



Multi-task learning based multi-energy load prediction in integrated energy system

Lulu Wang¹ · Mao Tan^{1,2} · Jie Chen³ · Chengchen Liao³

Accepted: 29 July 2022 / Published online: 16 August 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Accurate multi-energy load prediction plays a very crucial role in integrated energy system management. To address the load characteristics of strong relational coupling, volatility and uncertainty in user-side integrated energy systems, this paper proposes a multi-task learning based short-term multi-energy load prediction method. Load participation factor is proposed to portray the proportion of different loads in the total load demand, and multi-task learning method is introduced to deeply explore the coupling relationships among them. Then, the proposed method is validated on a real-world dataset. The results show that the method has higher prediction accuracy than the existing methods, and the prediction accuracy is improved by at least 1.3%.

Keywords Load prediction · Integrated energy system · Load participation factor · Multi-task learning

1 Introduction

Energy is the basic driving force of economic development in society. Various traditional energy systems, such as electricity, heating and natural gas, are designed and operated separately. It artificially severs the connectivity between different types of energy source and ignores the possibility for energy recycling, and thus to result in a serious waste of resources. However, integrated energy system (IES) with various energy supply and consumption systems is an

important trend in the energy transition. IES can effectively improve the efficiency of energy utilization, promote the coordination and optimization between systems, and realize the mutual complementation of multiple energy source. As an important topic in the demand side management, load forecasting of the user-side integrated energy system (UIES) has become a primary prerequisite for system planning and operational scheduling. Therefore, effective learning of coupled information from multiple types of energy sources is required to achieve accurate prediction of diverse energy demands [1–3].

Deep learning has become one of the most popular techniques at present [4, 5]. It has powerful ability to explain nonlinear complex structures. In recent years, the energy sector has shown a highly integrated development of information and physics. Meanwhile, data-driven load forecasting has become an important and widespread application of artificial intelligence in the energy field [6].

Some scholars have conducted research on single energy load forecasting for UIES. Due to different physical characteristics of the load, energy sources can usually be classified as electricity, cooling, heating and renewable energy. For common electricity [7, 8], cooling [9] and heating [10, 11] load forecasting, many studies focused on combining deep learning models with traditional methods, while considering the incorporation of integration mechanisms or optimization methods. In addition, for renewable energy, it tends to be more influenced by

✉ Mao Tan
mr.tanmao@gmail.com

Lulu Wang
201921002014@smail.xtu.edu.cn

Jie Chen
dchen123@126.com

Chengchen Liao
PeeledMango@foxmail.com

¹ Hunan National Center for Applied Mathematics, Xiangtan University, Yuhu District, Xiangtan, 411100, China

² Hunan Engineering Research Center of Multi-energy Cooperative Control Technology, Xiangtan University, Yuhu District, Xiangtan, 411100, China

³ School of Automation and Electronic Information, Xiangtan University, Yuhu District, Xiangtan, 411100, China

environmental factors. For example, Ref. [12] proposed a learning method based on nonparametric gradient boosting regression tree to forecast multi-site solar power from one to six hours in the future. However, the coupling relationship between different energy systems has not been considered in the above studies.

The application scenarios of multi-energy load forecasting (MELF) are more complex than single energy. Therefore, accurate forecasting is required by analyzing different energy usage information. Ref. [13] proposed an attention-based model for the temporal and nonlinear property of electrical loads. A hierarchical joint method based on classification and DBN was proposed for MELF [14]. In addition, several studies have shown that cooperative MELF can reduce the number of models and increase efficiency of model. A nonlinear autoregressive model was proposed for simultaneous prediction of electricity, cooling, and heating load in Ref. [15]. Moreover, the researchers tried to combine multiple models to design an integrated approach to improve the accuracy of the predictions. Combining ensemble learning and reinforcement learning methods, Ref. [16] proposed a dynamic integrated multivariate load forecasting method based on Q-learning reinforcement learning. It was considered that the high-dimensional temporal dynamic and cross-coupling characteristics of load sequences in Ref. [17].

Many traditional machine learning methods focus only on single-task learning [18]. However, there is a large amount of correlation information between different subtasks in the MELF problem, and single-task learning methods cannot fully utilize this information. As a result, it often leads to undertraining and underfitting during model training. To solve the above problem, Ref. [19] proposed a multi-task learning (MTL) mechanism, where each subtask can learn implicit information in parallel to support each other. In summary, MTL can facilitate multiple related tasks to learn together, so that each subtask can enhance the learning of other tasks to achieve the purpose of sharing information and finally obtain the optimal results of all tasks simultaneously.

Using the MTL mechanism, some studies have constructed a wide range of predictive models. Ref. [20] presented a model based on deep MTL and ensemble learning. Ref. [21] used a MTL approach to build a joint load forecasting model. Besides, in an industrial park of IES, a method for electricity, heating and gas loads based on DBN networks and multi-task regression is proposed [22]. From the point of horizontal and vertical interaction, Ref. [23] proposed a MTL prediction model based on parallel LSTM networks. Ref. [24] proposed an integrated load forecasting model based on bidirectional generative adversarial network (GAN) and transfer learning. It demonstrated the effectiveness of model migration. From these works we can find

that the MTL method can make full use of the fused information between different tasks, and improve the prediction accuracy of MELF. Nevertheless, the existing studies many focused more on model design while ignoring the lack of data samples and the issue of energy diversity.

In such a context, this paper proposes a short-term MTL-based MELF approach, and the main contributions can be summarized as follows.

- (1) This paper innovatively introduces participation factor in multi-energy load forecasting to represent the coupling between each kind of load and the total demand. It can reflect the relative change of each load in the whole system more effectively.
- (2) The correlation between weather factors and each load is analyzed and the weather matrix is further constructed. Based on feature engineering, the suitable features GHI, DP, PW, RH and TMP are selected for feature construction.
- (3) A method combining MTL with temporal convolutional network is proposed. The MTL approach fully mines the coupling relationship between different prediction subtasks, while the temporal convolutional network has good performance in the total load demand prediction.

The remainder of the paper is organized as follows: Section 2 introduces the UIES and the prediction problem of multiple time series. Section 3 analyzes the relevant load characteristics. Section 4 details the prediction model. Section 5 focuses on the analysis of experimental results. Section 6 summarizes our work.

2 Description of the problem

2.1 User-level integrated energy system

IES integrates multiple energy resources in a certain area. It achieves the coordination and optimization between different subsystems and the complementarity of multi-energy sources. It puts more emphasis on the development model of moving from individual energy sources to the joint dispatch of multiple types of resources. Then, it will help to meet diversified energy demands and improve utilization efficiency [25]. Based on geographical factors, energy characteristics and conversion difficulties, IES can be divided into three levels: inter-regional level, regional level and user-side level. Among them, the user-side systems tend to meet different energy demands of customers at the distribution and consumption levels, and the demand is generally influenced by customer behavior. Moreover, this behavior is generally regulated. Each energy source has different physical dynamic characteristics, which further

makes the multi-source load more unstable. Figure 1 shows the energy interaction of UIES. The system is centered on the power system, and it enables the cooperative management and interaction between different energy systems through energy supply devices, energy conversion devices and energy storage devices. Thus, it improves the economy and flexibility of system operation [26].

The dataset can be retrieved from the IES of Arizona State University's Tempe campus. The platform records the data of four campuses, and we choose the Tempe campus because of its large scale buildings and diverse energy types. After selecting the Tempe campus, the data can be downloaded easily according to different energy resource types and time scale needs by pressing the 'CSV Export' button. The region has hot summer and warm winter with abundant light and low rainfall, resulting in high demand for electricity and cooling energy from local customers. The experimental data are retrieved from the 2018 Campus Metabolism Project web platform with a temporal resolution of 1 hour [27]. Due to those data have different units of measure, KW, Ton/hrs, and mmBtu/hr, respectively, the units of the different loads were normalized before the experiments. According to the unit conversion formula $1 \text{ KW} = 0.284 \text{ Ton/hrs} = 0.0034 \text{ mmBtu/hr}$, the units of heating and cooling loads are standardized to KW. The graph of load demand curves is shown in Fig. 2. It can be seen that the curve shows a trend of high in the middle period and low on both sides, with some local randomness and volatility. The peak usage is mainly concentrated in summer. The heating load is stable and less volatile compared with the electricity load. Obviously, the climate change has a direct impact on the energy demand of customers and is distinctly seasonal in nature. Therefore, the climatic factors should be considered, which helps to achieve accurate multiple energy demand forecasting.

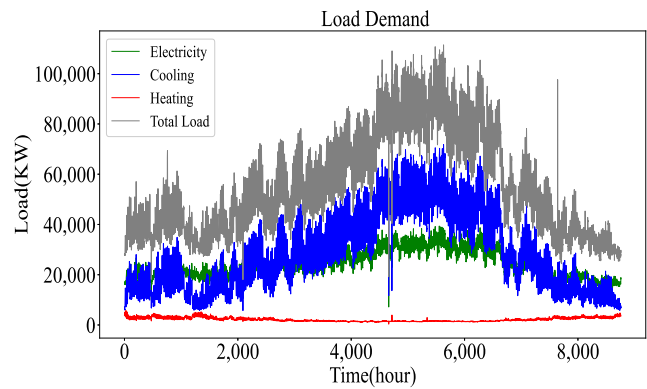


Fig. 2 The load demand curve of Tempe campus in 2018

2.2 Multiple time series problem

Time series forecasting plays a critical role in many fields. It is a technique for predicting the future states of an object by analyzing its historical data patterns. In UIES, a multi-energy load sequence is composed of load sequences from several different energy sources. Therefore, the multi-energy load forecasting problem is essentially a time series forecasting problem [28–30].

In many cases, the change pattern of the predicted object will be influenced by other factors. Hence, in addition to considering the change pattern of the object itself, the relationship between the object and other related factors should also be explored. From the perspective of time series analysis, the basic idea of multi-energy load forecasting problem is shown in Fig. 3, where t is the amount of time, which can be hours or minutes, etc. It is generally determined by the specific problem. Usually, on the basis of obtaining historical sample data y_t , it is necessary to forecast the future time period $T + P \leq t \leq T + Q$ based on

Fig. 1 The schematic diagram of the energy interaction of UIES

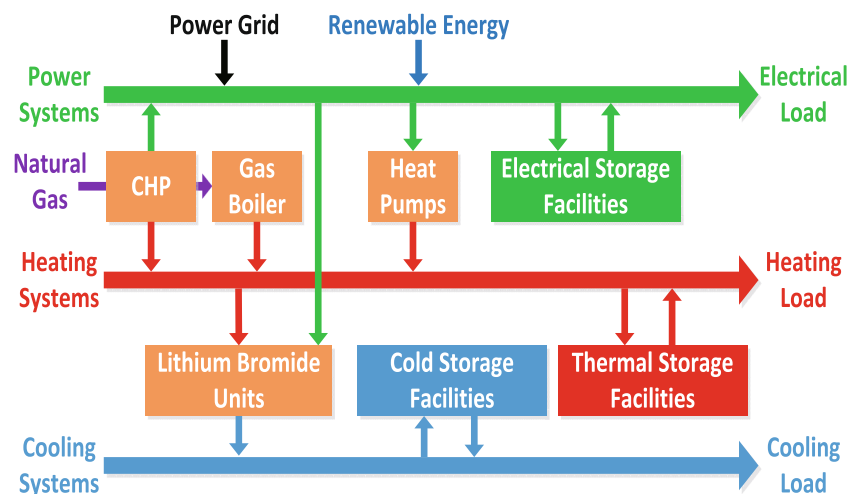
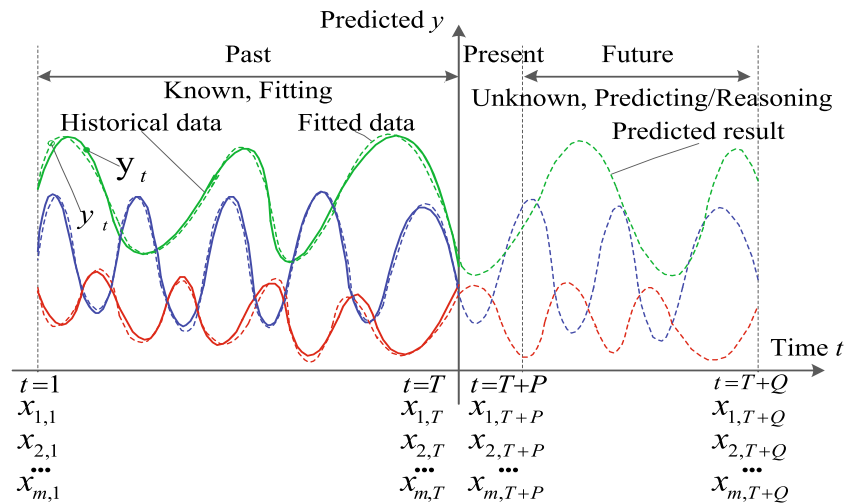


Fig. 3 The diagram of multiple time series forecasting problem



the development pattern of the object y , where $t \in [1, T]$, and T is the last period of known data. P and Q are a certain time point in the future, respectively.

In general, let the abstract expression of the prediction model be

$$y = f(S, X_t, t) \quad (1)$$

where X is a vector of m relevant factors; t is the time series number; y is the predicted object; and S is the parameter vector of this prediction model. The parameter $S = [s_1, s_2, \dots, s_n]$ characterizes the set of parameters to be determined for the prediction model, and n is the number of parameters. In Fig. 3, the factors of relevance for each time period are marked under the horizontal axis. Assuming that a total of m factors related to the predicted object y are collected, the vector $X = [x_1, x_2, \dots, x_m]^T$ is constructed. For $t \in [1, T]$, which is called the historical time period, the relevant factors take the values $X_t = [x_{1,t}, x_{2,t}, \dots, x_{m,t}]^T$. After the model is fitted to the known historical data y_t and x_t , y_t is obtained, during which the model digs deeper into the change pattern of the object and related factors. For $t \in [T + P, T + Q]$, which is called the prediction period, the predicted quantity y_t is obtained by using the established prediction model.

3 Characteristic analysis of multi-energy load

Prior to conducting research work, the load characteristics of the energy system should be analyzed in advance. Starting from the mechanisms of load components, it is necessary to reveal the regulation of intrinsic changes in the loads themselves. On this basis, this section will analyze the load participation factors, the coupling relations between loads, the correlation of load series itself, and the influence of weather factors. Among them, the data used for analysis

in this section are the same as Section 2, which come from the Campus Metabolism Project.

3.1 Load participation factor

The traditional MELF only considers the demand of various loads, and ignores the proportion of each load in the total load demand. By describing the proportion of different loads in the total demand, the variation of each energy demand can be depicted visually and clearly. Based on the understanding of the real load demand pattern, the coupling relationship between various energy sources can be further grasped. Therefore, the load participation factor (LPF) is introduced to describe the share of various loads in the total load.

The load participation factor are defined as follows,

$$LPF_e = \frac{L_e(t)}{T(t)} \quad (2)$$

$$LPF_c = \frac{L_c(t)}{T(t)} \quad (3)$$

$$LPF_h = \frac{L_h(t)}{T(t)} \quad (4)$$

where e , c , h denote electricity, cooling and heating loads, respectively; LPF_e , LPF_c , LPF_h denote the participation factors of electricity, cooling and heating loads, respectively; $L(t)$ denotes the load value at time t ; $T(t)$ denotes the total load value at time t , i.e., the sum of all load demands.

In order to analyze the specific usage of the various loads and their proportion in the total load, Fig. 4 shows the curves of participation factor for the Tempe campus in 2018. It can be observed that the heating load has the lowest share, but its sequence shows stronger volatility. In contrast, the electricity load starts with the highest share. As the season changes, the share of the cooling load gradually increases,

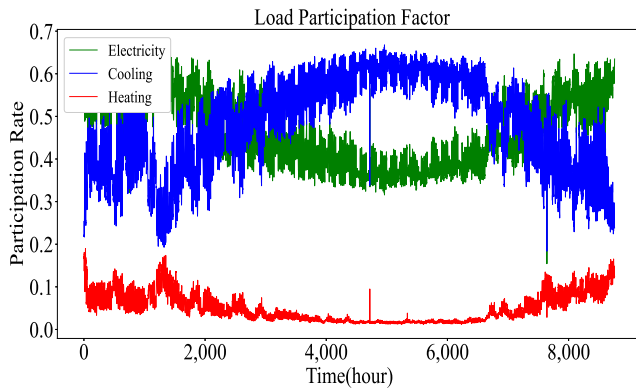


Fig. 4 The curves of load participation factors

which leads to a decrease in the share of the electricity and heating loads. This confirms the above analysis of the load demand situation, and it also shows that the load participation factor can describe the change pattern of load demand more intuitively.

3.2 The coupling relationship between loads

In addition, these loads are sensitive to climate. To further analyze the influence of the loads and their conversion relationship with each other, a scatter plot analysis of LPF is performed, as shown in Fig. 5. It shows that the electricity load and the cooling load have a negative correlation. Specifically, when the demand of electricity load increases, the demand of cold load decreases. Conversely, electricity and heating loads have a positive relationship. This indicates that there is a close correlation between load demands, where a change of one load will directly affect the other loads. This also verifies the existence of coupling relationship between loads.

$$x_t = \begin{bmatrix} x_{1,t}^e & x_{1,t}^c & x_{1,t}^h \\ x_{2,t}^e & x_{2,t}^c & x_{2,t}^h \\ \dots & \dots & \dots \\ x_{n,t}^e & x_{n,t}^c & x_{n,t}^h \end{bmatrix} \tag{5}$$

Based on the above analysis, a matrix x_t is constructed and listed in Eq. (5). It contains electricity $x_{n,t}^e$, cooling $x_{n,t}^c$ and heating $x_{n,t}^h$ LPF feature values at time t , where n represents the number of samples. In conclusion, from the perspective of LPF, it can portray the usage of different energy sources more comprehensively and intuitively, and it can also understand the conversion relationship between different energy sources more clearly.

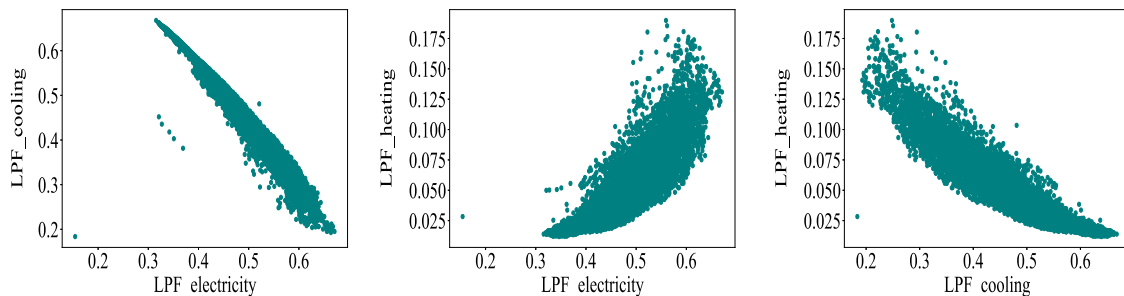
3.3 The correlation of load series

To investigate the temporal correlation of the load series, Figs. 6 and 7 show the autocorrelation function (ACF) and partial autocorrelation function (PACF) analysis of LPF for February and July 2018, respectively. The ACF and PACF measure the dependence of present samples on the past samples of the same series, which can be calculated by Eqs. (6) and (7), respectively.

$$\rho_k = \frac{\sum_{t=k+1}^N (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^N (y_t - \bar{y})^2}, 0 < k < N \tag{6}$$

$$\varphi_{kk} = \begin{cases} \rho_1, & k = 1 \\ \frac{\rho_k - \sum_{j=1}^{k-1} \varphi_{k-1,j} \rho_{k-j}}{1 - \sum_{j=1}^{k-1} \varphi_{k-1,j} \rho_j}, & k > 1 \end{cases} \tag{7}$$

In the above equations, N is the series length, y_t is the value at moment t of the series, \bar{y} is mean of the series, and $\varphi_{k,j} = \varphi_{k-1,j} - \varphi_{k-1,k-j} \varphi_{k-j}$, $j = 1, 2, \dots, k - 1$. In this case, the blue cone-shaped area is the confidence interval of 0.95. From the figures, the ACF analyses for the various load series has a significant periodicity. This indicates that each load series is strongly correlated with adjacent times and adjacent days. Meanwhile, the autocorrelation of the series gradually decreases when the time interval increases. It reflects the change pattern of the load series. By comparing the correlation analysis plots of February and July, it can be found that there are obvious differences between different seasons. These differences also demonstrate that the changes in weather environment can have an impact on customers' consumption behavior. However the behavior of individual customer is often



(a) Electricity and Cooling (b) Electricity and Heating (c) Cooling and Heating

Fig. 5 The scatter plot of participation factors for different loads

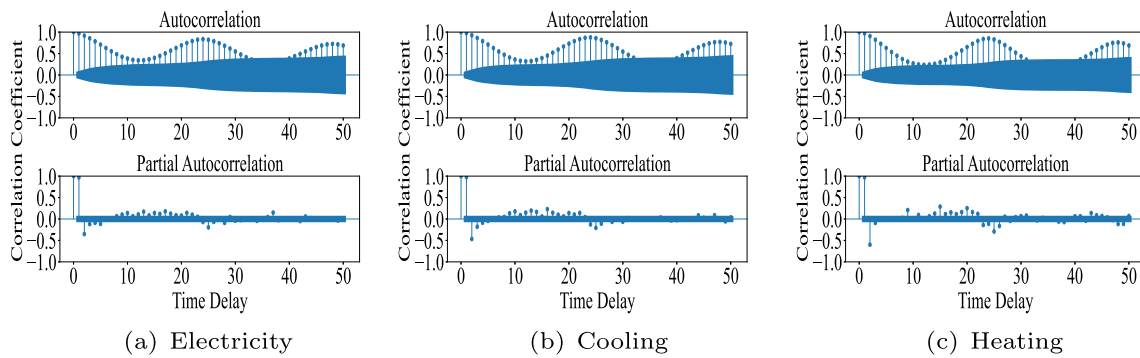


Fig. 6 The ACF and PACF of electricity, cooling and heating loads in February 2018

uncertain, which may make a significant impact on the load level of the overall system.

In addition, multi-source load has obvious time-series characteristics, and its load demand at the previous moment has a great influence on the load demand forecast at the next moment. At the same time, multi-energy load forecasting has certain weekly and daily characteristics, i.e., the load data at the moment to be forecast has a great correlation with the load of the previous week and the previous day. Therefore, three matrices x_t^e , x_t^c , x_t^h are constructed. They represent electricity, cooling and heating LPF feature matrix, respectively. Based on Eq. (5), x_t^W and x_t^D are created. The former represents the feature matrix for the week before the prediction moment t . The latter represents the feature matrix for the day before the prediction moment t . W and D denote the week and the day before the predicted day, respectively. x_t^W is calculated by shifting x_t forward by 7×24 time units, and x_t^D is calculated by shifting x_t forward by 24 time units.

$$x_t^e = \begin{bmatrix} x_{1,t}^e \\ x_{2,t}^e \\ \dots \\ x_{n,t}^e \end{bmatrix}, x_t^c = \begin{bmatrix} x_{1,t}^c \\ x_{2,t}^c \\ \dots \\ x_{n,t}^c \end{bmatrix}, x_t^h = \begin{bmatrix} x_{1,t}^h \\ x_{2,t}^h \\ \dots \\ x_{n,t}^h \end{bmatrix} \quad (8)$$

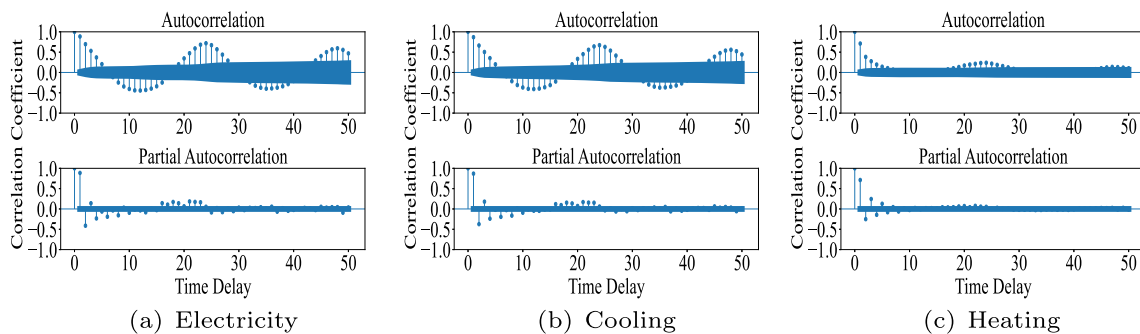


Fig. 7 The ACF and PACF of electricity, cooling and heating loads in July 2018

$$x_t^W = \begin{bmatrix} x_{1,t}^{W,e} & x_{1,t}^{W,c} & x_{1,t}^{W,h} \\ x_{2,t}^{W,e} & x_{2,t}^{W,c} & x_{2,t}^{W,h} \\ \dots & \dots & \dots \\ x_{n,t}^{W,e} & x_{n,t}^{W,c} & x_{n,t}^{W,h} \end{bmatrix}, x_t^D = \begin{bmatrix} x_{1,t}^{D,e} & x_{1,t}^{D,c} & x_{1,t}^{D,h} \\ x_{2,t}^{D,e} & x_{2,t}^{D,c} & x_{2,t}^{D,h} \\ \dots & \dots & \dots \\ x_{n,t}^{D,e} & x_{n,t}^{D,c} & x_{n,t}^{D,h} \end{bmatrix} \quad (9)$$

3.4 The impact of weather

Considering the impact of weather factors, climate information of one year is collected and shown in Fig. 8. Those data mainly include Global Horizontal Irradiance (GHI), Dew Point (DP), Precipitable Water (PW), Relative Humidity (RH), and Temperature (TMP). The data can be retrieved from the National Solar Radiation Database in the official website of the National Renewable Energy Laboratory [31]. To better understand the relationship between various loads and weather variations, correlations between different loads and between load and weather information were analyzed. Figure 9 shows the heat map of load and weather factors. The shades of color in the heat map correspond to the strength of the correlation, where the values indicate the magnitude of the correlation coefficient.

The coefficients between electricity, cooling and heating loads are -0.99, 0.88 and -0.94 respectively, which indicates that different loads have a strong correlation with each

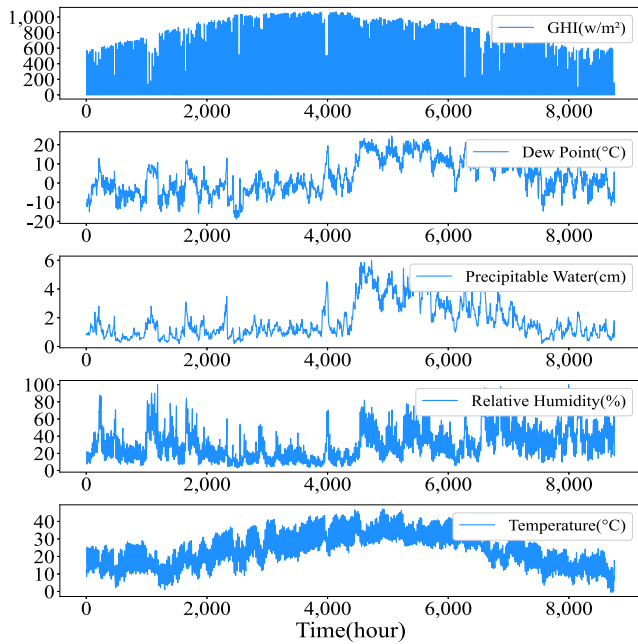


Fig. 8 The graph of climate data of Tempe campus in 2018

other. For the meteorological data, the correlation between temperature and load data is the highest, and its value can reach about 0.92. The second highest correlation is the precipitable water, while the lowest correlation is the relative humidity.

From the above analysis, it can be seen that there is a strong correlation between load and weather information, and that changes in climate in turn have a very important impact on the demand for different loads. The first step in setting up the input characteristics is to determine the composition and the necessary relevant information for each type of load. Then, the weather matrix E is constructed in Eq. (10). $h_{n,t}$ denotes the climate features at moment t ;

s_m denotes the climate features; m denotes the number of climate features.

$$E = \begin{bmatrix} h_{1,t}^{s1} & h_{1,t}^{s2} & \dots & h_{1,t}^{sm} \\ h_{2,t}^{s1} & h_{2,t}^{s2} & \dots & h_{2,t}^{sm} \\ \dots & \dots & \dots & \dots \\ h_{n,t}^{s1} & h_{n,t}^{s2} & \dots & h_{n,t}^{sm} \end{bmatrix} \quad (10)$$

Through the analysis above, it can be summarized as follows.

1. The load participation factor can portray the proportion of different loads in the total load, and it describes the law of load demand change more intuitively.
2. The demands of different loads are closely related. The conversion relationship between different energy sources verifies the existence of coupling relationship between loads.
3. There is a clear pattern of temporal changes in the load series, and the correlation of the series varies in different seasons.
4. The change of climate has an important influence on the demand of different loads.

4 Proposed model

Through the above analysis of UIES load characteristics, a prediction method based on MTL is proposed in combination with the special characteristics of multi-energy load demand variation. The whole prediction framework is shown in Fig. 10. Firstly, data pre-processing will be performed on the obtained raw multi-energy load data and its corresponding meteorological data. Secondly, the processed load data are divided into load participation factor data and total load data. Again, the load participation factor data feature matrix is constructed and relevant meteorological features are added to jointly form the input feature set. Then, the LSTM based multi-task learning network (LSTM-MTL) predicts the load participation factors, while the temporal convolutional network (TCN) predicts the total load demand. Finally, the outputs of LSTM-MTL and TCN network are combined to obtain the electricity, cooling and heating load prediction results.

4.1 Multi-task learning

The load participation factor reflects the proportion and variation of different loads. Effective load participation factor forecasting can accurately capture the changing patterns of different load demands and thus adjust the system to meet the energy consumption demand. Usually, the set of input and output features is the key to determine the performance capability of the model, where the input features are often various relevant attributes that affect

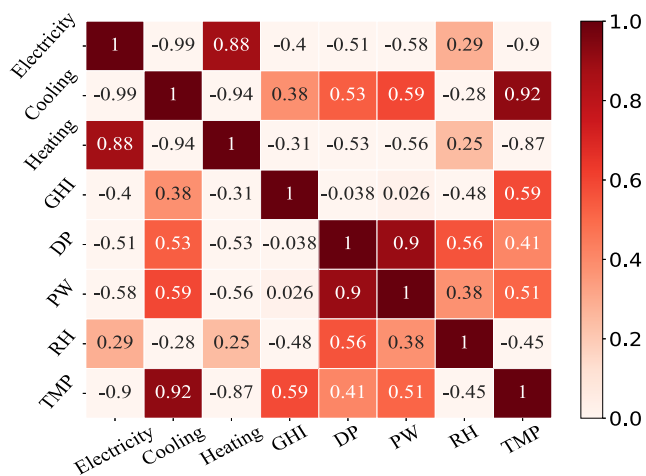


Fig. 9 The heat map of loads and weather factors

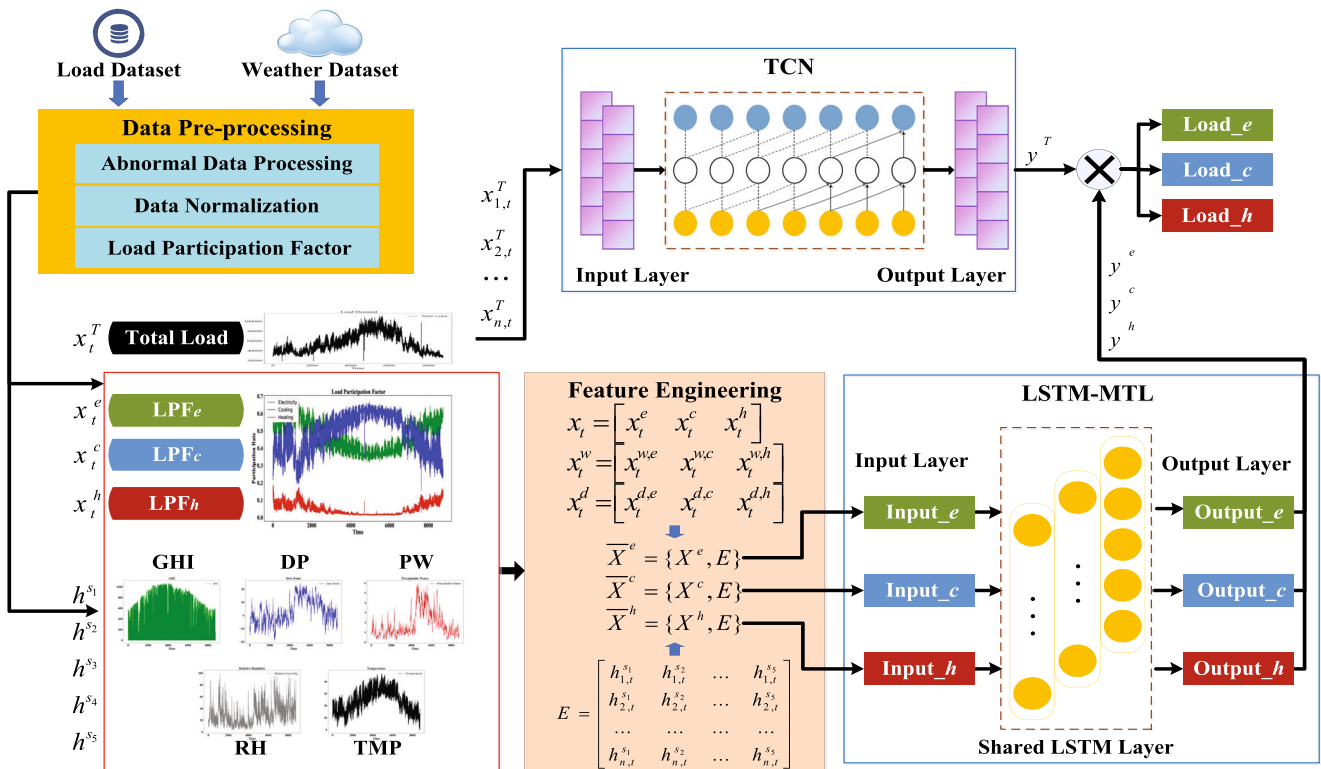


Fig. 10 The whole framework of the prediction model

load forecasting. When setting the input features, it is first necessary to determine the composition of various types of loads and the necessary related information, such as meteorological factors like temperature and rainfall.

$$\begin{aligned}
 X^e &= [x_{t-1}^W \ x_t^W \ x_{t+1}^W \ x_{t-1}^D \ x_t^D \ x_{t+1}^D \ x_{t-1}^e \ x_t^e] \\
 X^c &= [x_{t-1}^W \ x_t^W \ x_{t+1}^W \ x_{t-1}^D \ x_t^D \ x_{t+1}^D \ x_{t-1}^c \ x_t^c] \\
 X^h &= [x_{t-1}^W \ x_t^W \ x_{t+1}^W \ x_{t-1}^D \ x_t^D \ x_{t+1}^D \ x_{t-1}^h \ x_t^h]
 \end{aligned} \tag{11}$$

$$\begin{aligned}
 \bar{X}^e &= \{X^e, E\} & Y^e &= [x_{t+1}^e] \\
 \bar{X}^c &= \{X^c, E\} & Y^c &= [x_{t+1}^c] \\
 \bar{X}^h &= \{X^h, E\} & Y^h &= [x_{t+1}^h]
 \end{aligned} \tag{12}$$

In addition, multi-source load has obvious time-series characteristics, and its load demand at the previous moment has a great influence on the load demand forecast at the next moment. At the same time, multi-energy load forecasting has certain weekly and daily characteristics, i.e., the load data at the moment to be forecast has a great correlation with the load of the previous week and the previous day. Therefore, multi-energy load data X and meteorological data E are jointly selected as the input feature \bar{X} . The output feature Y is the actual multi-energy load data at the time to be measured. The input feature \bar{X} and the output feature Y together form the sample datasets $\{\bar{X}, Y\}$.

In UIES, different energy systems have their own physical characteristics. They are often coupled to each other in the energy flow through energy conversion devices. For example, a cogeneration units converts natural gas into electricity and heating, and an electric boilers can convert electricity into heating. Therefore, there is a strong coupling relationship between different energy systems. Energy systems generate a large amount of usage data during operation, and much energy conversion information is hidden in them. However, the characteristics of these conversion information are difficult to be generalized by traditional manual feature extraction methods. When multi-task learning is considered, it can effectively utilize the complex shared information of energy conversion. Also, it can help to train the model to extract high-level abstract features, which can describe the coupling relationship between loads more comprehensively and improve the prediction accuracy [32].

MTL is a classical inductive migration mechanism. Its main goal is to improve the generalization ability of a model by using information behind multiple related tasks [19]. The schematic diagram of MTL is shown in Fig. 11. The fusion layer is composed by LSTM network [33]. Specifically, it achieves the goal by training multiple learning tasks in parallel through a shared mechanism. It can learn one problem while learning and gaining knowledge from other related problems by using shared information, which helps

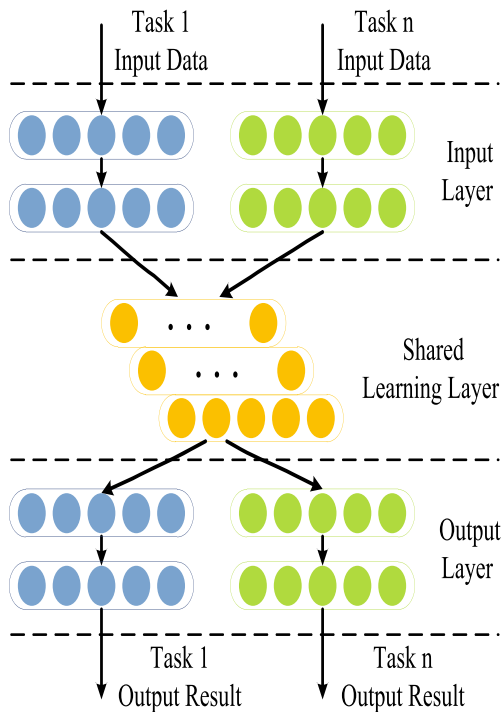


Fig. 11 The schematic diagram of multi-task learning

to dig out the correlation between different tasks and thus improve the accuracy of the respective tasks.

4.2 Temporal convolutional network

When predicting the load participation factors, the model is able to fully explore the coupling relationships between the loads to obtain the share of each load in the total load. However, in order to obtain the final predicted load value, the total load needs to be predicted. In fact, once time series data or time-series related data are mentioned, the neural network models that usually come to mind are recurrent neural networks (RNN), LSTM and GRU, etc. These models are built under the framework of RNN, because the recurrent autoregressive structure inherent in RNN is a good representation of time series data [28]. Meanwhile, CNN was first applied to image processing

and it is also used to deal with time-series problems. As a variant of CNN, temporal convolutional network (TCN) was proposed in Ref. [34]. It has been demonstrated that TCN also has good performance. The structure graph of TCN is shown in Fig. 12

TCN utilizes a unique dilation causal convolution to ensure causality, avoid future data leakage, and expand the perceptual field at the same time. It can achieve the overall perception of the specified length sequence data by adjusting the parameters of convolution kernel, number of convolution layers, and expansion coefficients, and ensure that the information length of each hidden layer is the same as the input sequence. It ensures that the sequence as a whole can exert influence on the deep network [35, 36].

5 Experiments

5.1 Evaluation index

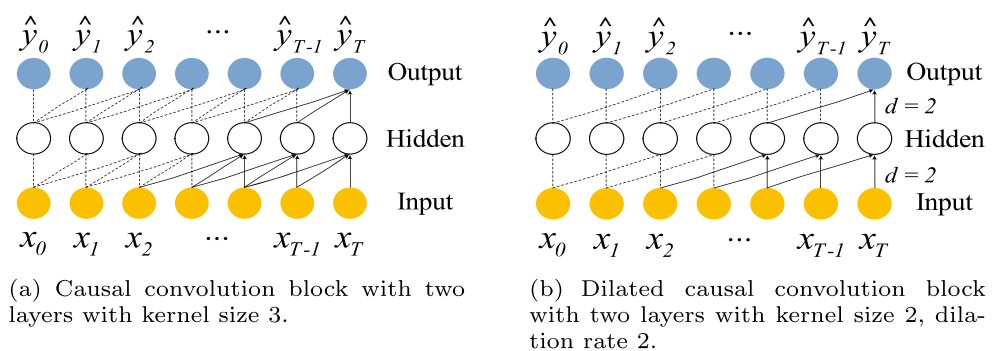
In order to evaluate the prediction effect of the model, this paper uses Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) to jointly evaluate the prediction error. The formulas for MAPE, MAE and RMSE are shown below.

At the same time, the constructed models need to predict multiple subtasks simultaneously. Given that each load has different characteristics, this often makes it difficult to set the same hyperparameters to satisfy the simultaneous optimal prediction effect of multiple types of loads. Therefore, the evaluation metrics are developed based on the loads that dominate in UIES. In addition to the Mean Accuracy (MA) metric, the Weighted Mean Accuracy (WMA) metric is also added to evaluate the performance of the prediction model. The expressions for MA and WMA are specified as below.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left\| \frac{y_i - \hat{y}_i}{y_i} \right\| \tag{13}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n \|y_i - \hat{y}_i\| \tag{14}$$

Fig. 12 The schematic diagram of causal convolution and dilated causal convolution structures in TCN



$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{15}$$

$$MA = 1 - MAPE \tag{16}$$

$$WMA = \beta_e MA^e + \beta_c MA^c + \beta_h MA^h \tag{17}$$

In the equations above, \hat{y}_i and y_i are the predicted and true values of the load at the i moment, respectively; n is the number of samples; MA^e , MA^c and MA^h are the MA values of the electricity, cooling, and heating loads, respectively; β_e , β_c and β_h are the weights occupied by the electricity, cooling, and heating loads, respectively. Considering that the studied UIES is dominated by electricity and cooling loads, the weight coefficients in Eq. (10) are taken as 0.3, 0.5 and 0.2, respectively [32].

5.2 Experimental data

The data recorded on the web platform of the Campus Metabolism project for 24 hours per day from January 2018 to December 2019, which has a temporal resolution of 1 hour, are used for the experiments [27]. Based on the prediction framework of Fig. 10, the training set, validation set and test set are divided according to 85%, 10% and 5% to predict the electricity, cooling and heating loads for the next 1 hour.

5.3 Parameters setting of models

During the course of the experiments, several time series forecasting models, including Multi-task Learning based on LSTM (MTL-LSTM) [32], combined convolutional neural network (CNN) and LSTM (CNN-LSTM) [37], DBN with three layers of Restricted Boltzmann Machine (RBM-DBN) and RFR with 100 regression trees, are used for comparison with the proposed model. The specific parameter settings of all models are shown in Table 1. It is worth noting that the MTL-LSTM, CNN-LSTM, RBM-DBN, and RFR models use load demand data for training and prediction, whereas our model uses LPF and total load data.

5.4 Experimental results

To analyze the prediction performance, the evaluation results are shown in Tables 2 and 3. For more visually displaying the prediction effect of the model, the prediction curves, scatter plot and percentage relative error plot are depicted in Figs. 13, 14, and 15, respectively.

In the process of deep learning modeling, k-fold cross-validation can be used for limited data in order to prevent problems such as overfitting during the training process. It can effectively use the experimental data so as to obtain

Table 1 Parameter settings of models

Models	Parameters Setting
MTL-LSTM [32]	$n_i=[32,32]$, $n_{lstm}=[96,96,96]$, $n_o=[32,32]$, $a_h=\text{sigmoid}$, $a_o=\text{linear}$, $op=\text{Adam}$
CNN-LSTM [37]	$n_f=[10,20]$, $s_k=[3,3]$, $s_p=[2,2]$, $w_d=[0.25,0.25]$, $a_{cnn}=\text{relu}$, $n_{lstm}=[24,16,32,16,16,16]$, $a_{lstm}=\text{sigmoid}$, $a_o=\text{linear}$, $op=\text{Adam}$
RBM-DBN	$n_h=[250,500,200]$, $a_h=\text{sigmoid}$, $a_o=\text{linear}$
RFR	$n_t=100$, $n_v=5$
Ours	$n_i=[32,32]$, $n_{lstm}=[96]$, $n_o=[32]$, $a_h=\text{sigmoid}$, $a_o=\text{linear}$, $n_f=14$, $s_k=4$, $s_d=[1,2,4,8]$, $op=\text{Adam}$

* n_i -number of input layer units, n_{lstm} -number of lstm layer units, n_o -number of output layer units, a_h -activation function of hidden layer, a_o -activation function of output layer, op -optimizer, n_f -number of filters, s_k -size of kernel, s_p -size of pool, w_d -weight of dropout, a_{cnn} -activation function of cnn layer, a_{lstm} -activation function of lstm layer, n_h -number of hidden layer units, n_t -number of trees, n_v -number of variables randomly sampled at each split, s_d -size of dilation factors

evaluation results with low bias and variance, and can be used as a model optimization metric for improving the performance of the model. In this paper, a 5-fold cross-validation was performed. By partitioning datasets, subsets K1 to K5 was formed. Those models were trained and validated on each of five subsets. The experimental results of the prediction model are presented in Tables 2 and 3, respectively. The experimental results reveal that the prediction effects of different model for electricity, cooling and heating load are different on the same dataset. At the same time, the prediction results of the same model for different loads have the same trend in different subsets. Specifically, the prediction effect of the proposed model and MTL-LSTM for electricity and cooling loads is better than that for heating loads. This is mainly due to the strong randomness of heating load data and its variable pattern, while electricity and cooling load data tend to have a certain trend of change and are relatively stable. Such an essential distinction makes it easier for the prediction models to uncover the change patterns of the data. In view of the differences between the models, it can be found that the models RBM-DBN and RFR are slightly less effective in predicting the cooling load, while the model CNN-LSTM has a better prediction effect for the electricity load. This is due to the fact that different divisions of dataset will have a certain degree of influence on the prediction results of the model. Especially, time is a crucial factor when considering the time series prediction problem. In this way, the order of data selection for the test set is bound to produce some noise to the model. Besides, the variability among the models themselves will be reflected in the datasets. Combining the prediction effects of K1 to K5, the WMA

Table 2 Accuracy evaluation of prediction model for MAPE, MAE and RMSE

K-Fold Subset	Models	MAPE			MAE			RMSE		
		Electricity	Cooling	Heating	Electricity	Heating	Cooling	Electricity	Heating	Cooling
K1	MTL-LSTM	3.337	4.893	3.952	535.019	502.978	125.618	706.990	656.446	164.564
	CNN-LSTM	6.383	21.164	6.824	1026.563	2202.586	194.439	1244.475	2620.579	243.174
	RBM-DBN	4.048	6.495	5.110	679.255	711.865	151.296	870.068	945.850	193.580
	RFR	6.874	8.503	6.860	1091.198	862.111	198.392	1410.399	1156.689	253.164
	Ours	2.298	3.091	3.495	385.192	326.057	105.679	528.538	451.123	142.154
K2	MTL-LSTM	2.480	2.633	7.871	621.852	1284.645	111.179	787.626	1624.860	122.498
	CNN-LSTM	5.910	7.832	4.684	1523.825	3767.628	68.098	1873.233	4435.814	85.762
	RBM-DBN	7.408	7.768	4.594	1854.822	3703.599	69.178	2125.084	4266.964	92.361
	RFR	5.313	5.308	3.654	1311.507	2566.275	53.305	1641.825	3275.843	70.799
	Ours	2.163	2.270	3.974	553.135	1088.559	57.611	742.726	1409.879	74.174
K3	MTL-LSTM	3.283	5.273	11.809	607.924	598.176	378.856	863.683	847.173	424.594
	CNN-LSTM	5.467	16.058	7.376	1006.577	1644.614	238.649	1321.801	2067.929	327.960
	RBM-DBN	5.640	13.862	8.145	1069.600	1259.802	269.800	1341.818	1531.776	338.096
	RFR	4.583	13.644	6.668	834.390	1332.259	212.070	1123.797	1632.364	284.679
	Ours	2.579	3.754	4.692	477.323	404.478	146.216	759.409	575.943	187.387
K4	MTL-LSTM	3.286	3.130	3.554	956.328	1056.509	60.421	1332.891	1405.827	81.524
	CNN-LSTM	11.220	10.477	12.421	3278.405	3439.804	205.305	3833.551	4167.426	229.512
	RBM-DBN	5.203	11.171	8.802	1383.137	4121.494	147.891	1698.254	4760.503	175.108
	RFR	6.518	8.978	4.759	1917.580	2910.416	84.254	2561.506	3610.438	114.551
	Ours	2.547	2.818	6.541	693.113	977.953	104.285	930.609	1326.021	139.021
K5	MTL-LSTM	3.611	4.296	3.983	891.143	1255.026	70.920	1061.307	1453.548	91.731
	CNN-LSTM	5.019	8.555	13.647	1205.575	2614.463	231.454	1529.389	3305.910	255.400
	RBM-DBN	3.604	5.134	6.190	884.433	1452.520	105.335	1167.687	1807.591	127.860
	RFR	4.227	7.212	5.963	1031.079	1955.520	102.544	1352.358	2360.047	127.293
	Ours	2.805	2.804	4.067	676.139	805.238	71.151	955.581	1015.321	91.354

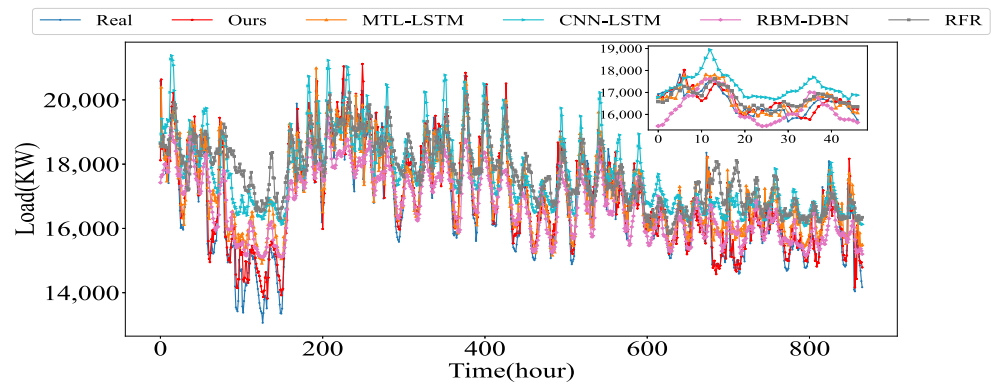
Table 3 The MA and WMA evaluation of prediction models

K-Fold Subset	Models	MA			WMA
		Electricity	Cooling	Heating	
K1	MTL-LSTM	0.96663	0.95107	0.96048	0.95762
	CNN-LSTM	0.93617	0.78836	0.93176	0.86138
	RBM-DBN	0.95952	0.93505	0.94891	0.94516
	RFR	0.93126	0.91497	0.93140	0.92315
	Ours	0.97702	0.96909	0.96505	0.97066
K2	MTL-LSTM	0.97520	0.97367	0.92129	0.96365
	CNN-LSTM	0.94090	0.92168	0.95316	0.93374
	RBM-DBN	0.92592	0.92232	0.95406	0.92975
	RFR	0.94687	0.94692	0.96346	0.95021
	Ours	0.97837	0.97730	0.96026	0.97422
K3	MTL-LSTM	0.96717	0.94727	0.88191	0.94017
	CNN-LSTM	0.94533	0.83942	0.92624	0.88855
	RBM-DBN	0.94360	0.86138	0.91855	0.89748
	RFR	0.95417	0.86356	0.93332	0.90469
	Ours	0.97421	0.96246	0.95308	0.96411
K4	MTL-LSTM	0.96714	0.96870	0.96446	0.96738
	CNN-LSTM	0.88780	0.89523	0.87579	0.88911
	RBM-DBN	0.94797	0.88829	0.91198	0.91093
	RFR	0.93482	0.91022	0.95241	0.92604
	Ours	0.97453	0.97182	0.93459	0.96519
K5	MTL-LSTM	0.96389	0.95704	0.96017	0.95972
	CNN-LSTM	0.94981	0.91445	0.86353	0.91487
	RBM-DBN	0.96396	0.94866	0.93810	0.95114
	RFR	0.95773	0.92788	0.94037	0.93933
	Ours	0.97195	0.97196	0.95933	0.96943

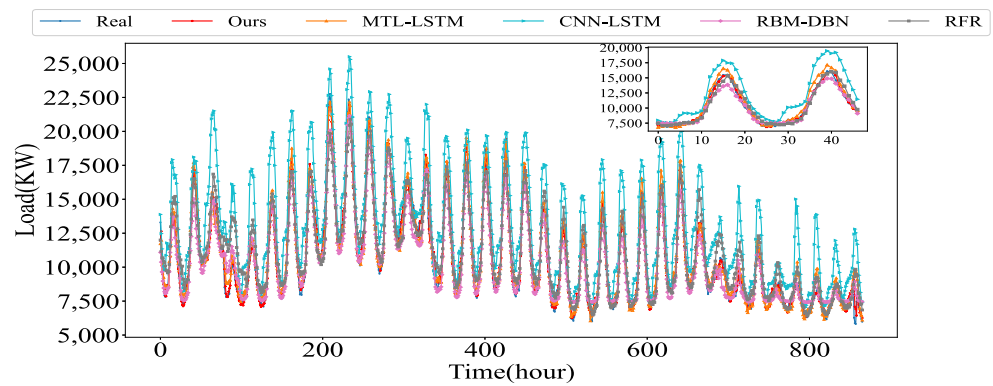
value of the proposed model is 0.96872, which has the best prediction accuracy. In general, the proposed model has a great prediction effects in K1 to K5. In the following, the experimental results of K1 are analyzed and illustrated in detail.

For electricity load, the MAPE values of CNN-LSTM and RFR models are 6.38% and 6.87%, respectively, while the value of the proposed model is only 2.29%. Compared with the worst RFR model, the prediction error of our model is reduced by 4.58%. Among the three load prediction curves, the cooling load curve is less volatile and more stable, while the opposite is true for the electricity and heating load curves. Their periodicity is also not obvious. However, for each load, the proposed model has better stability and the prediction curves are closer to the real load curves compared to other models. Specifically, the prediction curves of both the CNN-LSTM and RFR models are higher than the actual curves, which indicates that the predicted values of peak and valley are higher than

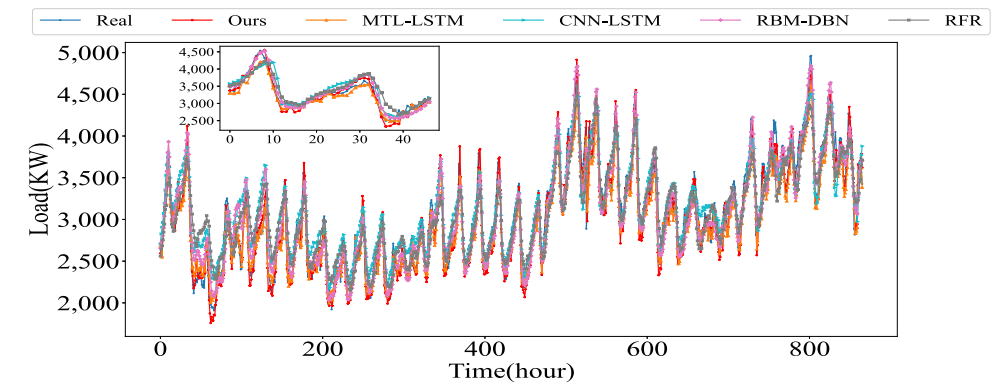
the actual values, but the predicted peaks of the CNN-LSTM model tend to deviate more from the real values. For cooling load, the prediction trends of several models are basically the same. This indicates that the models have better performance for time series with obvious regularity and small period variation, and the models can quickly dig out the characteristics of the series. However, a careful comparison reveals that the predicted peak values of CNN-LSTM model are too high compared with the true values, will result in large prediction errors. In addition, the more popular RBM-DBN in recent years has improved the prediction accuracy compared with CNN-LSTM and RFR models. However the MAPE, MAE and RMSE indexes of cooling load are still 3.404%, 385.8079 KW and 494.7266 KW respectively, higher than the proposed model. The prediction effect of the MTL-LSTM model is closest to the proposed model. The heating load prediction graph shows that the demand is much smaller than the demand of cooling and electricity load. The prediction effect of the proposed

Fig. 13 The prediction curves of different loads

(a) The curves of electricity load



(b) The curves of cooling load



(c) The curves of heating load

model is better than several other models for the peak point of load. The other models have better effect in predicting the overall trend of the load curve, but the predictions of the peak and valley values are often higher or lower than the real values, which results in the more prediction errors. The WMA of the proposed model reaches 97.066%, which are higher than other models, the minimum improvement is 1.3%. Through the above analysis, it can be seen that our model is all lower than the other four models in terms of prediction error and higher than them in terms of prediction accuracy.

In visual data analysis, scatter plots depict how well a model performs in prediction by plotting the degree of deviation between the overall distribution of prediction results and the true values. And it provides a more intuitive view of the strengths and weaknesses of the model. As shown in the scatter plots, the distribution of the prediction results of the proposed model is more aggregated, tighter and closer to the real load value, in which there are fewer outliers and the prediction results are real and reliable. Compared with the performance of the other models, it can be seen that the prediction results of the other models

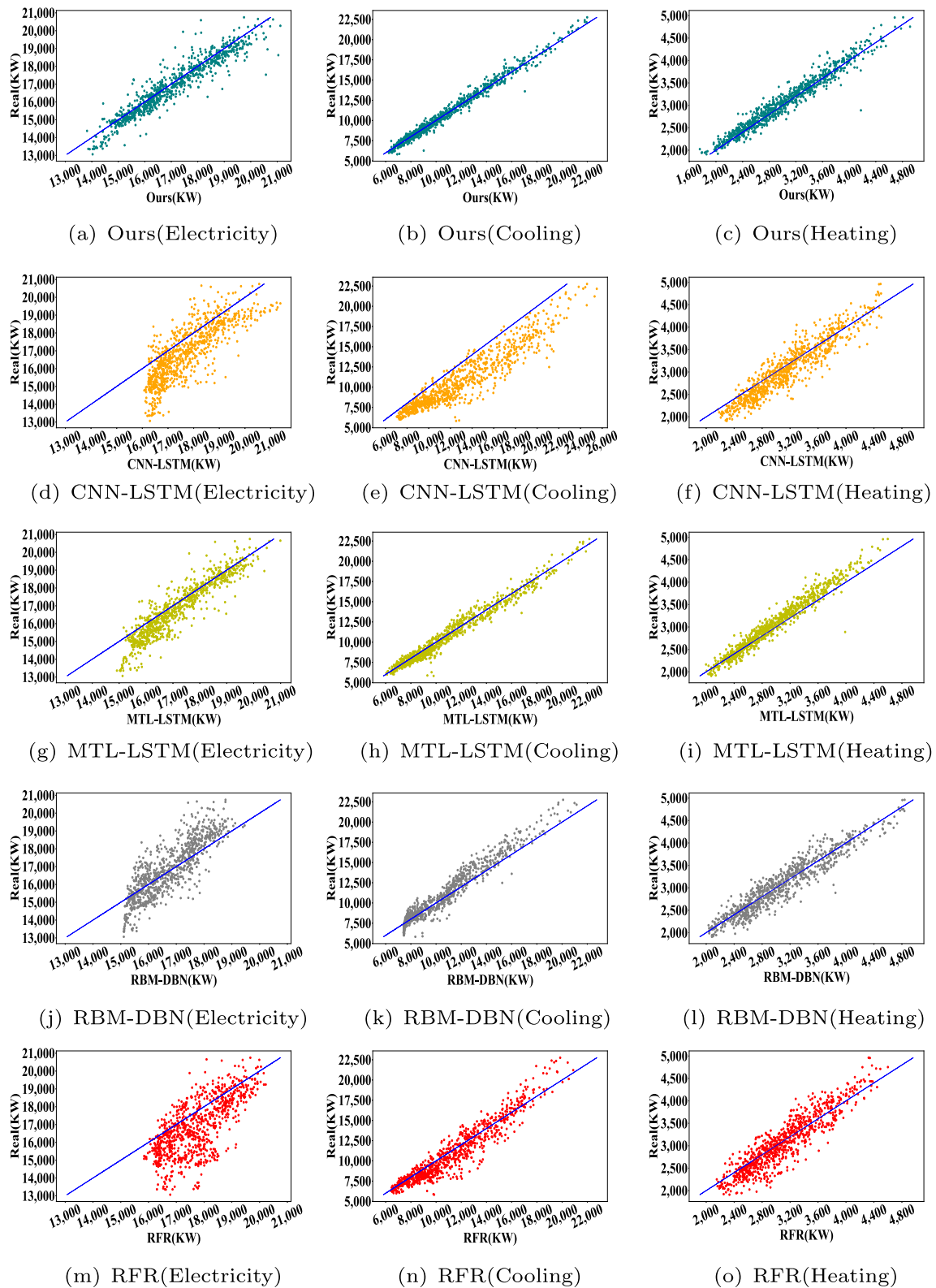
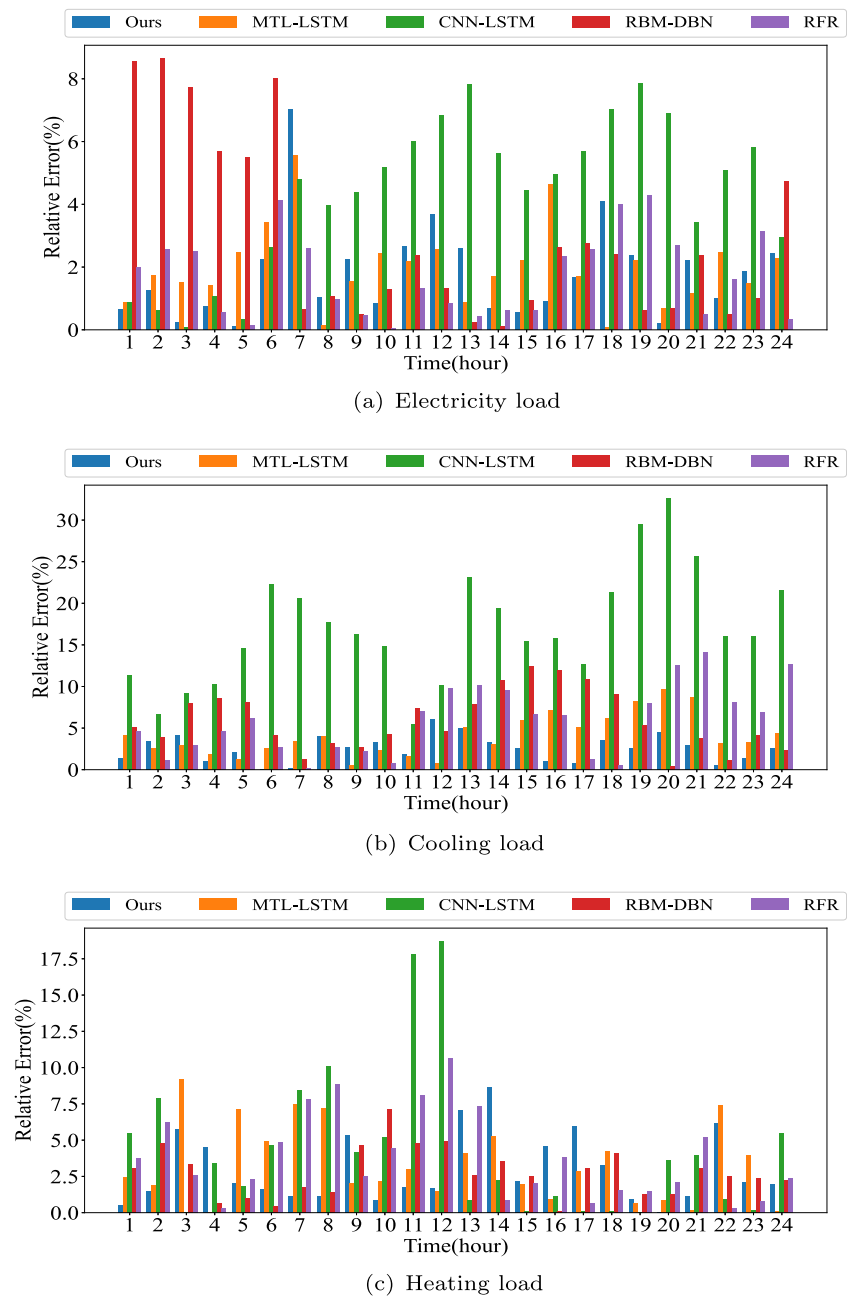


Fig. 14 The scatter plots of load forecasting

have more scattered distribution and are far from the true load values. Besides, there are even local deviation and

truncation phenomena, which show that their prediction effects are not good.

Fig. 15 The percentage relative error plots for different loads



In addition to the above analysis, Fig. 15 represent the percentage relative error plots of the one-day load forecasting results, respectively. The percentage relative error is defined as follows

$$\delta_t = \frac{\|x_t - \hat{x}_t\|}{x_t} \times 100\% \tag{18}$$

where x_t denotes the true value, \hat{x}_t denotes the predicted value, and t denotes the time series number. The load prediction results of a day in the test dataset is selected and the percentage relative error of different loads is calculated. It can be more intuitive to understand the prediction effect

of the model. From the Fig. 15, the relative errors of the proposed model always remain at a low level, while the relative errors of the CNN-LSTM and RBM-DBN models are more fluctuating and less stable, and the errors of the RFR and MTL-LSTM models are relatively low.

In order to verify the validity of the model, this paper analyzes the impact and effect of each part of the model. Based on this, the parts of the proposed model (MTL-LSTM.TCN) are replaced to form Model_1 (LSTM.TCN), Model_2 (MTL-LSTM.LSTM), and Model_3 (LSTM.LSTM), respectively. Among them, the LSTM network model is used for all the replacement

Table 4 Accuracy evaluation of ablation prediction model for MAPE, MAE and RMSE

Models	MAPE			MAE			RMSE		
	Electricity	Cooling	Heating	Electricity	Cooling	Heating	Electricity	Cooling	Heating
Model_1	2.441	3.402	3.771	415.682	362.097	115.575	570.185	497.679	152.961
Model_2	4.687	4.938	8.251	768.493	525.988	257.240	877.491	688.089	308.940
Model_3	4.535	3.739	3.838	750.191	389.279	116.099	870.474	525.141	154.126
Model_4	2.765	4.495	4.151	457.958	435.248	123.753	610.478	551.996	163.495
Ours	2.298	3.091	3.495	385.192	326.057	105.679	528.538	451.123	142.154

parts. The MAPE, MAE and RMSE evaluation metrics of the four models are listed in Table 4. As can be seen from the table, the MAPE metrics of the proposed model is 2.30%, 3.09%, and 3.50% for electricity, cooling, and heating loads, respectively. The electricity load has the lowest prediction error, while the heating load has the highest error. This situation is also reflected in the MAE and RMSE indexes. Meanwhile, comparing the other models, it can be found that Model_1 and Model_2 also have the lowest prediction error on the electricity load, followed by the cooling load, and the highest error is still the heating load. However, the prediction effect of Model_3 is different. The Model_3 has the lowest error for the cooling load and the highest error for the electricity load. From the above results, it can be found that the prediction results of models are different. Specifically, the focus is a problem of the model itself and it depends on the initial design of the model. Thereby, in order to evaluate the models more objectively, the MA and WMA evaluation metrics for the models is listed Table 5. The WMA index of the proposed model is 0.97066, and the values are higher than the other three models.

Changes in weather often have an impact on people's load usage, and thus it is necessary to consider climate change in multi-energy load forecasting. Therefore, this paper conducts a comparison experiment with or without including weather data. To ensure the fairness, the model settings of the proposed model and the comparative Model_4 are exactly same, where the data input of Model_4 does not include the meteorological data component. The

Table 5 The MA and WMA evaluation of ablation prediction model

Models	MA			WMA
	Electricity	Cooling	Heating	
Model_1	0.97559	0.96598	0.96229	0.96813
Model_2	0.95313	0.95062	0.91749	0.94474
Model_3	0.95465	0.96261	0.96162	0.96002
Model_4	0.97235	0.95505	0.95849	0.96093
Ours	0.97702	0.96909	0.96505	0.97066

results are shown in Tables 4 and 5. In terms of MAPE index, the prediction results of the proposed model are all lower than Model_4. This result reflects that meteorological data can help to improve the prediction performance of the model, and there is a close correlation between meteorological data and multivariate load data from the level of data analysis.

Based on this analysis, it can be seen that the proposed model has a better prediction effect. Among them, MTL-LSTM module uses the fusion layer to explore the coupling relationship between different loads, so that the model can better grasp the usage characteristics of different loads, which is more conducive to the demand prediction of different loads. At the same time, the results show that the prediction effect of TCN module is better than that of LSTM network, which reflects the reasonableness and effectiveness of the model design in this paper.

6 Conclusion

Given the problems such as strong volatility and stochasticity of load demand in UIES, multi-energy load forecasting puts forward higher requirements for the reliability and accuracy. In this paper, a short-term multi-energy load forecasting method based on MTL method is proposed. It well portrays the degree of load participation and deeply explores the complex coupling relationships among loads. By integrating and sharing the data information among different loads, the proposed method effectively improves the prediction accuracy. Comparing with current methods, the results show that the proposed MTL method has certain advantages, and the weighted average accuracy is improved by at least 1.3%, which is significant for the load prediction of integrated energy systems. Therefore, this method can provide an accurate and stable forecasting for UIES and support reliable decision making for market participants.

Acknowledgements This work was supported by the National Natural Science Foundation of China (61873222) and the Project of Hunan National Center for Applied Mathematics (2020ZYT003).

Data Availability The data that support the findings of this study are openly available at <http://cm.asu.edu/>, or from the corresponding author upon request.

References

- Wang J, Zhong H, Ma Z, Xia Q, Kang C (2017) Review and prospect of integrated demand response in the multi-energy system. *Appl Energy* 202:772–782
- Wang Y, Wang Y, Huang Y, Li F, Zeng M, Li J, Wang X, Zhang F (2019) Planning and operation method of the regional integrated energy system considering economy and environment. *Energy* 171:731–750
- Jia H, Mu Y, Yu X (2015) Thought about the integrated energy system in china. *Electric Power Construct* 36(1):16–25
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444
- Zhou S, He Y, Liu Y, Li C, Zhang J (2021) Multi-channel deep networks for block-based image compressive sensing. *IEEE Trans Multimed* 23:2627–2640. <https://doi.org/10.1109/TMM.2020.3014561>
- Cheng L, Yu T, Zhang X, Yin L (2019) Machine learning for energy and electric power systems: state of the art and prospects. *Autom Electr Power Syst* 43(1):15–43
- Chen L (2020) Designing a short-term load forecasting model in the urban smart grid system. *Appl Energy* 266:114850
- Chitalia G, Pipattanasomporn M, Garg V, Rahman S (2020) Robust short-term electrical load forecasting framework for commercial buildings using deep recurrent neural networks. *Appl Energy* 278:115410
- Fu G (2018) Deep belief network based ensemble approach for cooling load forecasting of air-conditioning system. *Energy* 148:269–282
- Geysen D, De Somer O, Johansson C, Brage J, Vanhoudt D (2018) Operational thermal load forecasting in district heating networks using machine learning and expert advice. *Energy Build* 162:144–153
- Nigitz T, Golles M (2019) A generally applicable, simple and adaptive forecasting method for the short-term heat load of consumers. *Appl Energy* 241:73–81
- Persson C, Bacher P, Shiga T, Madsen H (2017) Multi-site solar power forecasting using gradient boosted regression trees. *Sol Energy* 150:423–436
- Wu K, Wu J, Feng L, Yang B, Liang R, Yang S, Zhao R (2021) An attention-based cnn-lstm-bilstm model for short-term electric load forecasting in integrated energy system. *Int Trans Electr Energy Syst* 31(1):12637
- Zhou B, Meng Y, Huang W, Wang H, Deng L, Huang S, Wei J (2021) Multi-energy net load forecasting for integrated local energy systems with heterogeneous prosumers. *Int J Electr Power Energy Syst* 126:106542
- Powell KM, Sriprasad A, Cole WJ, Edgar TF (2014) Heating, cooling, and electrical load forecasting for a large-scale district energy system. *Energy* 74:877–885
- Ma M, Jin B, Luo S, Guo S, Huang H (2020) A novel dynamic integration approach for multiple load forecasts based on q-learning algorithm. *Int Trans Electr Energy Syst* 30(7):12146
- Wang S, Wang S, Chen H, Gu Q (2020) Multi-energy load forecasting for regional integrated energy systems considering temporal dynamic and coupling characteristics. *Energy* 195:116964
- Zhou S, Deng X, Li C, Liu Y, Jiang H (2022) Recognition-oriented image compressive sensing with deep learning. *IEEE Trans Multimed*. <https://doi.org/10.1109/TMM.2022.3142952>
- Caruana R (1997) Multitask learning. *Mach Learn* 28(1):41–75
- Wang X, Wang S, Zhao Q, Wang S, Fu L (2021) A multi-energy load prediction model based on deep multi-task learning and ensemble approach for regional integrated energy systems. *Int J Electr Power Energy Syst* 126:106583
- Tan Z, De G, Li M, Lin H, Yang S, Huang L, Tan Q (2020) Combined electricity-heat-cooling-gas load forecasting model for integrated energy system based on multi-task learning and least square support vector machine. *J Clean Prod* 248:119252
- Zhang L, Shi J, Wang L, Xu C (2020) Electricity, heat, and gas load forecasting based on deep multitask learning in industrial-park integrated energy system. *Entropy* 22(12):1355
- Wang B, Zhang L, Ma H, Wang H, Wan S (2019) Parallel lstm-based regional integrated energy system multienergy source-load information interactive energy prediction. *Complexity*, vol 2019
- Zhou D, Ma S, Hao J, Han D, Huang D, Yan S, Li T (2020) An electricity load forecasting model for integrated energy system based on bigan and transfer learning. *Energy Reports* 6:3446–3461
- Zeng M, Liu Y, Zhou P, Wang Y, Hou M (2018) Review and prospects of integrated energy system modeling and benefit evaluation. *Power Syst Technol* 42(6):1697–1708
- Yu X, Xu X, Chen S, Wu J, Jia H (2016) A brief review to integrated energy system and energy internet. *Trans China Electrotech Soc* 31(1):1–13
- Asu. Campus Metabolism. (2022). <http://cm.asu.edu/>. Accessed 26 Jan 2022
- Lara-Ben'itez P, Carranza-Garc'ia M, Riquelme JC (2021) An experimental review on deep learning architectures for time series forecasting. *arXiv:2103.12057*
- Du S, Li T, Yang Y, Horng S-J (2020) Multivariate time series forecasting via attention-based encoder–decoder framework. *Neurocomputing* 388:269–279
- Kang C, Xia Q, Liu M. (2017) Power system load forecasting (the 2nd edn).
- Nsrd Data Viewer [DB/OL]. (2022). America 26 Jan 2022. <https://maps.nrel.gov/nsrdb-viewer/>.
- Sun Q, Wang X, Zhang Y, Zhang F, Zhang P, Gao W (2021) Multiple load prediction of integrated energy system based on lstm-mtl. *Autom Electr Power Syst* 45(5):63–70
- Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780
- Bai S, Kolter JZ, Koltun V (2018) An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv:1803.01271*
- Lara-Ben'itez P, Carranza-Garc'ia M, Luna-Romera JM, Riquelme JC (2020) Temporal convolutional networks applied to energy-related time series forecasting. *Appl Sci* 10(7):2322
- Wan R, Mei S, Wang J, Liu M, Yang F (2019) Multivariate temporal convolutional network: a deep neural networks approach for multivariate time series forecasting. *Electronics* 8(8):876
- Zhu R, Guo W, Gong X (2019) Short-term load forecasting for cchp systems considering the correlation between heating, gas and electrical loads based on deep learning. *Energies* 12(17):3308

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.