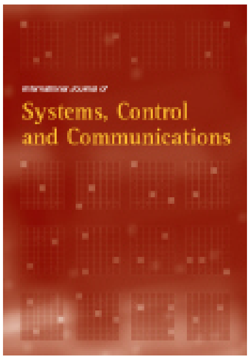


International Journal of Systems, Control and Communications

ISSN online: 1755-9359 - ISSN print: 1755-9340
https://www.inderscience.com/ijscc

Day-ahead hourly electricity load forecasting based on long short-term memory neural networks: a comparison study

Qianying Zhang, Junqiu Yuan, Xiang Liu, Li Sun

DOI: 10.1504/IJSCC.2025.10069426

Article History:

Received:	22 October 2024
Last revised:	25 November 2024
Accepted:	04 December 2024
Published online:	18 February 2025

Day-ahead hourly electricity load forecasting based on long short-term memory neural networks: a comparison study

Qianying Zhang, Junqiu Yuan, Xiang Liu and Li Sun*

Changzhou Power Supply Branch,
State Grid Jiangsu Electric Power Co., Ltd.,
Changzhou 213004, China
Email: zhangxiying0820@hotmail.com
Email: 15895002896@139.com
Email: 2925810656@qq.com
Email: yzsunli@foxmail.com
*Corresponding author

Abstract: Electricity load forecasting (ELF) is crucial for the economic planning of power systems. Due to its time-series essence, long short-term memory neural network (LSTM) is considered as a promising solution for ELF. This paper investigates the LSTM algorithms using different model inputs and structures. First, a univariate LSTM model is developed to train the relationship between the historical load and future load, based on which ELF is carried out. Second, a multi-variable LSTM model is proposed by incorporating the temperature data as another input, thus making the prediction results more reliable if day-ahead weather forecast is available. Finally, this multi-variable LSTM model is further extended to be of double layers for each cell. The error statistics show that the more complex two-layer LSTM reduces the RMSE, MAE, MAPE, IAE, and SD by 14.17%, 20.06%, 21.43% 20.08%, and 6.53% respectively, compared to the P_LSTM_T model.

Keywords: electricity load forecasting; ELF; long short-term memory; LSTM; day-ahead prediction.

Reference to this paper should be made as follows: Zhang, Q., Yuan, J., Liu, X. and Sun, L. (2025) 'Day-ahead hourly electricity load forecasting based on long short-term memory neural networks: a comparison study', *Int. J. Systems, Control and Communications*, Vol. 16, No. 1, pp.45–61.

Biographical notes: Qianying Zhang is an Engineer, received her Master's, and Supervisor of Technology Research and Development of Jiangsu Electric-Heat Coupling Integrated Energy System Engineering Research Center.

Junqiu Yuan is a Senior Engineer and technician, and has been engaged in dispatching, substation operation, marketing and other professional work, and in recent years has focused on the research of integrated energy services.

Xiang Liu is an Assistant Engineer, received his Bachelor's degree, and his research direction is integrated energy services. He is a technology researcher of Jiangsu Electric-Heat Coupling Integrated Energy System Engineering Research Center.

Li Sun is an Engineer and PhD candidate. His research interests are in integrated energy system optimisation and scheduling machine learning. He is a technology researcher of Jiangsu Electric-Heat Coupling Integrated Energy System Engineering Research Center.

1 Introduction

Electricity load forecasting (ELF) has received enormous attention and plays an increasingly vital role in the reliable operation and planning of power grid (Zhang et al., 2017b; Kramar and Alchakov, 2023). Forecasting is mainly divided into now forecasting, short-term forecasting, medium-term forecasting, and long-term forecasting. Long-term forecasts are suitable for forecasting financial markets and economic trends, while medium-term forecasts are suitable for forecasting seasonal commodity demand. Short-term forecasts, on the other hand, refer to forecasts within the next few days to weeks and are therefore well suited for forecasting electricity loads. Short-term ELF (i.e., one hour to one week in advance) (Mocanu et al., 2016), as an efficient guarantee for scheduling power generation and distribution scheme, guides electricity supply companies to reasonably arrange the electricity load, coordinate and make full use of renewable generations including solar energy and wind energy, so as to preserve the balance between the energy depletion and generation (Baskan et al., 2023). Electricity requirement of customers cannot be satisfied due to the low prediction load, which may even cause the power failure, while the electricity waste can be attributed to the high prediction load. The negative influence of ELF error is analysed in Ranaweera et al. (1997) and Douglas et al. (1998). Consequently, ELF with desired accuracy is becoming a pressing issue for power grid.

However, the complexity caused by the inherent nonlinear and stochastic characteristics of electricity load series has been an obstacle first to find the accurate model for ELF (Sadaei et al., 2019).

Irregular human energy depletion behaviour and various influencing factors (Ma et al., 2017) including climate conditions, increasing penetration of renewable generations and economic policy (Zhang et al., 2024) require to be simultaneously taken into consideration. Fortunately, the information related to the power depletion can be achieved through development and application of the smart meters, although the increase in the demand for energy consumption contributes to the exponential growth in power depletion data, which creates opportunities for ELF. Numerous technologies have been proposed and applied to predict the electricity load during the past decades. The theory of some traditional forecasting methods represented by autoregressive integrated moving average models (ARIMA) (Hongku, 2015; Borrero and Mariscal, 2023), season ARIMA (De Felice et al., 2015), regression methods (Guo et al., 2014), exponential smoothing methods (Wang et al., 2024; Sudheer and Suseelatha, 2015) and Kalman filter (Galka, 2024) is mainly based on the developed linear models, which limits their ability to learn and extract the strong nonlinearity and disorder of electricity load data. To address this deficiency, artificial intelligence technologies including support vector machine (SVM) (Zhang et al., 2017a), fuzzy logic (Barman et al., 2017), random forest (Chen et al., 2024) and artificial neural network (ANN) have been extensively applied in developing a more efficiency method for nonlinear ELF. Therein, ANNs represented by conventional back propagation neural network (BPNN) (Li and Huang, 2014) are commonly applied in ELF because of their powerful nonlinear mapping capability. However, ELF is only treated as a static regression problem in BPNN, while the time series load data is featured by the dynamic state (i.e., the current state of system are continuously associated with the previous information). Therefore, BPNN and other conventional neural networks simply model the nonlinear relationship between the input variables and output variables and fail in finding the inner correlation between the time series data, which results in their poor ability to provide ELF with high accuracy. Deep learning, as an extremely attractive technology in various research fields, conquers the weakness mentioned above and is then applied in studying time series prediction, whose tremendous capacity of data analysis, classification and prediction extremely fits the demand of data application in smart power grid. Therein, recurrent neural network (RNN) is proposed to predict time series load data. RNN-based ELF method developed in Shi et al. (2018) demonstrates the superiority over other methods. Among many variants of RNN, long short-term memory neural networks (LSTM) (Sherstinsky, 2020) compensates for the problems of gradient disappearance and explosion in RNN, which makes it a tremendous success in long sequence forecasting. ELF based on LSTM model offers competitive advantage and shows higher prediction accuracy in Jiao et al. (2018) and Kong et al. (2019). LSTM model is conceived as a promising tool in this paper to perform ELF. The prediction accuracy can be further improved since the temperature factor has not been taken into account in the LSTM model applied in Xu et al. (2018). Temperature is recognised as a main climate factor in electric energy consumption. Monteiro et al. (2018) detail that the increase in the electrical equipment such as air conditioning which consumes abundant electric energy on extreme cold and hot days, directly leads to the strong correlation between the power consumption and temperature. Therefore, the inclusion of temperature information certainly offers competitive advantage in day-ahead ELF. Furthermore, it is noted that the research efforts made in multi-layer LSTM model is relatively restricted, in spite of tremendous studies on single-layer LSTM model. Consequently, the issues on multilayer LSTM model for ELF remain open. The primary contributions of this paper are expressed as follows.

1. A univariate model using the basic LSTM algorithm is developed by training and extracting the relationship between the historical load and future load, based on which ELF can be realised.
2. A correlation analysis between the temperature and electricity load is conducted to lay a groundwork for the establishment of an advanced multi-variable LSTM model, thus the real-time temperature information is incorporated into the network input variables to achieve more reliable day-ahead forecasting accuracy.
3. The multi-variable LSTM model is further extended to be of a more complex double-layer structure, the efficacy of which is further discussed.
4. Several experiments are conducted to show and compare the results forecasted by the proposed basic and advanced LSTM models using different network inputs and structures. To confirm the superiority, BPNN model based on the same sampling data is developed to be compared with.

The remainder of the paper is organised as follows. Section 2 reviews the theoretical foundation of RNN and LSTM. The basic and advanced LSTM algorithms using different model inputs and structures are described in Section 3 in detail. Experimental results of the proposed models and further discussion for future research are presented in Section 4, while conclusions are summarised in Section 5.

2 Recurrent neural network-based LSTM

The basic introduction of RNN is explained in this section. Additionally, as a special type of RNN, LSTM applied in this paper is described in detail and shows superiority over the classical RNN in long time series data forecasting.

2.1 Recurrent neural network

Conventional neural networks such as BPNN implement prediction techniques by modelling the relationship between input and output variables. However, their application in time series data prediction is limited by their lack of ability to learn and capture internal correlations between serial data. To address this problem, RNNs have been proposed to achieve sequence-to-sequence connections and mapping by establishing recurrent connections between neurons (Wang et al., 2020). Unlike traditional neural networks in which neurons are independent of each other, RNNs are able to achieve the memory property through cyclic connections, making the current output not only dependent on the current input, but also influenced by previous information. This memory property (Lecun et al., 2015; Sutskever et al., 2014) gives RNNs a significant advantage when dealing with sequential data problems. The structure of RNN is shown in Figure 1.

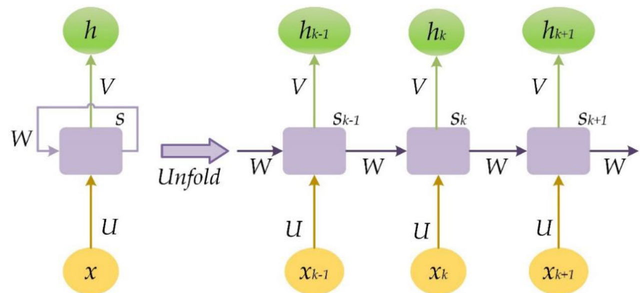
In the structure of RNN, x_k and h_k represent the sequence input and output of the network at time step k, respectively. The memory cell s_k is denoted as the hidden state of the network at time step k. Weights U, V and W connect the input layer to hidden layer, hidden layer to output layer and hidden layer to hidden layer, respectively. Parameters U, V and W are shared in each RNN cell, which illustrates that the parameters required to be learned and determined will be greatly reduced and thus training time will be shortened. The output h_k is functioned by previous and current input, which can be calculated by equations (1)–(2).

$$s_k = f\left(Ux_k + W \times s_{k-1}\right) \tag{1}$$

$$h_k = softmax\left(V \times s_k\right) \tag{2}$$

RNN's ability to predict sequence data can be attributed to the use of contextual information. For traditional RNN architectures, however, the contextual information that can be linked in practice is quite restricted, which denotes that the current RNN output is not sensitive to the distant information ahead. This deficiency directly leads to the gradient disappearance or explosion in RNN during back propagation and thus the long-range dependence cannot be maintained.

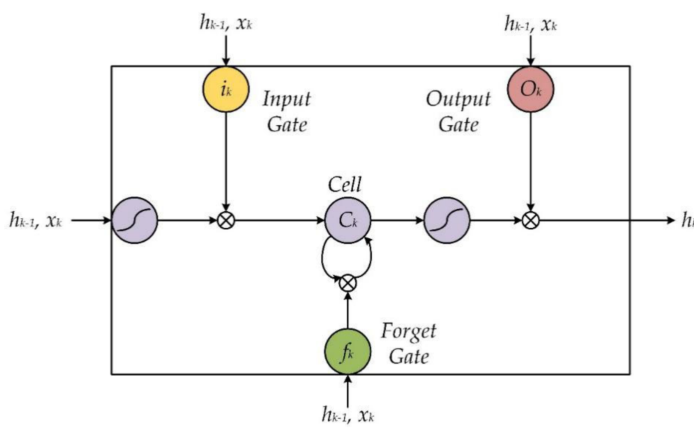
Figure 1 Structure of RNN (see online version for colours)



2.2 Long short-term memory neural network

RNN with LSTM cells is then developed to overcome the long-range dependence insufficiency problems to improve the traditional RNN architecture. Because of this superiority over the classical RNN, LSTM is popularised by researchers on long sequence prediction, including wind turbine power (Zhang et al., 2019), solar power (Wen et al., 2019) and battery state-of-charge (Chemali et al., 2018). The structure of LSTM cell is illustrated in Figure 2, where the input of memory cell state C_k is composed of the current input sequence x_k, previous output h_{k-1} and previous memory cell state C_{k-1}.

Figure 2 Structure of LSTM cell (see online version for colours)



It follows from Figure 2 that LSTM effectively manages the flow and storage of information by introducing three special 'gate' structures, as well as a memory cell state. These improvements allow LSTM to overcome the difficulties of traditional RNNs in dealing with long time interval dependencies. These gates including input gate i_k, output gate O_k, forget gate f_k allow the LSTM to erase or save new information to the memory cell. Forget gate is activated to delete some information in the memory cell and determine the remained information. Input gate is activated to calculate the information that required to be stored in the memory cell. Output gate is activated to control the information that will be transmitted to the following neurons or export the information of LSTM. Consequently, the ability to perform long short-term memory of LSTM can be realised by the activations of these gate unites. The work process of LSTM can be formulated by equations (3)–(8).

$$f_k = \sigma\left(W_{xf}x_k + W_{hf}h_{k-1} + b_f\right) \tag{3}$$

$$i_k = \sigma\left(W_{xi}x_k + W_{hi}h_{k-1} + b_i\right) \tag{4}$$

$$\tilde{C}_k = \tanh\left(W_{xc}x_k + W_{hc}h_{k-1} + b_c\right) \tag{5}$$

$$C_k = f_k \cdot C_{k-1} + i_k \cdot \tilde{C}_k \tag{6}$$

$$O_k = \sigma\left(W_{xo}x_k + W_{ho}h_{k-1} + b_o\right) \tag{7}$$

$$h_k = O_k \cdot \tanh\left(C_k\right) \tag{8}$$

where the hidden state h_k is set to a zero matrix at the initial time step. $\tilde{C}_k$ is a meaningless value for calculating the current memory cell state C_k. Sigmoid activation function σ limits the value between 0 and 1 to control the flow of the information, $\sigma = 0$ or $\sigma = 1$ denotes the complete abandon or reception of information. W is considered as the weight matrix between the two components and the bias of each gate b is added to enhance the network flexibility.

3 Modelling-based on LSTM

The proposed forecasting methodology in this section mainly includes three phrases. A univariate LSTM model based on the basic algorithm is first discussed with input feature selection during test. Second, a correlation analysis is conducted and then a multi-variable LSTM model is proposed by introducing the temperature variable into the network input to achieve more reliable day-ahead forecasting results. Finally, the effectiveness of extending the multivariate LSTM model to a more complex two-layer structure is explored.

3.1 Data preprocessing

The desired forecasting accuracy of the proposed models requires the experimental data preprocessed because of the network's sensitivity to data scale. Large difference in data scale will exert a negative influence on the training and prediction results of neural networks. Thus, the min-max normalisation is used to scale the experimental data between 0 and 1. The matrix X is alternatively written by X^{Δ} after normalisation.

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix}, X^{\Delta} = \begin{pmatrix} x_{11}^{\Delta} & \cdots & x_{1n}^{\Delta} \\ \vdots & \ddots & \vdots \\ x_{m1}^{\Delta} & \cdots & x_{mn}^{\Delta} \end{pmatrix} \tag{9}$$

$$x_{ij}^{\Delta} = \frac{x_{ij} - \min\limits_{1\le k\le n} x_{ik}}{\max\limits_{1\le k\le n} x_{ik} - \min\limits_{1\le k\le n} x_{ik}}, \qquad i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n \tag{10}$$

3.2 Univariate LSTM-based forecasting modelling

Basic LSTM algorithm-based univariate model is first developed to realise ELF by learning and capturing the relationship between the historical load data and future load data during the training phrase. However, it is worth mentioning that the prediction is conducted with the input feature selection during test. Real-time hourly electricity load can be recorded because of the development and installation of smart meters, which makes it possible to predict the current hourly electricity load with the previous hourly measurement. Therefore, historical load will be directly employed to predict the model outputs and update the current network. In other words, the test input in LSTM is defined as the previous load measurement at time step $k-1$ and thus electricity load at time step k can be predicted by the relationship that has been acquired.

For the case that electricity load requires to be forecasted a whole day in advance and thus the historical load measurements during the test process is not available. Instead of historical measurements, the prediction values should be fed into the test input, namely the model prediction output at previous time step $k-1$ is used as the input at time step k to predict the output at time step k.

To facilitate the presentation of experimental results, H_LSTM and P_LSTM are used to represent the univariate LSTM models with historical data and prediction data fed into the test input, respectively.

3.3 Multi-variable LSTM-based forecasting modelling

The accuracy of univariate LSTM model only with load information requires improved since the electricity load is subject to external influencing factors, which have a nonlinear relationship with the load. Temperature is recognised as the most important influencing variable, which directly affects the electric energy consumption generated from the electrical equipment such as air conditioning. To further demonstrate the correlation between the electricity load and temperature, hourly electricity load data and temperature data employed in the experiments are plotted in Figure 3, where the variation trend of power consumption is consistent with the temperature (circled portion of the figure). In order to clearly demonstrate the distribution characteristics of the training data, the probability density distributions of the load data and the temperature data are plotted in Figure 4 in this study, so that the correlation between the two can be visually compared. Similar to the single-layer LSTM model, the multilayer LSTM model also relies on the temperature and load data from the previous time step, as well as the temperature information from the current time step, to predict the load at the current time step.

Since the previous prediction outputs, other than the actual measurements, are used as the test input, univariate P_LSTM model inevitably makes poorer performance in day-ahead ELF than H_LSTM model, and it may even fail in day-ahead prediction. To achieve a more reliable day-ahead prediction accuracy, based on the above correlation analysis, an advanced multi-variable LSTM model is correspondingly developed by incorporating the temperature variable into the network input. The flowchart of the multi-variable LSTM model is presented in Figure 5. For similarly presentation purpose, H_LSTM_T and P_LSTM_T are used to represent the multi-variable forecasting models considering the temperature factor.

Figure 3 Hourly electricity load data and temperature data (see online version for colours)

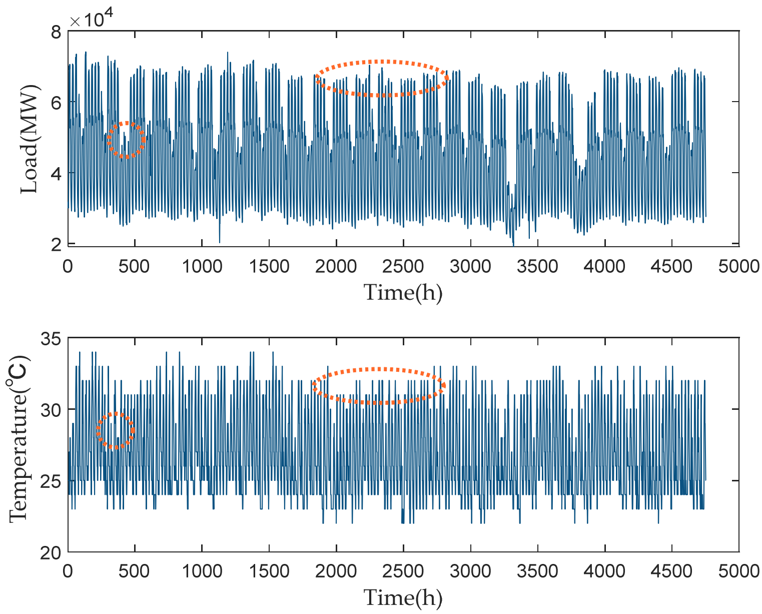


Figure 4 Probability density plots of load data and temperature data

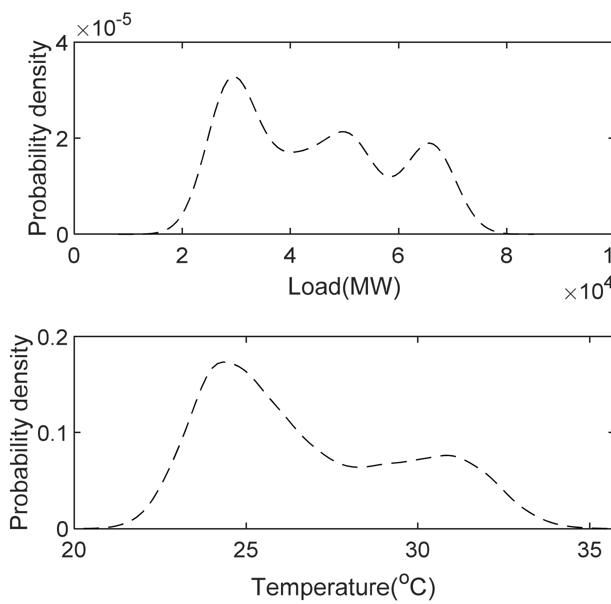
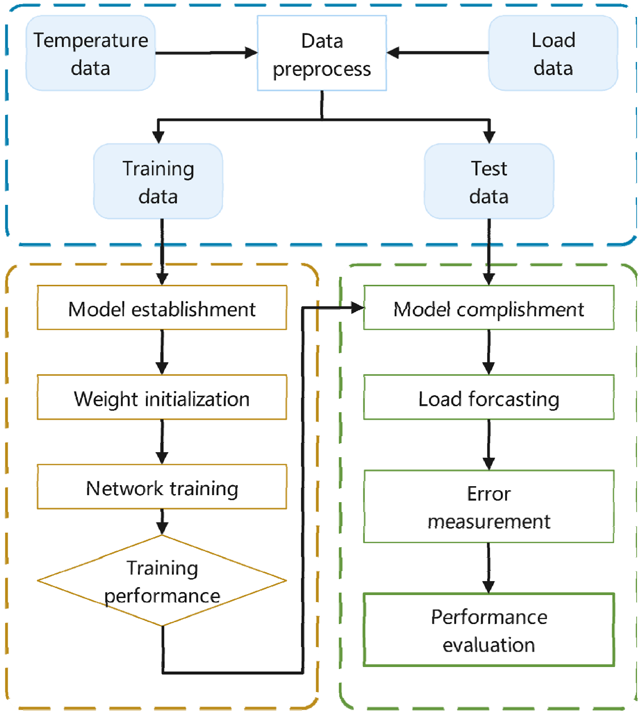


Figure 5 Flowchart of the multi-variable LSTM model (see online version for colours)



3.4 Double-layer LSTM-based forecasting modelling

As previous mentioned, the inclusion of temperature information in multi-variable LSTM model already overcome the deficiency of univariate model and can theoretically achieve more reliable accuracy in day-ahead ELF. Furthermore, the next meaningful and valid question is if the performance of the multi-variable model can be further improved by using a more complex network structure. For addressing this issue, the multi-variable LSTM model is further extended to be of double layers for each cell and thus the double-layer LSTM model with higher computational cost is developed.

4 Discussion

To investigate the feasibility and effectiveness of the basic and advanced LSTM algorithms using different model inputs and structures, several experiments based on the same sampling dataset are conducted to present the forecasting results and the corresponding evaluation indicators have been listed. For demonstrating the superiority of the proposed LSTM models, commonly-used traditional BPNN model is applied to be compared with.

4.1 Data description

The experimental data including hourly temperature data and electricity load data is collected from the power supply company of the City of Johor in Malaysia. The hourly electricity load data and temperature data employed in this paper are plotted in Figure 3. Hourly data of previous 198 days is selected as the training dataset and the following day is treated as the forecasted day to test the performance of the proposed LSTM models with different model inputs and structures, which are programmed on MATLAB.

4.2 Evaluation criteria

To intuitively compare the forecasting values of the proposed models with the measurements and quantitatively evaluate the performance, root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), integrated absolute error (IAE) and standard deviation (SD) are introduced and defined as follows:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N}(ypred - ytest)^2}{N}} \tag{11}$$

$$MAE = \frac{\sum_{t=1}^{N}\left|ypred - ytest\right|}{N} \tag{12}$$

$$MAPE = \frac{\sum_{t=1}^{N}\left|ypred - ytest\right| / ytest}{N} \times 100\% \tag{13}$$

$$IAE = \int_{t_1}^{t_n}\left|ypred - ytest\right| dt \tag{14}$$

$$SD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(ypred - ymean)^2} \tag{15}$$

where *ypred* and *ytest* are the predicted and true values of the test set, respectively N represents the forecasting steps. *ymean* represents the mean of the true values.

4.3 Experiment results

The hyperparameter settings for BPNN and LSTM are shown in Table 1. In this study, in order to perform a comparative analysis of univariate inputs versus multivariate inputs, we ensured that the hyperparameters of the BPNN and BPNN_T networks were kept the same. the number of neurons in the BPNN network was set to 128, the learning rate was set to 0.001, and the dropout layer parameter was set to 0.2, and the network was updated using Adam's optimiser. Similarly, for the LSTM network, we kept the same network parameter settings as for the LSTM_T, i.e., the hidden layer was set to 2 layers, the number of neurons in each layer was set to 60, and the dropout layer parameter was also set to 0.2. With this setup, we ensured that the two networks had the same initial conditions during the training and testing process, which allowed for a more comparable comparison of the results.

Table 1 Hyperparameters of BP and LSTM

Network	*Hyperparameter*	*Value*
BPNN	[number of neurons, learning rate, dropout and optimiser]	[128 0.001 0.2 adam]
LSTM	[number of hidden layers, number of neurons, dropout]	[1 100 0.2]
BP_T	[number of neurons, learning rate, dropout and optimiser]	[128 0.001 0.2 adam]
LSTM_T	[number of hidden layers, number of neurons, dropout]	[2 100 0.2]

LSTM is considered as the powerful prediction tool for the day-ahead ELF. To validate the superiority of LSTM model over the commonly-used conventional neural networks, BPNN model for ELF is employed to be compared with. For the fair comparison, all the proposed forecasting models are trained with the same dataset and predict the hourly electricity load of the same day. For the basic and advanced LSTM models with historical data fed into the test input, the forecasting results are shown in Figure 6 and the calculated evaluation indicators are listed in Table 2.

Table 2 Forecasting error of several ELF models with historical data fed into the test input

	RMSE	*MAE*	*MAPE*	*IAE*	*SD*
H_LSTM	1,437.11	1,069.89	2.34	25,676.63	15,535.98
H_BPNN	4,607.43	3,485.78	8.14	83,658.65	15,135.14
H_LSTM_T	1,178.55	876.58	1.96	21,038.00	14,590.92
H_BPNN_T	2,661.35	2,198.00	5.51	52,752.02	15,080.78

By comparing the prediction accuracies of the basic H_LSTM model and the advanced H_LSTM_T model, it can be clearly seen that the H_LSTM model shows satisfactory accuracy in the ELF prediction for each hour of the day before the present day, which demonstrates that the proposed LSTM model has a strong prediction capability for time series data. However, when temperature information is added as another input to the advanced H_LSTM_T model, the prediction errors are significantly reduced for both RMSE and MAE and MAPE, indicating that the advanced multivariate model significantly improves the output of the basic H_LSTM model and achieves more reliable prediction results. Specifically, the RMSE, MAE, and MAPE values were reduced by 17.99%, 18.07%, 16.24%, and 18.06% and 6.08%, respectively. In addition, the traditional BPNN model produced the largest prediction error regardless of whether the basic or advanced algorithms were used, indicating the limitations of the BPNN model in predicting sequential data.

For the basic and advanced LSTM models with prediction data fed into the test input, Figure 7 and Table 3 express the forecasting results of several models in detail. A similar conclusion can be obtained according to the analysis of previous paragraph. However, it is worth noting that the basic P_LSTM model only with prediction load information during test can achieve satisfactory hour-ahead prediction but fails in the day-ahead prediction, while the multi-variable P_LSTM_T model offers competitive advantage in day-ahead ELF because of the inclusion of temperature information. Specially, the value of RMSE, MAE and MAPE are significantly reduced by 62.07%, 59.77%, 61.88%, 59.76%, and 9.91% respectively, similarly indicating that the multi-variable LSTM model is more robust and enables an accurate day-ahead prediction.

Figure 6 Forecasting results of several ELF models with historical data fed into the test input (see online version for colours)

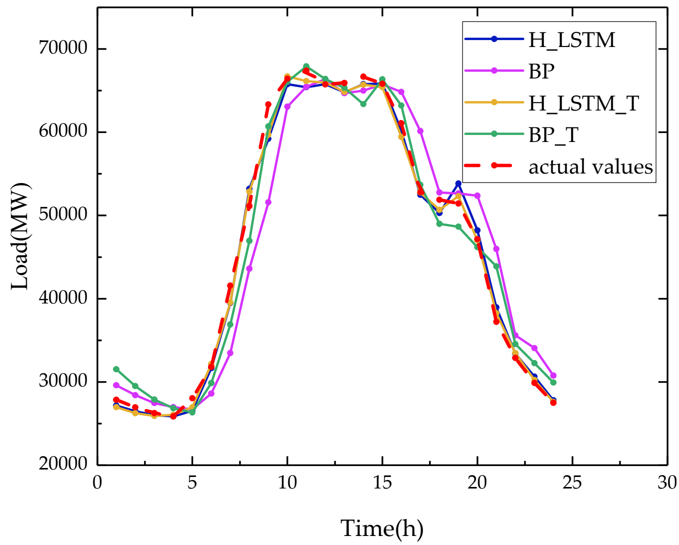


Table 3 Forecasting error of several ELF models with prediction data fed into the test input

	RMSE	*MAE*	*MAPE*	*IAE*	*SD*
P_LSTM	11,392.54	8,827.39	21.30	211,857.25	12,181.80
P_BPNN	19,853.76	15,958.55	46.86	383,005.22	12,701.15
P_LSTM_T	4,320.82	3,551.10	8.12	85,247.96	10,974.42
P_BPNN_T	6,954.59	5,151.33	12.10	123,631.86	11,878.01

Next experiment that is worth discussing is if the effectiveness of the more complex double-layer LSTM model can be verified. Since the multi-variable H_LSTM_T model performs excellently in day-ahead ELF, the multi-variable P_LSTM_T model is further extended to be of a more complex double-layer structure. Table 4 and Figure 8 confirm the effectiveness of the proposed double-layer LSTM model. Compared with P_LSTM_T model, the value of RMSE, MAE, MAPE, IAE and SD are reduced by 14.17%, 20.06%, 21.43%, 20.08%, and 6.53% respectively. As shown in Figure 9, when comparing the performance of different LSTM models, it can be found that both the bilayer LSTM and multivariate LSTM outperform the univariate single-layer LSTM in several evaluation metrics. Specifically, in terms of the root-mean-square error (RMSE), the bilayer LSTM reduces by 67.45% compared to the univariate single-layer LSTM, whereas the multivariate LSTM reduces by 14.18%. In terms of mean absolute error (MAE), the reduction is 67.84% for bilayer LSTM and 59.77% for multivariate LSTM. In addition, in terms of mean absolute percentage error (MAPE), the reduction of bilayer LSTM compared to univariate single-layer LSTM is as high as 70.04%, while multivariate LSTM also achieves a reduction of 61.8%. These results indicate that both bilayer LSTM and multivariate LSTM outperform univariate single-layer LSTM in terms of prediction performance.

Figure 7 Forecasting results of several ELF models with prediction data fed into the test input (see online version for colours)

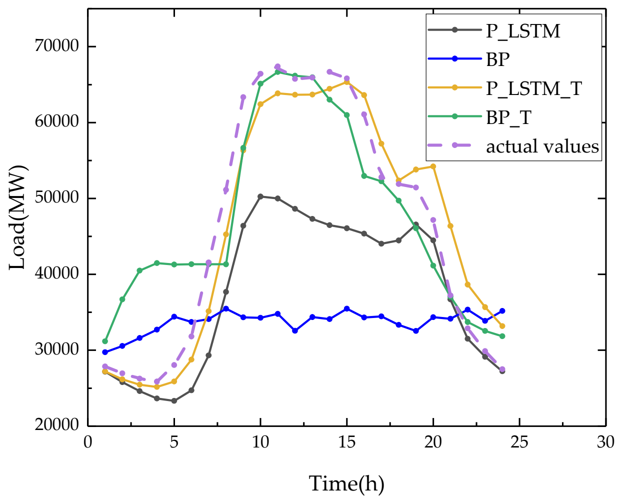
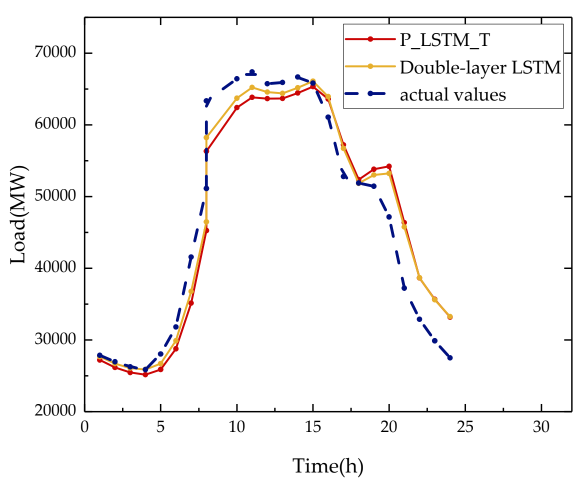


Table 4 Forecasting error of several ELF models.

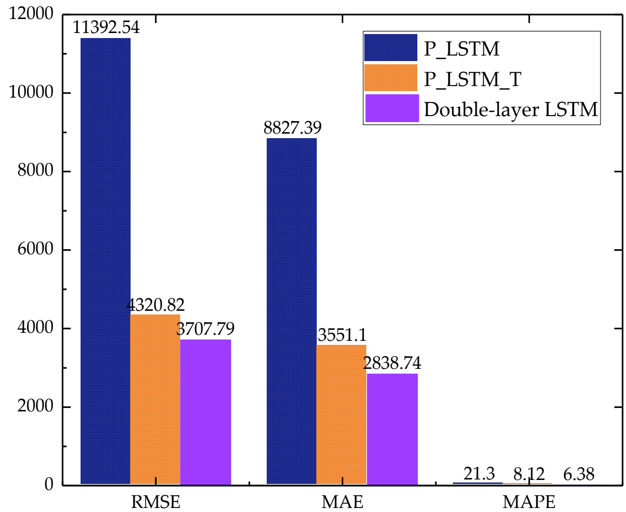
	RMSE	*MAE*	*MAPE*	*IAE*	*SD*
P_LSTM_T	4,320.82	3,551.10	8.12	85,247.96	14,974.42
Double-layer LSTM	3,707.79	2,838.74	6.38	68,129.76	13,996.31

Figure 8 Forecasting results of several ELF models (see online version for colours)



On the basis of above all comparisons of the forecasting results generated from the basic and advanced LSTM algorithms using different model inputs and structures, the superiority of LSTM model over the traditional BPNN model has been confirmed and the advancement of the multi-variable LSTM model is also validated by realising the reliable day-ahead prediction, not just hour-ahead prediction. Additionally, the error statistics shows a slight improvement of double-layer LSTM model.

Figure 9 Forecasting error of P_LSTM_T and double-layer LSTM models (see online version for colours)



4.4 Discussion

The basic and advanced LSTM algorithms using different model inputs and structures are proposed in this paper and the advancement and effectiveness of the proposed methods have been confirmed. However, some limitations on the proposed ELF methods require to be discussed here and will be used to guide where the future efforts can be devoted to.

1 Dataset scale requires to be enlarged to reflect the effect of photovoltaic power generation and wind power generation on electricity load due to the increasing penetration of renewable energy. Meanwhile, other meteorological factors such as humidity and wind speed can be taken into account.

2 In spite of the extensive literature on the single-layer LSTM model for ELF, research on the more complex multi-layer LSTM is relatively restricted and remains open.

5 Conclusions

LSTM is applied as the powerful tool for the day-ahead hourly ELF. The basic and advanced LSTM models using different network inputs and structures are proposed in this paper. LSTM-based forecasting modelling applied in ELF mainly includes three phrases. A univariate LSTM model where the test input is separated into the historical load and prediction load is first developed and ELF is achieved based on the trained relationship between the historical load and future load. Experimental results show that the basic H_LSTM model where the historical data is used in the test input realises the satisfactory day-ahead prediction, while the other P_LSTM model only achieves satisfactory hour-ahead prediction but fails in the day-ahead prediction. Second, based on the correlation analysis between the temperature and electricity load, a multi-variable LSTM model incorporating the temperature information into the network input is proposed, which achieves lower prediction error whether RMSE or MAE and MAPE are considered and enables an accurate day-ahead prediction for the basic P_LSTM model. Therefore, the advancement of multi-variable LSTM model can be verified. In addition, the superiority of LSTM model is confirmed by comparison with conventional BPNN model. Finally, a further significant finding is that the proposed double-layer LSTM model with higher computational cost shows the effectiveness in slightly improving the forecasting accuracy.

6 Subsequent studies

Subsequent studies will consider the collection of more comprehensive data, including meteorological factors such as humidity and cloudiness, to enable more refined load forecasting.

Acknowledgements

This work was supported by State Grid Jiangsu Electric Power Co., LTD. Science and Technology Project under grant J2023081.

References

Barman, M., Choudhury, N.B., Behera, S. et al. (2017) 'A fuzzy logic controller based mid-term load forecasting with renewable penetration in Assam, India', *ADBU Journal of Engineering Technology (AJET)*, Vol. 6, No. 3, p.006031207.

Baskan, D.E., Meyer, D., Mieck, S. et al. (2023) 'A scenario-based model comparison for short-term day-ahead electricity prices in times of economic and political tension', *Algorithms*, Vol. 16, No. 4, p.177.

Borrero, J.D. and Mariscal, J. (2023) 'Elevating univariate time series forecasting: innovative SVR-empowered nonlinear autoregressive neural networks', *Algorithms*, Vol. 16, No. 9, p.423.

Chemali, E., Kollmeyer, P.J., Preindl, M. et al. (2018) 'Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries', *IEEE Transactions on Industrial Electronics*, Vol. 65, No. 8, pp.6730–6739.

Chen, M., Yuan, J., Liu, D. et al. (2024) 'Effect of the light environment on image-based SPAD value prediction of radish leaves', *Algorithms*, Vol. 17, No. 1, p.16.

De Felice, M., Alessandri, A., Catalano, F. et al. (2015) 'Seasonal climate forecasts for medium-term electricity demand forecasting', *Applied Energy*, Vol. 137, pp.435–444.

Douglas, A.P., Breipohl, A.M., Lee, F.N. et al. (1998) 'Risk due to load forecast uncertainty in short term power system planning', *IEEE Transactions on Power Systems*, Vol. 13, No. 4, pp.1493–1499.

Galka, A. (2024) 'The weighted least-squares approach to state estimation in linear state space models: the case of correlated noise terms', *Algorithms*, Vol. 17, No. 5, p.194.

Guo, Y., Nazarian, E., Ko, J. et al. (2014) 'Hourly cooling load forecasting using time-indexed ARX models with two-stage weighted least squares regression', *Energy Conversion & Management*, Vol. 80, No. 2, pp.46–53.

Hongku, M. (2015) 'Research on ARIMA model parallelization in load prediction based on R language', *Power System Technology*, Vol. 39, No. 11, pp.3216–3220.

Jiao, R., Zhang, T., Jiang, Y. et al. (2018) 'Short-term non-residential load forecasting based on multiple sequences LSTM recurrent neural network', *IEEE Access*, Vol. 6, pp.59438–59448.

Kong, W., Dong, Z.Y., Jia, Y. et al. (2019) 'Short-term residential load forecasting based on LSTM recurrent neural network', *IEEE Transactions on Smart Grid*, Vol. 10, No. 1, pp.841–851.

Kramar, V. and Alchakov, V. (2023) 'Time-series forecasting of seasonal data using machine learning methods', *Algorithms*, Vol. 16, No. 5, p.248.

Lecun, Y., Bengio, Y. and Hinton, G. (2015) 'Deep learning', *Nature*, Vol. 521, No. 7553, p.436.

Li, L.J. and Huang, W. (2014) 'A short-term electricity load forecasting method based on BP neural network', *Applied Mechanics & Materials*, Vols. 494–495, pp.1647–1650.

Ma, W., Fang, S., Liu, G. et al. (2017) 'Modeling of district load forecasting for distributed energy system', *Applied Energy*, Vol. 204, pp.181–205.

Mocanu, E., Nguyen, H.P., Gibescu, M. et al. (2016) 'Deep learning for estimating building energy consumption', *Sustainable Energy, Grids and Networks*, Vol. 6, No. 6, pp.91–99.

Monteiro, C., Ramirez-Rosado, I.J., Fernandez-Jimenez, L.A. et al. (2018) 'New probabilistic price forecasting models: Application to the Iberian electricity market', *International Journal of Electrical Power & Energy Systems*, Vol. 103, pp.483–496.

Ranaweera, D.K., Karady, G.G., Farmer, R.G. et al. (1997) 'Economic impact analysis of load forecasting', *IEEE Transactions on Power Systems*, Vol. 12, No. 3, pp.1388–1392.

Sadaei, H.J., Silva, P.C., Guimaraes, F.G. et al. (2019) 'Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series', *Energy*, Vol. 175, pp.365–377.

Sherstinsky, A. (2020) 'Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network', *Physica D: Nonlinear Phenomena*, Vol. 404, p.132306.

Shi, H., Xu, M., Li, R. et al. (2018) 'Deep learning for household load forecasting – a novel pooling deep RNN', *IEEE Transactions on Smart Grid*, Vol. 9, No. 5, pp.5271–5280.

Sudheer, G. and Suseelatha, A. (2015) 'Short term load forecasting using wavelet transform combined with Holt-Winters and weighted nearest neighbor models', *International Journal of Electrical Power & Energy Systems*, Vol. 64, pp.340–346.

Sutskever, I., Vinyals, O. and Le, Q.V. (2014) *Sequence to Sequence Learning with Neural Networks. Advances in Neural Information Processing Systems*, arXiv:1409.3215.

Wang, F., Xuan, Z., Zhen, Z. et al. (2020) 'A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework', *Energy Conversion and Management*, Vol. 212, p.112766.

Wang, J., Jia, R., Zhao, W. et al. (2024) 'Application of the largest Lyapunov exponent and non-linear fractal extrapolation algorithm to short-term load forecasting', *Chaos Solitons & Fractals*, Vol. 45, No. 9, pp.1277–1287.

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

Xu, L., Li, C., Xie, X. et al. (2018) 'Long-short-term memory network based hybrid model for short-term electrical load forecasting', *Information – An International Interdisciplinary Journal*, Vol. 9, No. 7, p.165.

Zhang, J., Yan, J., Infield, D. et al. (2019) 'Short-term forecasting and uncertainty analysis of wind turbine power based on long short-term memory network and Gaussian mixture model', *Applied Energy*, Vol. 241, pp.229–244.

Zhang, X., Wang, J., Zhang, K. et al. (2017a) 'Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by cuckoo search algorithm', *Electric Power Systems Research*, Vol. 146, No. 2, pp.270–285.

Zhang, Y., Zheng, Y., Meng, K. et al. (2017b) 'Flexible operation planning scheme considering wind power generation forecasting uncertainties', *Electric Power Components and Systems*, Vol. 45, No. 5, pp.465–475.

Zhang, X., Wang, J., Zhang, K. et al. (2024) 'Multi-criteria decision support system for automatically selecting photovoltaic sets to maximise micro solar generation', *Algorithms*, Vol. 17, No. 7, p.274.