CIT. Journal of Computing and Information Technology, Vol. 31, No. 3, September 2023, 151–166
doi: 10.20532/cit.2023.1005730

Refining Short-Term Power Load Forecasting: An Optimized Model with Long Short-Term Memory Network

Sile Hu[1, 2], Wenbin Cai[3], Jun Liu[3], Hao Shi[4] and Jiawei Yu[4]

[1]Zhejiang University, Hangzhou, Zhejiang, China
[2]Inner Mongolia Power (Group) Co., Ltd., Hohhot, Inner Mongolia, China
[3]Inner Mongolia Power Electric Power Economic and Technological Research institute, Hohhot, Inner Mongolia, China
[4]Beijing Tsintergy Technology Co., Ltd., Beijing, China

Short-term power load forecasting involves the stable operation and optimal scheduling of the power system. Accurate load forecasting can improve the safety and economy of the power grid. Therefore, how to predict power load quickly and accurately has become one of the urgent problems to be solved. Based on the optimization parameter selection and data preprocessing of the improved long short-term memory (LSTM) network, the study first integrated particle swarm optimization (PSO) algorithm to achieve parameter optimization. Then, combined with convolutional neural network (CNN), the power load data were processed to optimize the data and reduce noise, thereby enhancing model performance. Finally, simulation experiments were conducted. The PSO-CNN-LSTM model was tested on the GEFC dataset and demonstrated stability of up to 90%. This was 22% higher than the competing CNN-LSTM model and at least 30% higher than the LSTM model. The PSO-CNN-LSTM model was trained with a step size of 1.9×10^4, the relative mean square error was 0.2345×10^{-4}. However, when the CNN-LSTM and LSTM models were trained for more than 2.0×10^4 steps, they still did not achieve the target effect. In addition, the fitting error of the PSO-CNN-LSTM model in the GEFC dataset was less than 1.0×10^{-7}. In power load forecasting, the PSO-CNN-LSTM model's predicted results had an average absolute error of less than 1.0% when compared to actual data. This was an improvement of at least 0.8% compared to the average absolute error of the CNN-LSTM prediction model. These experiments confirmed that the prediction model that combined two methods had further improved the speed and accuracy of power load prediction compared to traditional prediction models, providing more guarantees for safe and stable operation of the power system.

ACM CCS (2012) Classification: Computing methodologies → Machine learning → Machine learning algorithms → Regularization

Applied computing → Operations research → Forecasting

Keywords: Genetic Algorithm, BP Neural Network, Short Term Power Load, Prediction Model

1. Introduction

Power load forecasting (PLF) is an important component of power system operation. Accurate load forecasting can provide a reliable basis for actual operation and ensure the safe and stable operation of the power system. Therefore, it should facilitate efficient and accurate predictions for short-term power load (PL). This prediction can help power companies and energy management departments make wiser decisions in scheduling and arranging power production, to ensure the power system operates safely and stably [1–2]. In addition, PLF can help users make more reasonable electricity consumption plans to reduce energy waste and cost expenditures. However, PLF faces many challenges, such as large load fluctuations and the significant impact of environmental factors on predictive data, which pose great challenges to PLF [3–5]. Since the 1970s, experts have extensively researched short-term PLF and established various forecasting models. Short-term PLF can be classified into two types: classical and artificial intelligence. Classical PLF mainly includes time-series, regression analysis, and grey prediction methods. The modeling of these methods is relatively simple, and the prediction accuracy is relatively low. Artificial intelligence-based power prediction mainly includes support vector machines, artificial neural networks, long short-term memory (LSTM), *etc.* These prediction methods may suffer from issues such as local optimal solutions and missing historical data, which can negatively impact prediction accuracy. Despite some limitations, LSTM is widely used in the PLF due to its strong ability to process sequence data. Therefore, improving LSTM has become one of the current research hotspots. To address issues such as local optimal solutions, data redundancy, and parameter optimization in LSTM, particle swarm optimization (PSO) and convolutional neural network (CNN) are used to improve the original model, and a high-precision PSO-CNN-LSTM PLF model is designed. This research is expected to further improve the accuracy and stability of short-term PLF, providing guarantees for the safe and stable operation of the power system. The main innovation of this paper is centered on the dataset and processing methods of PLF. First, the data is processed to remove redundancies and compensate for missing information. Then, LSTM is used for PLF. To improve its prediction accuracy, a combined model, namely PSO-CNN-LSTM, is constructed. The model combines PSO's optimization ability with CNN's local feature extraction, addressing the limitations of CNN-LSTM and PSO-LSTM models. The research content mainly includes three parts. Firstly, it introduces the research of domestic and foreign experts and scholars on PLF, outlining the application of PSO and LSTM. Secondly, an overview of the process of short-term PLF is provided and each step is analyzed, with a focus on the construction of the PSO-CNN-LSTM prediction model. The third part is to conduct experimental verification and analysis on the constructed PSO-CNN-LSTM prediction model's performance. Nowadays, short-term PLF is an important means for a power grid department to predict power demand in the coming days. Its accuracy directly affects the safety, reliability, and economy of power system operation. Therefore, the accuracy and effectiveness of short-term PLF are extremely important. Therefore, many researchers have conducted relevant research on short-term PLF. Chafi and Afrakhte proposed a method for short-term PLF using neural networks and PSO to reduce the impact of parameters on performance, using networks with optimized parameters for prediction. This method was tested on the Iranian power grid. These results reflected that the proposed method could accurately predict PL [6]. Traditional load forecasting methods had limitations due to insufficient or missing data. M. Gilanifar *et al.* proposed an improved MTL algorithm for this. It assumed similar effects of environmental and traffic conditions on electricity consumption to improve short-term load forecasting methods. The effectiveness of this method has been demonstrated through real cases. Compared with other MTL methods, the new method was significantly superior to traditional prediction methods [7]. A. Tudose *et al.* put forward a CNN-LSTM forecasting model to address the issues of feature extracting and predicting accuracy. This model fused feature vectors as input to LSTM and used it for forecasting. Furthermore, the model was applied to predict the actual load data. These experiments confirmed that it had higher predicting accuracy [8]. A. Io *et al.* proposed a migration prediction method based on data correlation to accurately predict the PL requirements of buildings. They obtained data from different regions of the world to achieve more successful predictions with limited data. The application results in actual energy systems confirmed that this method had significant predictive advantages [9]. R. Patel *et al.* explored the resources required to operate power plants, rotational reverse planning, generator scheduling, and other applications. They used well-defined machine learning methods called recurrent neural network (RNN) and LSTM to predict future PL. These results confirmed that this method could accurately predict future PLs [10].

In PLF, the performance research of LSTM and PSO has become a focus of attention. Many scholars have focused on the research of wideband and narrowband hybrid control algorithms and have achieved many remarkable research results. A. Amiri *et al.* used PSO to adjust fuzzy parameters to improve the adaptability and flexibility of the control chart. These results confirmed that this method had significant advantages and improvement effects [11]. Chafi and Afrakhte proposed neural network models to load forecasting and used PSO to optimize the parameters of the neural network. These results confirmed that this method had good accuracy and predictive ability in short-term load forecasting [6]. W. Lu *et al.* constructed a comprehensive model to improve the accuracy of stock price prediction. By training historical stock price data, this model could learn patterns and trends in price sequences and be used to predict future changes in stock prices [12]. Liang and Zhang collected a large amount of ship trajectory information using AIS data and applied it to LSTM for training and prediction. To improve the performance of LSTM, they also used PSO to optimize its parameter settings. These results confirmed that this method could accurately predict the future trajectory of ships [13]. To greatly improve the accuracy and precision of short-term PLF, R. Wang had improved the similar day category screening method based on a time period neural network model. They divided the load into 7 time periods using the regional load characteristics. Based on real-time meteorological data, the prediction model could provide load values for the predicted day. These results confirmed that the prediction accuracy of different types of days could reach over 96% [14].

In summary, although many experts have designed and developed various models to improve the accuracy and precision of short-term PLF, research on short-term PLF based on improved LSTM is still quite rare. To further improve the accuracy of PLF, a short-term PLF based on PSO-CNN-LSTM is proposed, which utilizes the advantages of improved LSTM in prediction and provides more reference experience for improving the accuracy and effectiveness of short-term PLF.

2. Construction of PLF Based on PSO-CNN-LSTM

This section mainly analyzes the process of building a short-term PL and provides explanations for each step of the process. Then the construction process and operational methods of PSO-CNN-LSTM are analyzed. Traditional LSTM is first introduced, and then PSO is introduced to address the shortcomings of LSTM parameter optimization. On this basis, CNN is introduced to preprocess PL data to improve the final predicting accuracy.

2.1. The Construction Process and Evaluation Method of PLF

PLF is a key task in the operation and planning of the power system, which can estimate the electricity consumption for some point in time [15]. Accurate load forecasting can help power companies optimize generation plans, dispatch energy resources, and improve power supply reliability, making it is also crucial for participants in the electricity market. Figure 1 shows the general process of PLF.

In Figure 1, PLF first needs to determine an appropriate prediction target, *e.g.* selecting PL for the next month as the prediction target. Subsequently, a suitable prediction method is determined. Then, historical data of the corresponding detection targets are collected and preprocessed. Afterwards, a prediction model is established, which is LSTM-PLF. Finally, the LSTM-PLF is improved using CNN and PSO, and the prediction results are obtained through example validation. In PLF, Figure 2 shows the process of data preprocessing [16].

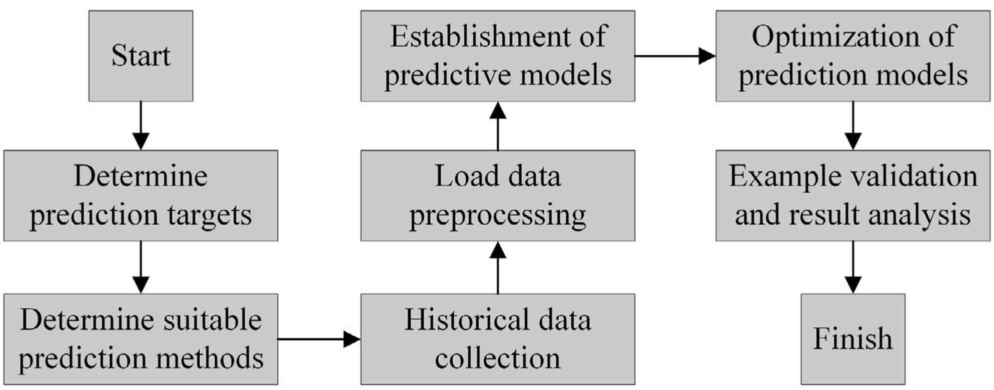


Figure 1. Flow chart of power load forecast.

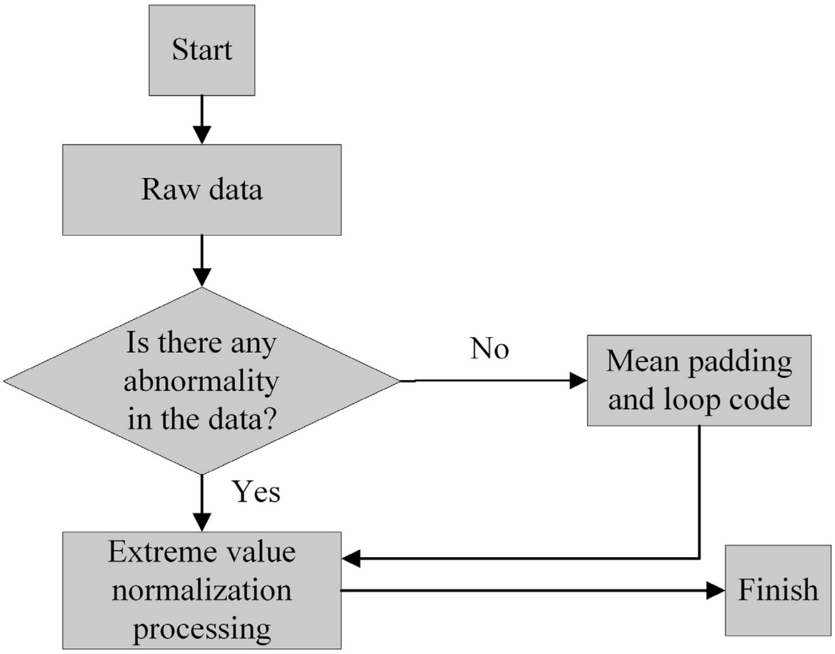


Figure 2. Power load data preprocessing flowchart.

For the selection of PL raw data, one year's PL data and meteorological characteristic data from a certain region are selected. For the missing data, the study uses cyclic code and mean filling methods to ensure its integrity, as presented in equation (1).

$$\bar{X} = \frac{\sum_{i=1}^{n} \gamma X_i}{n_i} \tag{1}$$

In equation (1), X_i is the data value. n_i is the amount of data. γ is the basis for determining whether the data are filled in. $\gamma = 1$ indicates that it needs to be filled in. $\gamma = 2$ indicates that no padding is required. After supplementing the missing data, the dimensional differences of the added data should be removed, and, therefore, a normalization method should be used to process the data [17]. Considering that PL and meteorological data are single-dimensional data, the experiment selects the maximum and minimum normalization to process the data in equation (2).

$$X_s = \frac{X - X_{\min}}{X_{\max} - X_{\min}}(\max - \min) + \min \tag{2}$$

In equation (2), $X_{\max}$ and $X_{\min}$ are the maximum and minimum values within the set range, respectively. max and min are 0.9 and 0.1, respectively. X_s is the normalized value. Taking European PL data from January 1998 to December 1998 as an example, the 3D grid graph module is used to display the changes in data filling and normalization processing in Figure 3.

In Figure 3(a), the cyclic code and mean filling method fill in the missing data to form a coherent data line. After preprocessing the data, the construction of a prediction model begins. Considering the temporal nature of the collected data, LSTM was selected as the basic prediction model, and PSO and CNN were used to improve the LSTM prediction model. The general evaluation indicators for constructed prediction models are Sum of Squares Error (SSE), Mean Square Error (MSE), Root Mean Square Error (RMSE), standard deviation, and Mean Absolute Error (MAE) [17]. Equation (3) shows the calculation method for variance.

$$SSE = \sum_{i=1}^{n} \omega_i (y_i - y_j)^2 \tag{3}$$

In equation (3), n represents the number of samples. y_i is the actual data. y_j is the fitted data, and

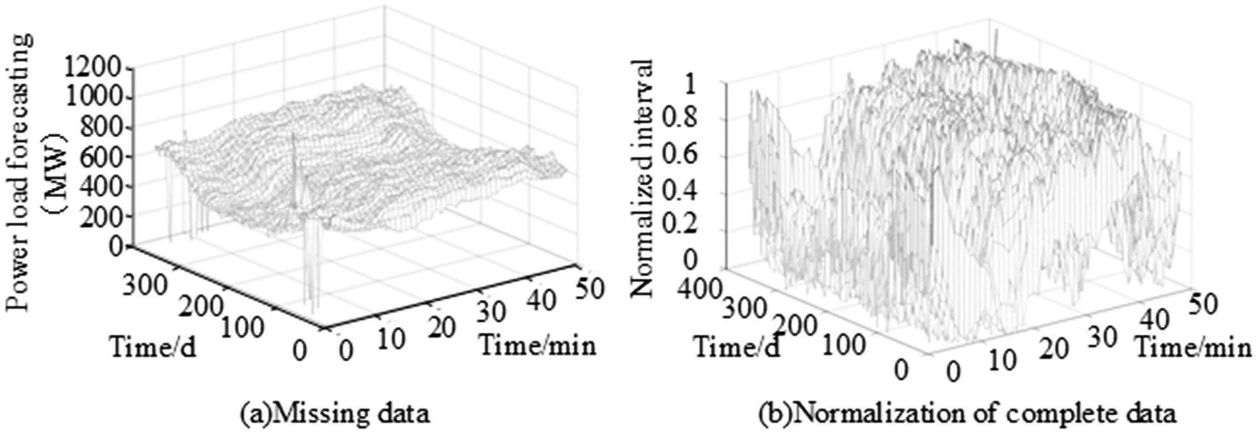


Figure 3. Comparison of 3D grid images before and after data processing.

$\omega i \geq 0$. The closer *SSE* approaches 0, the better the fitting degree of the prediction model and the higher the prediction accuracy [18]. Equation (4) is MSE.

$$MSE = \frac{\sum_{i=1}^{n} \omega_i (y_i - y_j)^2}{n} \tag{4}$$

Similarly, the closer *MSE* approaches 0, the higher the model accuracy and the smaller the error. Equation (5) defines RMSE.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \omega_i (y_i - y_j)^2}{n}} \tag{5}$$

Like *MSE* and *SSE*, the closer *RMSE* approaches 0, the higher the accuracy. Equation (6) represents MAE.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| y_i - y_j \right| \tag{6}$$

MAE can accurately reflect the model's prediction error, and a small difference indicates a higher model accuracy. Equation (7) represents the standard deviation.

$$\begin{cases} \sigma = \sqrt{\dfrac{1}{n} \sum_{i=1}^{n} (y_i - \mu)^2} \\ \mu = \dfrac{1}{n} (y_1 + y_2 + \ldots y_n) \end{cases} \tag{7}$$

2.2. Construction and Improvement Strategy of PLF Based on LSTM

LSTM is an upgraded form of RNN that can selectively receive data information and quickly process and save it. Figure 4 shows the complete LSTM.

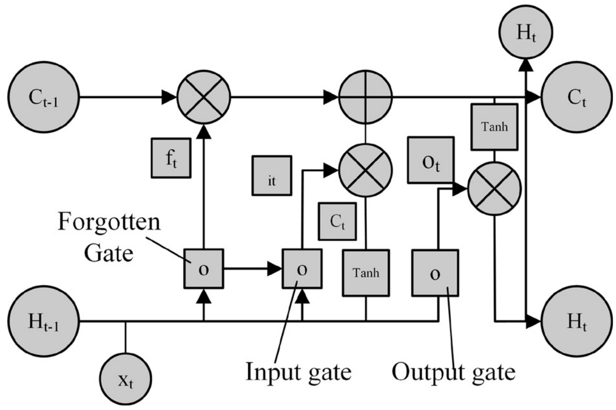


Figure 4. Schematic diagram of LSTM structural model structure.

According to Figure 4, LSTM mainly consists of three parts. The input and output gates are responsible for receiving and transmitting information between memory units, respectively. Furthermore, the forgetting gate is utilized to forget the information in the memory unit by blocking the transmission of information layer by layer. LSTM is mainly divided into three stages during runtime: "forget", "selective memory", and "output". The forget stage is mainly used to delete information that is no longer needed, update the memory state, and provide space for new inputs in equation (8).

$$f_t = \sigma\left(W_{hf} h_{t-1} + W_{xf} x_t + b_f\right) \tag{8}$$

In equation (8), f_t is the forget output, W_{hf} and W_{xf} are weight matrices, and b_f is the bias term. f_t is used to determine whether the previous state value should be forgotten or memorized. The selective memory stage is determined by the activation level of the input gate and forget gate. The input gate controls the entry of new information into memory, while the forget gate controls whether old memories are retained. Equation (9) represents the output i of the output gate.

$$i_t = \sigma\left(W_{hi} h_{t-1} + W_{xi} h_t + b_i\right) \tag{9}$$

Equation (10) represents the state of selecting the memory stage.

$$\begin{cases} \overline{C}_t = \tanh\left(W_{hc} h_{t-1} + W_{xc} x_t + b_c\right) \\ C_t = (C_{t-1} \otimes f_t) \oplus (\overline{C}_t \otimes i_t) \end{cases} \tag{10}$$

In equation (10), $\overline{C}_t$ represents a temporary state, and C_t is the new unit status obtained after updating the temporary unit status. In the output stage, tanh activation function is used to change the output of the previous unit to obtain the output h_t. Equation (11) is the specific calculation.

$$o_t = \sigma(W_{xo} h_{t-1} + W_{xo} x_t + b_o) \tag{11}$$

In equation (11), h_{t-1} and x_t are the two inputs of the output gate, and σ is activation function *sigmoid*. After activating *sigmoid*, the numerical values between [0, 1] are obtained, and the magnitude of the values represents the degree of preservation and forgetting. For the information to be forgotten, its memory state will be updated to a smaller value, gradually disappearing [19]. In this way, LSTM can remember the most important information when processing sequential data and forget information that is no longer needed. Due to its characteristics, LSTM is very suitable for solving the problems of predicting time series with temporal order and predicting discontinuous time series. The general LSTM is far from meeting the requirements of PL short-term prediction and requires further optimization. There are currently two optimization approaches, one is "self-optimization", which mainly focuses on optimizing the parameters and function settings of the model itself [20]. This optimization can be achieved by repeating multiple experiments to determine the optimal parameters of the model. Another approach is "algorithm assisted optimization", which utilizes algorithms and mathematical models to improve and optimize the basic model. This study uses PSO and CNN to optimize the solution of LSTM. PSO has extremely high convergence speed and optimization ability, which can help obtain the optimal values for the two important parameters, namely, the number of hidden layer neurons and learning rate in LSTM [21]. Usually, particles in PSO update themselves through two outliers: individual outlier *pBEST* and global optimal solution *gBest*. After several iterations, the particle's own velocity and position are continuously updated to seek the optimal solution. Figure 5 shows the specific process of PSO.

In Figure 5, it is assumed that there are n particle populations $X = (X_1, X_1, ..., X_n)$ undergoing random initialization motion in D-dimensional space, and the spatial position of each particle is represented by $X = (x_{i1}, x_{i1}, ..., x_{iD})^D$. It is assumed that the velocity of the i-th particle in the particle swarm is V_i, the extreme value of the individual particle is p_i, and the extreme value of the particle population is P_g, which are calculated in equation (12).

$$\begin{cases} V_i = \left(V_{i1}, V_{i2}, ..., V_{iD}\right)^T \\ p_i = \left(p_{i1}, p_{i2}, ..., p_{iD}\right)^T \\ p_g = \left(p_{g1}, p_{g2}, ..., p_{gD}\right)^T \end{cases} \tag{12}$$

After calculating the extreme values of individual and population particles, the speed and position of the particles are updated. Equation (13) represents the update of particle position:

$$X_{id}^{k+1} = X_{id}^{k} + V_{id}^{k+1} \tag{13}$$

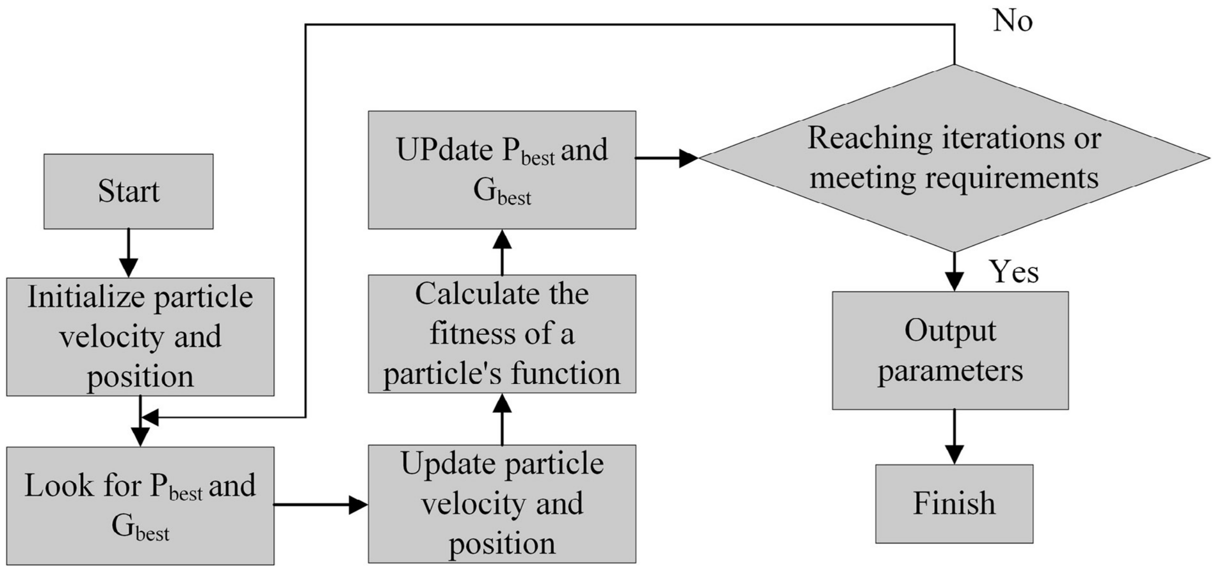


Figure 5. Particle swarm optimization algorithm process.

Equation (14) represents the update of particle velocity V_{id}^{k+1}.

$$V_{id}^{k+1} = \omega V_{id}^{k} + c_1 r_1 \left(p_{id}^{k} - X_{id}^{k}\right) + c_2 r_2 \left(p_{gd}^{k} - X_{gd}^{k}\right) \quad (14)$$

In equation (14), ω is the inertia weight, k is iteration, c_1, c_2 are non-negative learning factor constants, and r_1, r_2 are random numbers within [0, 1]. ω reflects the degree to which the model seeks the optimal solution and is calculated using adaptive methods in equation (15).

$$\begin{cases} \omega = \omega_{\min} - \dfrac{(\omega_{\max} - \omega_{\min})(f - f_{\min})}{f_{avg} - f_{\min}}, \ f \le f_{avg} \\ \omega = \omega_{\max}, \ f > f_{avg} \end{cases} \quad (15)$$

In equation (15), $\omega_{\max}$ and $\omega_{\min}$ are the maximum and minimum values of ω, respectively, f is particle's current fitness value, and $f_{\min}$ and f_{avg} are the minimum and average fitness values of all current particles. CNN can help LSTM handle longer time series and more dimensional problems. Due to the large amount of information in PL data, the model generates many useless signals during prediction. Therefore, one-dimensional CNN is used for feature extraction of PL data [22]. This can reduce the noise and instability caused by excessive data, thereby improving the accuracy of the model for short-term PLF. Considering the advantages of LSTM, PSO, and CNN, a PSO-CNN-LSTM short-term power model is constructed, as shown in Figure 6 [23].

Figure 6 shows how PSO-CNN-LSTM constructs a short-term power model. Historical PL and weather data are collected and preprocessed. CNN extracts features from the input weather data, which are then fused with the PL data to obtain a comprehensive feature representation. Then, the constructed LSTM is used to model the comprehensive features to learn long-term dependencies and sequence features in time series data. PSO is subsequently used to adjust and optimize the parameters to improve prediction accuracy. Finally, historical data is used to train PSO-CNN-LSTM, and the resulting model is used to predict PL for a future period, generating corresponding results. The optimization parameters in the operation of the PSO-CNN-LSTM short-term PLF model mainly include learning rate, number of hidden layers, number of hidden layer neurons, and loss value. The learning rate ranges from 0.01 to 0.001. To reduce model complexity, the number of layers and nodes in the model is set to the minimum, and the hidden layer is set to 2 layers. The optimization range for the number of neurons is 2n, and the loss function is shown in equation (5).

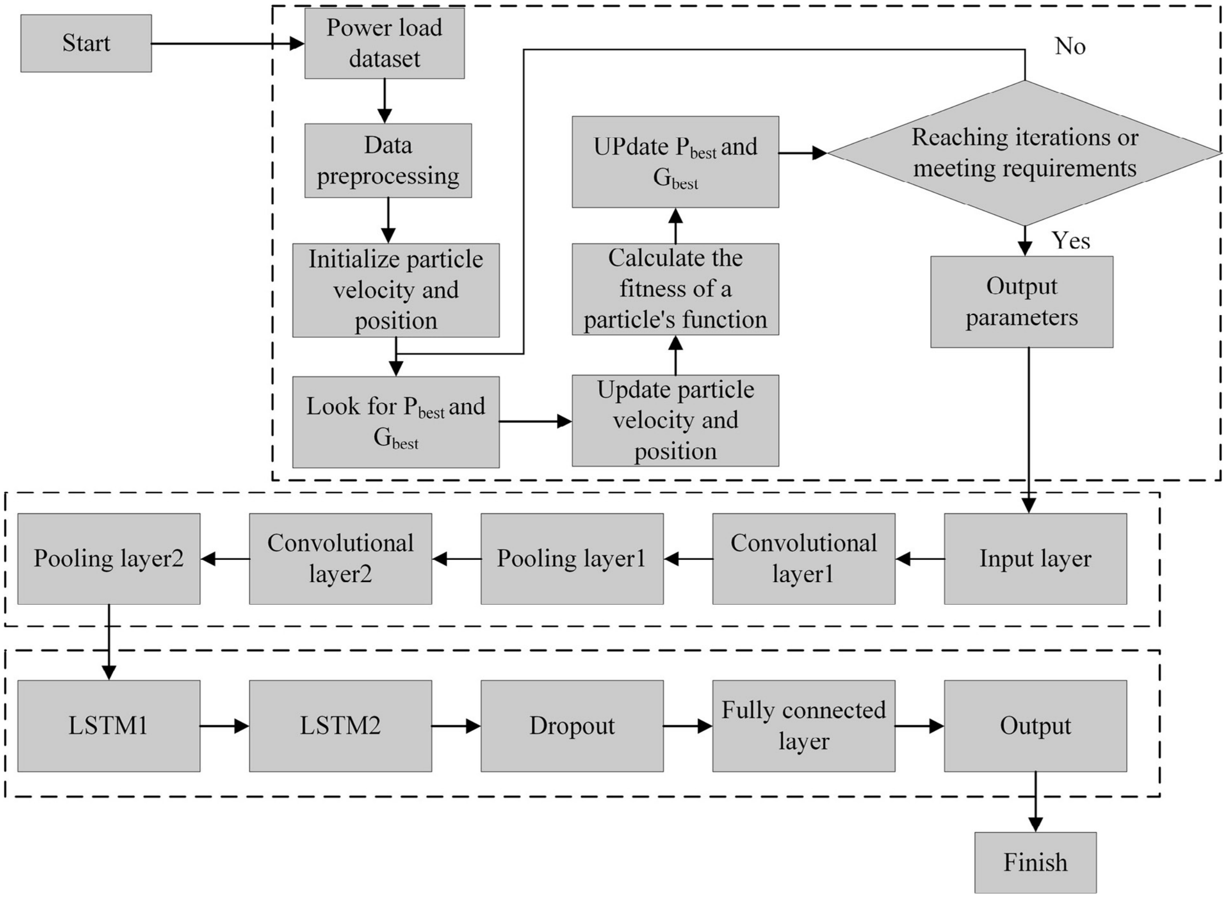


Figure 6. PSO-CNN-LSTM process.

3. Performance and Application Analysis of Short-Term PLF Based on PSO-CNN-LSTM

This section mainly analyzes the performance and application of PSO-CNN-LSTM short-term PLF. Firstly, a testing environment was constructed. Then, the stability convergence, RMSE, fitting error, Relative Error (RE), and Absolute Error (AE) of LSTM, CNN-LSTM, and PSO-CNN-LSTM were analyzed. In practical application analysis, data from different time intervals and regions were selected for testing to verify the prediction accuracy of PSO-CNN-LSTM.

3.1. Performance Analysis of Short Term PLF Based on PSO-CNN-LSTM

To verify the accuracy and effectiveness of the constructed PSO-CNN-LSTM short-term PLF, simulation experiments were conducted to determine the performance of the constructed model. Table 1 shows the computer hardware.

This study selected LSTM, CNN-LSTM, and PSO-CNN-LSTM for performance comparison. To ensure the accuracy of the final experimental results and eliminate randomness, the iterations of each algorithm were set to 1000 times. PL data were selected from January to December 2021 from the Global Energy Forecasting Competition (GEFC) and IEEE Power and Energy Society (IEEE) datasets. PL data from January to April were used as the training set and the PL data from May to December were used as the testing set to train the model. Then the performance of the model was tested. The purpose of stability testing for the model is to ensure consistent output results even with slight changes in input data, making it more robust. Figure 7 shows the results of the stability testing for the model.

In Figure 7, as the experimental iterations increased, the stability of all three algorithms gradually increased. Figure 7 (a) shows the test of model stability in GEFC dataset. As the iteration progressed, CNN-LSTM and

Table 1. Simulation experiment computer hardware.

Equipment	Model
CPU	Intel Core
Internal storage	32G
Caliche	256GBSSD
Graphics card	NVIDIA GeForce GTX 1060
Operating system	Win10

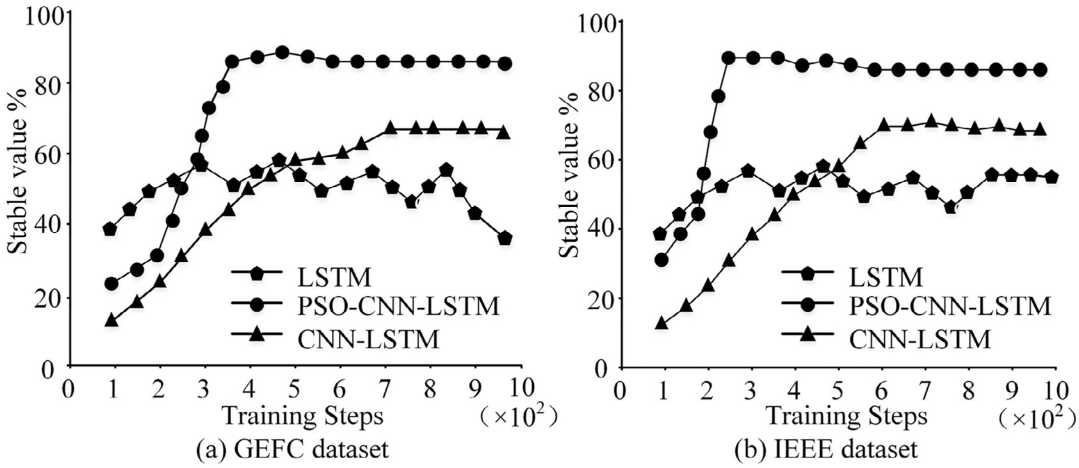


Figure 7. Stability convergence of three models.

PSO-CNN-LSTM's stability values converged. After 720 iterations, CNN-LSTM gradually stabilized with a stable value of 68%. PSO-CNN-LSTM tended to stabilize around 350 times, with a stable value of up to 90%. LSTM exhibited tortuous fluctuations, with stable values ranging from 40% to 60%. Figure 7 (b) shows the test of model stability for the IEEE dataset. As the iteration progressed, three models' stability converged. LSTM tended to stabilize after 850 iterations, with a stable value of 58%. CNN-LSTM tended to stabilize after 620 iterations, with a stable value of 70%. PSO-CNN-LSTM tended to stabilize after 200 iterations, with a stable value of 92%. Therefore, the iterations of PSO-CNN-LSTM tended to stabilize in different datasets sooner than for other approaches. So, the simulation model constructed by PSO-CNN-LSTM had better stability and higher computational efficiency. After testing the stability of the model, next step size training was performed on the three algorithms to obtain RMSE in Figure 8.

By analyzing the training step curves of three algorithm models in Figure 8, the RMSE of PSO-CNN-LSTM decreased as the training step size continued to increase. During the process, it experienced two fluctuations, one from 0 to 1000, with a rapid decrease in RMSE, and the second at around 2000 steps, with a gradual decrease. In the end, it achieved a training effect with RMSE of 0.2345×10^{-4} at the step size of 1.9×10^{4}, which was basically consistent with the target effect. However, CNN-LSTM and LSTM still did not achieve the target effect after reaching 2.0×10^{4} training steps. So, they required more calculation steps and longer calculation time. Based on these data, the above two datasets were used to analyze the fitting error of the training results of the three models. The fitting error of the model reflects the difference between the predicted results and the actual results, and the performance and accuracy of the model can be evaluated. The results are shown in Figure 9.

Figure 9 (a) confirmed that as the training step size continued to increase, the Mu of LSTM fluctuated several times and did not converge. The fitting error values of PSO-CNN-LSTM and CNN-LSTM both experienced a fluctuation and eventually stabilized. The Mu of PSO-CNN-LSTM was 1.0×10^{-7}, while the Mu of CNN-LSTM was 1.0×10^{-6}. In Figure 9 (b), as the training step size increased, the fitting errors of three algorithms eventually stabilized. The difference was that the Mu of ML showed several fluctuations, and Mu eventually tended to stabilize at 1.0×10^{-6}. The Mu of PSO-CNN-LSTM and CNN-LSTM began to rapidly decrease, then became flat, and finally stabilized. Furthermore, the Mu of PSO-CNN-LSTM was 1.0×10^{-8}, which was still smaller than that of CNN-LSTM. Through the above analysis, the stability and effectiveness of PSO-CNN-LSTM were higher than of the other two models [24]. Finally, based on the public dataset Global Energy Forecasting Competition, three models were used to predict the electricity load from 2010 to 2020. The predicted results are shown in Table 2.

According to the results in Tables 2 and 3, the prediction results of PSO-CNN-LSTM were relatively close to the original data, with a confidence level of 98.3% and a P of 0.223, indicating that there was no significant difference within the confidence interval. The prediction results of LSTM showed significant differences from the original data, with a confidence interval of $P < 0.05$. Thus, PSO-CNN-LSTM could accurately predict short-term PL. When deploying prediction models in practice, the optimal parameters were set to solve hyper-parameters and improve model accuracy. When deploying prediction models in practice, the actual data usually contain missing values, outliers, or biases, which may have an adverse impact on the stability performance of the model. The constructed model used cyclic code and mean filling methods to pre-process the data, ensuring the integrity of the data and improving the stability of the model. In addition, actual deployment typically required models to process large amounts of data in real-time or quasi-real-time. CNN was used to extract data features, reducing the complexity of data processing, and improving the accuracy of model predictions.

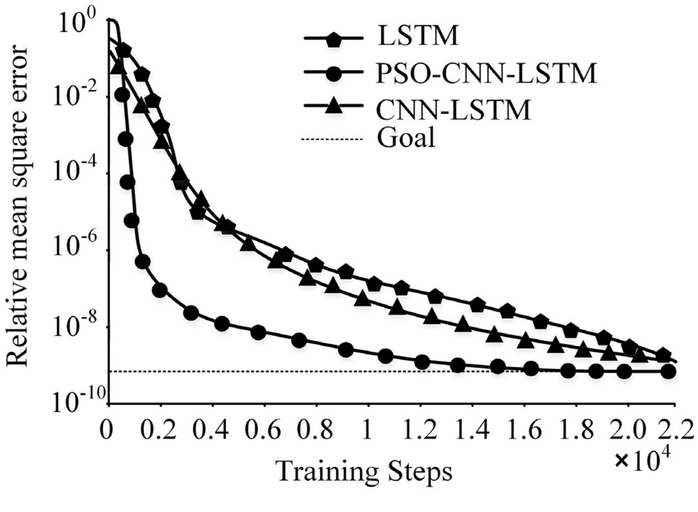
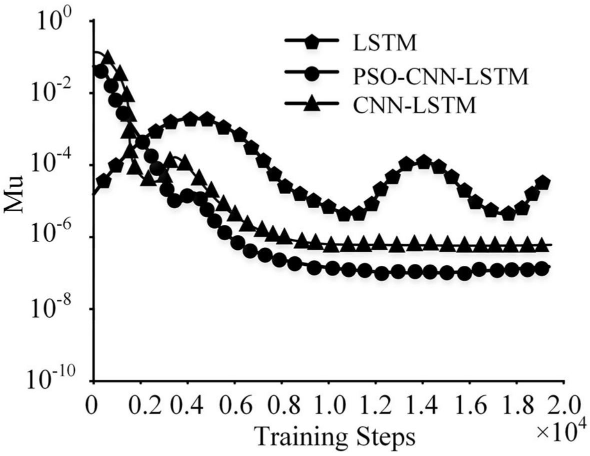
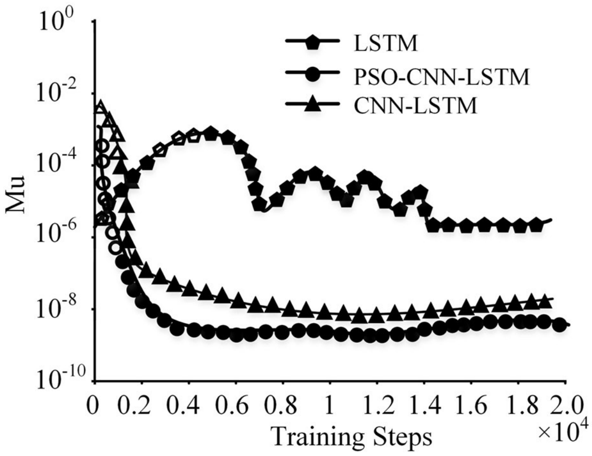


Figure 8. Comparative analysis of RMSE of three models.



(a) GEFC dataset



(b) IEEE dataset

Figure 9. Analysis of fitting errors of three models.

Table 2. Electricity load forecast results for 2010–2020.

Years	Serial Number	Raw data (ten thousand kWh)	Predicted data		
			LSTM	CNN-LSTM	PSO-CNN-LSTM
2010	1	226	259	236	239
2011	2	245	203	234	564
2012	3	298	356	326	310
2013	4	343	410	365	356
2014	5	386	456	423	402
2015	6	415	489	435	435
2016	7	446	502	486	476
2017	8	498	563	534	512
2018	9	528	589	556	563
2019	10	598	658	523	623
2020	11	678	756	541	702

Table 3. Testing indicators.

Index	Confidence level d	Posterior error C	Small error probability	*P* value
LSTM	80.33%	0.86	0.65	$P < 0.05$
CNN-LSTM	92.34%	0.13	0.86	$P = 0.185$
PSO-CNN-LSTM	98.36%	0.09	0.96	$P = 0.223$

3.2. Application Effectiveness Based on PSO-CNN-LSTM-PLF

This study tested the PSO-CNN-LSTM and CNN-LSTM models using historical PL data from A and B during the first and third quarters, and compared them with actual PL change data in Figure 10.

From Figure 10, the predicted results of PSO-CNN-LSTM in different regions and periods were more consistent with actual data than the other two models. Therefore, using PSO to optimize CNN-LSTM effectively improved the accuracy of PLF. This improvement achieved very good prediction results both at individual moments and as a whole. This indicated that using PSO could not only effectively compensate for the shortcomings of CNN-LSTM prediction, but also effectively improve the application value of LSTM in PLF. Finally, an analysis was conducted on RE and AE generated by the predicted data in Figure 11.

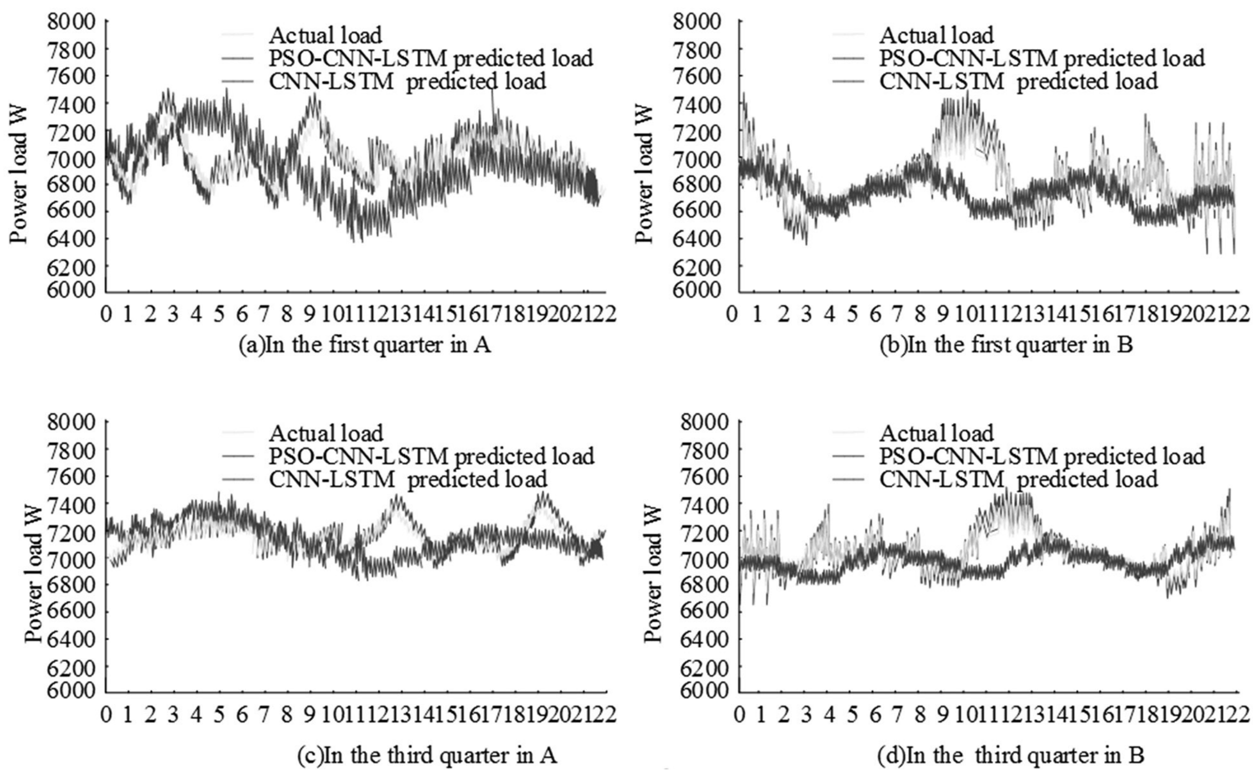
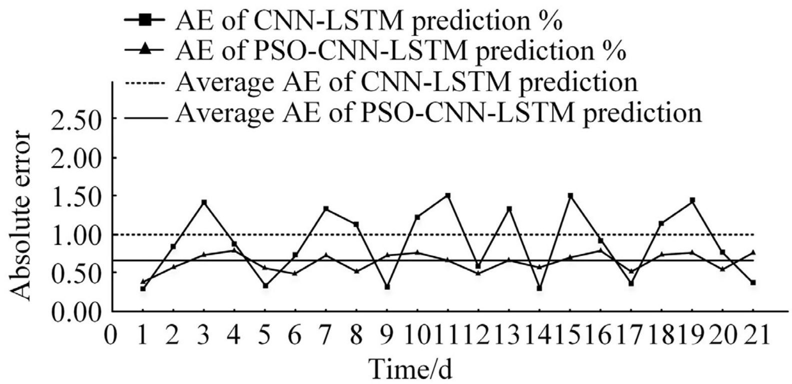
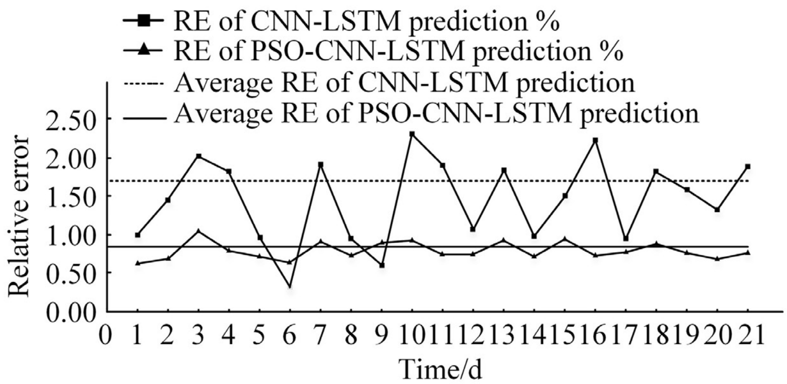


Figure 10. Power load prediction results of two models.



(a) AE of CNN-LSTM and PSO-CNN-LSTM prediction



(b) RE of CNN-LSTM and PSO-CNN-LSTM prediction

Figure 11. RE and AE Analysis of two models.

According to Figure 11 (a), the fluctuation of AE predicted by CNN-LSTM was significant, with an average predicted AE value of 1.0%. The overall AE predicted by PSO-CNN-LSTM was relatively flat, with small fluctuations, and the average predicted AE value was 0.6%.

In Figure 11 (b), CNN-LSTM had significant fluctuations in RE, with an average predicted RE value of 1.6%. The overall AE predicted by PSO-CNN-LSTM was relatively flat, with relatively small fluctuations, and the average predicted AE value was 0.8%. Therefore, for RE and AE, PSO-CNN-LSTM had smaller results than for CNN-LSTM. Therefore, using PSO-CNN-LSTM significantly improved the prediction accuracy of LSTM, greatly reducing the AE and RE caused by the limitations of LSTM itself.

4. Conclusion

Accurate short-term PLF can help improve the safety, economy, and reliability of power systems. The research aims to further improve the accuracy and speed of short-term PLF by designing a PSO-CNN-LSTM power load forecasting model. Firstly, by introducing PSO to accelerate the convergence rate and stability of LSTM, the optimal parameters were found. Subsequently, CNN was introduced to extract the data features of power compliance, reducing noise and instability. Finally, the performance analysis and prediction results of three prediction models, LSTM, CNN-LSTM, and PSO-CNN-LSTM were compared. These results confirmed that when PSO-CNN-LSTM prediction model was iterated 350 times using the GEFC dataset, the performance stability reached 90%. CNN-LSTM gradually stabilized after 720 iterations, with a stable value of 68%, while LSTM had no stable value. For the IEEE dataset, PSO-CNN-LSTM had a performance stability of 92% after 200 iterations. LSTM tended to stabilize after 850 iterations, with a stable value of 58%. CNN-LSTM tended to stabilize after 620 iterations, with a stable value of 70%. In error analysis, the RMSE of PSO-CNN-LSTM for training step size of 1.9×10^4 was 0.2345×10^{-4}, achieving the desired training effect. The other two models still did not reach the target value at the end of training. The fitting error of PSO-CNN-LSTM in GEFC was 1.0×10^{-7} when training for 2.0×10^4 steps. The fitting error for the IEEE dataset was 1.0×10^{-8}. Comparing the predicted results with actual data, the predicted curves of PSO-CNN-LSTM were basically consistent with the actual data curves, having an AE of 1.0% and a RE of 0.8%. The above results confirmed that the prediction performance of PSO-CNN-LSTM was much higher, indicating that PSO-CNN-LSTM had extremely high stability and could achieve relatively accurate prediction results in short-term PLF. However, there are also some shortcomings in the presented research methods. Firstly, during model testing, only two commonly used datasets were utilized, which may lead to overfitting or underfitting issues. This can limit the generalization ability of predictive models and result in poor performance in actual usage scenarios. Secondly, the PSO-CNN-LSTM model requires multiple parameter settings, including the parameters of the PSO algorithm, the structure and parameters of the CNN and LSTM. Adjusting these parameters will be a complex and time-consuming process. Further research is necessary to address the aforementioned issues.

References

[1] C. Fan *et al.*, "Multi-objective LSTM Ensemble Model for Household Short-term Load Forecasting", *Memetic Computing*, vol. 14, no. 1, pp. 115–132, 2022. http://dx.doi.org/10.1007/s12293-022-00355-y

[2] R. Zhang *et al.*, "A Similar Day Based Short Term Load Forecasting Method Using Wavelet Transform and LSTM", *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 17, no. 4, pp. 506–513, 2021. http://dx.doi.org/10.1002/tee.23536

[3] Y. Pei *et al.*, "Load Forecasting of Refrigerated Display Cabinet Based on CEEMD–IPSO–LSTM Combined Model", *Open Physics*, vol. 19, no. 1, pp. 360–374, 2021. http://dx.doi.org/10.1515/phys-2021-0043

[4] W. J. Zhang *et al.*, "Short-term Power Load Forecasting Using Integrated Methods Based on Long Short-term Memory", *Science China-Technological Sciences*, vol. 63, no. 4, pp. 614–624, 2020. http://dx.doi.org/10.1007/s11431-019-9547-4

[5] V. Veeramsetty *et al.*, "Short-term Electric Power Load Forecasting Using Random Forest and Gated Recurrent Unit", *Electrical Engineering*, vol. 104, no. 1, pp. 307–329, 2022. http://dx.doi.org/10.1007/s00202-021-01376-5

[6] Z. S. Chafi and H. Afrakhte, "Short-Term Load Forecasting Using Neural Network and Particle Swarm Optimization (PSO) Algorithm", *Mathematical Problems in Engineering*, vol. 21, no. 2, pp. 1–10, 2021. http://dx.doi.org/10.1155/2021/5598267

[7] M. Gilanifar *et al.*, "Multitask Bayesian Spatiotemporal Gaussian Processes for Short-Term Load Forecasting", *IEEE Transactions on Industrial Electronics*, vol. 67, no. 6, pp. 5132–5143, 2020. http://dx.doi.org/10.1109/TIE.2019.2928275

[8] A. Tudose *et al.*, "Neural Networks Application in Short-Term Load Forecasting", *UPB Scientific Bulletin, Series C: Electrical Engineering*, vol. 83, no. 2, pp. 231–240, 2021. Retrived from https://www.researchgate.net/publication/353670061_Neural_Networks_Application_in_Short-Term_Load_Forecasting

[9] I. Ozer *et al.*, "A Combined Deep Learning Application for Short Term Load Forecasting", *Alexandria Engineering Journal*, vol. 60, no. 4, pp. 3807–3818, 2021. http://dx.doi.org/10.1016/j.aej.2021.02.050

[10] R. Patel *et al.*, "Electrical Load Forecasting Using Machine Learning Methods, RNN and LSTM", *Xi'an Dianzi Keji Daxue Xuebao/Journal of Xidian University*, vol. 14, no. 4, pp. 1376–1386, 2020. http://dx.doi.org/10.37896/jxu14.4/160

[11] A. Amiri *et al.*, "Adaptive Shewhart Control Charts Under Fuzzy Parameters with Tuned Particle Swarm Optimization Algorithm", *Journal of Industrial Integration and Management*, vol. 8, no. 2, pp. 241–276, 2021. http://dx.doi.org/10.1142/S2424862221500226

[12] W. Lu *et al.*, "A CNN-LSTM-Based Model to Forecast Stock Prices", *Complexity*, vol. 2020, pp. 1–10, 2020. http://dx.doi.org/10.1155/2020/6622927

[13] Y. Liang and H. Zhang, "Ship Track Prediction Based on AIS Data and PSO Optimized LSTM Network", *International Core Journal of Engineering*, vol. 6, no. 5, pp. 23–33, 2020. http://dx.doi.org/10.6919/ICJE.202005_6(5).0005

[14] R. Wang *et al.* "Electric Short-term Load Forecast Integrated Method Based on Time-segment and Improved MDSC-BP", *Systems Science & Control Engineering: An Open Access Journal*, vol. 9, no. 1, pp. 1–7, 2020. http://dx.doi.org/10.1080/21642583.2020.1843088

[15] Z. Zhuang *et al.*, "A Reliable Short-Term Power Load Forecasting Method Based on VMD-IWOA-LSTM Algorithm", *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 17, no. 8, pp. 1121–1132, 2022. http://dx.doi.org/10.1002/tee.23603

[16] Y. J. Song, "Blockchain-based Power Trading Process", *Journal of System and Management Sciences*, vol. 9, no. 3, pp. 78–91, 2019. http://dx.doi.org/10.21742/ijhit.2019.12.2.02

[17] D. Monniaux and C. Six, "Formally Verified Loop-Invariant Code Motion and Assorted Optimizations", *ACM Transactions on Embedded Computing Systems*, vol. 22, no. 1, pp. 3–27, 2022. http://dx.doi.org/10.1145/3529507

[18] D. Y. Lee, "A Study on Smart Factory Efficiency of Small and Medium-sized Enterprises Based on Electricity and Electronics", *Journal of System and Management Sciences*, vol. 12, no. 2, pp. 431–442, 2022.

[19] S. Al-Marzouki *et al.*, "Estimation of Finite Population Mean Under PPS in Presence of Maximum and Minimum Values", *AIMS Mathematics*, vol. 6, no. 5, pp. 5397–5409, 2021. http://dx.doi.org/10.3934/MATH.2021318

[20] A. A. Permana *et al.*, "Machine Learning Approaches for Precision Medicine: A Review", *Journal of System and Management Sciences*, vol. 13, no. 4, pp. 416–432, 2023. http://dx.doi.org/10.33168/JSMS.2023.0425

[21] Q. Cui *et al.*, "Comprehensive Evaluation of Electric Power Prediction Models Based on D-S Evidence Theory Combined with Multiple Accuracy Indicators", *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 3, pp. 597–605, 2022. http://dx.doi.org/10.35833/MPCE.2020.000470

[22] R. Hanif *et al.*, "A Study of Time Series Forecasting Enrollments Using Fuzzy Interval Partitioning Method", *Journal of Computational and Cognitive Engineering*, vol. 2, no. 2, pp. 143–149, 2022. http://dx.doi.org/10.47852/bonviewJCCE2202159

[23] J. W. Choi and E. R. Jeong, "A Multi-output Convolutional Neural Network-based Distance and Velocity Estimation Technique", *Journal of Logistics, Informatics and Service Science*, vol. 9, no. 1, pp. 11–25, 2022. http://dx.doi.org/10.14704/web/v19i1/web19302

[24] X. Wang *et al.*, "Fake Node Attacks on Graph Convolutional Networks", *Journal of Computational and Cognitive Engineering*, vol. 1, no. 4, pp. 165–173, 2022. http://dx.doi.org/10.47852/bonviewJCCE2202321

Received: October 2023
Revised: December 2023
Accepted: December 2023

Contact addresses:
Sile Hu
Zhejiang University
Hangzhou
Zhejiang
China
Inner Mongolia Power (Group) Co., Ltd.
Hohhot
Inner Mongolia
China
e-mail: 12110087@zju.edu.cn

Wenbin Cai
Inner Mongolia Power Electric Power Economic and Technological Research institute
Hohhot
Inner Mongolia
China
e-mail: caiwenbin@impc.com.cn

Jun Liu
Inner Mongolia Power Electric Power Economic and Technological Research institute
Hohhot
Inner Mongolia
China
e-mail: 15848085500@163.com

Hao Shi
Beijing Tsintergy Technology Co., Ltd.
Beijing
China
e-mail: shihao@tsintergy.com

Jiawei Yu*
Beijing Tsintergy Technology Co., Ltd.
Beijing
China
e-mail: Yjw_971106@163.com
*Corresponding author

SILE HU graduated from the Hong Kong Polytechnic University with a Master's degree in 2010. He is now studying at Zhejiang University and is a PhD student. He currently serves as the Director of the Planning Development Department at Inner Mongolia Power Company. His research interests include power system planning, operation, and flexible regulation capability analysis.

WENBIN CAI graduated from China Agricultural University in 2000 with a bachelor's degree. Currently, he serves as the Director of the Planning Review Department at Inner Mongolia Electric Power Economic and Technological Research Institute Branch of Inner Mongolia Electric Power (Group) Co., LTD. His research focuses on power system planning.

JUN LIU holds a bachelor's degree from the Inner Mongolia University of Technology, where he graduated in 2004. Currently, he serves as the Deputy General Manager at Inner Mongolia Electric Power Economic and Technological Research Institute Branch of Inner Mongolia Electric Power (Group) Co., LTD. His research interests lie in power grid planning and energy development.

HAO SHI holds a Master's degree from the University of Edinburgh, UK, where he graduated in 2020. Currently, he works as a Product Manager at Beijing Tsintergy Technology Co., LTD. His research focuses on power system planning and operation.

JIAWEI YU holds a Master's degree from the University of Edinburgh where he graduated in 2021. Currently, he works as a Product Specialist at Beijing Tsintergy Technology Co., LTD. His research focuses on power system supply-demand balance analysis and strategic planning.