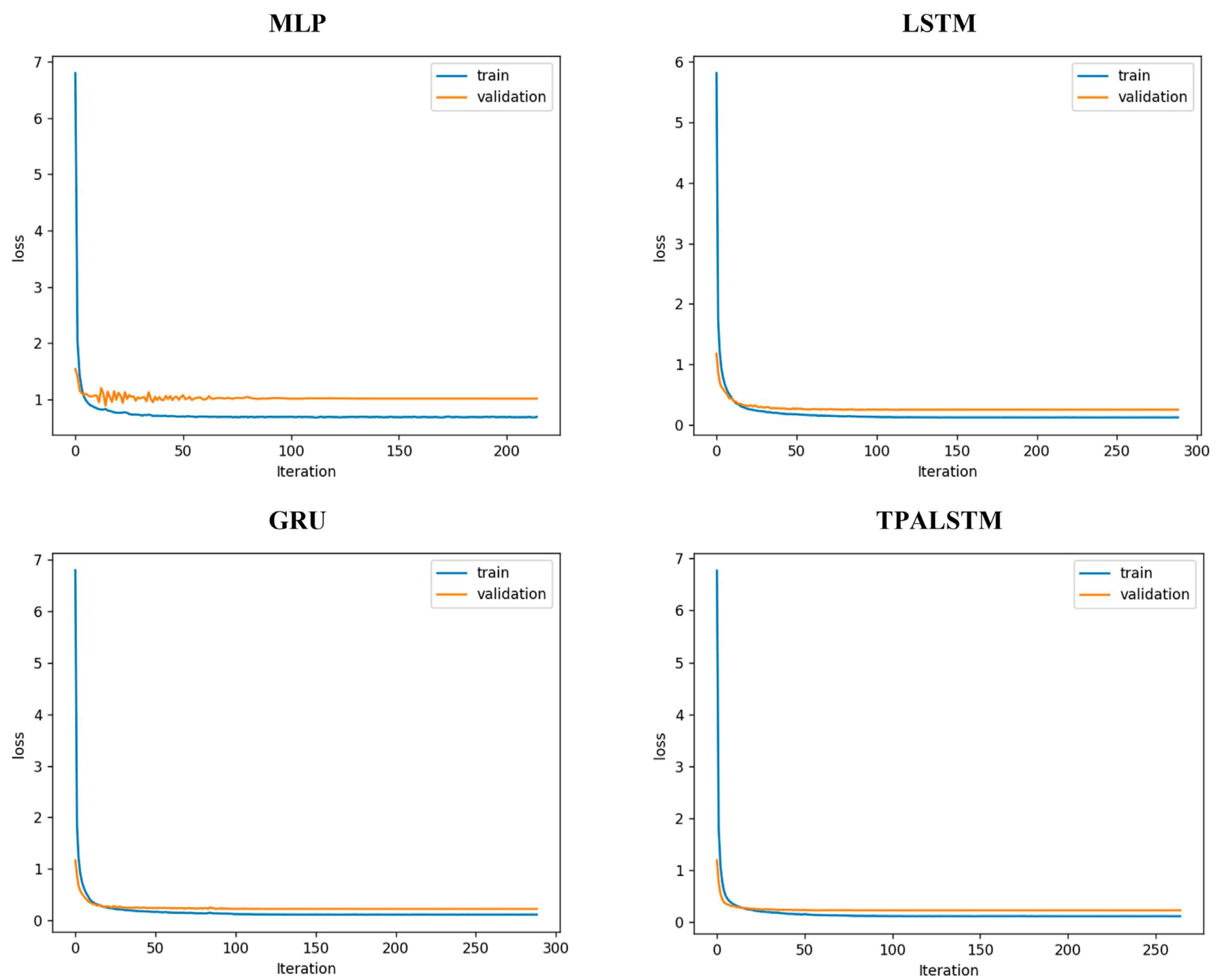

FIGURE 4
Comparison of model training processes.

mathematical expression for this transformation is as follows,

$$\tilde{x}_i^k = \frac{x_i^k - \bar{x}_i}{\sigma(x_i)} \tag{10}$$

where x_i^k and $\tilde{x}_i^k$ represent the original and transformed values of the k th sample in the i th feature of data, respectively; $\bar{x}_i$ and $\sigma(x_i)$ represent the mean and standard deviation of the i th feature of data, respectively.

After all data is transformed, the models can be trained, with the mean squared error (MSE) as the training loss function, as follows,

$$Loss = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{11}$$

where N is the total number of samples for training; y_i and $\hat{y}_i$ are true value and prediction of the i th sample.

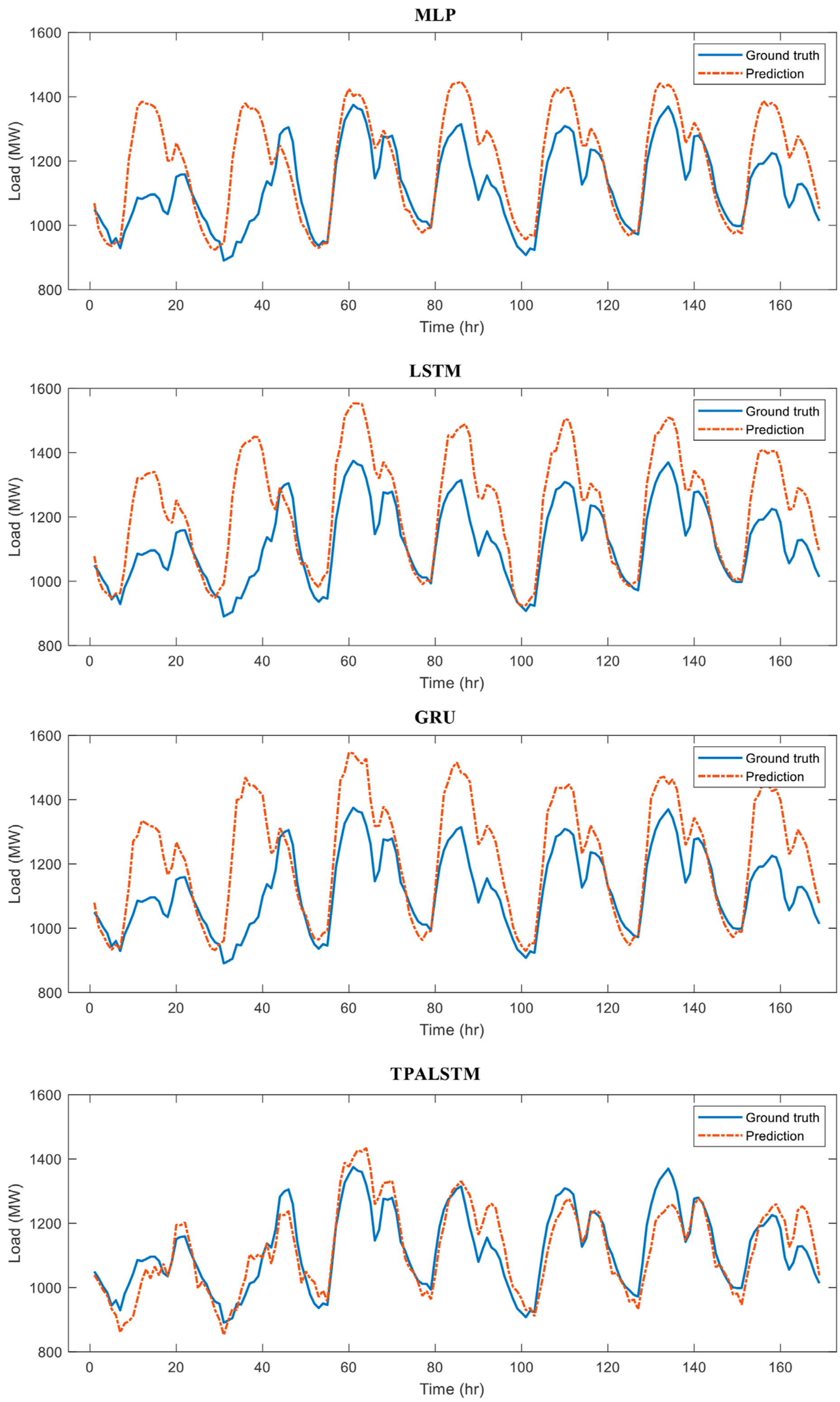

FIGURE 5
Comparison of model forecasting results.

The training process of the four models is shown in Figure 4. It is clear that all models converge at the end of the training.

For visualization, the last week of data in the test set is selected to compare the predictive performance over the four models. These predicted results are denormalized to the normal scale, as shown in Figure 5. It is evident that the MLP model, which only focuses on climate information, economic factors, and date information, cannot accurately predict the load profile. The LSTM and GRU, two

TABLE 2 Statistics of model performance.

Model	MAPE (%)	MAE (MW)	RMSE (MW)
MLP	10.54	124.24	156.64
LSTM	8.56	96.41	148.85
GRU	8.10	91.30	144.18
TPA-LSTM	**4.41**	**51.43**	**73.73**

The bold values indicates the best performance.

special types of RNNs, can fit the load profile to a certain extent based on the extraction of external factors, but still have significant errors. However, the TPA-LSTM model can not only capture the changes in the load itself but also pay attention to the impact of external factors on the load. The attention mechanism can focus on high-value features, thereby accurately predicting the electric load.

To comprehensively evaluate the performance of the four models, the mean absolute percentage error (MAPE), mean absolute error (MAE), and root mean squared error (RMSE) are used as indicators to statistically analyze the predictions of the four models on the test set, as follows,

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{12}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{13}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{14}$$

The statistical results are shown in Table 2. It can be seen that the proposed TPA-LSTM model demonstrates superior predictive performance in all indicators. The reason is that this model not only takes into account the external factors for load variations, but also considers the influence from the historical load profile. More importantly, it employs the TPA mechanism that can identify valuable features and eliminate irrelevant ones to improve the model performance.

In summary, the model proposed in this article processes historical load data based on the TPA mechanism to establish a baseline for load forecasting. It also uses fully-connected layers to extract the impact of external factors (such as regional climate information, economic factors, and date information on the load), thereby accurately predicting the load curve. According to the MAPE statistics, the model has an error of only 4.41%, making it suitable for use in actual load forecasting models.

5 Conclusion

Accurate load forecasting is crucial for ensuring the stable operation of smart grids. This study introduces a short-term load forecasting approach utilizing the TPA mechanism to fulfill the goal. First, the LSTM network is applied to process historical load time-series data, while the TPA mechanism is incorporated to extract temporal feature correlations, thereby enhancing the model's learning capability. Second, FCLs are employed to analyze external factors such as climate, economy, and dates, investigating their influence on future load patterns and establishing a high-precision forecasting model. Last, the proposed method is simulated and compared through a realistic dataset from Panama. The simulation results demonstrate that the proposed load forecasting approach achieves the lowest errors in terms of MAPE, MAE, and RMSE indicators, displaying the closest alignment with the actual load values. Thus, this method holds significant potential for practical load forecasting applications.

It is worth noting that the proposed load forecasting method still has two unresolved challenges. As a result, future work will focus on the integration of various sampling frequency meters into load forecasting methods, as well as the development of highly robust load forecasting techniques that are able to handle unconventional emergencies.

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

WL and JR proposed the research methods, wrote the code for experiments, and wrote the manuscript. YX and QW conceived the overall structure and framework of the article. JL, RW, and JZ provided constructive discussions and technical support for the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported in part by the National Natural Science Foundation of China under Grant 72171206, 71931003, and 72061147004; in part by the Shenzhen Institute of Artificial Intelligence and Robotics for Society; in part by the Shenzhen Key Lab of Crowd Intelligence Empowered Low-Carbon Energy Network under Grant ZDSYS20220606100601002.

Conflict of interest

Authors WL, YX, QW, JL, and RW were employed by the Shenzhen Power Supply Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Akorede, M., Hizam, H., and Pouresmaeil, E. (2010). Distributed energy resources and benefits to the environment. *Renew. Sustain. energy Rev.* 14, 724–734. doi:10.1016/j.rser.2009.10.025

Candanedo, L., Feldheim, V., and Deramaix, D. (2017). Data driven prediction models of energy use of appliances in a low-energy house. *Energy Build.* 140, 81–97. doi:10.1016/j.enbuild.2017.01.083

Chitsaz, H., Shaker, H., Zareipour, H., Wood, D., and Amjady, N. (2015). Short-term electricity load forecasting of buildings in microgrids. *Energy Build.* 99, 50–60. doi:10.1016/j.enbuild.2015.04.011

Gilanifar, M., Wang, H., Sriram, L. M. K., Ozguven, E. E., and Arghandeh, R. (2019). Multitask Bayesian spatiotemporal Gaussian processes for short-term load forecasting. *IEEE Trans. Industrial Electron.* 67 (6), 5132–5143. doi:10.1109/tie.2019.2928275

Hartvigsson, E., Jakobsson, N., Taljegard, M., and Odenberger, M. (2021). Comparison and analysis of GPS measured electric vehicle charging demand: The case of western Sweden and seattle. *Front. Energy Res.* 9, 730242. doi:10.3389/fenrg.2021.730242

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9, 1735–1780. doi:10.1162/neco.1997.9.8.1735

Hossain, E., Khan, I., Un-Noor, F., Sikander, S., and Sunny, M. (2019). Application of big data and machine learning in smart grid, and associated security concerns: A review. *IEEE Access* 7, 13960–13988. doi:10.1109/access.2019.2894819

Li, Z., Wu, L., Xu, Y., and Zheng, X. (2021). Stochastic-weighted robust optimization based bilayer operation of a Multi-energy building microgrid considering practical thermal loads and battery degradation. *IEEE Trans. Sustain. Energy* 13 (2), 668–682. doi:10.1109/tste.2021.3126776

Li, Z., Wu, L., Xu, Y., Wang, L., and Yang, N. (2023a). Distributed tri-layer risk-averse stochastic game approach for energy trading among multi-energy microgrids. *Appl. Energy* 331, 120282. doi:10.1016/j.apenergy.2022.120282

Li, Z., Xu, Y., Wang, P., and Xiao, G. (2023b). Coordinated preparation and recovery of a post-disaster multi-energy distribution system considering thermal inertia and diverse uncertainties. *Appl. Energy* 336, 120736. doi:10.1016/j.apenergy.2023.120736

Lusis, P., Khalilpour, K., Andrew, L., and Liebman, A. (2017). Short-term residential load forecasting: Impact of calendar effects and forecast granularity. *Appl. Energy* 205, 654–669. doi:10.1016/j.apenergy.2017.07.114

Massana, J., Pous, C., Burgas, L., Melendez, J., and Colomer, J. (2015). Short-term load forecasting in a non-residential building contrasting models and attributes. *Energy Build.* 92, 322–330. doi:10.1016/j.enbuild.2015.02.007

Menezes, A., Cripps, A., Buswell, R., Wright, J., and Bouchlaghem, D. (2014). Estimating the energy consumption and power demand of small power equipment in office buildings. *Energy Build.* 75, 199–209. doi:10.1016/j.enbuild.2014.02.011

Ruan, J., Liu, G., Qiu, J., Liang, G., Zhao, J., He, B., et al. (2022a). Time-varying price elasticity of demand estimation for demand-side smart dynamic pricing. *Appl. Energy* 322, 119520. doi:10.1016/j.apenergy.2022.119520

Ruan, J., Liang, G., Zhao, J., Qiu, J., and Dong, Z. (2022b). An inertia-based data recovery scheme for false data injection attack. *IEEE Trans. Industrial Inf.* 18, 7814–7823. doi:10.1109/tii.2022.3146859

Ruan, J., Fan, G., Zhu, Y., Liang, G., Zhao, J., Wen, F., et al. (2023a). Super-resolution perception assisted spatiotemporal graph deep learning against false data injection attacks in smart grid. *IEEE Trans. Smart Grid*, 1–17. doi:10.1007/s10734-023-01031-x

Ruan, J., Liang, G., Zhao, J., Lei, S., He, B., Qiu, J., et al. (2023b). Graph deep learning-based retail dynamic pricing for demand response. *IEEE Trans. Smart Grid*, 1–17. doi:10.1007/s10734-023-01031-x

Ruan, J., Yang, C., Wang, Q., Wang, S., Liang, G., Zhao, J., et al. (2023c). Assessment of spatiotemporally coordinated cyberattacks on renewable energy forecasting in smart energy system. *Appl. Energy* 347, 121470. doi:10.1016/j.apenergy.2023.121470

Ruzic, S., Vuckovic, A., and Nikolic, N. (2003). Weather sensitive method for short term load forecasting in electric power utility of Serbia. *IEEE Trans. Power Syst.* 18, 1581–1586. doi:10.1109/tpwrs.2003.811172

Virote, J., and Neves-Silva, R. (2012). Stochastic models for building energy prediction based on occupant behavior assessment. *Energy Build.* 53, 183–193. doi:10.1016/j.enbuild.2012.06.001

Wang, H., Ruan, J., Wang, G., Zhou, B., Liu, Y., Fu, X., et al. (2018). Deep learning-based interval state estimation of AC smart grids against sparse cyber attacks. *IEEE Trans. Industrial Inf.* 14, 4766–4778. doi:10.1109/tii.2018.2804669

Wang, H., Ruan, J., Ma, Z., Zhou, B., Fu, X., and Cao, G. (2019a). Deep learning aided interval state prediction for improving cyber security in energy internet. *Energy* 174, 1292–1304. doi:10.1016/j.energy.2019.03.009

Wang, H., Ruan, J., Zhou, B., Li, C., Wu, Q., Raza, M., et al. (2019b). Dynamic data injection attack detection of cyber physical power systems with uncertainties. *IEEE Trans. Industrial Inf.* 15, 5505–5518. doi:10.1109/tii.2019.2902163

Wen, J., Xian, Z., Chen, K., and Luo, W. (2022). A novel forward operator-based Bayesian recurrent neural network-based short-term net load demand forecasting considering demand-side renewable energy. *Front. Energy Res.* 10, 963657. doi:10.3389/fenrg.2022.963657

Yang, Y., Qiu, J., Ma, J., and Zhang, C. (2021). Integrated grid, coal-fired power generation retirement and GESS planning towards a low-carbon economy. *Int. J. Electr. Power Energy Syst.* 124, 106409. doi:10.1016/j.ijepes.2020.106409

Yang, Y., Qiu, J., Zhang, C., Zhao, J., and Wang, G. (2022). Flexible integrated network planning considering echelon utilization of second-life of used electric vehicle batteries. *IEEE Trans. Transp. Electrification* 8, 263–276. doi:10.1109/tte.2021.3068121

Yu, W., Wen, G., Yu, X., Wu, Z., and Lü, J. (2014). Bridging the gap between complex networks and smart grids. *J. Control Decis.* 1, 102–114. doi:10.1080/23307706.2014.885293

Zhang, C., Qiu, J., Yang, Y., and Zhao, J. (2021). Trading-oriented battery energy storage planning for distribution market. *Int. J. Electr. Power Energy Syst.* 129, 106848. doi:10.1016/j.ijepes.2021.106848