



OPEN ACCESS

EDITED BY

Xin Ning,
Chinese Academy of Sciences (CAS),
China

REVIEWED BY

Tomasz Górski,
University of Gdansk, Poland
Luyang Hou,
Beijing University of Posts and
Telecommunications (BUPT), China
Juvenal Rodriguez,
Autonomous University of Queretaro,
Mexico
Ebrahim Elsayed,
Mansoura University, Egypt

*CORRESPONDENCE

Xuhan Li,
✉ littlecat0601@163.com

RECEIVED 24 May 2023

ACCEPTED 14 June 2023

PUBLISHED 26 July 2023

CITATION

Li X (2023), CNN-GRU model based on
attention mechanism for large-scale
energy storage optimization in smart
grid.
Front. Energy Res. 11:1228256.
doi: 10.3389/fenrg.2023.1228256

COPYRIGHT

© 2023 Li. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](#). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

CNN-GRU model based on attention mechanism for large-scale energy storage optimization in smart grid

Xuhan Li*

School of Electrical Engineering, Hebei University of Technology, Tianjin, China

Introduction: Smart grid (SG) technologies have a wide range of applications to improve the reliability, economics, and sustainability of power systems. Optimizing large-scale energy storage technologies for smart grids is an important topic in smart grid optimization. By predicting the historical load and electricity price of the power system, a reasonable optimization scheme can be proposed.

Methods: Based on this, this paper proposes a prediction model combining a convolutional neural network (CNN) and gated recurrent unit (GRU) based on an attention mechanism to explore the optimization scheme of large-scale energy storage in a smart grid. The CNN model can extract spatial features, and the GRU model can effectively solve the gradient explosion problem in long-term forecasting. Its structure is simpler and faster than LSTM models with similar prediction accuracy. After the CNN-GRU extracts the data, the features are finally weighted by the attention module to improve the prediction performance of the model further. Then, we also compared different forecasting models.

Results and Discussion: The results show that our model has better predictive performance and computational power, making an important contribution to developing large-scale energy storage optimization schemes for smart grids.

KEYWORDS

CNN, GRU, attentional mechanisms, smart grid, large-scale energy storage optimization solutions

1 Introduction

Smart Grid is a new type of power grid based on information technology, automation technology, and energy technology (Estévez-Bén et al., 2020b). Through monitoring, scheduling, and controlling the whole process of power production, transmission, distribution, and use, it realizes intelligent management and optimization of the power system Alvarez-Diazcomas et al. (2019), thus improving its reliability, economy, and sustainability. Energy storage technology is an important part of the smart grid, which can increase the power system's flexibility and dispatchability, reduce the system's peak-to-valley difference, and improve the efficiency and stability of power grid operation. Deep learning is a neural network-based machine learning method that can automatically extract features from data and perform pattern recognition and prediction by learning from a large amount of data López et al. (2018). In a large-scale energy storage optimization

scheme for smart grids, deep learning models can be used to predict the load of the power system and the price of electricity to develop an optimal energy storage strategy.

At this stage, a common deep learning-based large-scale energy storage optimization scheme for smart grids is that: First, perform data preprocessing: collect historical load and electricity price data of the power system and perform preprocessing, including data cleaning, normalization, and temporalization (Rezaeimozafer et al., 2022); second, perform deep learning modeling: use deep learning models such as Recurrent Neural Network (RNN) (Abedi and Kwon, 2023) or Long Short-term memory (LSTM) (Zhao et al., 2022) to Second, deep learning modeling is performed: using deep learning models such as Recurrent Neural Network (RNN) or Long Short-term memory (LSTM) to learn and predict historical load and electricity price data. The prediction results are compared with the actual data to evaluate the accuracy and stability of the model; thirdly, energy storage optimization is formulated: according to the prediction results of the deep learning model, the optimal energy storage strategy is formulated, including the selection of the best energy storage equipment, energy storage capacity, charging and discharging strategy. At the same time, considering the changes in the power market and the economics of energy storage equipment, the optimal energy storage plan is formulated by considering the costs and benefits; finally, real-time control: according to the energy storage plan, the energy storage equipment is controlled and dispatched in real-time to ensure the stable operation of the power Estévez-Bén et al. (2019) system (Lai et al., 2022). The models commonly used for learning and forecasting historical load and power price data are time series model: time series model is a statistical model based on time series data, which can be used to predict future trends and seasonal changes, but there are strict assumptions on data smoothness, autocorrelation, which need to meet certain preconditions, may not be flexible enough for nonlinear and complex relationships, prediction The disadvantages include limited accuracy. In the power system, the commonly used time series models include the ARIMA model (Zhang et al., 2022), seasonal ARIMA model Han et al. (2022), exponential smoothing model (Zheng and Jin, 2022); regression model: regression model predicts future trends by establishing the relationship between load or power price and some related factors, but it may not be flexible enough for non-linear and complex relationships and has limited prediction accuracy. In the power system, these factors may include temperature, humidity, weather, economic indicators, etc. Commonly used regression models include linear regression (Choo and Go, 2022), polynomial regression (Jena and Ray, 2023), ridge regression (Ahmad et al., 2022); neural network model Abedi and Kwon (2023): neural network model is a nonlinear model that predicts future trends by learning historical data, which can handle large amounts of data and nonlinear relationships and is suitable for complex prediction tasks, but the training process requires a lot of computational resources and time. In a power system, the commonly used neural network models include feedforward neural networks, recurrent neural networks, long and short-term memory networks; decision tree model (Mostafa et al., 2022): decision tree model predicts future trends by constructing a tree-like structure, which can handle nonlinear and interactive relationships, and is suitable for complex prediction tasks, but there may be overfitting problems for high-dimensional data and noisy data. The commonly used decision

tree models in power systems include the CART (Classification and Regression Tree) algorithm, random forest (Estévez-Bén et al., 2020a).

Based on the advantages and disadvantages of the above models, this paper proposes a prediction model that combines an attention mechanism-based convolutional neural network (CNN) (Lu et al., 2022) and a recurrent gated unit (GRU) (Xiao et al., 2023). First, the historical data and other predictive indicators are input into the CNN for convolutional processing to extract their important features. Then the output results are passed through the GRU network, which can effectively perform long-term time series prediction, which is an improved version of the LSTM model, and then go through the attention mechanism (Li et al., 2022) module to reasonably assign weights and optimize the model through autonomous learning, and finally form a CNN-GRU-AM model to predict the historical load and power price of the power system, and then develop a reasonable large-scale energy storage optimization scheme.

The contribution points of this paper are as follows.

- Compared with traditional financial time series and regression models, the attention mechanism can automatically focus on historical data that have a decisive impact on the current prediction results by learning important features in the historical data. Compared with other models requiring manual selection and feature extraction, the attention mechanism can automatically discover important features and improve prediction accuracy.
- Compared with decision tree models, CNN-GRU models combine convolutional neural networks and gated recurrent unit networks, which can handle sequential and spatial data. This allows the model to consider both the time-series characteristics of historical load or electricity prices and their spatial distribution characteristics, thus improving the prediction accuracy.
- Compared with deep learning models, CNN-GRU models have better generalization ability and can be applied to different electricity systems and scenarios. Compared with other models that need to be retrained in each scenario, the CNN-GRU model can apply the learned knowledge to new scenarios through migration learning and other methods, thus improving the efficiency and effectiveness of the model.

In the rest of this paper, we present recent related work in Section 2. Section 3 offers our proposed methods: overview, convolutional neural networks; gated recurrent unit (GRU); attention mechanism. Section 4 presents the experimental part, details, and comparative experiments. Section 5 concludes.

2 Related work

Optimizing energy storage systems in smart grids is key to improving energy efficiency, reducing costs, and ensuring a reliable energy supply. Researchers have used various techniques to optimize energy storage systems, including reinforcement learning, LSTM networks, and multi-objective optimization. These techniques

offer different advantages and disadvantages but aim to find the optimal trade-off between competing goals. By leveraging these technologies, energy providers, consumers, and policymakers can develop more efficient and effective energy storage systems to meet the needs of various stakeholders and contribute to a more sustainable energy future. In the rest of this section, three works on optimizing energy storage systems, namely, reinforcement learning, LSTM network, and multi-objective optimization, will be introduced.

2.1 Reinforcement learning for energy storage optimization

This work uses reinforcement learning algorithms to make energy storage systems operate more efficiently and reliably in smart grids (Cao et al., 2020). In this field of study, researchers have proposed various optimization methods to improve the efficiency and reliability of energy storage systems. Reinforcement learning is an algorithm widely used in machine learning, which trains agents to make decisions through trial-and-error interactions. Researchers have recently used reinforcement learning to optimize energy storage systems in smart grids. For example, Liu et al. (2020) proposed a reinforcement learning-based algorithm for optimizing energy storage scheduling in microgrids. Wang and Hong (2020) used a deep reinforcement learning algorithm to optimize the operation of energy storage in renewable energy systems. These studies demonstrate that reinforcement learning can effectively optimize energy storage systems in smart grids.

An energy storage scheduling algorithm based on deep reinforcement learning is at the heart of the research. The algorithm uses an intelligent agent to learn how to make decisions about storing and releasing energy for maximum economic efficiency and reliability. The algorithm's input includes historical load data, energy prices, and weather forecasts. Based on these inputs, the agent learns how to make optimal decisions. During training, the agent interacts with the environment and continuously adjusts its policy to maximize cumulative reward. The algorithm takes a value-based approach, where a value function is used to assess the value of each state, and based on that value, the decision to store and release energy is made.

The study evaluates the algorithm's performance by conducting simulation experiments in an experimental environment. The results show that the algorithm can significantly improve the efficiency and reliability of energy storage systems. Compared with traditional methods, the algorithm can reduce energy costs and improve energy usage efficiency. In addition, the algorithm can adapt to different environments and needs and has high practicability and adaptability.

2.2 Multi-objective optimization for energy storage systems

Multi-objective optimization is a powerful technique that can help find the best compromise between multiple conflicting goals Terlouw et al. (2019). In the context of smart grid energy

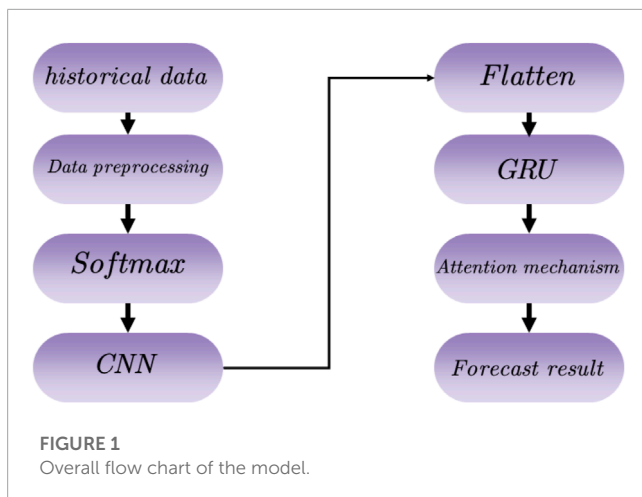
storage systems, multiple objectives often need to be considered simultaneously, such as maximizing renewable energy, minimizing energy costs, and ensuring a reliable energy supply. Multi-objective optimization can be used to find the optimal trade-off between these objectives, resulting in more efficient and effective energy storage systems.

Several studies have used multi-objective optimization to optimize energy storage systems in smart grids. One such study by Li et al. (2018) proposed a multi-objective optimization model for energy storage dispatch in microgrids. The model considers several objectives, such as minimizing energy costs, maximizing the use of renewable energy, and ensuring a reliable energy supply. Another study by Yang et al. (2021) Uses a multi-objective optimization algorithm to design the capacity and configuration of energy storage systems in community microgrids. The study considered several goals: minimizing energy costs, reducing greenhouse gas emissions, and ensuring a reliable energy supply. These studies demonstrate that multi-objective optimization can effectively optimize energy storage systems in smart grids. By considering multiple objectives simultaneously, multi-objective optimization can help to find more efficient and effective energy storage systems to meet the needs of various stakeholders, including energy suppliers, consumers, and the environment.

2.3 Long short-term memory (LSTM) networks for energy storage optimization

A Long Short-Term Memory (LSTM) network is a neural network architecture designed for processing sequential data (Balakumar et al., 2023). The architecture of an LSTM network is based on the concept of a recurrent neural network (RNN), capable of processing sequential data by maintaining a state that is updated at each time step of the sequence.

Several recent studies have used LSTM networks to optimize energy storage systems in smart grids. One such study by Li F. et al. (2021a) Used LSTM-based models to predict building energy consumption and optimize energy storage dispatch. The model inputs historical energy consumption data and weather-related variables and outputs an optimal energy storage dispatch strategy for buildings. The study results show that the LSTM-based model can accurately predict energy consumption and optimize energy storage scheduling, thereby reducing energy costs. Another study by Zhao et al. (2022) Used LSTM-based models to predict wind power output and optimize energy storage operations in wind-solar-battery microgrids. The model takes historical wind and solar power generation output data and battery state-of-charge (SOC) data as input, and outputs an optimal battery charging and discharging strategy. The study found that the LSTM-based model could accurately predict wind power output and optimize battery operation, increasing renewable energy utilization and reducing energy costs. These studies demonstrate that LSTM networks can effectively predict energy consumption and optimize energy storage dispatch in smart grids. The ability of LSTM networks to process time series of data makes them a powerful tool for energy management and optimization in smart grids.



3 Methodology

In this section, we will introduce the method proposed in this paper in detail, first an overview of the method, and then introduce the CNN network, GRU network and attention mechanism used in this paper.

3.1 Overview of our network

In this paper, a CNN-GRU model based on the attention mechanism is proposed to predict the load and power price of the power system so as to formulate the optimal energy storage strategy and effectively solve the problems of missing historical data and complicated operation of previous deep learning models in long period sequences, while improving the accuracy of the prediction model, the flow chart of the model is shown in [Figure 1](#).

First, select the appropriate historical data input, then perform data preprocessing and normalization processing to make the operation speed of the CNN network faster. Then, the data enters the CNN network for feature extraction. Since the one-dimensional CNN network is selected here, it is more suitable for the time series prediction; after the flattening layer, the data is subjected to dimensionality reduction processing. After that, the data is subjected to the GRU layer for feature extraction and enters the attention layer. Weight distribution is performed through independent learning, the CNN-GRU structure is optimized, and finally, output predicted results [Feng et al. \(2022\)](#). The CNN-GRU-AM model consists of three parts: the CNN module, the GRU module, and the attention module. Each of the three components plays its advantage to complete the load and power price prediction of the power system. The overall structure of the model is shown in [Figure 2](#).

3.2 CNN model

Convolutional Neural Network (CNN) is a neural network architecture widely used in computer vision tasks. CNN is a feed-forward neural network that consists of multiple convolutional, pooling, and fully-connected layers. The main feature of CNN is

the ability to capture local features and spatial structures in images and to learn more abstract feature representations. This is achieved using convolutional layers, which convolve the input image with filters (kernels or feature detectors). The convolution operation can effectively reduce the number of network parameters and preserve the local information in the input image. Pooling layers are usually followed by convolutional layers, which can further reduce the size of feature maps and keep their essential features. Convolutional neural networks can be divided into one-dimensional CNN and multi-dimensional CNN according to dimension. One-dimensional CNN is mainly used to process time series data, while multi-dimensional CNN is primarily used to recognize text, image, and video data. Therefore, this paper uses one-dimensional CNN to process the time series of historical data of Smart Grid. Its model structure diagram is shown in [Figure 3](#):

In [Figure 3](#), we have established a network consisting of an input layer, a 2-layer 1D convolution, a 2-layer 1D max pooling, a flattening layer, and an output layer for the nonlinear and complex data characteristics of a one-dimensional time series. Dimensional CNN network [Kattenborn et al. \(2021\)](#). The convolutional layer is used to extract temporal features; ReLU is the activation function, and the highest pooling layer reduces the output dimension of the convolutional layer to extract the most significant features; the flattening layer converts the multi-dimensional feature vector into a one-dimensional feature vector, as Global feature extraction. CNN can effectively and automatically extract and learn nonlinear features of one-dimensional time series, which can remove time series features of different scales by alternately stacking convolutional layers and top pooling layers.

The formula of convolution is shown in [Formula \(1\)](#):

$$g(i) = \sum_{x=1}^m \sum_{y=1}^n \sum_{z=1}^p a_{x,y,z} w_{x,y,z}^i + b^i, \quad i = 1, 2, \dots, q \quad (1)$$

Where: $g(i)$ is used to represent the i -th feature map, a is used to represent the input data, i is used to represent the first convolution kernel, b represents the bias, and x, y, z represents the three dimensions of the input.

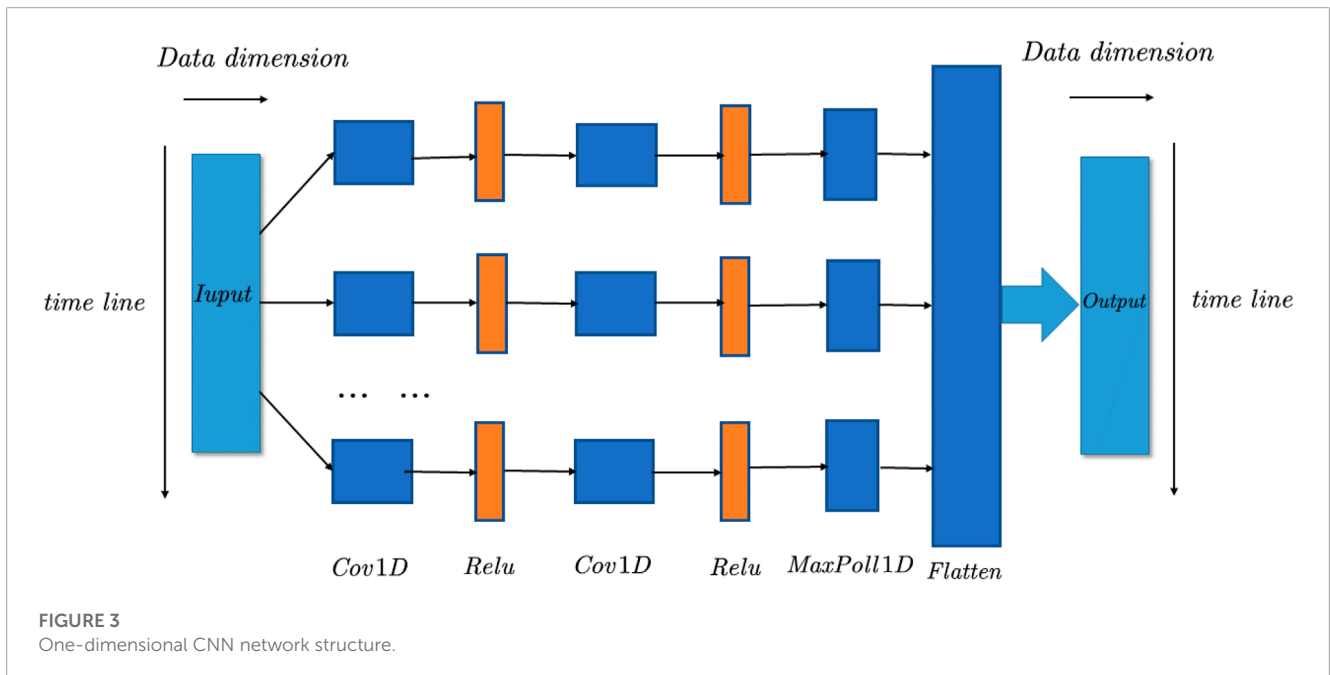
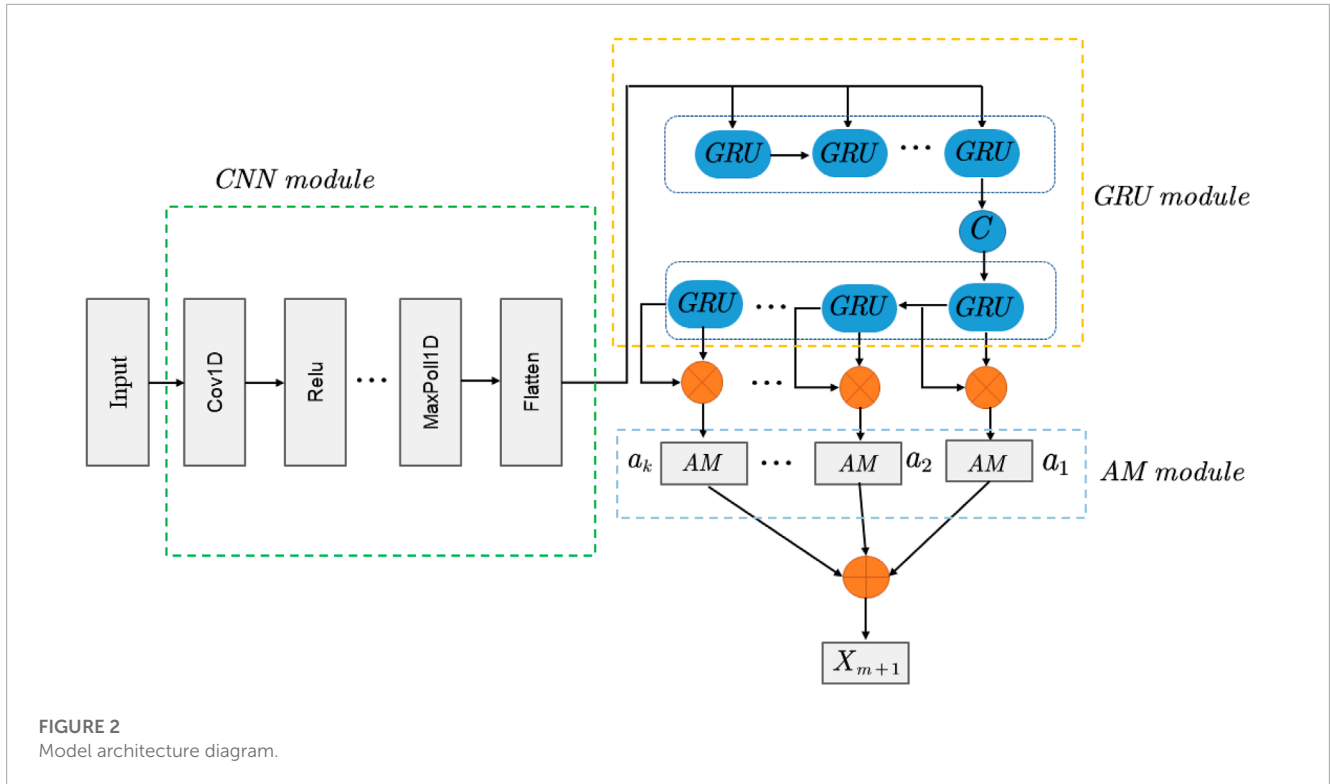
After the convolution operation, the activation Li number is usually used to realize the nonlinear transformation, and *RELU* is used as the activation function in this paper.

$$y(i) = f(g(i)) = \max\{0, g(i)\}, \quad i = 1, 2, \dots, q \quad (2)$$

Considering the one-dimensional characteristics of track data, this paper uses a one-dimensional convolutional neural network to process one-dimensional data. The convolution formula is shown in [Formula \(3\)](#):

$$g(i) = \sum_{x=1}^m a_x w_x^i + b^i, \quad i = 1, 2, \dots, q \quad (3)$$

The features of the financial time series are extracted in the form of a sliding window with a step size of l . The elements of all windows will be fused through the pooling layer to retain the main features of the time series. Then the output will be put into the next convolutional layer for further processing—extraction of high-dimensional features. The final result of the CNN module is obtained by flattening layers and fully connecting layers.



3.3 GRU model

A gated Recurrent Unit (GRU) (Li W. et al., 2021b) is a recurrent neural network (RNN) structure similar to long short-term memory (LSTM). Compared with LSTM, GRU has fewer parameters simpler structure, and the training efficiency of the model is higher, so it runs faster on some tasks and is an advanced variant of the LSTM model. GRUs are designed to alleviate the vanishing gradient problem

when training long-term sequences. It has been widely used in natural language processing (NLP) tasks, such as language modeling, machine translation and text generation, and other modeling tasks of sequence data. The structure diagram of GRU is shown in Figure 4.

Compared with the three gate structures of the input gate, output gate, and forget the entrance of the LSTM structure, the GRU structure only has two gate structures a reset gate and an update gate. The reset gate controls the influence of the hidden state of the last

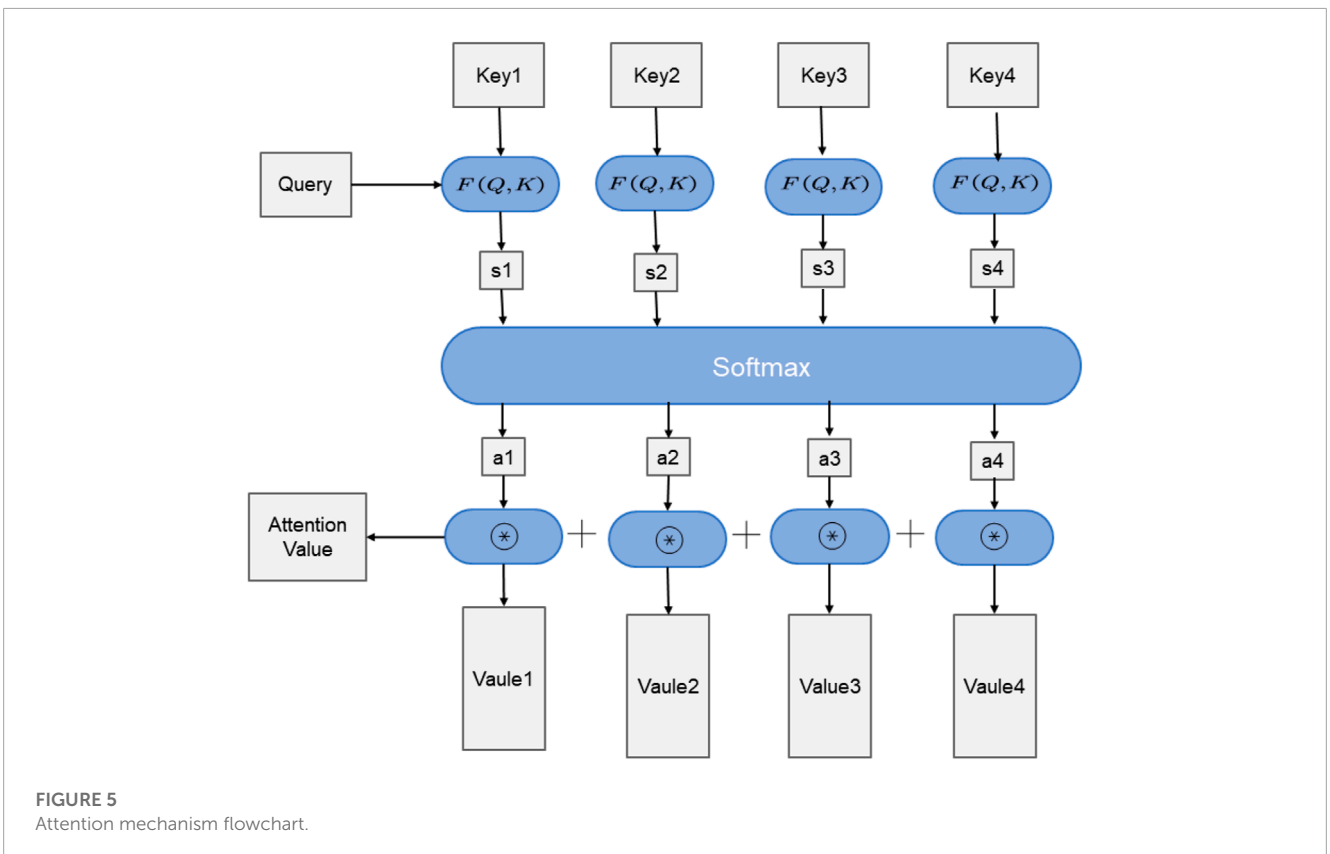
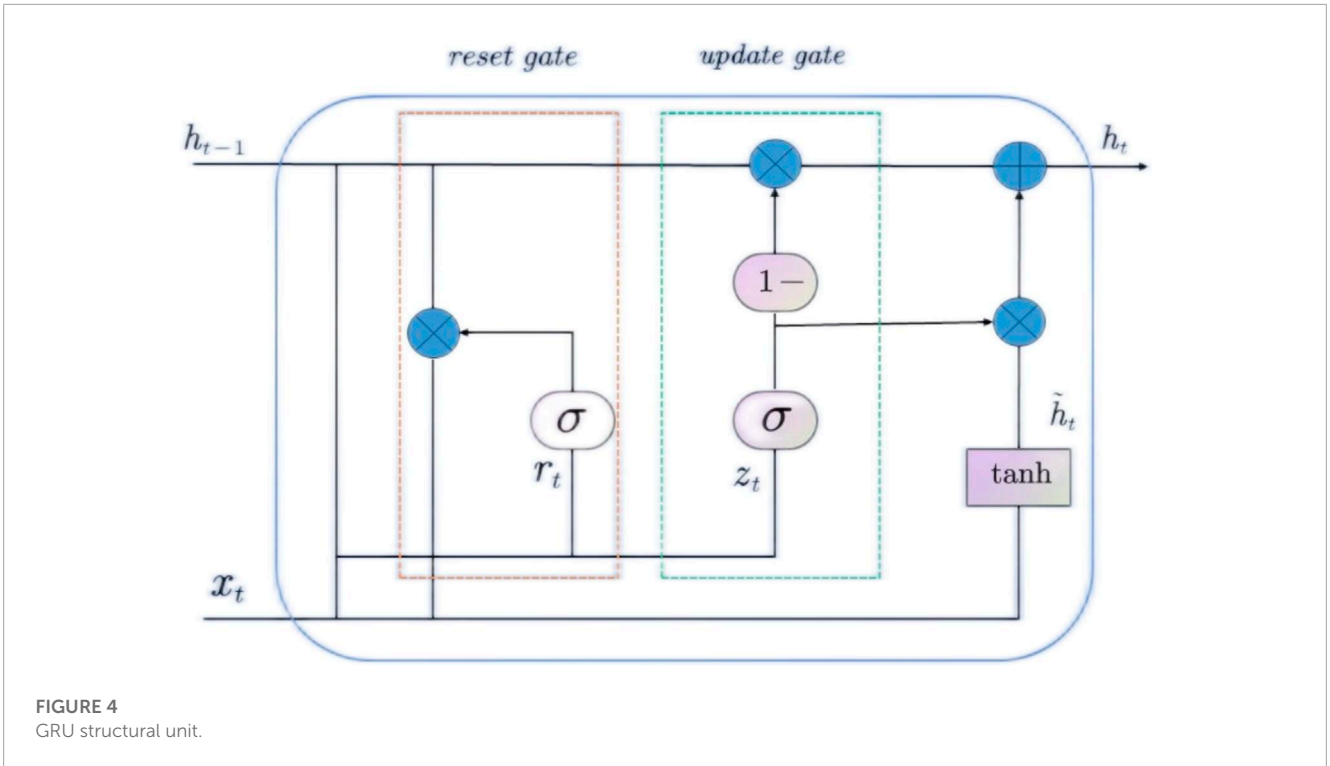


TABLE 1 Smart grid index classification table.

Index	Value	Code	Classification
Index1	Peak-to-valley difference	C1	Load
Index2	Load rate	C2	Load
Index3	Energy utilization	C3	Energy utilization and efficiency
Index4	Power generation efficiency	C4	Energy utilization and efficiency
Index5	Electricity market price	C5	Electricity market
Index6	Smart grid operational cost	C6	Electricity market
Index7	Grid reliability index	C7	Grid stability
Index8	Power system response time	C8	Grid stability
Index9	Power system controll ability	C9	Grid stability
Index10	Power system stability	C10	Grid stability
Index11	Renewable energy consumption	C11	Emissions
Index12	Emissions	C12	Emissions
Index13	Energy security	C13	Energy
Index14	Market transparency	C14	Energy
Index15	Energy storage efficiency	C15	Energy

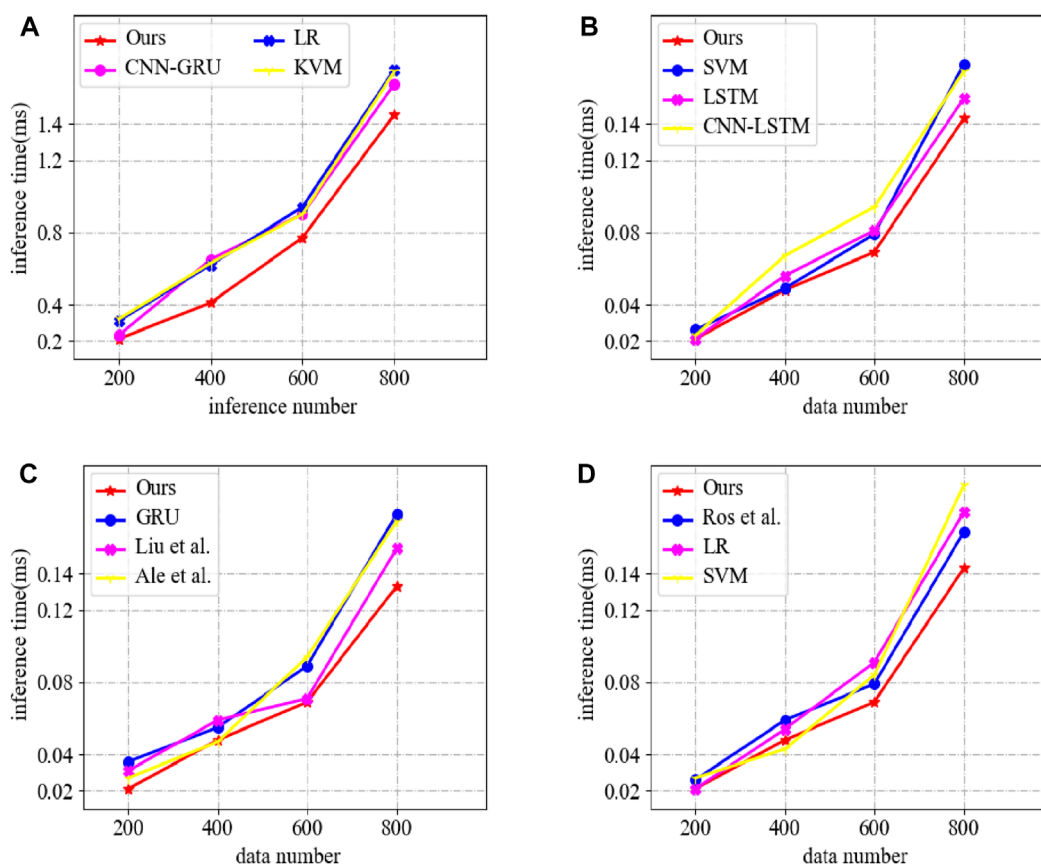


FIGURE 6 Comparison of inference time of different models. (A) corresponds to Dataset 1 from the EIA dataset, (B) corresponds to Dataset 2 from the EPEX dataset, (C) corresponds to Dataset 3 from the IEA dataset, and (D) corresponds to Dataset 4 from the CNEE dataset.

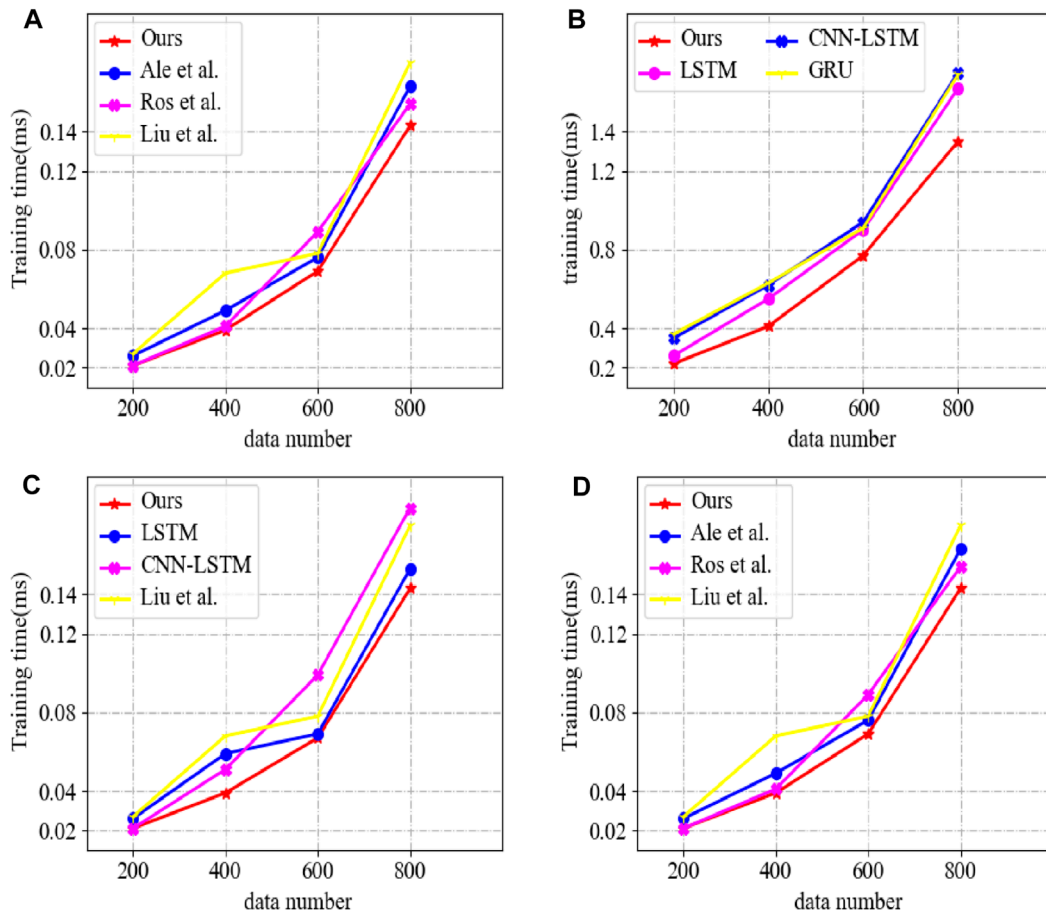


FIGURE 7 Comparison of inference time of different models. (A) corresponds to Dataset 1 from the EIA dataset, (B) corresponds to Dataset 2 from the EPEX dataset, (C) corresponds to Dataset 3 from the IEA dataset, and (D) corresponds to Dataset 4 from the CNEE dataset.

moment on the input of the current moment, and the update gate controls the power of the hidden state of the last moment on the output of the present moment. GRU also uses a candidate hidden form to compute the hidden state at the current moment.

The formula of the GRU model is as follows:

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \tag{4}$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1}) \tag{5}$$

$$\tilde{h}_t = f(W_h x_t + U_h (r_t \odot h_{t-1})) \tag{6}$$

$$h_t = z_t \odot \tilde{h}_t + (1 - z_t) \odot h_{t-1} \tag{7}$$

In the formula: r_t is the reset gate; z_t is the update gate; x_t is the input information; σ is the sigmoid activation function; h_{t-1} is the state of the hidden layer at the previous moment; f is the tanh activation function; \tilde{h}_t is the input x_t , and the previous hidden layer; \odot is the element-wise multiplication.

3.4 Attention mechanism

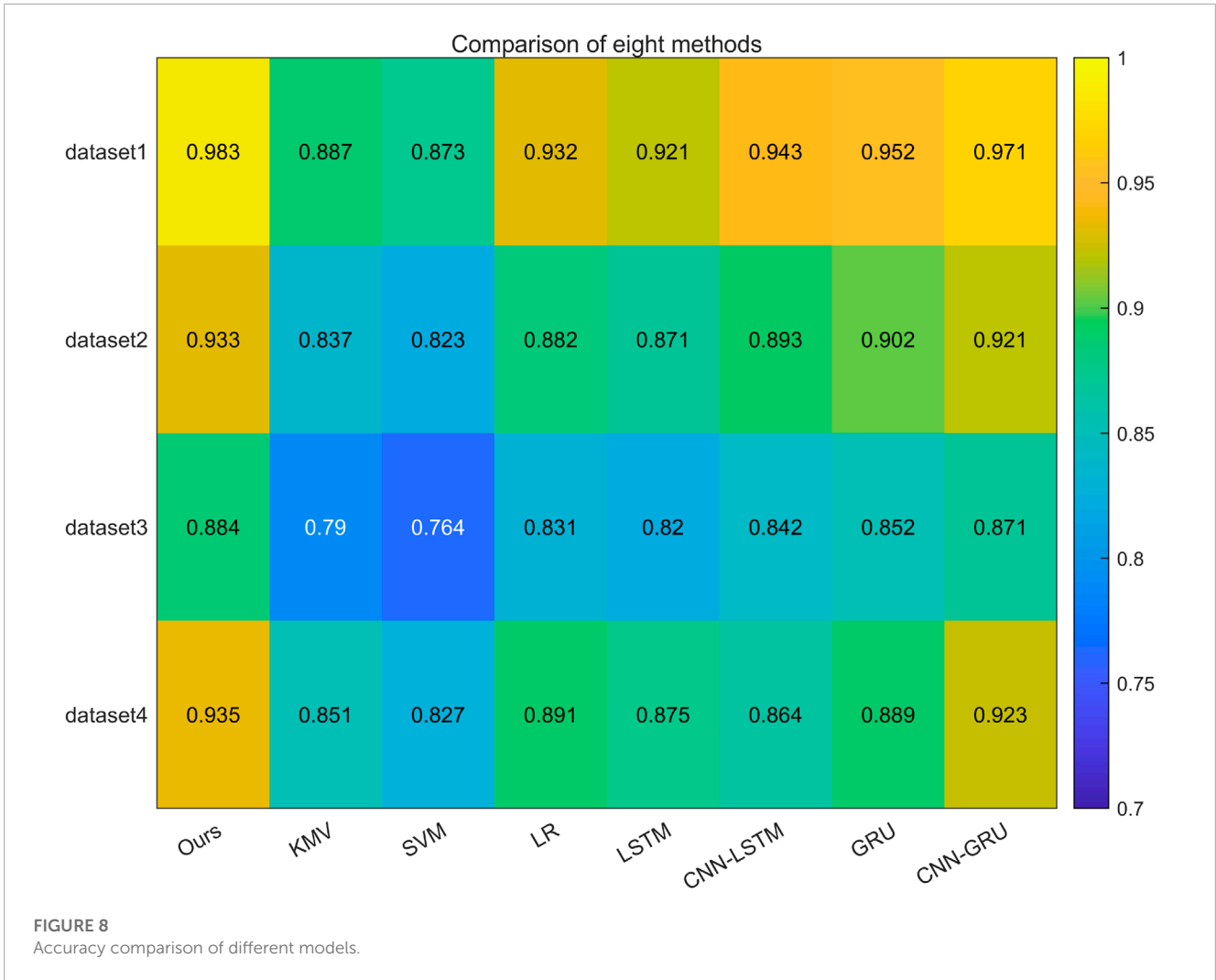
The attention mechanism is a mechanism that simulates the perception process of human vision or hearing, and its primary

function is to select relevant information from the input data and assign it to the corresponding output (Li et al., 2023). In machine learning, attention mechanisms often process sequential data, such as natural language sentences or time series data. Attention mechanisms are widely used in natural language processing, especially machine translation tasks (Guo et al., 2022). It is considered one of the essential means to improve the performance of machine translation. At the same time, the attention mechanism has also been applied in image processing and speech recognition. Its structure diagram is shown in Figure 5.

In natural language processing tasks, the attention mechanism can help the model focus on the part relevant to the current task in encoding the input. Typically, the input data is in sequences, such as sentences or paragraphs of text. During encoding, the model progressively reads each element in the series and generates the corresponding hidden state. On this basis, the attention mechanism can help the model pay attention to other factors related to the current element for better prediction and decision-making.

The formula of the attention mechanism is as follows:

$$a_i = \frac{\exp(s(h_i, h_t))}{\sum_{i=1}^N \exp(s(h_i, h_t))} \tag{8}$$



$$O = \sum_{i=1}^N a_i h_i \tag{9}$$

Where: $s(h_i, h_t)$ is the score of the feature vector. The larger the score, the greater the attention. a_i is the weight value of the i th input feature in the attention mechanism, which also is the ratio of the score of the feature vector to the whole. Then all vectors are summed and averaged to get the final vector o .

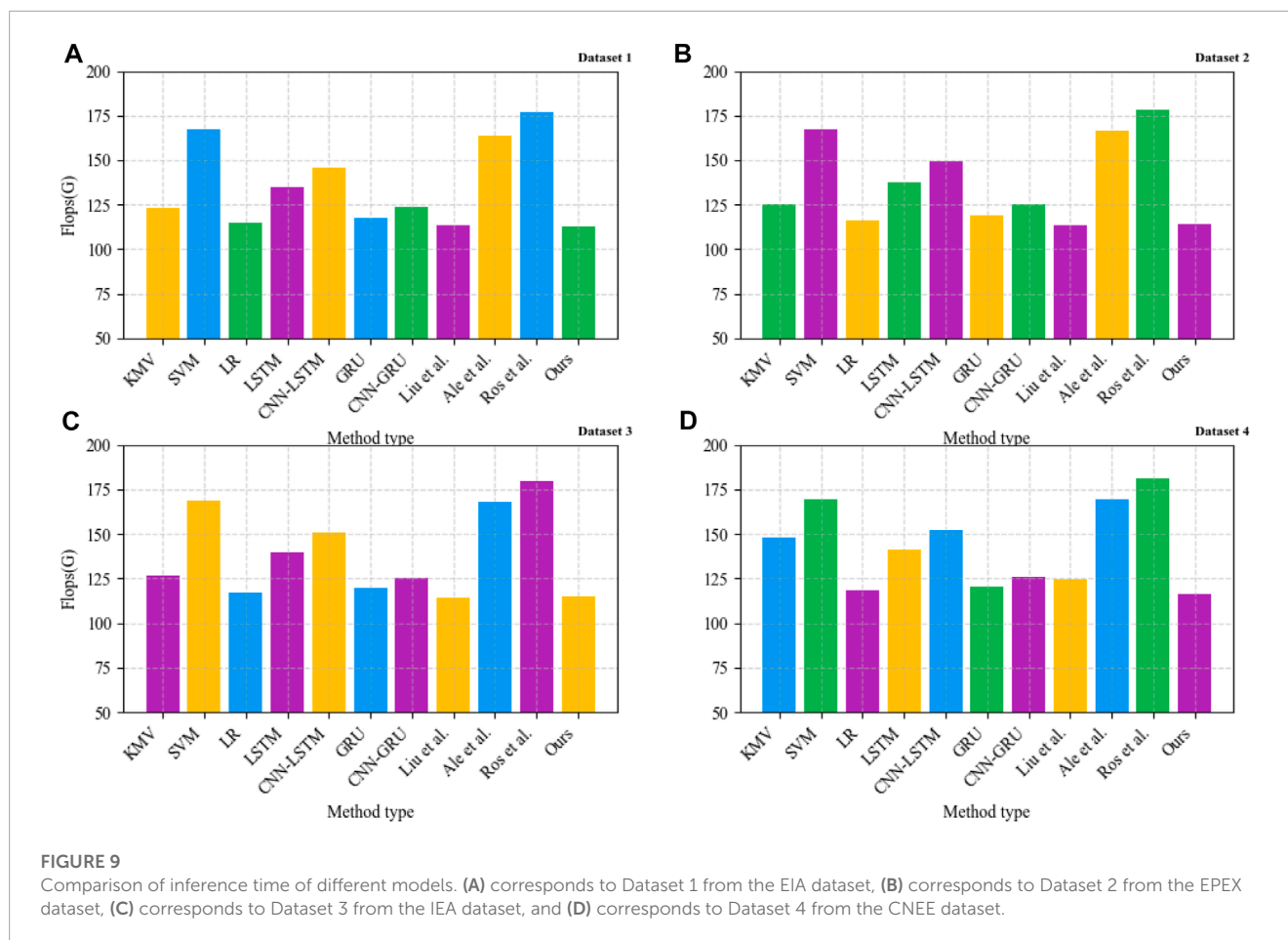
This formula can be understood as selecting other elements related to the current component according to the query vector in encoding the input sequence and then performing the weighted average of their value vectors. This weighted average method can make the model pay more attention to the information related to the current task while ignoring the information irrelevant to the study, thereby improving the model's performance.

4 Experiment

In this section, we will introduce our laboratory, firstly, the selection of data sets, followed by the experimental settings and results analysis and summary.

4.1 Datasets

The U.S. Energy Information Administration (EIA) is an independent statistical agency of the U.S. government that collects, analyzes, and publishes data, reports, and analyzes (Gao et al., 2022) on energy, electricity, and the environment. EIA is one of the nation's leading energy intelligence agencies, and its data and analysis significantly impact the U.S. government, businesses, and the public. It aims to provide accurate and comprehensive energy information to support national energy decision-making and planning and reliable energy information to policymakers, industry, academia, and the public. EIA is responsible for collecting data on energy production, consumption, and transportation across the United States, and performing statistics and analysis on these data, generating various energy data reports and databases, analyzing and predicting energy trends: EIA uses the collected energy data to conduct various analyzes and forecasts, including energy prices, supply, and demand trends, environmental impacts, etc. EIA provides the government and the public with information about future energy trends; supports policy formulation: EIA's data and analysis greatly influence the energy policy formulation of the U.S.



federal government and state governments and can provide the government with a key basis for energy decision-making.

The European Power Exchange (EPEX) is one of the largest power exchanges in Europe, established in 2008 and headquartered in (Mahilong et al., 2022) in Brandenburg, Germany. EPEX is a trading platform created by merging power exchanges in several European countries, aiming to provide more efficient trading and clearing services for the entire European power market. In addition to providing trading and clearing services, EPEX is committed to providing reliable market data and analysis for the European electricity market. EPEX releases data such as real-time prices and trading volumes of the electricity market and various market analysis reports daily, providing an important decision-making basis for market participants. In short, EPEX is one of Europe's largest electricity trading platforms, providing efficient, transparent, and secure trading and clearing services for the entire European electricity market and providing market participants with reliable market data and analysis.

The International Energy Agency (IEA) (Hattori et al., 2022) aims to promote the rational use of energy, energy security, and sustainable environmental development and provide support and advice for formulating and implementing global energy policies. The main task of the International Energy Agency is to provide member countries with support and advice on energy policy and technology and to promote the rational use and sustainable development of

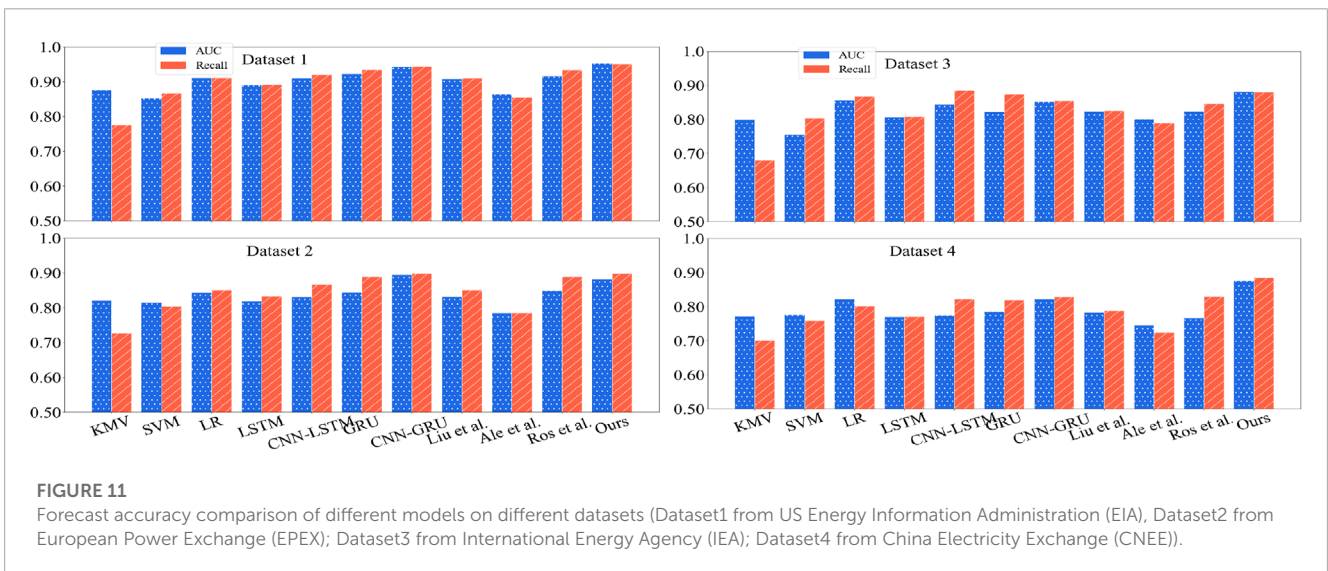
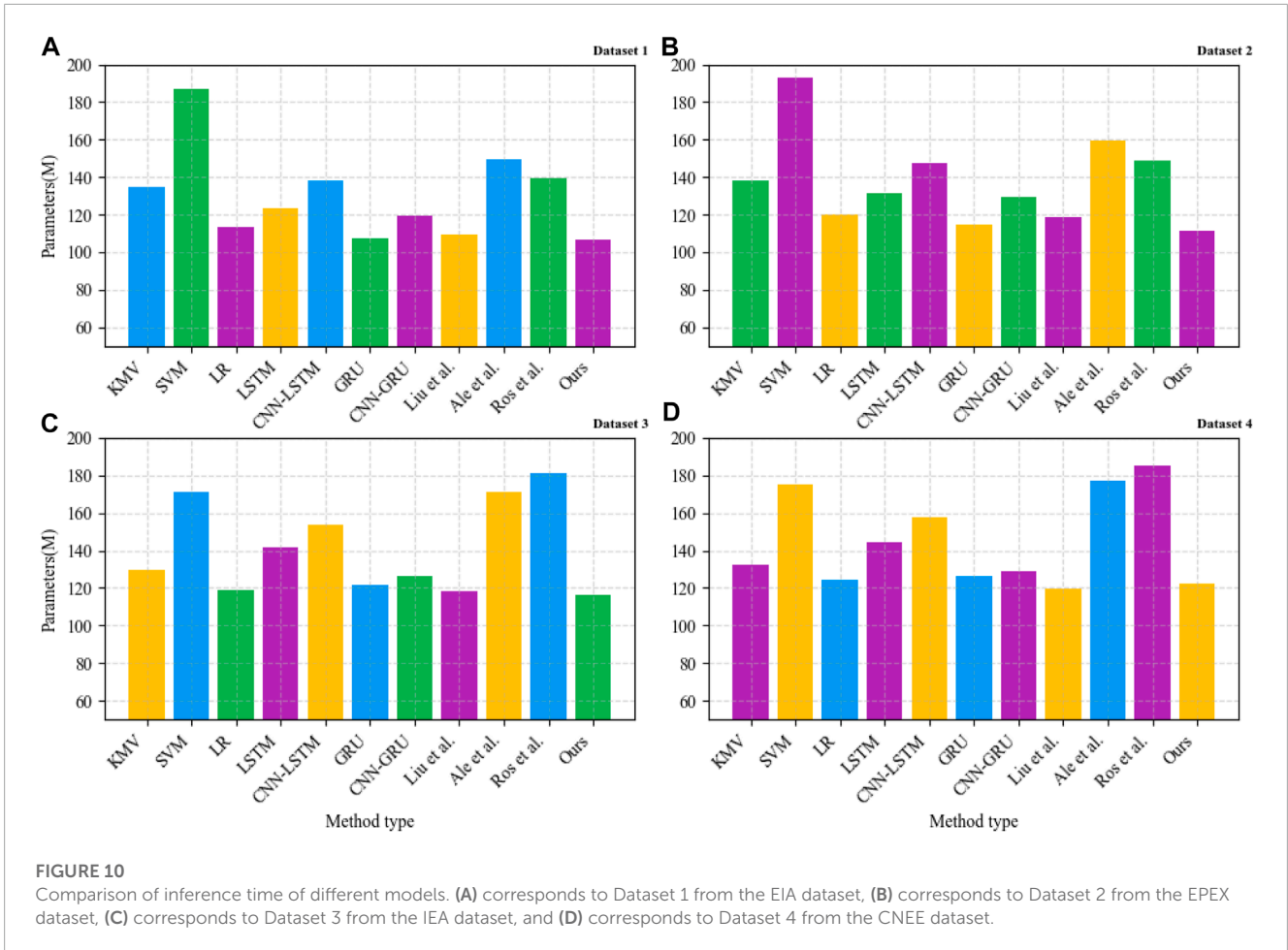
global energy. The International Energy Agency is an independent international organization dedicated to supporting and advising member countries on energy policy and technology to promote the rational use and sustainable development of global energy. The work of the International Energy Agency covers many energy fields and has great influence and importance.

China Electricity Council (Juárez Coronado, 2022) is one of the core trading platforms of China's electricity market and an important part of China's electricity market reform: market transparency and competitiveness. CNEE is one of the core trading platforms of China's electricity market, undertaking functions such as electricity trading, market supervision, information release, and market analysis. The establishment and operation of CNEE will play an important role in promoting the reform of my country's electricity market, promoting the optimal allocation and utilization of electricity resources, and improving the transparency and competitiveness of the electricity market.

We selected 15 indicators from these data as the input set for the model, as shown in Table 1.

Impact of the model on the intensity of communication between electricity market participants:

1. The impact of the model on information collection and analysis: The CNN-GRU model based on the attention mechanism can better capture and analyze patterns and trends in historical data, thereby increasing the sensitivity of market participants to market



changes. This will help facilitate the exchange of information among market participants, including information on trading strategies and risk management [Górski \(2023\)](#).

2. The impact of the model on information transmission and reception: By predicting future energy demand and renewable energy generation, the CNN-GRU model based on the attention

mechanism can help market participants better plan and coordinate their trading behavior [Zhao et al. \(2023\)](#). This will help facilitate the transfer and receipt of information between market participants, including information on the execution of orders and trades.

3. The impact of technical diversity on information exchange: Besides the CNN-GRU model based on the attention mechanism,

other models and technologies can be used to predict information, such as energy demand and production. These different models and techniques may have different effects on the exchange of information among market participants. Therefore, this study could compare and analyze different models and techniques to assess their impact on information exchange.

4.2 Experimental details

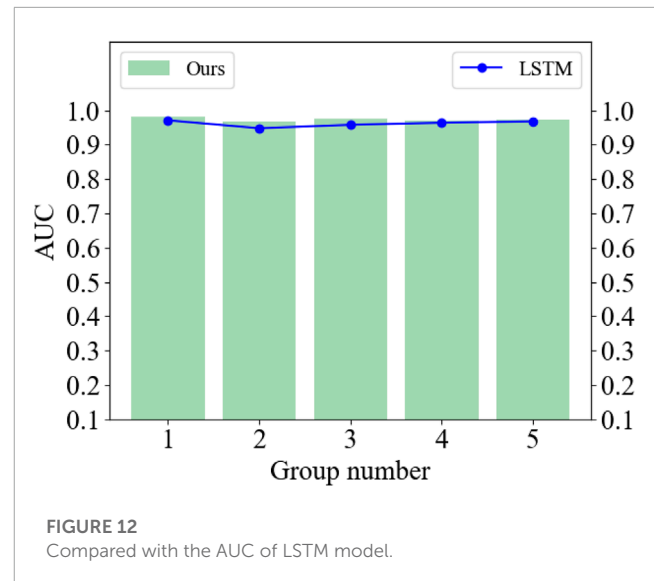
In this paper, to test the performance of our model, we conduct several sets of experiments to compare the performance of our model. First, we conduct the inference time and training time comparison experiments to test the basic performance of our model and the accuracy of the model to predict the load and price of the power system, which is an important criterion to test the performance of the model, for this purpose, we will conduct several sets of experiments to compare the prediction accuracy of different models and our model on different data sets, and the prediction ability of the same data set with different levels of complexity. For this purpose, we will conduct several experiments to compare the prediction accuracy of different models and our model on different data sets and the prediction ability of data of different complexity in the same data set. Then we also compare the number of parameters and operations between different models. We also conduct a comparison between AUCs to test the performance of our model.

4.3 Experimental results and analysis

In [Figure 6](#), we compare the inference time of LR, CNN-GRU, KVM, SVM, GRU, Liu et al., [Alegria et al. \(2013\)](#); [Rostampour et al. \(2019\)](#) and our model for a total of 11 models with the different numbers of inferences. However, as the number of inferences keeps increasing, the advantage of our model in terms of computing speed becomes more and more obvious, which is due to the attention mechanism that can autonomously learn to assign weights and give more attention to more important data features, so that the deep learning model has the more computational power to compute these more important features, thus increasing the inference speed of the model.

In [Figure 7](#), we compare the training time of LR, CNN-LSTM [Wu et al. \(2021\)](#), KVM, SVM, GRU [Chi and Chu \(2021\)](#), Liu et al. ([Faraji et al., 2022](#)); [Alegria et al. \(2013\)](#); [Rostampour et al. \(2019\)](#) models, and our model in one dataset with different complexity levels, because the structure of GRU is simpler compared to LSTM, so its operation speed is faster and its training time is also faster. Our model incorporates an attention mechanism to allocate the number of operations better, and therefore, its Sunlen time is also better than the standard GRU model.

In [Figure 8](#), since the data for load and power price forecasting of power systems are always nonlinear and complex, the prediction accuracy for complex data is one of the important indicators of credit risk forecasting models. As shown in the figure, we compare the prediction accuracy of different models and our model on data sets with different levels of complexity.



In [Figure 9](#), we compare the magnitude of the operations of different models. We can see that for the LSTM model ([Li et al., 2023](#)), the operation speed is lower due to the good improvements made by the GRU model. Still, when comparing the traditional LR logistic regression model ([Zhu et al., 2016](#)), the operations are lower than ours because the data input to the LR model has generally been preprocessed. Still, the LR logistic regression model cannot perform nonlinear data prediction, so the performance of our model is better than the other models from an overall perspective.

For example, process 1 is the operation process of our model. First, use the CNN network to extract features from historical data, input the output results into the GRU network, and then introduce the attention mechanism to improve the model's performance.

In [Figure 10](#), to further compare the performance of our model, we also compared the parameters of CNN-GRU ([Faraji et al., 2022](#)), SVM ([Yao et al., 2022](#)), KVM ([Li et al., 2016](#)), and our model. The results show that due to the feature extraction of the CNN module and the unique dual structure of the GRU module, and the attention mechanism introduced, our model, compared with the other three models, has a significant reduction in the number of parameters and improves the performance of the model.

In [Figure 11](#), to verify the generalizability of our model, we separately use U.S. Energy Information Administration (EIA): the EIA provides historical load and electricity price data for the U.S. electricity market; these data are available from the EIA's website; European Power Exchange (EPEX): the EPEX provides historical load and electricity price data for the European electricity market; International Energy Agency (IEA): IEA provides historical load and electricity price data for the global electricity market. China Electricity Exchange (CNEE): CNEE provides historical load and electricity price data of China's electricity market. Recall and AUC comparisons are made on four datasets, and the results show that our model is more accurate in predicting the load and electricity price of the power system than the other models in the four different datasets, showing good applicability to both linear and nonlinear data.

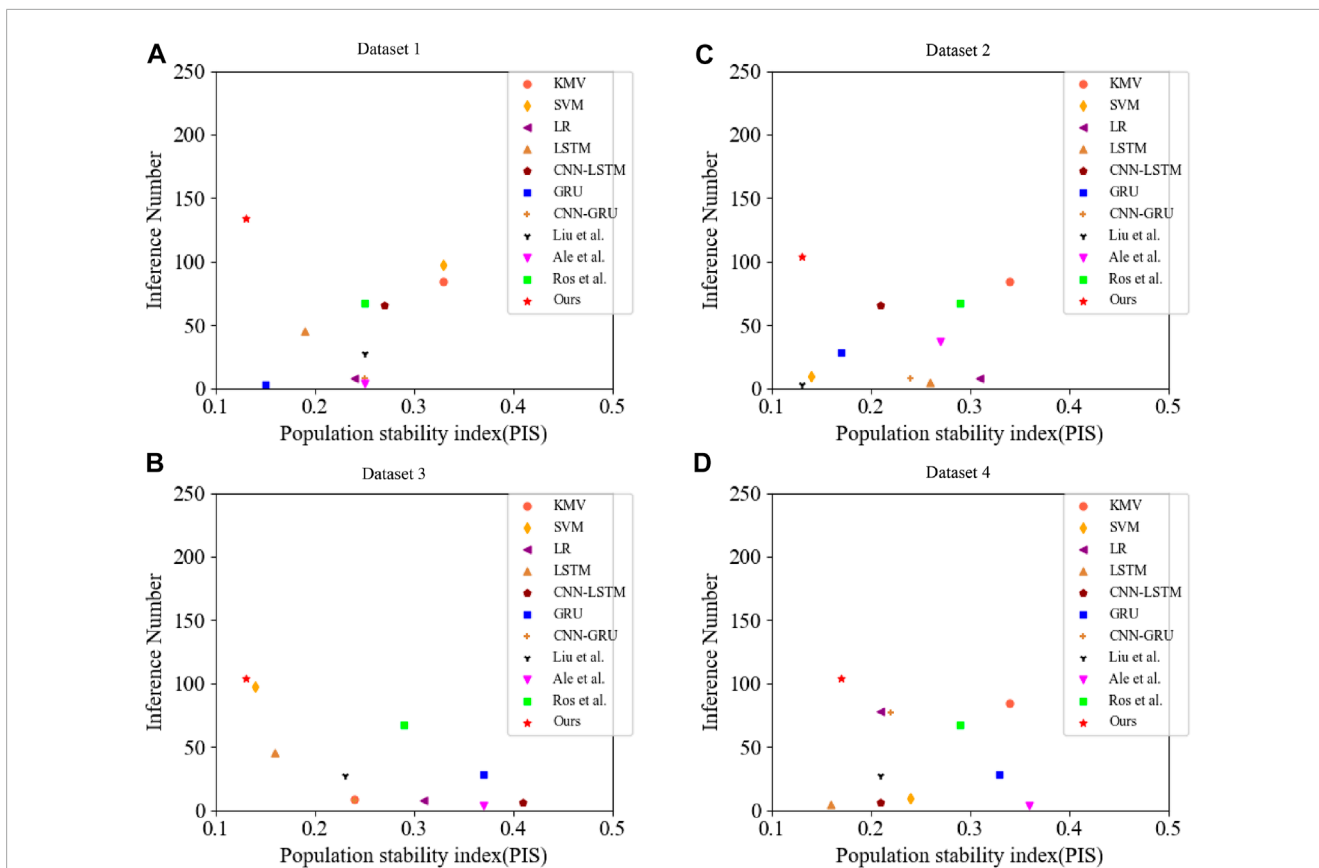


FIGURE 13 Inference time comparison of different models. (A) corresponds to Dataset 1 from the EIA dataset, (B) corresponds to Dataset 2 from the EPEX dataset, (C) corresponds to Dataset 3 from the IEA dataset, and (D) corresponds to Dataset 4 from the CNEE dataset.

TABLE 2 Summary comparison of different models.

Method	Accuracy ↑	Recall ↑	Flops(G) ↓	Parameters(M) ↓	AUC ↑	PIS ↓
KMV Li et al. (2016)	0.887	0.776	123.3	134.5	0.876	0.34
SVM Yao et al. (2022)	0.873	0.867	167.3	187.2	0.853	0.14
LR Zhu et al. (2016)	0.932	0.912	114.7	113.6	0.912	0.31
LSTM Li et al. (2023)	0.921	0.892	134.5	123.3	0.891	0.26
CNN-LSTM Wu et al. (2021)	0.943	0.921	145.6	137.9	0.911	0.21
GRU Chi and Chu, (2021)	0.952	0.935	117.5	107.3	0.923	0.17
CNN-GRU Faraji et al. (2022)	0.971	0.944	123.7	119.6	0.943	0.24
Liu et al. Liu et al. (2016)	0.931	0.911	113.4	109.1	0.908	0.13
Ale et al. Alegria et al. (2013)	0.871	0.855	163.7	149.6	0.865	0.27
Ros et al. Rostampour et al. (2019)	0.954	0.934	176.7	139.2	0.917	0.29
Ours proposed model	0.983	0.951	112.4	106.3	0.953	0.1

In Figure 12, to further verify the prediction accuracy of our model, AUC was made between our model and the LSTM model, and the data in the Diane database was selected as the data. The results show that the ACU value of our model is better than that of the LSTM model. We can further test the predictive ability of our model.

In Figure 13, to compare the operational stability of the models, we compare the population stability index (PSI), an index used to assess the stability of two samples for different models with a different numbers of inferences. PSI is often used to assess the stability of the performance of a risk model over time or across different data sets. It is usually considered that if the

```

Data: EIA, EPEX, IEA, CNEE input: The train  $D$ ,
the dataset for transfer learning  $D_T$ , the
parameters  $\theta$ , Value =  $b$ , Query =  $a$ 
output: The trained CNN-GRU  $Model_{Ours}$ 
Random initialization  $W, b$ ; while The error
rate of the neurat network model on the
vatisation set  $v$  no longer decreases do
for epoch in iterationsdo
Query =  $f(\theta_{now})$ ; if  $Value_{now} < Value_i$  then
 $\theta_i = \theta_t$ ;
Query $t$  = Query $t-1$ ; end
end
 $\theta_{now} = Adam(\theta_{now}, D_T)$ ; end
Load the best model with  $b$ ;
Return the trained model  $Model_{Ours}$ 

```

Algorithm 1 : An algorithm with caption.

PSI is less than 0.1, it indicates that the difference between the two samples is acceptable; if the PSI is greater than 0.25, it indicates that the difference between the two samples is large and the performance stability of the model may need to be reassessed. Our model maintains good stability in a larger number of inferences.

In Table 2, we summarize the prediction accuracy, parameter amount, and calculation amount of the eight models mentioned above so that we can compare the performance of our models more intuitively. We can show that our models, in these aspects, have good performance.

5 Discussion and limitations

In this paper, we propose a CNN-GRU model based on an attention mechanism to investigate the optimization scheme of large-scale energy storage in a smart grid to effectively predict the load and power price of the power system and to develop the optimal energy storage strategy. At the same time, compared with other prediction models, the problem of missing historical data is effectively solved, and the model's operational efficiency and prediction accuracy are improved. The GRU network is an excellent variant of LSTM, which can effectively perform long-time series prediction. The structure of GRU is simpler than the LSTM network, with an easier training process and faster operation rate. It thoroughly explores the smart grid's large-scale energy storage optimization scheme.

However, our model still has some drawbacks; although the GRU network is used instead of the LSTM network, the complexity of the deep learning network still exists, and the computation speed is more complicated; In contrast, our model combines three modules, and the number of parameters has increased. In addition, the problem of missing historical data has been improved but has yet to solve completely. Subsequently, migration learning can be applied to optimize the model further.

6 Conclusion

With the increasing energy demand and environmental protection awareness, the smart grid large-scale energy storage optimization scheme is getting more and more attention. This scheme mainly realizes the regulation and optimization of the power system through energy storage technology and intelligent control means to improve the reliability, safety, and economy of the power system. Smart grid large-scale energy storage optimization scheme can realize the balance of power system load through energy storage technology, reduce the peak-valley difference of power system, relieve the pressure of power system, and improve the stability and reliability of power system; meanwhile, it can also realize the regulation and optimization of power system through energy storage technology, reduce the operation cost of power system, and improve the economy and benefit of power system; then it can also realize the regulation and optimization of power system through energy storage technology, reduce the operation cost of power system, and improve the economy and benefit of power system; then it can also realize the regulation and optimization of power system through Energy storage technology can be used to store and utilize renewable energy, promote the development and application of renewable energy, reduce energy consumption and carbon emissions, and help protect the environment and sustainable development; finally, the smart grid large-scale energy storage optimization program can realize the regulation and optimization of the power system through intelligent control means, promote energy transformation and upgrading, and promote the clean, low-carbon, efficient and intelligent development of energy. Therefore, the smart grid large-scale energy storage optimization program has an important significance.

In summary, our work can provide a better large-scale energy storage optimization scheme for smart grids, thus improving the reliability, safety, and economy of power systems, promoting the development of renewable energy, promoting energy transformation and upgrading, and making an individual contribution to the sustainable development of human beings.

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

XL contributed to the conception and design of the study. XL organized the database. XL performed the statistical analysis. XL wrote the first draft of the manuscript. XL wrote sections of the manuscript.

Conflict of interest

The author declares 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

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abedi, S., and Kwon, S. (2023). Rolling-horizon optimization integrated with recurrent neural network-driven forecasting for residential battery energy storage operations. *Int. J. Electr. Power & Energy Syst.* 145, 108589. doi:10.1016/j.ijepes.2022.108589
- Ahmad, T., Madonski, R., Zhang, D., Huang, C., and Mujeeb, A. (2022). Data-driven probabilistic machine learning in sustainable smart energy/smart energy systems: Key developments, challenges, and future research opportunities in the context of smart grid paradigm. *Renew. Sustain. Energy Rev.* 160, 112128. doi:10.1016/j.rser.2022.112128
- Alegria, E., Brown, T., Minear, E., and Lassetter, R. H. (2013). Certs microgrid demonstration with large-scale energy storage and renewable generation. *IEEE Trans. Smart Grid* 5, 937–943. doi:10.1109/tsg.2013.2286575
- Alvarez-Diazcomas, A., López, H., Carrillo-Serrano, R. V., Rodríguez-Reséndiz, J., Vázquez, N., and Herrera-Ruiz, G. (2019). A novel integrated topology to interface electric vehicles and renewable energies with the grid. *Energies* 12, 4091. doi:10.3390/en12214091
- Balakumar, P., Vinopra, T., and Chandrasekaran, K. (2023). Deep learning based real time demand side management controller for smart building integrated with renewable energy and energy storage system. *J. Energy Storage* 58, 106412. doi:10.1016/j.est.2022.106412
- Cao, J., Harrold, D., Fan, Z., Morstyn, T., Healey, D., and Li, K. (2020). Deep reinforcement learning-based energy storage arbitrage with accurate lithium-ion battery degradation model. *IEEE Trans. Smart Grid* 11, 4513–4521. doi:10.1109/tsg.2020.2986333
- Chi, D.-J., and Chu, C.-C. (2021). Artificial intelligence in corporate sustainability: Using lstm and gru for going concern prediction. *Sustainability* 13, 11631. doi:10.3390/su132111631
- Choo, B. L., and Go, Y. I. (2022). Energy storage for large scale/utility renewable energy system—an enhanced safety model and risk assessment. *Renew. Energy Focus* 42, 79–96. doi:10.1016/j.ref.2022.05.001
- Estévez-Bén, A. A., Alvarez-Diazcomas, A., Macias-Bobadilla, G., and Rodríguez-Reséndiz, J. (2020a). Leakage current reduction in single-phase grid-connected inverters—A review. *Appl. Sci.* 10, 2384. doi:10.3390/app10072384
- Estévez-Bén, A. A., Alvarez-Diazcomas, A., and Rodríguez-Reséndiz, J. (2020b). Transformerless multilevel voltage-source inverter topology comparative study for pv systems. *Energies* 13, 3261. doi:10.3390/en13123261
- Estévez-Bén, A. A., López Tapia, H. J. C., Carrillo-Serrano, R. V., Rodríguez-Reséndiz, J., and Vázquez Nava, N. (2019). A new predictive control strategy for multilevel current-source inverter grid-connected. *Electronics* 8, 902. doi:10.3390/electronics8080902
- Faraji, M., Nadi, S., Ghaffarpassand, O., Homayoni, S., and Downey, K. (2022). An integrated 3d cnn-gru deep learning method for short-term prediction of pm_{2.5} concentration in urban environment. *Sci. Total Environ.* 834, 155324. doi:10.1016/j.scitotenv.2022.155324
- Feng, Z.-k., Huang, Q.-q., Niu, W.-j., Yang, T., Wang, J.-y., and Wen, S.-p. (2022). Multi-step-ahead solar output time series prediction with gate recurrent unit neural network using data decomposition and cooperation search algorithm. *Energy* 261, 125217. doi:10.1016/j.energy.2022.125217
- Gao, Y., Fang, C., and Zhang, J. (2022). A spatial analysis of smart meter adoptions: Empirical evidence from the us data. *Sustainability* 14, 1126. doi:10.3390/su14031126
- Górski, T. (2023). Integration flows modeling in the context of architectural views. *IEEE Access* 11, 35220–35231. doi:10.1109/access.2023.3265210
- Guo, M.-H., Xu, T.-X., Liu, J.-J., Liu, Z.-N., Jiang, P.-T., Mu, T.-J., et al. (2022). Attention mechanisms in computer vision: A survey. *Comput. Vis. Media* 8, 331–368. doi:10.1007/s41095-022-0271-y
- Han, M., Johari, F., Huang, P., and Zhang, X. (2022). *Generating hourly electricity demand data for large-scale single-family buildings by a decomposition-recombination method*. Energy and Built Environment.
- Hattori, T., Nam, H., and Chapman, A. (2022). Multilateral energy technology cooperation: Improving collaboration effectiveness through evidence from international energy agency technology collaboration programmes. *Energy Strategy Rev.* 43, 100920. doi:10.1016/j.esr.2022.100920
- Jena, C. J., and Ray, P. K. (2023). Power management in three-phase grid-integrated pv system with hybrid energy storage system. *Energies* 16, 2030. doi:10.3390/en16042030
- Juárez Coronado, J. E. (2022). *Diseño de investigación del estudio de factibilidad técnica para el desarrollo de una infraestructura de carga vehicular de acceso público en Guatemala*. Universidad de San Carlos de Guatemala. Ph.D. thesis.
- Kattenborn, T., Leitloff, J., Schiefer, F., and Hinz, S. (2021). Review on convolutional neural networks (cnn) in vegetation remote sensing. *ISPRS J. photogrammetry remote Sens.* 173, 24–49. doi:10.1016/j.isprsjprs.2020.12.010
- Lai, C. S., Chen, D., Zhang, J., Zhang, X., Xu, X., Taylor, G. A., et al. (2022). Profit maximization for large-scale energy storage systems to enable fast ev charging infrastructure in distribution networks. *Energy* 259, 124852. doi:10.1016/j.energy.2022.124852
- Li, F., Yu, X., Tian, X., and Zhao, Z. (2021a). “Short-term load forecasting for an industrial park using lstm-rnn considering energy storage,” in *2021 3rd asia energy and electrical engineering symposium (AEEES)* (IEEE), 684–689.
- Li, L., Liu, P., Li, Z., and Wang, X. (2018). A multi-objective optimization approach for selection of energy storage systems. *Comput. Chem. Eng.* 115, 213–225. doi:10.1016/j.compchemeng.2018.04.014
- Li, L., Yang, J., and Zou, X. (2016). A study of credit risk of Chinese listed companies: Zpp versus kmv. *Appl. Econ.* 48, 2697–2710. doi:10.1080/00036846.2015.1128077
- Li, M., Zhang, Z., Lu, M., Jia, X., Liu, R., Zhou, X., et al. (2023). Internet financial credit risk assessment with sliding window and attention mechanism lstm model. *Teh. Vjesn.* 30, 1–7. doi:10.17559/TV-20221110173532
- Li, W., Wu, H., Zhu, N., Jiang, Y., Tan, J., and Guo, Y. (2021b). Prediction of dissolved oxygen in a fishery pond based on gated recurrent unit (gru). *Inf. Process. Agric.* 8, 185–193. doi:10.1016/j.inpa.2020.02.002
- Li, Y., Wei, X., Li, Y., Dong, Z., and Shahidehpour, M. (2022). Detection of false data injection attacks in smart grid: A secure federated deep learning approach. *IEEE Trans. Smart Grid* 13, 4862–4872. doi:10.1109/tsg.2022.3204796
- Liu, M., Li, W., Wang, C., Polis, M. P., Li, J., et al. (2016). Reliability evaluation of large scale battery energy storage systems. *IEEE Trans. Smart Grid* 8, 2733–2743. doi:10.1109/tsg.2016.2536688
- Liu, Y., Zhang, D., and Gooi, H. B. (2020). Optimization strategy based on deep reinforcement learning for home energy management. *CSEE J. Power Energy Syst.* 6, 572–582. doi:10.17775/CSEEJPES.2019.02890
- López, H., Rodríguez-Reséndiz, J., Guo, X., Vázquez, N., and Carrillo-Serrano, R. V. (2018). Transformerless common-mode current-source inverter grid-connected for pv applications. *IEEE Access* 6, 62944–62953. doi:10.1109/access.2018.2873504
- Lu, H., Yang, X., and Guo, H. (2022). Smart energy storage dispatching of peak-valley load characteristics based-convolutional neural network. *Comput. Electr. Eng.* 97, 107543. doi:10.1016/j.compeleceng.2021.107543
- Mahilong, N., Sarangan, V., Bichpuriya, Y., Prajapati, A., and Rajagopal, N. (2022). “Trading strategy for renewable energy sources in day-ahead and continuous intraday market,” in *2022 IEEE PES innovative smart grid technologies-asia (ISGT asia)* (IEEE), 444–448.
- Mostafa, N., Ramadan, H. S. M., and Elfarrow, O. (2022). Renewable energy management in smart grids by using big data analytics and machine learning. *Mach. Learn. Appl.* 9, 100363. doi:10.1016/j.mlwa.2022.100363
- Rezaeimozafar, M., Monaghan, R. F., Barrett, E., and Duffy, M. (2022). A review of behind-the-meter energy storage systems in smart grids. *Renew. Sustain. Energy Rev.* 164, 112573. doi:10.1016/j.rser.2022.112573
- Rostampour, V., Jaxa-Rozen, M., Bloemendal, M., Kwakkel, J., and Keviczky, T. (2019). Aquifer thermal energy storage (ates) smart grids: Large-scale seasonal energy storage as a distributed energy management solution. *Appl. Energy* 242, 624–639. doi:10.1016/j.apenergy.2019.03.110
- Terlouw, T., AlSkaif, T., Bauer, C., and Van Sark, W. (2019). Multi-objective optimization of energy arbitrage in community energy storage systems using different battery technologies. *Appl. Energy* 239, 356–372. doi:10.1016/j.apenergy.2019.01.227
- Wang, Z., and Hong, T. (2020). Reinforcement learning for building controls: The opportunities and challenges. *Appl. Energy* 269, 115036. doi:10.1016/j.apenergy.2020.115036
- Wu, J. M.-T., Li, Z., Herencsar, N., Vo, B., and Lin, J. C.-W. (2021). A graph-based cnn-lstm stock price prediction algorithm with leading indicators. *Multimed. Syst.* 29, 1751–1770. doi:10.1007/s00530-021-00758-w

Xiao, H., Pu, X., Pei, W., Ma, L., Ma, T., Wen, L., et al. (2023). One-off low-dose ct screening of positive nodules in lung cancer: A prospective community-based cohort study. *IEEE Trans. Smart Grid* 177, 1–10. doi:10.1016/j.lungcan.2023.01.005

Yang, Z., Ghadamyari, M., Khorramdel, H., Alizadeh, S. M. S., Pirouzi, S., Milani, M., et al. (2021). Robust multi-objective optimal design of islanded hybrid system with renewable and diesel sources/stationary and mobile energy storage systems. *Renew. Sustain. Energy Rev.* 148, 111295. doi:10.1016/j.rser.2021.111295

Yao, G., Hu, X., and Wang, G. (2022). A novel ensemble feature selection method by integrating multiple ranking information combined with an svm ensemble model for enterprise credit risk prediction in the supply chain. *Expert Syst. Appl.* 200, 117002. doi:10.1016/j.eswa.2022.117002

Zhang, P., Zong, X., Cao, Y., and Zhao, Y. (2022). “Multi-energy storage evolution model of regional integrated energy system based on load forecasting,” in *2022 IEEE/IAS industrial and commercial power system asia (I&CPS asia)* (IEEE), 725–731.

Zhao, N., Zhang, H., Yang, X., Yan, J., and You, F. (2023). Emerging information and communication technologies for smart energy systems and renewable transition. *Adv. Appl. Energy* 9, 100125. doi:10.1016/j.adapen.2023.100125

Zhao, S., Zhang, C., and Wang, Y. (2022). Lithium-ion battery capacity and remaining useful life prediction using board learning system and long short-term memory neural network. *J. Energy Storage* 52, 104901. doi:10.1016/j.est.2022.104901

Zheng, X., and Jin, T. (2022). A reliable method of wind power fluctuation smoothing strategy based on multidimensional non-linear exponential smoothing short-term forecasting. *IET Renew. Power Gener.* 16, 3573–3586. doi:10.1049/rpg2.12395

Zhu, Y., Xie, C., Sun, B., Wang, G.-J., and Yan, X.-G. (2016). Predicting China's sme credit risk in supply chain financing by logistic regression, artificial neural network and hybrid models. *Sustainability* 8, 433. doi:10.3390/su8050433