Science and Technology for Energy Transition **79**, 57 (2024)

https://doi.org/10.2516/stet/2024062

Available online at:
stet-review.org

Emerging Advances in Hybrid Renewable Energy Systems and Integration
Muhammad Zakarya, Santosh Tirunagari, Jinguang Han and Peiying Zhang (Guest Editors)

REGULAR ARTICLE OPEN ACCESS

Net load forecasting method in distribution grid planning based on LSTM network

Ye Yuan*, Xinping Yuan, Haiyan Wang, Ming Tang, and Mengyu Li

Yunnan Power Grid Co., Ltd., Information Center, 105 Yunda West Road, Guandu District, Kunming City 650500, Yunnan Province, China

Received: 4 June 2024 / Accepted: 26 July 2024

Abstract. Distribution grid planning involves multiple nodes, lines, equipment, and other elements. Due to the large scale of the system, there are complex interactions in space. The net load is affected by the load changes of different nodes. If this spatial complexity is not fully considered, the net load prediction results may be inaccurate. Therefore, in order to ensure the effect of net load forecasting, a method of net load forecasting in distribution grid planning based on the Long Short-Term Memory (LSTM) network is proposed. This method fully considers the characteristics of distribution grid planning and constructs a net load forecasting model for distribution grid planning based on the LSTM network. This model selects the 3σ criterion detects and corrects the singular values in the historical load data, and obtains the reasonable maximum time series results of each day; The adaptive noise complete set empirical mode decomposition method is used to decompose the sequence results and generates Intrinsic Mode Function (IMF) components of each time series; According to the component results, a load forecasting model based on LSTM network is constructed, and the initial learning rate and cell number parameters of LSTM network are optimized by improving the Pelican optimization algorithm to improve the precision of load forecasting of LSTM network. The test results show that the method can detect singular values in the data and weaken the impact of grid planning on netload forecasting; It can effectively complete the decomposition of historical load data, and each component after decomposition will not be aliased; The prediction error of net load is less than 1.25%, which can provide a reliable basis for grid planning of distribution network.

Keywords: LSTM network, Distribution network, Grid planning, Medium net load forecasting, Time series, Parameter optimization.

1 Introduction

Based on the original power supply area, the grid planning of the distribution network is to divide the power supply grid and power supply unit deeply into two levels to form a three-level power supply network. The operation of the distribution network is solidified according to the results of the hierarchical division. Grid planning mainly adopts the top-down approach, starting with low voltage and then proceeding to high voltage planning for distribution network planning [1], with the core goals of meeting the actual power demand of users, improving the details of distribution network planning, and enhancing the operation and maintenance management level of the distribution network. In the process of grid planning of the distribution network, according to the load forecasting results [2], the construction planning of power equipment such as high-voltage substations and medium-voltage lines can be effectively completed, ensuring the operability and applicability of the distribution network planning scheme [3]. Among them, the actual load supplied by the distribution network to users is called the net load. This load is adjusted with the change in users' electricity consumption, which is an important indicator to measure the operation status of the distribution network [4]. Therefore, when the distribution network conducts grid planning, in order to ensure the rationality of the planning scheme, the stability of the distribution network operation after the planning, and the power supply quality [5], real-time prediction of the distribution network's net load is needed to improve the accurate management of net load [6], to better achieve the absorption and optimal utilization of renewable energy. However, in the grid planning of the distribution network, due to the large scale of the distribution network system, numerous nodes, lines and equipment, and complex interactions, the net load has significant volatility and randomness, resulting in poor accuracy of prediction results, affecting the rationality of grid planning of distribution network.

Reference [7] decomposes the original load through variable decomposition, forms the intrinsic mode function, and

* Corresponding author: yeyuan0513@163.com

selects the days, hours, and historical data as well as the decomposition results as the input data of the four-level wavelet decomposition model to carry out the load forecasting; however, the four-level wavelet decomposition model needs to process and decompose a large amount of data, which may lead to too much computation and is not suitable for real-time requirements. However, the four-level wavelet decomposition model needs to process and decompose a large amount of data, which may lead to too much computation and is not favorable to real-time requirements. Reference [8], in order to ensure the planning effect of the distribution network, to realize long-term spatial load prediction, uses the hierarchical trend method to predict the peak load of each region, to determine the prediction interval, and finally uses Monte Carlo for load prediction. However, the Monte Carlo method is a stochastic simulation method, and the results are affected by the random number generation algorithm, the number of simulations, etc., and there may be a certain degree of error. Reference [9] mainly focuses on short-term load forecasting, combines the changing characteristics of renewable energy and high penetration, determines the relevant variables of load impact, and uses machine learning to predict the nonlinear load. However, because load forecasting is a dynamic process, the machine learning model may not have enough generalization ability in the face of the new situation, resulting in unstable prediction results. Reference [10], in order to improve the planning effect of the distribution network and realize the short-, medium- and long-term load prediction at the same time, combining the multilayer perception and statistical methods, based on the operating data of distribution lines, to complete the load prediction of distribution network. However, the parameter adjustment of the multilayer perception model may be more cumbersome, and it requires appropriate parameter settings to achieve better prediction results.

Long Short-Term Memory (LSTM), the core part of the network is the memory unit, which can accurately capture a certain key information [11] and save the information at certain time interval. Therefore, the network has good processing ability for time series data [12]. That is, the memory unit of LSTM can learn the complex relationships and patterns in the data during the training process, including the interaction between different nodes, lines, and equipment in space. This solves the negative impact of the grid planning characteristics of the distribution network on netload forecasting, thus improving the accuracy of net load forecasting. Therefore, in order to realize the net load forecasting in the grid planning of the distribution network, this paper proposes the net load forecasting method in the grid planning of the distribution network based on the LSTM network.

2 Characterization of distribution grid planning

In the grid-based planning of distribution networks, netload prediction needs to take into account the complex interactions among multiple nodes, lines, equipment, etc. in the system. The net load is affected by the load changes at different nodes, and the load transfer between nodes, transmission characteristics of lines, and characteristics and constraints of equipment all have an impact on the net load. If this spatial complexity is not fully considered, the net load forecast results may not truly reflect the actual load situation, thus affecting the accuracy and effectiveness of planning. Therefore, it is necessary to consider the spatial structure of the distribution network system and the interactions between various elements in the net load forecast, in order to improve the accuracy and reliability of the forecast. Therefore, the correlation between the three-level network of the distribution grid is shown in Figure 1.

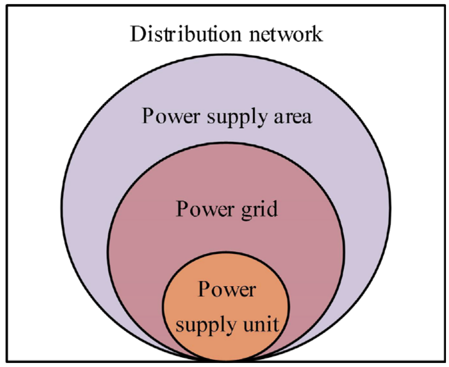

Fig. 1. The correlation between the grid-based three-level networks of the distribution network.

In Figure 1, the power supply area is divided based on the planned annual load density as the primary criterion, combined with the regional power consumption level, economic status, and other factors. Concerning the Technical Guidelines for Distribution Network Planning and Design, the power supply area is normally divided into five categories: A^{+}, A, B, C, D. The power supply grid is a mutually independent grid formed by further division based on the power supply area and in combination with the planning demand, operation status, marketing serviceability, operation and maintenance emergency repair ability of the planned area. On the basis of this grid, the load density and load characteristics are combined for in-depth division, and the relatively independent units formed are called power supply units, which are usually divided by towns.

Rational planning of the distribution grid is based on the results of load forecasting while taking into account the difficulty of reconstruction of the current distribution network and the surrounding environment and other factors, to ensure the stability of the distribution network after planning and short- and long-term adaptability. Distribution grid planning needs to ensure that there is no overlap or leakage in the grid, to meet the management needs of the distribution network operation. Power supply unit division is the basis of distribution grid planning. In the planning process, while considering the administrative division of the region, it is necessary to combine the location of the planning site, transmission lines, and substations, as well as factors such as the level of utilization [13]. The rationality of the division of the power supply unit area is particularly important; it should not be too large or too small. The division needs to be completed according to the density of the load situation so that the division of the unit can be regarded as a constantly adjusted process. Therefore, the division of the unit can also be regarded as a process of constant adjustment and modification, and the planning scheme is constantly adjusted to ensure the rationality of the program.

3 Methods

According to the grid planning details of the distribution network analyzed in the above subsections, the load density and load characteristics are the important basis for determining the rationality of the grid planning scheme of the distribution network, and the power supply unit adjusts the planning scheme according to the load conditions. In order to ensure the grid planning effect of the distribution network, it is necessary to accurately master the load conditions at different stages in the planning process [2, 14]. Based on this, from three aspects of net load data processing, time series net load forecasting based on LSTM network, and LSTM network parameter optimization, the net load forecasting algorithm for distribution grid planning is designed to provide a reliable basis for distribution grid planning.

3.1 Netload data processing

After the grid planning of the distribution network, there will be some strange data in the load time series data of the power supply unit, because the base load of the power supply unit is small and fluctuates greatly [15, 16], which will cause these data to change during the acquisition process and form strange data. Therefore, if the data is directly used in the net load forecasting, the singular data in the data will reduce the forecasting accuracy. The 3σ standard is based on the principle of normal distribution, which assumes that in a normal distribution, almost all data will fall within the range of the mean plus or minus three standard deviations (i.e. 3σ). Data points beyond this range are considered outliers because their deviation from most data is too large, which may be caused by measurement errors, equipment failures, or other atypical factors. Therefore, to improve the accuracy and reliability of net load forecasting, the 3σ criterion is adopted in this paper to perform singular value detection on the historical load measurement data of power supply units in grid planning of distribution networks, in order to determine the maximum reasonable value of daily net load for each planned power supply unit [17].

3.1.1 Singular value detection in historical load-measured data of power supply units

When the 3σ criterion processes the singular value of the data, the criterion needs to detect the singular value data first. If the measured data set of the historical load of the power supply unit in the grid planning of the distribution network is denoted by $X = [X_1, X_2, \ldots, X_s]$, where the amount of data is denoted by s, the pth data in the sample is denoted by X_p $(p = 1, 2, \ldots, s)$, then the process of detecting singular values is described as follows:

Step 1: Calculate the mean value of X which is $\bar{X}$, calculates the residual error of each data in X according to the calculation results of $\bar{X}$, and the calculation formula is:

$$\xi_p = X_p - \bar{X} \tag{1}$$

where: ξ_p indicates that the pth residual error values correspond to the individual data.

Step 2: Calculate the standard deviation of X which is σ, and the calculation formula is:

$$\sigma = \sqrt{\frac{1}{s}\sum_{p=1}^{s}\left(\bar{X} - X_p\right)^2} \tag{2}$$

Step 3: Calculate the residual error of each historical load data in X according to formula (1), and obtaining the absolute values of each ξ_p, the residual error ξ_p result of each data is compared with 3σ, and judged according to formula (3):

$$\left|\xi_p\right| > 3\sigma \tag{3}$$

If formula (3) is satisfied, it indicates that the corresponding data of ξ_p is odd data, and the opposite is normal data.

Complete the singular value data detection in X according to the above steps, the load time series is obtained by detecting the singular values in the data and correcting the detected singular values, and the maximum value of the load is extracted, and the reasonable maximum time series results are constructed for each day.

3.1.2 Maximum value time series decomposition

After obtaining the reasonable maximum time series results of each day according to the above steps, the results have a certain degree of random volatility, which will reduce the forecasting performance and accuracy of the net load. CEEMDAN (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise) is a signal processing method aimed at decomposing nonlinear and non-stationary signals into multiple Intrinsic Mode Functions (IMFs) with different scales and frequency characteristics. The CEEMDAN method is an extension of EMD (Empirical Mode Decomposition), which decomposes signals into multiple IMFs and a residual term through a series of decomposition and reconstruction processes, solving the problem of mode aliasing that occurs during the decomposition process of EMD. This method has shown excellent performance in the field of signal processing, especially in processing time series data with complex nonlinear and non-stationary characteristics. Therefore, in order to improve the decomposition accuracy, solve the problem of modal aliasing, and enhance the anti-noise performance, the CEEMDAN method is used in this paper to decompose the results into multiple IMF components at different time scales.

White noise has multiple sets of independent sequence characteristics, which can help decomposition algorithms amplify the uncorrelation degree of difficult-to-separate modes (such as modes with similar frequency characteristics) in the original data using the characteristics of white noise, thereby extracting the original inseparable modes and effectively suppressing mode aliasing. Therefore, in order to obtain the reasonable maximum time series result $g(t)$ for each day, a white noise sequence containing a standard normal distribution is added to $g(t)$, then:

$$g^l(t) = (g + \beta_0 \varepsilon^l)t \tag{4}$$

where: $l = 1, 2, \ldots, Q$; β_0 represents the adaptive coefficient of IMF component decomposition; $g^l(t)$ represents the time series of net load after performing l times noise addition; the $\varepsilon^l(t)$ represents the noise sequence after performing l times noise addition.

The EMD is performed for each new load time series, and the average value of the first IMF component of all time series is calculated. On this basis, the first IMF component and the first residual component of CEEMDAN are obtained, the two are denoted by $g_{\mathrm{IMF1}}(t)$ and $\xi_l(t)$ respectively. The formulae for the two are:

$$f_{\mathrm{IMF1}}(t) = \frac{1}{Q}\sum_{i=1}^{Q} I_1^l \tag{5}$$

$$\xi_1(t) = g^{l(t)} - g_{\mathrm{IMF1}}(t) \tag{6}$$

where: I_1^l represents the first IMF component acquired after the l times of EMD.

After calculate $\xi_l(t)$ based on the above formula, the new sequence $\xi_l(t) + \beta_1\gamma_1[\varepsilon^l(t)]$ is generated by adding the white noise sequence containing the standard normal distribution. Of which $\gamma_1[\cdot]$ represents the calculation operator of the first IMF component. The EMD is performed on each new sequence generated to calculate the average value of the first IMF component of all sequences, and then the second IMF component $g_{\mathrm{IMF2}}(t)$ is obtained, which is calculated by the following formula:

$$g_{\mathrm{IMF2}}(t) = \frac{1}{Q}\sum_{l=1}^{Q} \gamma_1\{\xi_1(t) + \beta_1\gamma_1[\varepsilon^l(t)]\}. \tag{7}$$

Calculate the remaining stages to obtain the mth residual component of CEEMDAN and the $m+1$th of IMF components, the two are denoted by $\xi_m(t)$ and $g_{\mathrm{IMF}m+1}(t)$ respectively. The formula for both are:

$$\xi_m(t) = \xi_{m-1}(t) - g_{\mathrm{IMF}m}(t) \tag{8}$$

$$g_{\mathrm{IMF}m+1}(t) = \frac{1}{Q}\sum_{l=1}^{Q} \gamma_1\{\xi_m(t) + \beta_m\gamma_m[\varepsilon^l(t)]\}. \tag{9}$$

Using equations (8) and (9), the computation of all residual components is calculated and the number of extreme points for that component is determined, and when that number is less than two, the continuation of the decomposition is stopped.

If the algorithm stops decomposition, CEEMDAN obtains M IMF components, and $g(t)$ finally forms M IMF component and a residual component $R(t)$. At this time, the formula for $g(t)$ is:

$$g(t) = \sum_{m=1}^{M} g_{\mathrm{IMF}m}(t) + R(t). \tag{10}$$

The processing of $g(t)$ can be accomplished by the above steps, and obtain multiple IMF component results, and using the results as the forecast data of the net load of the power supply unit in the grid planning of the distribution network.

3.2 Time series net load forecasting based on LSTM network

Although the singular value data has been removed through the above processing, because grid planning of distribution networks needs to consider the net load at different time points [18, 19], singular data with different influence weights will also be generated at this time, distorting the data distribution and causing the model to be disturbed during the training and forecasting process, thus affecting the accuracy of the net load forecasting. Therefore, according to the results of multiple IMF components obtained in the above sections, the special network structure of the LSTM network is used for time series net load forecasting. Better prediction results can be obtained through long-term and short-term information-dependent learning in time series; The long-term and short-term memory units in the network are used to accurately capture the complex features inside the time series data and improve the accuracy of the net load forecasting.

The LSTM network has a feedback connection structure, which can complete the learning of data characteristics through repeated learning and long-term and short-term units to achieve the net load prediction. The network structure is shown in Figure 2.

Based on the time series of the daily reasonable maximum values of the historical load of the power supply unit, the LSTM network constructs a net load forecasting model for each IMF component obtained from the decomposition. It then predicts the maximum value of each day's net load of the power supply unit using each model built, so as to obtain the maximum annual net load of the power supply unit and realize net load forecasting in the grid planning of the distribution network.

If the input of each IMF component contains a total of k time steps, at the t moment, the input vector of the long and short memory cells at that moment is denoted by x_t, and the implicit layer state is denoted by h_t, and that it can output LSTM cells; For cell status, it denoted by c_t, it can complete LSTM cell state calculation, and the detailed prediction process of LSTM network is as follows:

Step 1: Forgetting gate output results updated

The net load in grid planning of distribution networks has obvious temporal characteristics. Through the forget gate, LSTM networks can selectively retain important information from past time steps while ignoring irrelevant information, effectively handling long-term dependencies in time series data. This is particularly important for applications that require memory and understanding of data over long time spans. Meanwhile, time series data may contain noise and outliers, which may have adverse effects on the prediction results. The forgetting gate can filter out unimportant information by controlling the degree of information

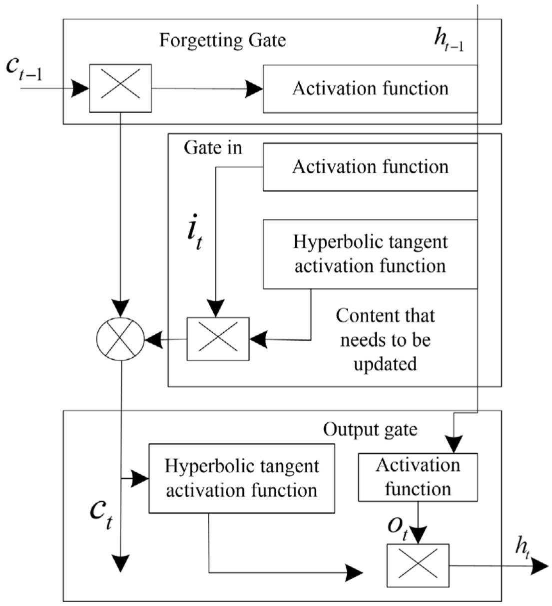

Fig. 2. LSTM network structure.

forgetting, thereby improving the accuracy of predictions. Update the formula to:

$$f_t = \eta[w_f g(t) + \omega_f h_{t-1} + b_f] \tag{11}$$

where: w_f denotes the weight of the forgetting gate cycle; the ω_f denotes the influence weight of the odd data input to the forgetting gate; the h_{t-1} indicates the stats of the forgetting gate at the moment $t - 1$. b_f denotes the bias vector; the η denotes the activation function.

Step 2: Input gate two parts output result update

The output results of the input gate are denoted by i_t and $\tilde{c}$ respectively, the formula for the two is as follows:

$$\begin{cases} i_t = \eta[w_i h_{t-1} + \omega_i g(t) + b_i] \\ \tilde{c} = \tanh\left[w_c g(t) + \omega_c h_{t-1} + b_c\right] \end{cases} \tag{12}$$

where: w_i and ω_i and b_i represent the loop weights, input weights, and bias vectors of i_i; the w_c and ω_c and b_c represent the loop weights, input weights, and bias vectors of $\tilde{c}$; the tanh denotes the hyperbolic tangent activation function, so that it lies between $(-1, 1)$. This design aims to capture complex nonlinear relationships in temporal data. The tanh function not only compresses the output value to the $(-1, 1)$ interval, stabilizing the learning process of the network but also enables the model to express positive and negative changes in the data. This feature is particularly important in applications such as net load forecasting, as it allows LSTM to more accurately capture the trend of load increase and decrease, thereby improving the accuracy of forecasting.

Activation function η is able to complete the transformation of the input quantities so that they form probability values, the function is given by:

$$\eta(x) = \frac{1}{1 + e^{-1}}. \tag{13}$$

The main role for tanh is to realize the state vector filtering and form the output information, which is calculated by the formula:

$$\tan h[g(t)] = \frac{\sin\ h[g(t)]}{\cos\ h[g(t)]} \tag{14}$$

where: $h[g(t)]$ indicates the input vector state.

Step 3: Cell state updates

The updated formula for c_t is:

$$c_t = f_t c_{t-1} + i_t \tilde{c}_t. \tag{15}$$

Step 4: Output gate output result updates

The output result of the output gate is denoted by o_t, the updated formula is:

$$o_t = \eta[w_o g(t) + \omega_o h_{t-1} + b_o] \tag{16}$$

$$h_t = o_t \times \tan h(c_t) \tag{17}$$

where: w_o and ω_o and b_o denote the cyclic weights, input weights and bias vectors of the output gate.

Step 5: Current moment net load forecast output updates

If the net load forecast output result at the current moment is denoted by $\widehat{y}_t$, the updated formula is:

$$\widehat{y}_t = \eta(Zh_t + c) \tag{18}$$

where: z represents the internal characteristic matrix of the time series of IMF components.

3.3 LSTM network parameter optimization

In the process of load forecasting of the LSTM network, the initial learning rate of the LSTM network and the number of cell units affect the forecasting effect of the network. Therefore, in order to ensure better accuracy of load forecasting, help optimize model performance, and improve the accuracy and stability of time series load forecasting, this paper uses an improved Pelican optimization algorithm to optimize parameters and obtain the best parameter values, The net load forecast is realized according to this parameter.

The basic Pelican algorithm may prematurely converge to local optima during the search process while ignoring the existence of global optima. The dynamic selection strategy can adjust in real-time or near real-time based on changes in system status, external environment, or user needs. This strategy can quickly respond to these changes, ensuring that the optimization process always remains consistent with the actual situation. In order to avoid the possibility of local optima and improve the global search ability of the algorithm, a dynamic selection strategy is adopted to optimize the basic Pelican algorithm. Based on the progress of the search and the quality of the current solution, the search direction and step size are dynamically adjusted to form an improved Pelican optimization algorithm.

When defining the initial learning rate and number of cell units of the LSTM network as the search space, and using the improved algorithm for parameter optimization, the search space is defined as a lake, and the fitness value to be optimized is defined as a fish. Obtaining optimal solutions in lakes through the behavior of pelicans fishing; in this process, in order to better determine the new target location, the new target location is updated through a dynamic selection strategy. The improved Pelican optimization algorithm has the ability of random initialization and parameter adaptive adjustment and can achieve multiple benchmark optimization. The improved Pelican optimization algorithm can be divided into three stages in the process of LSTM network parameter optimization, namely the initialization stage, the search stage, and the location stage.

In the net load prediction of grid-based planning of distribution networks, LSTM networks may fall into local optima. When the population diversity is high, the algorithm is more likely to jump out of the local optimal solution, avoiding the dilemma of premature convergence. In load forecasting, local optimal solutions often only reflect the laws of partial load changes, while global optimal solutions can more comprehensively reflect the complexity and dynamics of load changes. Chaos mapping can generate sequences with ergodicity and randomness, and the characteristics of these sequences can be used to initialize or update individuals in the population, thereby increasing the diversity of the population. Therefore, in the initialization process, the Logistic map is selected to complete the chaotic mapping of the optimal individuals in the population and generate a random number sequence $y_n(t)$ that meets the mapping rule. The sequence represents the location where the search individual is initialized. In the search phase, Pelicans search for potential food sources in the search space in a random way, that is, the parametric solution of the network; At this time, the position of the parametric solution is generated randomly, and the position update between pelicans is completed through cyclic iteration. In the positioning phase, the pelican realizes the positioning of the parameter solution and starts hunting. In this process, the dynamic selection strategy is introduced. Typical selection strategies include two types, namely, reverse learning strategy (F1 strategy) and Cauchy mutation perturbation strategy (K1 strategy). Among them, the reverse learning strategy (F1 strategy) selects individuals with closer distances than the initial individuals of the population, making each one step closer to the optimal solution, thereby improving the convergence speed of all individuals in the population. Through reverse learning, the model may be able to identify and correct errors in the learning process faster, thereby improving learning efficiency. The Cauchy mutation perturbation strategy (K1 strategy) is a commonly used mutation operation strategy in optimization algorithms, which generates a new candidate solution vector by randomly perturbing each dimension of the solution vector. By perturbing individuals (i.e. the parameter configuration of LSTM networks) through Cauchy mutation, new individuals can be generated and more diversity can be introduced into the population. This helps to avoid premature convergence and increases the chances of finding the global optimal solution. In order to choose the best strategy for the target position update, the paper is determined by the calculation of the selection probability P_s, and the calculation formula is:

$$P_s = -\exp\left(1 - \frac{t}{T_{\max}}\right)^{20} + \theta \tag{19}$$

where: θ indicates the adjustment parameter. $T_{\max}$ and t denote the current iteration number and the maximum iteration number, respectively.

After calculating the result of P_s according to formula (19), it is compared with the random number matrix r, which satisfies the standard uniform distribution, if $r < P_s$, select F1 strategy as the target location update strategy; If $r \geq P_s$, K1 strategy is selected as the target location update strategy.

After determining the selection strategy, the target position is updated by combining the greedy rule, and the update rule is formulated as:

$$\widehat{X} = \begin{cases} X_{i,j,t+1}, f\left(X_{i,j,t+1}\right) < f(\widehat{X}) \\ \widehat{X}, f\left(X_{i,j,t+1}\right) \geq f(\widehat{X}) \end{cases} \tag{20}$$

where: $X_{i,j,t+1}$ indicates that under the $t + 1$ iteration, in the j dimension, the position of the ith pelican. $\widehat{X}$ denotes the globally optimal solution; the $f(\cdot)$ denotes the distribution probability density function.

According to the above steps, the parameter solution corresponding to the location of the global optimal solution is the optimal parameter value of the LSTM network. Use this parameter value for net load forecasting to improve the accuracy of load forecasting.

4 Results and discussion

To verify the net load prediction effectiveness of the forecasting method in distribution network grid planning, the distribution network of a certain region is used as the test object. The scope of distribution network grid planning for this region covers the entire county area, including urban and rural areas. The base year for planning is 2020, the planning level year is 2024, and the outlook extends up to 2032 (the saturation year). Therefore, it is necessary to ensure that the grid planning results meet the short-term and long-term development needs of the distribution network. The current planning results for the region are shown in Figure 3.

In Figure 3, the classification of the five categories A^+, A, B, C, and D is determined according to the load density. If the load density is indicated by ρ, $\rho \geq 30$, it belongs to category A^+; When $15 \leq \rho < 30$, it belongs to category A; When $6 \leq \rho < 15$, it belongs to category B; When $1 \leq \rho < 6$, it belongs to category C; When $0.1 \leq \rho < 1$, it belongs to category D.

According to the statistics of historical load data in this region, the maximum overall load in the 5 years preceding the base year was 409.6 MW, with an annual growth rate of 13.36%. Due to the continuous increase in urban and rural industrial construction, this load is still gradually

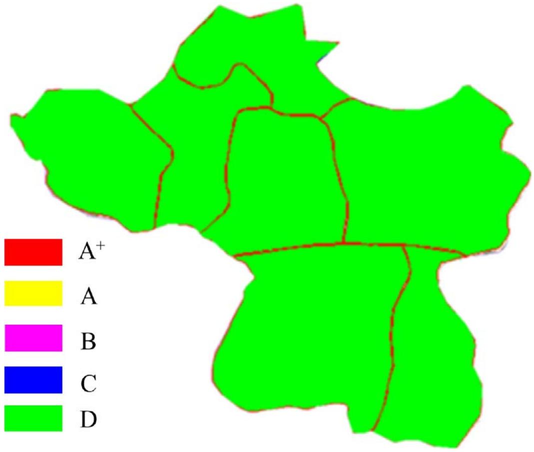

Fig. 3. Planning results of the current power supply area in the study area.

increasing. The urban load density is approximately 1.91 MW/km^2, and the rural load density is approximately 0.20 MW/km^2. When planning the distribution grid in this region, it is necessary to fully consider load growth to ensure that the planning scheme meets the construction and development needs of the distribution network. Therefore, to ensure the effectiveness of grid planning for the distribution network in this region, the method presented in this paper is used to forecast the net load during planning, obtain accurate load results, and provide a reliable basis for planning. During the test, the historical load data used were the load results collected during the operation of the distribution network. The sampling period for this data was 6 min. A total of 1095 groups of historical load data were selected to test the 240 historical load data points of the power supply unit.

Parameter setting: the relevant parameters of the LSTM network, improved Pelican optimization algorithm, and CEEMDAN method are shown in Table 1.

With long and short-term memory network (LSTM) as the core, combine 3σ criteria and data processing scheme, build the net load forecasting model in distribution grid planning based on the LSTM network. The whole model includes three parts, namely data processing, load forecasting, and parameter optimization. The detailed forecasting steps of the model are as follows:

Step 1: Distribution network operation data processing

As long as the grid planning of the distribution network is completed according to the historical load data information, other operation data, grid structure, and transmission line power supply range data of the area to be planned, due to the large amount of data required and the large number of categories, it is necessary to process the data after integrating these data, to avoid the LSTM network from, Mode collapse and gradient disappearance occur.

In the data processing part, to ensure the effectiveness of data processing, this paper selects the 3σ criterion and the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) method for processing 0historical load data. First, the 3σ criterion is used to detect the singularity of the load data, and the detection results are modified to obtain the load time series. The maximum value of the load is extracted, and reasonable maximum value time series results for each day are constructed. The CEEMDAN method is then used to decompose these results into multiple IMF components.

Step 2: Load forecasting

Based on the IMF components obtained by decomposition in step 1, the LSTM prediction network model is constructed according to its characteristics, and the prediction results of each IMF component are linearly superimposed, that is, the net load prediction results are obtained.

Step 3: Parameter optimization

When the LSTM network is forecasting the net load, the initial learning rate of the network and the number of cell units directly affect the forecasting effect of the network. Therefore, in order to ensure the accuracy of the network for forecasting the net load, the improved Pelican optimization algorithm is used to optimize the above parameters, obtain the best parameter results, and use the best parameters in the LSTM network to forecast the net load, Get the best prediction results.

According to the above steps, the prediction can be completed. Before the net load prediction in the paper, in order to ensure the accuracy of net load prediction, it is necessary to process the historical load data, detect the singular values in the data, and correct them, therefore, in order to analyze the singular value detection effect of this method, the standard deviation σ is set to 0.15, and use the 3σ criteria to detect singular values in data and obtain the detection results of singular values, as shown in Figure 4.

After analyzing the test results in Figure 4, after the prediction method detects the singular values through the 3σ criterion, it can accurately determine the singular values in the data by calculating the residual error ξ_p results of the data, and the detection results can clearly present the distribution of singular values, i.e., whether they are continuous singular values or single-point singular values, which avoids the impact of distribution grid planning on the net load prediction. Corrections are made for different singular values to ensure the reliability of the data and to provide accurate data sources for subsequent forecasts.

After the singular value processing of the historical load data, in order to improve the forecasting efficiency of the net load, the processed historical load data is decomposed to obtain the IMF component results of each time series. To verify the decomposition effect of this method, the orthogonality index μ_{IO} of each IMF component and the original signal is selected in this paper as an evaluation indicator, applying the criterion of less than 0.001. The smaller the value μ_{IO}, the better the decomposition accuracy of the method. The formula for μ_{IO} is:

$$\mu_{\mathrm{IO}} = \sum_{n=1}^{N}\left\{\sum_{j=1}^{n+1}\sum_{i=1,i\neq 1}^{n+1}\frac{\chi_{\mathrm{IMF}j}(t)\chi_{\mathrm{IMF}i}(t)}{[x(t)]^2}\right\} \tag{21}$$

Table 1. Details of relevant parameters.

Parameter category	Parameter	Numerical value
LSTM network	Number of neurons	50
	Number of cell units	136
	Initial learning rate	0.0101
	Maximum number of cycles	220
Improved pelican optimization algorithm	Population size	35
	Maximum Number Of Iterations	200
	Solving step size	10
	Output length	1
CEEMDAN method	Decomposition mode number	12
	Penalty factor	3000
	Noise tolerance	0
	Convergence criterion	1×10^{-7}
	Initial center frequency	0
	DC Component	0

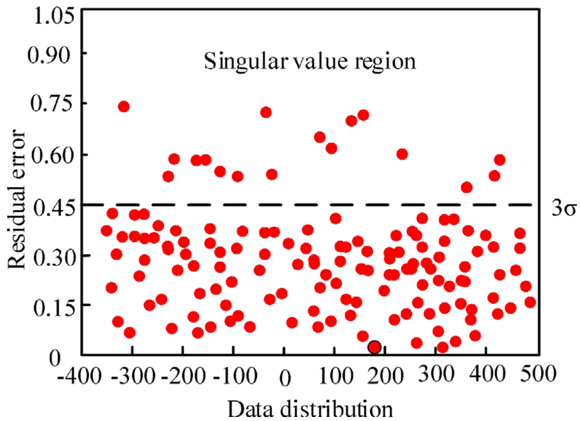

Fig. 4. Singular value detection results in historical load data.

where: $\chi_{\mathrm{IMF}j}(t)$ and $\chi_{\mathrm{IMF}i}(t)$ represent the decomposition error of IMF components i and IMF components j respectively. $x(t)$ represents the original data. N indicates the amount of data.

After calculating the data decomposition according to this formula, the orthogonality index μ_{IO} results of each IMF component and the original signal, due to the limited space, the test results of only 10 IMF components are randomly presented, as shown in Table 2.

After analyzing the test results in Table 2, it is concluded that under different data volumes, the orthogonality index μ_{IO} of each IMF component and the original signal obtained after the decomposition processing with the method in the paper, the results are all below 0.001, and the maximum value is 0.00092. Therefore, the method in the paper can complete the data decomposition well, and the decomposed components are smooth without aliasing, which provides a reliable basis for the subsequent net load prediction. This is because the method used in this article adopts the CEEMDAN method. CEEMDAN is an improved EMD technique that effectively solves the modal aliasing problem in traditional EMD methods by adding limited white noise to the original signal to assist in the decomposition process. CEEMDAN can adaptively decompose complex signals into a series of IMF components with different frequency scales, which are more stable and independent of each other.

In order to verify the effect of the prediction method in the distribution network grid planning for the net load prediction, in the acquisition of data, to obtain the peak load and the valley load of the two states of the operating data, using the method in the paper on the net load prediction results, to obtain weekdays and weekends in the net load test results, as shown in Figure 5.

After analyzing the test results in Figure 5, it is concluded that the net load prediction on weekdays and weekend days using the method in the paper can realize the net load results in both peak load and valley load states, and the trends of peak load and valley load on weekdays are close to each other, showing regular changes on the whole, while the changes of peak load and valley load on weekend days show irregular fluctuation changes. The prediction errors of the net loads on weekdays and weekends are below 1.25%, which is high and meets the application requirements. This is because the method in the article combines the adaptability of LSTM networks, refined data decomposition, modeling of differences between weekdays and weekends, and the application of optimization algorithms to achieve accurate prediction of peak and valley loads on weekdays and weekends, and shows different trends and prediction accuracy, meeting the application requirements of grid planning in distribution networks.

In order to visually verify the applicability of the method in the paper, the net load prediction is carried out through the method in the paper, and the distribution grid planning is carried out in the area based on the prediction results, and the planning results are obtained, and the results are compared with the planning results of the

Table 2. Test results of 10 IMF components.

IMF component	Historical load data volume/group		
1	200	600	1000
2	0.00074	0.00055	0.00074
3	0.00069	0.00076	0.00058
4	0.00081	0.00069	0.00073
5	0.00077	0.00081	0.00092
6	0.00086	0.00077	0.00086
7	0.00065	0.00084	0.00081
8	0.00073	0.00088	0.00078
9	0.00064	0.00062	0.00055
10	0.00089	0.00091	0.00091

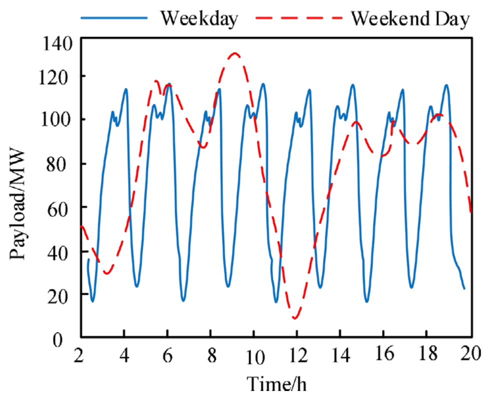

Fig. 5. Netload test results on weekdays and weekends.

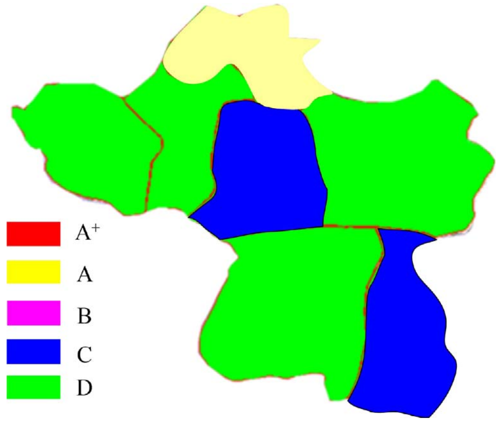

Fig. 6. Grid planning results of distribution network.

Table 3. Correlation coefficient experimental results.

Time stamp	Actual load (MW)	Benchmark model prediction (MW)	LSTM model prediction (MW)
2023/1/1	1200	1180	1205
2023/1/2	1250	1220	1248
2023/1/3	1180	1150	1182
2023/1/4	1320	1290	1318
2023/1/5	1270	1250	1272
...	...	...	...
2023/1/29	1150	1130	1155
2023/1/30	1290	1270	1292
2023/1/31	1300	1280	1295

current power supply area of the study area in Figure 4, to measure the applicability of the method in the paper, and the planning results are shown in Figure 6.

After analyzing the test results in Figure 6, it is concluded that the grid planning of the distribution network in the study area based on the prediction method in the paper can be combined with the net load prediction results to complete the grid planning of the distribution network, and the results of the planning can be accurately divided into different areas, which is a more comprehensive consideration of the power supply demand in the region, as well as the short- and long-term construction and development of the distribution network. This is because the grid planning of the distribution network involves multiple nodes, lines, and equipment, and there are complex spatial interactions. The method in the article can fully consider the spatial complexity and capture the mutual influence and linkage effects of load changes between different regions by constructing a prediction model based on the LSTM network. This allows for a more comprehensive consideration of the power supply demand and influencing factors inside and outside the area when dividing the area, thus obtaining more reliable experimental results.

To further verify the reliability of the proposed prediction method, experiments were conducted using the correlation coefficient (R) as an indicator. The benchmark model selected for the comparison group is the autoregressive integral moving average model. During the experiment, 30 days of test set data were selected to evaluate the predictive performance. Table 3 shows the experimental results of the predicted and actual values of some test set data.

After analyzing the test results in Table 3, it was found that the correlation coefficient between the predicted values of the benchmark model and the actual values is 0.87. The correlation coefficient between the predicted and actual values of the LSTM model is 0.96. From this, it can be seen that in complex and ever-changing load forecasting scenarios, benchmark models often find it difficult to accurately capture the dynamic characteristics and trends of load changes, resulting in low prediction accuracy. The LSTM model, through its unique gating mechanism and long-term memory ability, can effectively capture complex patterns and long-term dependencies in time series data, thereby providing more accurate prediction results. In this example, the correlation coefficient of the LSTM model is significantly higher than that of the benchmark model, indicating higher prediction accuracy. Therefore, the prediction results of the LSTM network are not only numerically closer to the actual load but also maintain high prediction accuracy when the load fluctuates greatly, which is of great significance for power grid scheduling and energy management.

5 Conclusion

Netload forecasting is the basis of distribution grid planning, and its forecasting accuracy directly affects the rationality, long-term, and development ability of the planning scheme. The distribution grid planning is divided according to the load density, which is determined according to the load forecasting results. In order to ensure the effect of distribution grid planning, this paper proposes a net load forecasting method in distribution grid planning based on the LSTM network. This method fully considers the demand and characteristics of distribution grid planning, uses the LSTM network model to forecast the net load in power supply units, and optimizes the parameters of the network model to obtain more accurate prediction results. According to the load forecasting results, the grid planning scheme of the distribution network can be adjusted to ensure the rationality of the grid planning scheme of the distribution network to the greatest extent.

The characteristics and patterns of time series data may vary in different application scenarios. Therefore, building a predictive model with good generalization ability and robustness to adapt to various application scenarios is a challenging task. Therefore, in the future, prediction and optimization systems that can process time series data in real-time or near real-time will be developed. These systems need to be able to quickly respond to changes in the external environment and automatically adjust prediction and optimization strategies to adapt to dynamically changing needs.

Funding

The authors received no financial support for the research.

Conflicts of interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and publication of this article.

Data availability statement

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Author contribution statement

Xinping Yuan contributed to the conception of the study, Ye Yuan wrote the first draft of the manuscript and worked on the coding of tables and figures. Haiyan Wang contributed to the conception and design of the study. Ming Tang and Mengyu Li helped perform the analysis with constructive discussions. All the authors read the manuscript and approved the final manuscript.

Ethics approval

No ethical approval was required as it did not involve the collection or analysis of data involving human or animal subjects.

References

1 Keck F., Jütte Silke, Lenzen M., Li M. (2022) Assessment of two optimisation methods for renewable energy capacity expansion planning, *Appl. Energy* **306**, Pt. A, 117988.
2 López M., Sans C., Valero S. (2022) Automatic classification of special days for short-term load forecasting, *Electr. Power Syst. Res.* **202**, 107533.
3 Chu Z., Lakshminarayana S., Chaudhuri B., Teng F. (2023) Mitigating load-altering attacks against power grids using cyber-resilient economic dispatch, *IEEE Trans. Smart Grid* **14**, 4, 3164–3175.
4 Gorka J., Roald L. (2022) Efficient representations of radiality constraints in optimization of islanding and de-energization in distribution grids, *Electr. Power Syst. Res.* **213**, 1–8.
5 Khosravi N., Echalih S., Baghbanzadeh R., Hekss Z., Hassani R., Messaoudi M. (2022) Enhancement of power quality issues for a hybrid AC/DC microgrid based on optimization methods, *IET Renew. Power Gener.* **16**, 8, 1773–1791.
6 Song X.H., Wag P., Niu D.X. (2023) Short-term power load forecasting based on heterogeneous data from multiple sources, *Comput. Simul.* **40**, 09, 59–65.
7 Ahajjam M.A., Licea D.B., Ghogho M., Kobbane A., Yan J. (2022) Experimental investigation of variational mode decomposition and deep learning for short-term multi-horizon residential electric load forecasting, *Appl. Energy* **326**, 492–506.
8 Evangelopoulos V.A., Georgilakis P.S. (2022) Probabilistic spatial load forecasting for assessing the impact of electric load growth in power distribution networks, *Electr. Power Syst. Res.* **207**, 107847.
9 Rafati A., Joorabian M., Mashhour E., Shaker H.R. (2022) Machine learning-based very short-term load forecasting in microgrid environment: evaluating the impact of high penetration of PV systems, *Electr. Eng.* **104**, 4, 2667–2677.
10 Kim J.H., Kim C.H., Lee B.S. (2023) A study on the development of long-term hybrid electrical load forecasting model based on MLP and statistics using massive actual data considering field applications, *Electr. Power Syst. Res.* **221**, 109415.
11 Ruíz-Guirola David E., Rodríguez-López Carlos A., Montejo-Sánchez Samuel, Souza R.D., López Onel L.A., Alves H. (2022) Energy-efficient wake-up signalling for machine-type devices based on traffic-aware long short-term memory prediction, *IEEE Internet Things J.* **9**, 21, 21620–21631.
12 Ahmad T., Zhang D. (2022) A data-driven deep sequence-to-sequence long-short memory method along with a gated recurrent neural network for wind power forecasting, *Energy* **239**, Pt. B, 122109.
13 Alizadeh Bidgoli M., Ahmadian A. (2022) Multi-stage optimal scheduling of multi-microgrids using deep- learning artificial neural network and cooperative game approach, *Energy* **239**, Pt. B, 122036.
14 Moutevelis D., Roldan-Perez J., Prodanovic M., Sanchez-Acevedo S. (2022) Bifurcation analysis of active electrical distribution networks considering load tap changers and power converter capacity limits, *IEEE Trans. Power Electr.* **37**, 6, 7230–7246.
15 Dumas J., Wehenkel A., Lanaspeze D., Cornélusse Bertrand, Sutera A. (2022) A deep generative model for probabilistic energy forecasting in power systems: normalizing flows, *Appl. Energy* **305**, 117871.
16 Vaishya S.R., Abhyankar A.R., Kumar P. (2023) A novel loss sensitivity based linearized OPF and LMP calculations for active balanced distribution networks, *IEEE Syst. J.* **17**, 1, 1340–1351.
17 Azeroual M., Boujoudar Y., Aljarbouh A., Moussaoui H.E., Markhi H.E. (2022) A multi-agent-based for fault location in distribution networks with wind power generator, *Wind Eng.* **46**, 3, 700–711.
18 Al-Qaness M.A.A., Ewees A.A., Fan H., Abualigah L., Elaziz M.A. (2022) Boosted ANFIS model using augmented marine predator algorithm with mutation operators for wind power forecasting, *Appl. Energy* **314**, 118851.
19 Bendaoud N.M.M., Farah N., Ahmed S.B. (2022) Applying load profiles propagation to machine learning based electrical energy forecasting, *Electr. Power Syst. Res.* **203**, 107635.