Electric Power Systems Research 259 (2026) 113149

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

A novel multi-source spatial-temporal forecasting network for power prediction of electric vehicle charging stations

Zhenhua Zhou [a], Bozhen Jiang [a], Qin Wang [a,b,*]

[a] Department of Electrical and Electronic Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong Special Administrative Region
[b] Research Institute for Smart Energy, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong Special Administrative Region

A R T I C L E I N F O

Keywords:
Electric vehicle charging load forecasting
Spatial-temporal modeling
Attention mechanism
CNN-BiLSTM

A B S T R A C T

Accurate power demand prediction for Electric Vehicle (EV) charging stations is critical for smart grid stability. However, most existing methods focus predominantly on the temporal dependencies of single-station historical data, often neglecting the complex spatial correlations between different stations. To address this, this paper proposes a novel Multi-Source Spatial-Temporal Forecasting Network (MSSTON) for high-precision EV charging load forecasting. First, we design a Spatial-Temporal Network (STN) by embedding Convolutional Neural Network (CNN) blocks directly into Bi-directional Long Short-Term Memory (Bi-LSTM) units. This architecture facilitates the deep fusion of local spatial features and long-term temporal dependencies. Second, to fully leverage multi-source data, a Multi-Source Attention Mechanism (MSAM) is introduced. This mechanism dynamically weighs the importance of diverse data sources, effectively filtering noise and enhancing the extraction of high-correlation spatial features. Validated on the Boulder EVCS dataset, experimental results demonstrate that MSSTON achieves superior predictive performance with an root mean squared error (RMSE) of 0.7705 and an R-squared (R^2) of 0.9868. The proposed method significantly outperforms traditional LSTM and hybrid CNN-BiLSTM baselines, exhibiting exceptional robustness and generalization ability across different geographical locations.

1. Introduction

With the increasing severity of global energy crises and environmental pollution issues, electric vehicles (EV) have garnered widespread attention as a clean and efficient mode of transportation [1]. Charging stations, serving as critical infrastructure for energy replenishment of electric vehicles, directly impact the popularization and promotion of EV. Accurately predicting the power demand of charging stations is of significant importance for optimizing the layout of charging stations, enhancing the efficiency of power grid scheduling, and reducing operational costs. This not only ensures that the charging needs of electric vehicle users are met but also mitigates the impact on the power grid, facilitating its stable operation[2],[3]. Consequently, conducting research on the power prediction of electric vehicle charging stations holds substantial theoretical and practical value for promoting the sustainable development of the electric vehicle industry.

Approaches to electric vehicle charging station power prediction can be categorized into two main types: physics-based models and data-driven models. The physics-based model approach is a method that predicts power demand based on the physical characteristics of electric vehicles and charging stations [4],[5]. This method typically requires an in-depth understanding of the internal physical processes of charging stations, and involves establishing mathematical models to describe these processes. The main steps of the physics-based model approach include model development, parameter estimation, model solving, and prediction evaluation [6]. Model development: Based on the physical characteristics of electric vehicles and charging stations, mathematical models are established to describe power demand. Models may include battery charging models, user behavior models, grid interaction models, etc. Model development must consider various influencing factors, such as battery capacity, charging rate, user preferences, grid constraints, etc. Parameter estimation: Estimate parameters within the model, such as internal resistance of the battery, user charging habits, etc. Parameter estimation can be conducted through experimental measurement, data fitting methods, requiring substantial experimental data and computational resources. Model solving: Numerical methods are used to solve the model and obtain predicted values for power demand. Model solving may involve complex mathematical issues such as differential equations, optimization problems, requiring specialized mathematical knowledge and computational tools. Prediction evaluation: Evaluate the

* Corresponding author.
E-mail address: qin-ee.wang@polyu.edu.hk (Q. Wang).

https://doi.org/10.1016/j.epsr.2026.113149
Received 30 December 2025; Received in revised form 25 March 2026; Accepted 13 April 2026
Available online 25 April 2026

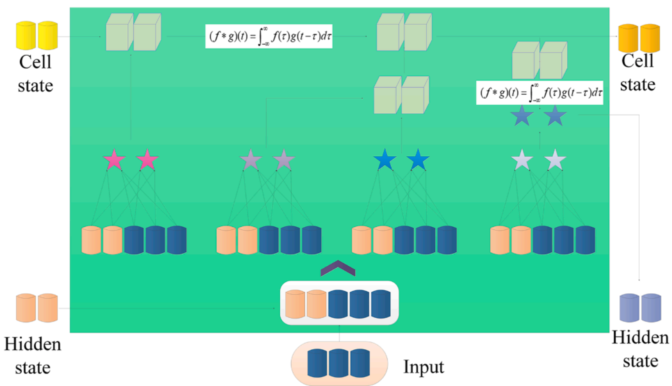

Fig. 1. Structure of STN.

model's predictive performance on actual data, using metrics like Mean Squared Error (MSE), Mean Absolute Error (MAE) to measure prediction accuracy. Adjustments and optimizations are made to the model based on evaluation results. However, this method requires extensive parameter estimation, and model construction and solving can be quite complex.

Data-driven methods are approaches that predict electric vehicle charging station power demand based on historical data and statistical principles. This method typically does not rely on an in-depth understanding of the internal physical processes of charging stations, but instead identifies patterns and regularities by analyzing a large amount of historical data, overcoming the limitations of the physics-based approach [7]. The main steps of the data-driven method include data collection, feature engineering, model selection, model training, and prediction evaluation.

Guo et al. [8] conducted day-ahead prediction of EV charging demand by deep learning. VMD was utilized to decompose charging time series data to obtain explicit trend and dual attention mechanism was applied to improve influence of importance historical time points. Sreekumar and Lekshmi [9] proposed a data-slotting data preprocessing method, which can handle non-linear and highly stochastic charging station data more effectively compared to traditional data processing techniques. Furthermore, models such as Random Forest are utilized to predict the power of charging stations, with the least mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE) serving as model evaluation criteria to measure the predictive performance of the model. Cruz et al. [10] used statewide transportation planning model to predict power demand for EV charging station. The methodology forecasts the volume of EV that pass the vicinity of a charging station and the number of EV that will stop at the charging station to charge every hour of the day. Cabrera-Tobar et al. [11] proposed a robust-explicit model for demand response of EV charging station. A day-ahead energy planning was designed to reduce environmental impact. Researchers have approached the problem of data-driven electric vehicle (EV) charging station power prediction from various perspectives.

In addition to the methods based on regression, there are also regression-based approaches widely applied to EV charging station power prediction, achieving significant results.

Regression-based methods have gained popularity due to their straightforward nature, efficiency, and effectiveness in short-term load forecasting [12,13]. These techniques involve modeling the relationship between input variables like historical charging data, weather conditions, and driving patterns, and the output variable, which includes charging demand, through regression analysis. A study referenced as [14] utilizes a mixed-integer nonlinear programming model coupled with a neural network-based charging demand forecasting approach to achieve enhanced prediction accuracy and performance. Lee et al. [15] have applied a Long Short-Term Memory (LSTM) recurrent neural network to forecast short-term load at electric vehicle charging stations (EVCS).

To address missing data in the dataset and enhance prediction accuracy, an innovative imputation method is also proposed. The proposed model has demonstrated superior performance over traditional methods. Moreover, an LSTM was introduced in Ref. [16] for short-term prediction of EV charging demand, outperforming ARIMA and Multilayer Perception (MLP) models across various scenarios. Additionally, a synergistic learning technique that integrates Monte Carlo Simulation (MCS) with LSTM neural network is established in Ref. [17] to enhance the prediction accuracy of EV charging demand. A hybrid KAN-ANN based model has been employed in Ref. [18] for enhancing load forecasting using big data.

However, there are shortcomings in the aforementioned studies. Most research focuses only on predicting the daily demand of a single station without considering the broader spatial patterns, which neglects the overall network load analysis that is crucial for regional and urban infrastructure planning and market design. Additionally, there is no analysis of the spatial distribution between different time periods of a single station (many days). Therefore, spatiotemporal combined analysis of the spatial distribution patterns of data across multiple regions and multiple stations has been conducted. Shang et al. [19] proposed an explainable spatiotemporal multi-task learning for EV charging demand prediction, simultaneously learning multi-dimensional charging demand features at the station level.Tian et al. [20]introduced an encoder-decoder architecture underpinned by split-pyramid attention to resolve issues of low spatial resolution, enabling the model to process multi-scale data and capture temporal dynamics effectively. Wang et al. [21] proposed an adaptive spatiotemporal graph recurrent network for short-term electric vehicle charging demand prediction, embedding the recurrent unit in a gated recurrent graph convolution to extract spatiotemporal features from charging station data. However, the aforementioned spatiotemporal integrated methods directly mix multiple charging station data into the data-driven model to extract temporal and spatial features without exploring and utilizing the correlation between multiple charging station data. Wang et al. [22] utilized a heterogeneous spatiotemporal graph convolutional network to predict EV charging demand. The heterogeneous spatiotemporal graph convolutional network learned the spatial correlations between charging regions by constructing heterogeneous graphs. Subsequently, they proposed an adaptive spatiotemporal graph recurrent network for short-term electric vehicle charging demand prediction, embedding the recurrent unit in a gated recurrent graph convolution to extract spatiotemporal features from charging station data [23].

Data-driven approaches typically identify patterns within massive historical datasets without relying on complex physical modeling. However, existing studies often exhibit significant limitations. Most research focuses on predicting the daily demand of a single station in isolation, ignoring broader spatial patterns and the overall network load analysis crucial for regional infrastructure planning. While some spatiotemporal methods exist, they often simply concatenate data from multiple stations into the model. This approach fails to explicitly explore the dynamic correlations between multiple stations, leading to a loss of structural spatiotemporal information.

To this end, this paper proposed a novel multi-source spatial temporal forecasting network (MSSTON) to realize power prediction of EV charging station. Firstly, a spatial temporal network (STN) that integrates recurrent neural units and convolutional blocks is proposed, which enables the STN to extract temporal and spatial information from different time periods of data through seamless integration. Secondly, in order to fully utilize multi-source data, a multi-source attention mechanism (MSAM) is proposed, which can focus on the correlation between multi-source data, improve the model's ability to extract spatial features and generalization performance. Finally, the effectiveness of the

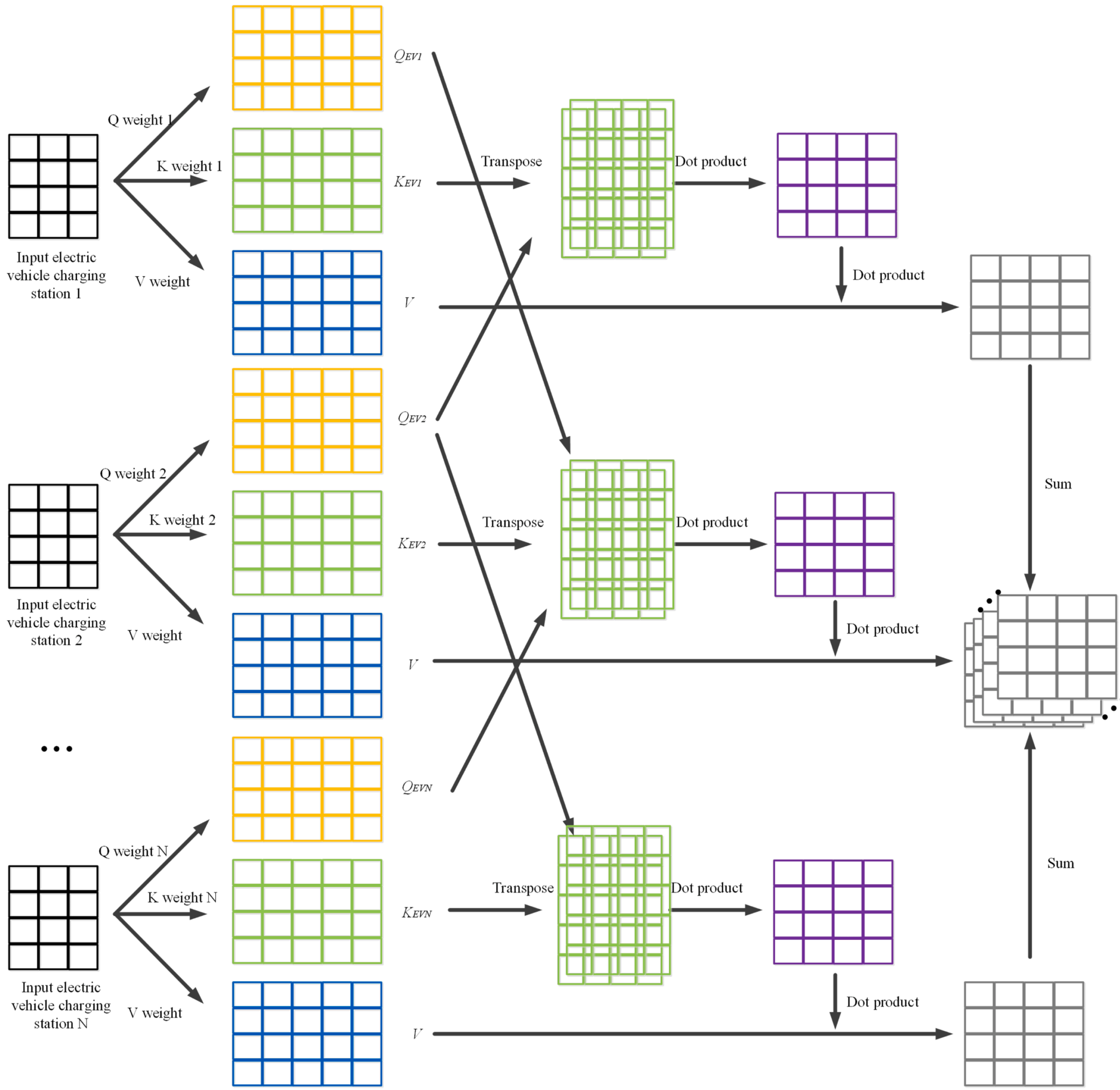

Fig. 2. Structure of MSAM.

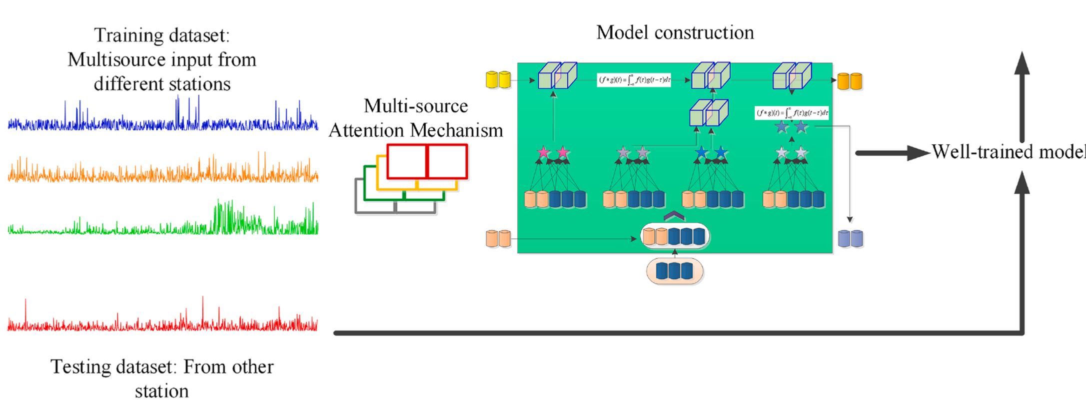

Fig. 3. The prediction flow chart of the proposed method in electric vehicles charging stations.

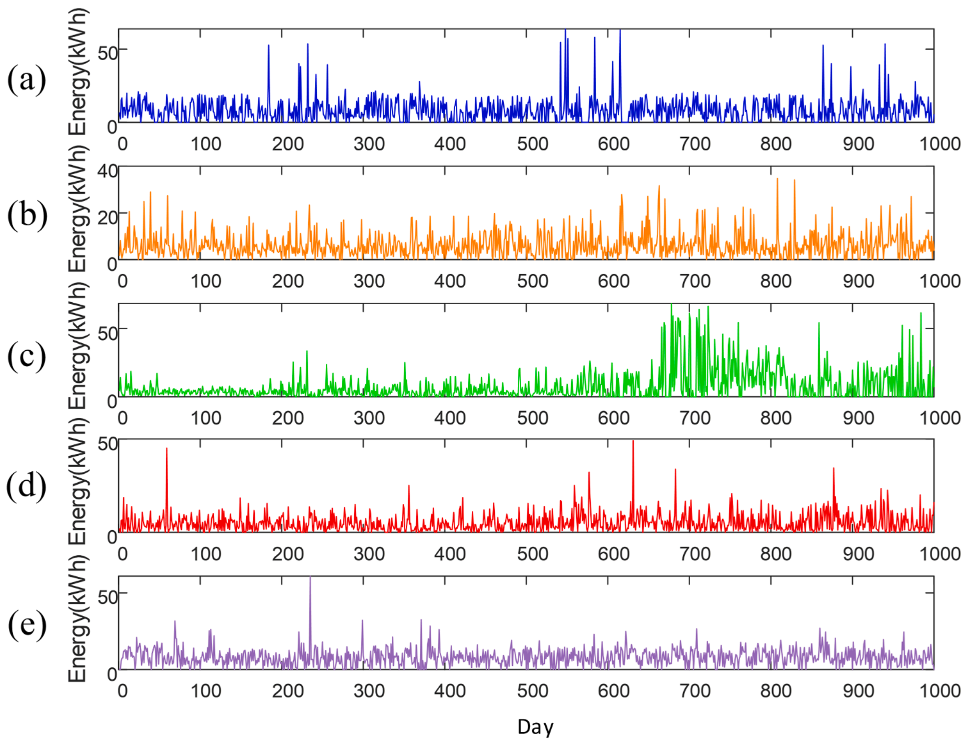

Fig. 4. The historical data sequence diagram of EV charging stations.

proposed method was validated on the open-source dataset of EV charging stations. The experimental results show that MSSTON has the ability of spatial temporal prediction, and can accurately predict the power of electric vehicle charging station, which can provide the basis for ensuring grid reliability.

The main contributions of this paper are summarized as follows:

1. A novel multi-source spatiotemporal forecasting network is proposed to achieve precise power prediction of electric vehicle (EV) charging stations. The proposed model focuses on the spatial and temporal features of power prediction when dealing with sequential data, fully leveraging the hidden spatiotemporal information in sequential data. By employing attention mechanisms to extract correlations between sequential data of different EV charging stations, the model can effectively utilize the rich spatiotemporal coupling features of multi-source data.
2. A spatiotemporal network based on bidirectional long short-term memory (Bi-LSTM) units and convolutional neural network blocks is proposed. The network integrates the temporal feature extraction capability of Bi-LSTM with the local spatial feature attention ability of CNNs, enhancing the model's capability to extract spatiotemporal features from EV charging station sequential data. The network focuses on changes and patterns in both temporal and spatial aspects of data to obtain more comprehensive features.
3. To fully utilize multi-source data, a multi-source attention mechanism (MSAM) is proposed, which can focus on the correlation between multi-source data, improving the model's ability to extract spatial features and generalization performance.

The remaining sections of the paper are organized as follows. The second section introduces the related work and foundational theories of the proposed method. The third section details the structure and prediction process of the proposed method. The fourth section conducts experiments and analyzes the results. The fifth section summarizes the conclusions of the paper.

2. Preliminaries

This section successively introduces the research and theories on Bi-LSTM and CNN in the context of electric vehicle (EV) charging station power prediction, followed by an explanation of the mechanism and characteristics of the attention mechanism.

2.1. Bi-LSTM

Bidirectional Long Short-Term Memory (Bi-LSTM) is a special type of recurrent neural network structure that can process both forward and backward temporal sequences. Its core lies in using two separate LSTM layers to handle the input sequence from both the forward and backward directions. This structure enables the Bi-LSTM to capture context information more comprehensively, making it particularly useful for tasks that require understanding context, such as natural language processing, speech recognition, and other tasks that necessitate understanding of the surrounding text. Below is the computation process of the forward LSTM unit:

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t+1} + b_i) \tag{1}$$

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \tag{2}$$

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \tag{3}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{4}$$

$$h_t = o_t \odot \tanh(C_t) \tag{5}$$

Where i_t is the output of the input gate at time t, W_{ix} is the weight of input temporal data, W_{ih} represents the weight of hidden units, b_i, b_j and b_o are bias, f_t is the output of forget gate, o_t is the output of output gate, C_t represents memory state, $\tilde{C}_t$ is the candidate memory state. h_t is hidden state, $\sigma(\odot)$ is activation function operation. W_{fx} is the weight of current input in forget gate, W_{fh} is the weight of hidden units in forget gate, W_{ox} is the weight of current input in output gate, W_{oh} is the weight of hidden units in output gate, $\odot$ is dot product operation.

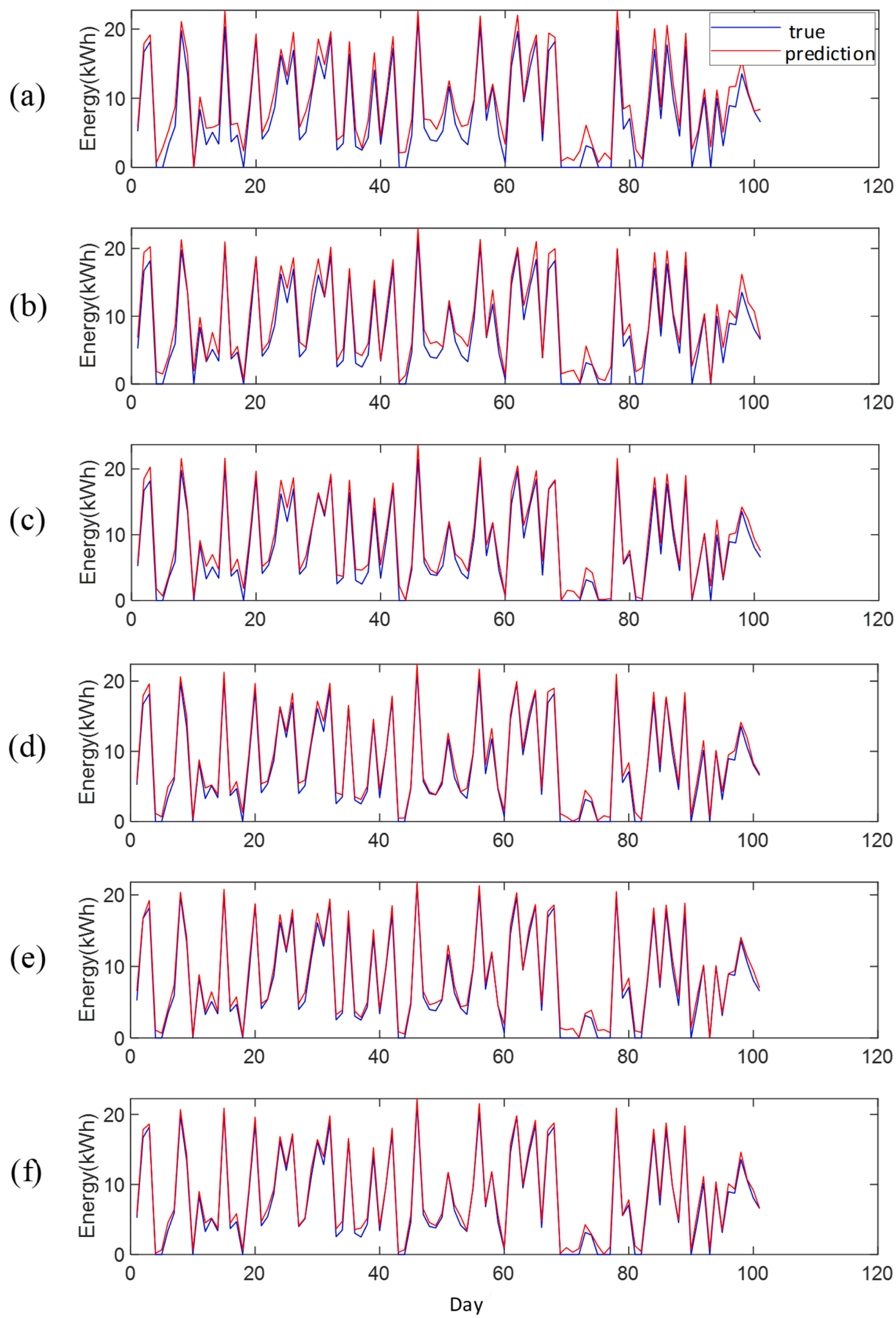

Fig. 5. Prediction results of different methods(One-Step-Ahead). (a) LSTM (b) GRU (c) CNN-BiLSTM (d) DAS-GCN (e) ASTGRN (f) MSSTON.

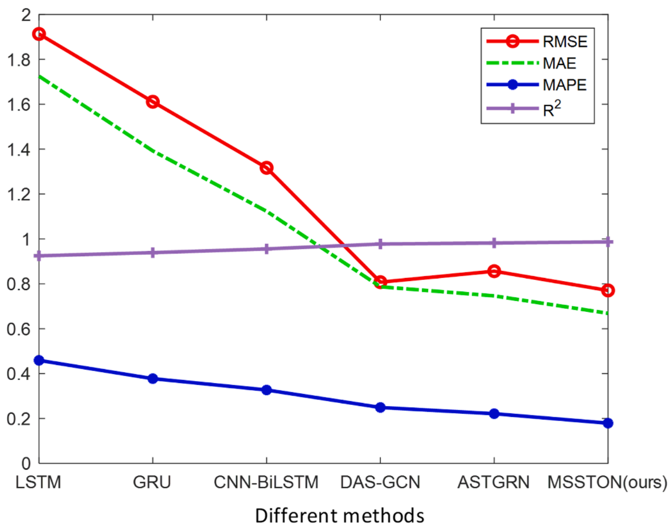

Fig. 6. Evaluation results of different methods (One-Step-Ahead).

The following is the calculation process of the backward LSTM unit, which is different from the direction of the forward processing sequence:

$$i_t' = \sigma(W_{ix}' x_t + W_{ih}' h_{t-1}' + b_i') \tag{6}$$

$$f_t' = \sigma(W_{fx}' x_t + W_{fh}' h_{t-1}' + b_f') \tag{7}$$

$$o_t' = \sigma(W_{ox}' x_t + W_{oh}' h_{t-1}' + b_o') \tag{8}$$

$$C_t' = f_t' \odot C_{t-1}' + i_t' \odot \tilde{C}_t' \tag{9}$$

$$h_t' = o_t' \odot \tanh(C_t') \tag{10}$$

Where the meaning of variables is the same as Eqs. (1)–(5), but it is the calculation formula of reverse LSTM.

LSTM has achieved significant results in power prediction for EV charging stations. Boulakhbar et al. [24] studied the application of models such as artificial neural network (ANN), recurrent neural network (RNN), and long short-term memory (LSTM) in predicting the demand for EV charging stations. The research results proved that LSTM has advantages in predicting the temporal data of EV charging stations. Bakumar et al. [25] presented Deep Learning (DL) - based Bidirectional Long Short Term (Bi LSTM) models for predicting feed wise EC for distribution subsidiaries. Zamee et al. [26] used LSTM and Bi LSTM as comparative experimental methods, demonstrating that deep learning techniques based on LSTM have solid performance in the field of temporal power prediction for EV charging stations. With the powerful temporal prediction capability of LSTM, significant achievements have been made in the research of power prediction for EV charging stations.

2.2. 1D-CNN

1D-Convolutional Neural Network (1D-CNN) are neural network architectures specifically designed to handle one-dimensional data sequences, such as time series data or text data. 1D-CNN extract local features by sliding convolution kernels over data, which can capture patterns and structures in the data, such as frequency features in audio signals or phrases in text. The key components of 1D-CNN include the convolutional layer, which is the core of 1D-CNN and extracts features from input data through convolution operations. The convolutional layer uses a set of learnable filters (convolution kernels) that slide over the input data to calculate the weighted sum of local regions. Activation function: Nonlinear activation functions, such as ReLU (Rectified Linear Unit), are typically applied after convolution operations to introduce non-linear properties, enabling the network to learn more complex patterns. Pooling layer: The pooling layer (such as max pooling) is used to reduce the dimensionality of feature maps, reduce computational complexity, and improve the robustness of the model to input changes. Fully connected layer: At the end of the network, there are usually one or more fully connected layers used to map extracted features to output categories

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau) g(t-\tau)\, d\tau \tag{11}$$

Where f is input signal, g is convolutional kernel, t is time index. τ is the integral variable, $*$ is the convolutional operation.

Zhou et al. [27] combined CNN and LSTM to achieve load prediction for EV charging stations, and utilized CNN's ability to focus on spatial features to achieve multi-objective optimization of charging stations. Mekkaoui et al. [28] used graph convolutional neural networks to extract temporal spatial features, and improved the accuracy of EV charging station power prediction by leveraging the ability of convolutional neural networks to extract spatial features.

2.3. Attention mechanism

Attention mechanism is a key technology in deep learning, widely used in fields such as natural language processing, image recognition, and speech recognition. It assigns different weights to different inputs, allowing the model to focus on key parts of the input sequence, thereby improving the performance and interpretability of the model. The calculation of attention mechanism can be summarized as the following steps: Calculate attention score:

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \tag{12}$$

Where Q is Query matrix, K is Key matrix, V is value matrix, d is the dimension of the key vector used to scale the dot product result to prevent gradient vanishing.

Normalization: Use the softmax function to convert scores into probability distributions, ensuring that the sum of all weights is 1, to achieve soft selection of different positional information.

Weighted Sum: Multiply the normalized weights by V to obtain the final attention output.

The principle of attention mechanism: Firstly, the correlation between the query and the key is calculated, and high correlation will result in greater weight. Then, softmax ensures that the weights are positive and the sum is 1, achieving soft selection for different positional information. The introduction of attention mechanism enables the model to process sequential data more flexibly and effectively, improving prediction accuracy by focusing on the most important parts of the input.

3. Proposed method

The multi-source spatial temporal forecasting network (MSSTON) consists of spatial temporal network (STN) and multi-source attention mechanism (MSAM). STN integrates bi-directional long short-term memory (Bi-LSTM) units and convolutional blocks, which enables the STN to extract temporal and spatial information from different time periods of data through seamless integration. MSAM focus on the correlation between multi-source data, improve the model's ability to extract spatial features and generalization performance

3.1. Spatial temporal network

In existing research, the combination of CNN and BiLSTM is often done in series or parallel, fusing the spatial and temporal features that the two networks are good at extracting together through concatenation or weighting, thus utilizing spatiotemporal features. However, this combination method is only a simple concatenation and cannot organically combine temporal and spatial features, thus breaking the structure of temporal and spatial features in the data and losing the original connection between temporal and spatial features in the data. The spatial temporal network (STN) proposed in this article adopts an embedding approach to design a temporal spatial feature extraction structure, which

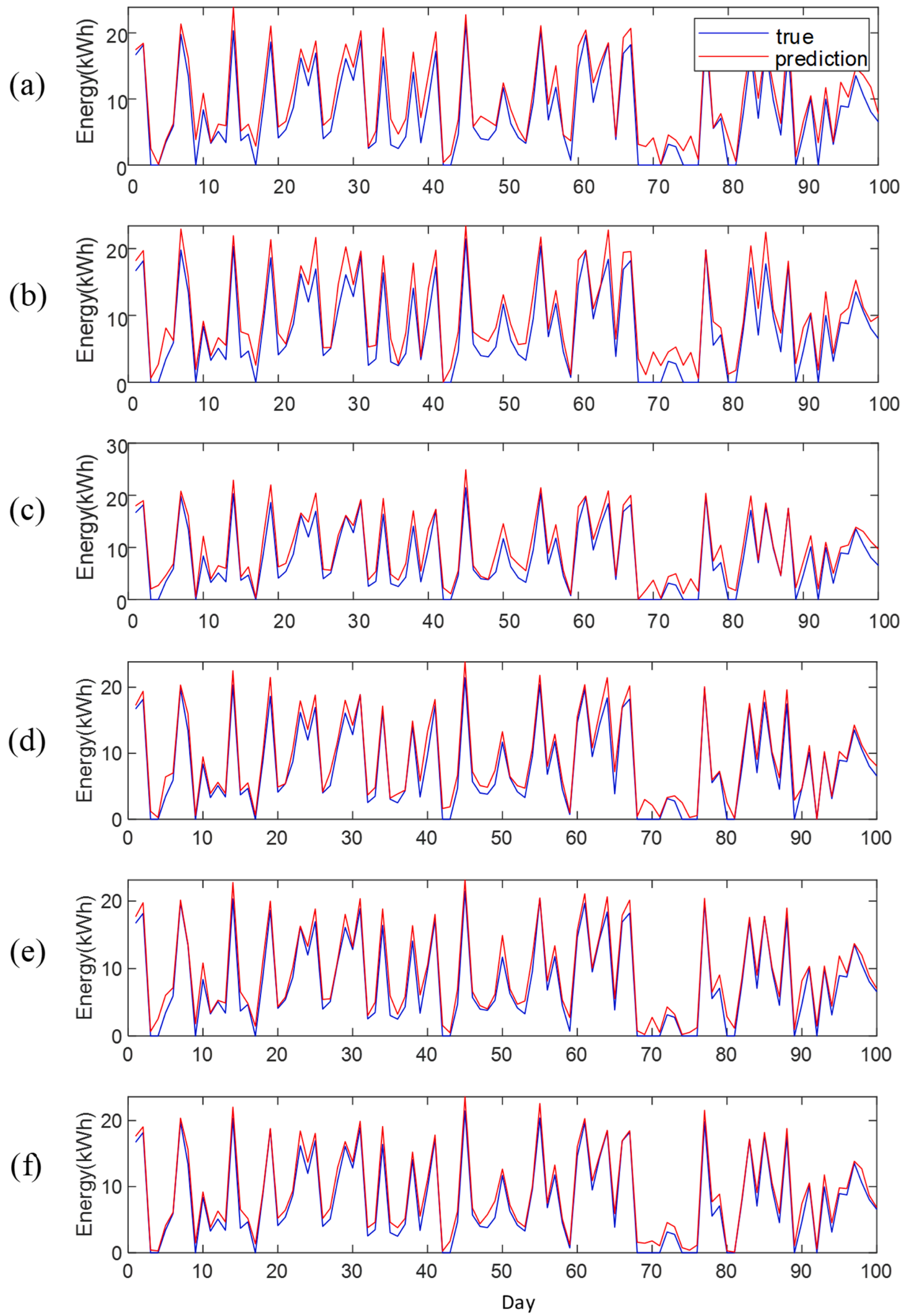

Fig. 7. Multi-step-ahead prediction results of different methods. (a) LSTM (b) GRU (c) CNN-BiLSTM (d) DAS-GCN (e) ASTGRN (f) MSSTON.

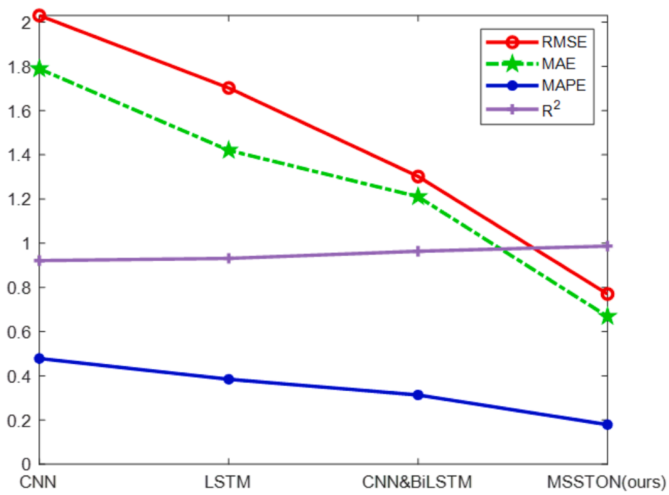

Fig. 8. Measurement index results of ablation study.

integrates CNN into the temporal units of BiLSTM. This enables BiLSTM to extract temporal relationship features from data while mining spatial attributes between data at different time points. The structural design is shown in Fig. 1. In principle, compared with Eqs. (4) and (5), STN replaces the inherent feature output mode of LSTM with convolution operator, and integrates the extraction of spatial features into the temporal feature mining process of LSTM, rather than simple fusion. For each layer of STN structure, the calculation formula is as follows:

$$C_t = (f_t \odot C_{t-1}) * (i_t \odot \tilde{C}_t) \tag{13}$$

$$h_t = o_t \odot \tanh(C_t * K_{sp}) \tag{14}$$

where $*$ is the convolutional operation, the kernel of convolutional operation is set to 3, and the padding is set to 1. C_t and C_{t-1} represent the cell state at time t and $t-1$, respectively. $\tilde{C}_t$ denotes the candidate memory content, which is generated using a convolutional layer to extract spatial features from the input. In Eq. (14), We introduce a learnable spatial convolution kernel K_{sp} (denoted as $\tilde{h}$ in preliminary designs) to filter the cell state C_t. This allows the hidden state h_t to encode local spatial correlations before being gated by the output gate o_t.

3.2. Multi-source attention mechanism (MSAM)

In order to fully utilize the data from multiple EV charging stations, the MSAM is proposed to extract correlation information from multi-source data using data from different stations as the query matrix, key matrix, and value matrix for calculating scores through attention mechanism, different combinations are formed between them. To balance and fully utilize the data, the data from different stations are placed in different matrix positions, and the MSAM module is designed accordingly. As shown in Fig. 2, the multiplication operation between query matrix and key matrix can better reflect the information relationship between data, so we use data from different sites to participate in the operation. The process of calculating attention score is as follows:

$$\text{Attention}(Q, K, V) = \sum_{j=1}^{n} \sum_{\substack{i=1 \\ i \neq j}}^{n} \text{softmax}\left(\frac{Q_{EVi} K_{EVj}^{\top}}{\sqrt{d}}\right) V \tag{15}$$

Where Q_{EVi} is Query matrix of EV charging station i, K_{EVj} is Key matrix of EV charging station j, V is value matrix, d is the dimension of the key vector, n is the number of multi-source EV charging stations.

3.3. Multi-source spatial temporal forecasting network

The multi-source spatiotemporal forecasting network (MSSTON) consists of a spatiotemporal network (STN) and a multi-source attention mechanism (MSAM). As shown in Fig. 3, the MSAM extracts the spatial-temporal correlation features between multi-source data, then inputs these correlated features into the hierarchical STN modules to extract spatial features, and finally feeds the spatial-temporal features into a fully connected layer-based prediction module to complete the feature output. The Mean Squared Error (MSE) is used as the loss function to update the model parameters between the actual and predicted values, enabling the model to possess power prediction capabilities. The prediction process of MSSTON is as follows:

Step 1: First, divide the dataset into training and testing sets. The training set includes data from multiple EV charging stations, uniformly segmented according to the prediction demand.

Step 2: Then, construct the MSSTON model, select a set of hyperparameters, input the multi-source training sets into the model simultaneously to extract features, complete the prediction task, and obtain a well-trained model.

Step 3: Finally, input the testing set into the well-trained model. Since the testing set does not require multi-site data input to the model simultaneously, the MSAM is removed at this time, leaving only the STN to test the data. If the results are not satisfactory, adjust the model parameters and proceed to Step 2.

4. Case study

4.1. Dataset description

The Electric Vehicle Charging Station Data (EVCSD) is from city of boulder in Colorado State. The dataset shows the energy use, length of charging time, gasoline savings and greenhouse gas emission reductions from all city-owned electric vehicle (EV) charging stations. Data from multiple addresses, including 2280 Junction PL, 900 Baseline Rd, 1745 14th street, 1770 13th St, and so on. The data collection time is from January 1st, 2018 to November 30th, 2023, and the sampling interval is 1 day. We use raw data to complete training and testing directly.

We select Energy as the direct prediction object and use historical data to predict power consumption for EV charging stations. To meet the needs of single step prediction and multi-step prediction, the raw data is segmented and 200 historical data are input as training samples into the model for training. The historical energy data of EV charging stations in different places are shown in Fig. 4.

To ensure the fairness of the results, several methods are tested on 100 power data points. Test the data for the location of "2280 junction PL".

4.2. One-step-ahead prediction

To demonstrate the effectiveness of the proposed method, some comparative experiments were designed. Long short-term memory (LSTM), gated recurrent unit (GRU), CNN-BiLSTM [27], dual-adaptive spatio-temporal graph convolutional network (DAS-GCN) [28], adaptive spatial-temporal graph recurrent network (ASTGRN) [23] are selected as a comparison method. Among them, CNN-BiLSTM, DAS-GCN, and ASTGRN maintain the same structure and training method as the original literature. DAS-GCN employed a parallel GCN architecture. ASTGRN leveraged graph neural network modules to capture spatial interdependence across diverse charging stations and incorporates recurrent modules to learn time-series characteristics. Both LSTM and GRU adopt a three-layer structure, with 300, 200, and 500 neural units in each hidden layer, followed by two fully connected layers to obtain the final prediction results. As shown in Table 1, several methods use the Adam optimizer to update model parameters, with a training epoch set to 500, and the learning rate adjusted based on past experience and multiple experimental attempts. To ensure the fairness of the results, several methods are tested on 100 power data points. The testing set data is from the location of "2280 junction PL". The model is built based on PyTorch 2.10 framework and Python 3.10, which runs on Windows 11 with 16GB RAM.

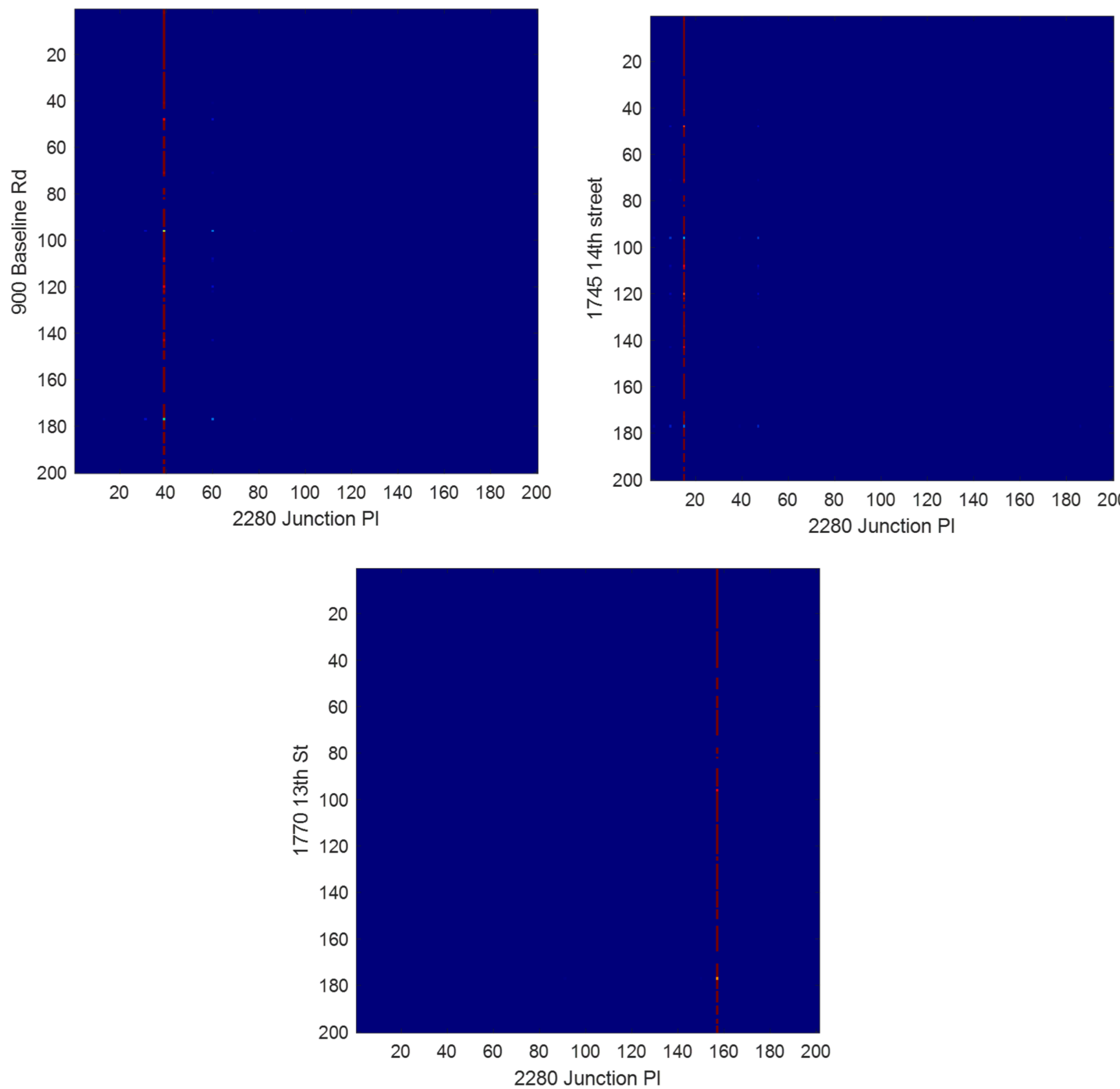

Fig. 9. Spatial correlation heatmap.

Table 1
Key hyperparameters of different methods.

Method	Epoch	Optimizer	Learning rate	Batch size
LSTM	500	Adam	0.0001	16
GRU	500	Adam	0.001	32
CNN-BiLSTM	500	Adam	0.0001	16
DAS-GCN	500	Adam	0.0001	16
ASTGRN	500	Adam	0.0001	32
MSSTON	500	Adam	0.0001	16

The power prediction results of different methods are shown in Fig. 5 . Comparing Fig. 5 a), b), and f), it can be seen that the fitting effect in Fig. 5 f) is significantly better than the methods in Fig. 5 a) and b). The results show that the proposed method has better power prediction performance than LSTM and GRU. This is because MSSTON has higher feature extraction ability than LSTM and GRU alone, not only focusing on temporal features, but also utilizing spatial features from multiple charging station data. The fitting effect of the true and predicted values in Fig. 5 (d-f) is better than that in Fig. 5 (a-c), and the results show that the method that integrates spatiotemporal features is more effective than the method that extracts purely temporal features, proving the superiority of the combination of spatial-temporal features. In order to clearly compare the predictive performance of different methods, the performance of the model was evaluated using root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and R-squared (R2) as metrics. The calculation results are shown in the Table 2 and Fig. 6. From Table 2, it can be seen that MSSTON has the smallest RMSE, MAE, and MAPE results, significantly better than LSTM, GRU, and CNN-BiLSTM methods, and comparable to DAS-GCN and ASTGRN in performance, but still lower than DAS-GCN and ASTGRN. The results indicate the superiority of the proposed method in power prediction, which can fully utilize the spatiotemporal features in multi-source data and achieve high-precision prediction results. Compared with CNN-BiLSTM and MSSTON, the prediction effect of the combination of series structure is worse than that of MSSTON, indicating that MSSTON has more advantages in spatial-temporal feature extraction. From the indicator R2, it can be seen that MSSTON has the highest value, indicating that the proposed method has the best fit between the predicted results and the true values, and also proving that the power prediction accuracy of the proposed method is higher than other methods. It can be clearly seen from the graph that the RMSE, MAE, and MAPE indicators show a downward trend, and the proposed method has the smallest value, indicating the best power prediction performance.

Table 2
Evaluation results of different methods (one-step-ahead).

Method	RMSE (kWh)	MAE (kWh)	MAPE (%)	R^2
LSTM	1.9132	1.7251	0.4590	0.9248
GRU	1.6111	1.3922	0.3775	0.9389
CNN-BiLSTM	1.3170	1.1232	0.3270	0.9554
DAS-GCN	0.8074	0.7868	0.2491	0.9773
ASTGRN	0.8562	0.7463	0.2213	0.9818
MSSTON	**0.7705**	**0.6687**	**0.1792**	**0.9868**

Table 3
Statistical analysis of different methods.

Method	Avg. RMSE (kWh)	variance
LSTM	1.8197	0.0137
GRU	1.6294	0.0040
CNN-BiLSTM	1.3727	0.0033
DAS-GCN	0.8320	0.0016
ASTGRN	0.8448	0.0017
MSSTON	**0.7694**	**0.0011**

In order to further prove the effectiveness of the proposed method and ensure the fairness of the results. All methods were repeated for ten times to calculate the variance and average value of the predicted RMSE. As can be seen from Table 3, the average RMSE and variance of MSSTON are the smallest. On the one hand, it shows that the prediction accuracy of MSSTON is higher than that of other models, on the other hand, it proves that the prediction stability of MSSTON is better.

4.3. Multi-step-ahead prediction

In addition to verifying the real-time prediction capability of the proposed method (which proves the results of single step prediction), it is also necessary to demonstrate the method's ability to predict future power at multiple times. Therefore, in order to further verify the effectiveness of the proposed method, similar to Part B, five comparative methods including LSTM, GRU, CNN BiLSTM, DAS-GCN, Graph Neural Network (GNN), and ASTGRN are still utilized to design comparative experiments. Several methods use the Adam optimizer to update model parameters, with a training epoch set to 500, and the learning rate adjusted based on past experience and multiple experimental attempts. In order to ensure the fairness of the results, several methods were validated through multi-step prediction on 100 power data. Test the data for the location of "2280 junction PL".

The power prediction results of different methods are shown in Fig. 7. Compare Fig. 7 a), b) and f) It can be seen that the goodness of fit in Fig. 7 f) is significantly better than Fig. 7 a) and b) Method. The results show that the proposed method has better power prediction performance than LSTM and GRU. This is because MSSTON has higher feature extraction ability than LSTM and GRU alone, not only focusing on temporal features, but also utilizing spatial features from multiple charging station data. The alignment between predicted and actual values in Fig. 7 (d-f) is better than that in Fig. 7 (a-c), the results show that the method that integrates spatiotemporal features performs better than the method that extracts purely temporal features, demonstrating the superiority of combining spatiotemporal features.

In order to clearly compare the predictive performance of different methods, the performance of the model was evaluated using root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and R-squared (R2) as metrics. The calculation results are shown in Table 4. From Table 4, it can be seen from the table that the effect of GNN is slightly better than that of LSTM and GRU, which is due to its advantage in mining data association. However, compared with CNN-LSTM, its effect is poor, which may be that CNN-LSTM pays more attention to temporal and spatial features than GNN.

Table 4
Evaluation results of different methods on multi-step-ahead prediction.

Method	RMSE (kWh)	MAE (kWh)	MAPE (%)	R^2
LSTM	2.4500	2.0993	0.6182	0.8569
GRU	2.0957	1.7991	0.5426	0.8953
GNN	2.0286	1.7632	0.5174	0.8975
CNN-BiLSTM	2.0143	1.7551	0.4875	0.9033
DAS-GCN	1.3333	1.0904	0.3133	0.9576
ASTGRN	1.4730	1.2245	0.3642	0.9483
MSSTON	**1.3003**	**1.0208**	**0.3093**	**0.9586**

Table 5
Complexity evaluation of different methods.

Method	Training time (s)	Parameter count (MB)	Inference time (s)
LSTM	0.76	1.10	0.13
GRU	0.69	1.06	0.11
CNN-BiLSTM	1.35	2.10	0.19
DAS-GCN	2.64	3.22	0.56
ASTGRN	2.78	3.78	0.64
MSSTON	**2.41**	**3.02**	**0.51**

In addition, taking training time and parameter count as indicators to evaluate the complexity of different methods and examine the efficiency of real-time deployment. Among them, the training time is measured by a batch of training. As can be seen from Table 5, the complexity of the proposed method is higher than that of LSTM, GRU and CNN-BiLSTM classical neural networks. However, considering the accuracy results and the current development speed of computer hardware, the weak speed difference will soon be narrowed. The MSSTON takes only 0.51s to generate 24-hour forecast. Therefore, the proposed method has the advantages of implementation and deployment.

4.4. Ablation study

To rigorously verify the specific contributions of each component within the proposed framework, an ablation study was conducted by systematically removing modules. The results are summarized in Table 6 and in Fig. 8. The analysis focuses on three comparative perspectives: single-mechanism models versus hybrid models, and the specific impact of the attention mechanism.

First, regarding the single-mechanism baselines, the pure Bi-LSTM model demonstrates significantly better performance than the pure CNN model. This suggests that for EV charging load forecasting, temporal dependencies and long-term historical patterns are more critical than static local features. The CNN model, while capable of extracting local trends, struggles to capture the continuous time-series dependencies inherent in charging behaviors. Second, comparing CNN, Bi-LSTM, and CNN&BiLSTM, it can be seen that CNN&BiLSTM (a parallel structure) has better power prediction performance, high accuracy, and good stability. The results demonstrate that the combination of spatial and temporal features can improve the feature extraction ability of the model, thereby enhancing the accuracy of power prediction for EV charging stations.

Finally, and most crucially, the proposed MSSTON outperforms the CNN&BiLSTM baseline across all metrics. This performance gain validates the necessity of the Multi-Source Attention Mechanism (MSAM). Without MSAM, the model treats all neighboring stations and input features with equal importance, potentially introducing noise from low-correlation sources. The MSAM successfully addresses this by assigning dynamic weights to different data sources, effectively "filtering out redundant information and forcing the model to focus on the most relevant spatial temporal patterns. Consequently, MSSTON achieves not only higher accuracy but also greater robustness against data fluctuations.

Table 6
Ablation study results.

Method	RMSE (kWh)	MAE (kWh)	MAPE (%)	R^2
CNN	2.0302	1.7891	0.4785	0.9213
BiLSTM	1.9132	1.7251	0.4590	0.9248
CNN&BiLSTM	1.3021	1.2105	0.3134	0.9637
MSSTON	**0.7705**	**0.6687**	**0.1792**	**0.9868**

4.5. Physical interpretation

In order to prove the interpretability of the proposed method in extracting spatial temporal features at the physical level, we implemented visualization. As shown in Table 9, the attention scores between different sites and the same site are not the same, which also shows that the data between different sites are related, but not completely consistent. Therefore, it is necessary to make full use of data from different sites. MSAM can integrate the relevance between different sites, and establish and use the relationship between source data at different times and places.

5. Conclusion

This paper addresses the challenge of accurate ultra-short-term load forecasting for EV charging stations by proposing a novel Multi-Source Spatial-Temporal Forecasting Network (MSSTON). The study makes several key contributions to the field of smart grid management:

1. Deep Feature Integration: By embedding convolutional blocks directly into Bi-LSTM units, the proposed Spatiotemporal Network (STN) achieves a seamless fusion of local spatial features and long-term temporal dependencies, overcoming the limitations of loosely coupled hybrid models.
2. Dynamic Correlation Modeling: The introduction of the Multi-Source Attention Mechanism (MSAM) enables the model to dynamically prioritize information from neighboring stations based on their real-time relevance. This effectively solves the problem of noise interference typically found in multi-source data inputs.
3. Superior Empirical Performance: Extensive experiments on the real-world Boulder EVCS dataset demonstrate that MSSTON consistently achieves the lowest RMSE and MAE compared to state-of-the-art baselines (including LSTM, GRU, and standard CNN-BiLSTM). The model exhibits exceptional stability, particularly when predicting sharp load peaks and troughs.

In summary, MSSTON provides a reliable and precise tool for grid operators to anticipate charging demands, thereby facilitating better load dispatching and grid stability.

In the future, we will explicitly model with spatial features to give the model interpretability. The seasonal variability and types of EV load will also be used to verify the effectiveness of the model.

CRediT authorship contribution statement

Zhenhua Zhou: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Bozhen Jiang:** Writing – review & editing, Writing – original draft, Software, Methodology; **Qin Wang:** Supervision, Methodology.

Acknowledgements

This work was supported in part by the University Grants Committee's Collaborative Research Fund (CRF) 2025/26 (Contract No. C5132-25YF) and in part by the Otto Poon Charitable Foundation Research Institute for Smart Energy (RISE) (Project ID: P0060534).

Data availability

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Y. Li, L. Song, Y. Hu, et al., Load profile inpainting for missing load data restoration and baseline estimation, IEEE Trans. Smart Grid. 15 (2) (2024) 2251–2260.
[2] J. Feng, Y. Yao, Z. Liu, Developing an optimal building strategy for electric vehicle charging stations: automaker role, Environ. Dev. Sustainability 27 (5) (2025) 12091–12151.
[3] B. Cao, K. Liu, G. Wu, et al., A self-supervised evaluation approach of insulation condition for vehicle cable terminals using hypergraph neural network with dynamic features, IEEE Trans. Ind. Inf. 21 (2025) 1–11.
[4] T. Ji, S. Wang, J. Wang, J. Zhang, Beyond technology: role of technology-organization-environment-human factors in autonomous driving adoption for new energy vehicles, J. Retailing Consum. Serv. 88 (2026) 104552.
[5] Y. Zhang, J. Song, L. He, et al., Performance study of Tesla valve-based direct cooling thermal management system for batteries, Appl. Therm .Eng. 278 (2025) 127207.
[6] C.L. Tian, Y.C. Liu, G.Q. Zhang, Y.L. Yang, Y. Yan, C.D. Li, Transfer learning based hybrid model for power demand prediction of large-scale electric vehicles, Energy 300 (2024) 131461.
[7] Y. Zhang, J. Li, M. Yu, et al., Dominant factor identification and fast optimization of carnot battery by integrating SHAP and physics-guided neural network, Appl. Energy 401 (2025) 126641.
[8] H.X. Guo, L.X. Chen, Z.C. Wang, L. Li, Day-ahead prediction of electric vehicle charging demand based on quadratic decomposition and dual attention mechanisms, Appl. Energy 381 (2025) 125198.
[9] A.V. Sreekumar, R.R. Lekshmi, Electric vehicle charging station demand prediction model deploying data slotting, Results Eng. 24 (2024) 103095.
[10] M. Cruz, E.Y. Rineh, S.A.L. Fong, R.L. Cheu, Z.Q. Song, Using statewide transportation planning model to forecast demand for electric vehicle charging at stations along intercity highways, Int. J. Transp. Sci. Technol. (2024) 282–299.
[11] A. Cabrera-Tobar, N. Blasuttigh, A.M. Pavan, G. Spagnuolo, Demand response of an electric vehicle charging station using a robust-explicit model predictive control considering uncertainties to minimize carbon intensity, Sustainable Energy Grids Netw. 38 (2024) 101381.
[12] H. Liu, S. Zhou, W. Gu, et al., Fast dynamic identification algorithm for key nodes in distribution networks with large-scale DG and EV integration, Appl. Energy 388 (2025) 125608.
[13] C. Gujjarlapudi, D. Sarkar, S.K. Gunturi, Data driven machine learning models for short term load forecasting considering electrical vehicle load, Energy Storage (2024) e467.
[14] J. Liu, G. Lin, C. Rehtanz, S. Huang, Y. Zhou, Y. Li, Data-driven intelligent EV charging operating with limited chargers considering the charging demand forecasting, Int. J. Electr. Power Energy Syst. 141 (2022) 108218.
[15] B.-S. Lee, H. Lee, H. Ahn, Improving load forecasting of electric vehicle charging stations through missing data imputation, Energies 13 (2020) 18.
[16] S. Wang, C. Zhuge, C. Shao, P. Wang, X. Yang, S. Wang, Short-term electric vehicle charging demand prediction: a deep learning approach, Appl. Energy 340 (2023) 121032.
[17] A. Garrison, M. Rashid, N. Chen, A synergistic learning based electric vehicle charging demand prediction scheme, in: Proc. SoutheastCon 2023, 2023, pp. 5–10.
[18] B. Jiang, Y. Wang, Q. Wang, H. Geng, A hybrid KAN-ANN based model for interpretable and enhanced short-term load forecasting, IEEE Trans. Artif. Intell. (2026). https://doi.org/10.1109/TAI.2026.3656248
[19] Y.T. Shang, D. Li, Y. Li, S. Li, Explainable spatio-temporal multi-task learning for electric vehicle charging demand prediction, Appl. Energy 384 (2025) 125460.
[20] R. Tian, J. Wang, Z. Sun, J. Wu, X. Lu, L. Chang, Multi-scale spatial-temporal graph attention network for charging station load prediction, IEEE Access 13 (2025) 29000–29017.
[21] S.Y. Wang, Y. Li, C.F. Shao, P.X. Wang, A.X. Wang, C.X. Zhuge, An adaptive spatio-temporal graph recurrent network for short-term electric vehicle charging demand prediction, Appl. Energy 383 (2025) 125320.
[22] S.Y. Wang, A. Chen, P.X. Wang, C.X. Zhuge, Predicting electric vehicle charging demand using a heterogeneous spatio-temporal graph convolutional network, Transp. Res. Part C Emerg. Technol. 153 (2023) 104205.
[23] S.Y. Wang, Y. Li, C.F. Shao, P.X. Wang, A.X. Wang, C.X. Zhuge, An adaptive spatio-temporal graph recurrent network for short-term electric vehicle charging demand prediction, Appl. Energy 383 (2025) 125320.
[24] M. Boulakhbar, M. Farag, K. Benabdelaziz, T. Kousksouk, M. Zazi, A deep learning approach for prediction of electrical vehicle charging stations power demand in regulated electricity markets: the case of morocco, Cleaner Energy Syst.ems 3 (2022) 100039.
[25] P. Bakumar, S.K. Ramu, T. Vinopraba, Optimizing electric vehicle charging in distribution networks: a dynamic pricing approach using internet of things and bi-directional LSTM model, Energy 294 (2024) 130815.

[26] M.A. Zamee, D.J. Han, H. Cha, D.J. Won, Self-supervised online learning algorithm for electric vehicle charging station demand and event prediction, J. Energy Storage 71 (2023) 108189.

[27] J.S. Zhou, Y. Xiang, X. Zhang, Z. Sun, X.F. Liu, J.Y. Liu, Optimal self-consumption scheduling of highway electric vehicle charging station based on multi-agent deep reinforcement learning, Renew. Energy 238 (2025) 121982.

[28] D.E. Mekkaoui, M.A. Midoun, A. Smaili, B. Feng, M.Z. Talha, Y.M. Shen, Probabilistic dual-adaptive spatio-temporal graph convolutional networks for forecasting energy consumption dynamics of electric vehicle charging stations, Comput. Electr. Eng. 122 (2025) 109976.