International Journal of Advanced Engineering Research and Science (IJAERS)
Peer-Reviewed Journal
ISSN: 2349-6495(P) | 2456-1908(O)
Vol-13, Issue-4; Apr, 2026
Journal Home Page Available: https://ijaers.com/
Article DOI: https://dx.doi.org/10.22161/ijaers.134.6

Research on Power Prediction Method for Distributed Photovoltaic Power Generation Systems Based on LSTM Optimized by Grey Wolf Optimizer

Kai-Ming Chen, Yan-Zuo Chang, Yan-Xiao Jia, Yu-Xuan Chen, Hong-Rui Yang, Wen-Min Wen, Zi-Rui He, Jie-Zhen Yang, Yong-Qing Wang, Zheng-Kuan Deng, Guan-Hong Xie

[1]School of Energy and Power Engineering, Guangdong University of Petrochemical Technology, Maoming, Guangdong 525000, China
[*]Corresponding author : 2785215630@qq.com

Received: 23 Feb 2026,

Received in revised form: 25 Mar 2026,

Accepted: 29 Mar 2026,

Available online: 08 Apr 2026

©2026 The Author(s). Published by AI Publication. This is an open-access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

***Keywords*— Distributed photovoltaic system, long short-term memory (LSTM), grey wolf optimizer (GWO), time series..**

***Abstract*— *Accurate prediction of the output power of distributed photovoltaic (PV) systems is crucial for achieving efficient renewable energy integration and ensuring stable grid operation. Given that the power output of distributed PV systems is significantly influenced by meteorological factors and exhibits strong randomness and volatility, this study takes a distributed PV system in a region of Guangdong Province as the research object and constructs a PV power generation calculation model based on real meteorological data and system parameters. For the prediction approach, traditional time series methods are first employed as a benchmark for comparison. Subsequently, a GWO-LSTM model is proposed, in which the Grey Wolf Optimizer (GWO) is used to optimize the hyperparameters of the Long Short-Term Memory (LSTM) model. The experimental evaluation employs mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE) as performance metrics. The results indicate that the MSE of the GWO-LSTM model is reduced by approximately 85% compared with traditional methods, while the RMSE and MAE are reduced to around 38% and 33% of those of the traditional methods, respectively. This model demonstrates significantly higher prediction accuracy than conventional time series approaches, verifying the effectiveness and superiority of using GWO to optimize LSTM hyperparameters in distributed PV power forecasting.*

I. INTRODUCTION

With the continuous advancement of global energy structure transformation and the "carbon peak and carbon neutrality" goals, solar energy, as a clean and renewable energy source, has seen its share in power systems steadily increase [1, 2]. Distributed photovoltaic (PV) generation systems have become a key direction in the development and utilization of renewable energy due to their advantages such as flexible installation and on-site consumption. However, the output power of PV generation is significantly influenced by meteorological conditions (e.g., irradiance, temperature, cloud cover, weather type), exhibiting strong randomness, intermittency, and volatility [3, 4]. When large-scale PV generation is integrated into the grid, such uncertainty poses serious challenges to the safe and stable operation of the power system, increasing the complexity of grid dispatch and the cost of reserve capacity [2]. Therefore, accurately predicting the output power of distributed PV systems holds significant theoretical value and practical importance for optimizing grid dispatch, enhancing renewable energy integration capability, and ensuring the economic and stable operation of power systems.

PV power forecasting methods can be classified into three categories based on technical approaches: physical methods, statistical methods, and hybrid methods [5]. Physical methods rely on numerical weather prediction (NWP) and physical models of PV systems, requiring precise system parameters and meteorological data, leading to high modeling complexity. Statistical methods, on the other hand, extract patterns from historical data, mainly including traditional time series analysis.

Traditional time series methods, such as autoregressive integrated moving average (ARIMA) and exponential smoothing (ES), perform forecasting by analyzing the autocorrelation and trends in historical power data. These methods are algorithmically simple and computationally efficient, but they struggle to capture the complex nonlinear relationships between meteorological factors and PV power output [7, 8]. Zhao et al., for instance, employed an ARIMA model to forecast the output power of distributed PV systems and analyzed the prediction errors under different weather types [3].

With the development of artificial intelligence technologies, deep learning methods have demonstrated significant advantages in the field of PV power forecasting. Long Short-Term Memory (LSTM), as an improved variant of Recurrent Neural Networks (RNNs), effectively addresses the vanishing gradient problem through the introduction of gating mechanisms and is capable of capturing long-term and short-term dependencies in time series [4, 6]. A review by Kumari et al. indicates that deep learning methods possess unique strengths in handling the nonlinear and dynamic aspects of PV power forecasting, with the Autoencoder-LSTM model achieving an R^2 score of up to 99.98% [1]. However, the predictive performance of LSTM models is highly dependent on the setting of hyperparameters (e.g., number of neurons, learning rate, batch size). Traditional empirical tuning or grid search methods are inefficient and prone to getting trapped in local optima [7, 8]. To optimize the selection of model hyperparameters, researchers have introduced swarm intelligence optimization algorithms. The Grey Wolf Optimizer (GWO), a novel metaheuristic algorithm, mimics the hunting behavior of grey wolf packs to achieve global optimization, characterized by fast convergence and the ability to avoid local optima [5, 9]. Existing studies have applied the combination of GWO and LSTM to PV power forecasting: Wang et al. proposed a GWO-LSTM model, achieving a prediction accuracy of 95% under extreme weather conditions [10]; Yang et al. constructed a hybrid XGBoost-LSTM-GWO model and introduced an extended uncertainty index (EUI) to evaluate prediction reliability, reducing RMSE by two orders of magnitude compared to conventional LSTM [7]; Liu et al. utilized GWO to optimize the hyperparameters of Bidirectional LSTM (BiLSTM), and further enhanced prediction accuracy and model robustness by incorporating feature selection and signal decomposition methods [5].

Based on the above analysis, current research on PV power forecasting still exhibits the following limitations: First, most studies rely on data from a single meteorological station or experimental validation in specific regions, leaving the generalizability of models across different regions and diverse climatic conditions untested. Second, while the effectiveness of the combined GWO-LSTM model has been demonstrated, existing research primarily focuses on improving prediction accuracy, with insufficient analysis of the model's predictive stability under extreme weather conditions. Third, relatively few studies address power forecasting specifically for distributed PV systems (as opposed to centralized PV power plants). The output characteristics of distributed systems differ from those of centralized ones, necessitating tailored modeling approaches.

To address these issues, this study takes a rooftop distributed PV system in a region of Guangdong Province as the research object and constructs a distributed PV power generation calculation model based on real meteorological data (temperature, cloud cover, ultraviolet intensity, sunshine duration, weather conditions) and PV system parameters. Regarding the prediction methodology, this study first employs traditional time series methods (moving average and exponential smoothing) as benchmarks for comparison. Second, an LSTM neural network model is constructed to leverage its capability to handle nonlinear features in time series for PV power forecasting. Finally, to address the challenge of LSTM hyperparameter selection, the Grey Wolf Optimizer is introduced to automatically optimize the number of neurons, learning rate, and batch size of the LSTM, establishing a GWO-LSTM hybrid prediction model.

By comparing the performance of the three methods on three evaluation metrics—mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE)—this study verifies the effectiveness and superiority of the GWO-LSTM model in distributed PV power forecasting, providing technical support for accurate prediction of distributed PV systems and optimization of grid dispatch.

II. CONSTRUCTION OF BASIC DATA FOR THE DISTRIBUTED PV POWER GENERATION SYSTEM ENERGY SUPPLY MODEL

This study generates photovoltaic power generation data using a formula-based approach, based on meteorological data and distributed photovoltaic system parameters from a specific region in Guangdong Province.The font size for **heading is 11 points bold face** and **subsections with 10 points and not bold.** Do not underline any of the headings, or add dashes, colons, etc.

2.1 Acquisition of Meteorological Data

Daily meteorological data for a location in Guangdong Province in 2025 were obtained from the China Meteorological Administration Meteorological Data Center and the official website of the Meteorological Bureau. The data include temperature (maximum/minimum), cloud cover, ultraviolet (UV) intensity, sunshine duration, and weather conditions (sunny, cloudy, overcast, rainy).

Data Preprocessing： Feature Engineering

Categorical Variable Encoding: Label encoding was applied to "weather conditions" to convert discrete categories into numerical values recognizable by the model.

Feature Splitting: The "temperature (maximum/minimum)" feature was split into two separate features: "maximum temperature" and "minimum temperature".

Data Standardization

Due to significant differences in the scales and ranges of the meteorological features (e.g., sunshine duration ranges from 5.0 to 10.0 hours, cloud cover ranges from 21 to 99), min-max normalization was employed to map all feature variables to the [0,1] interval using the following formula:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

where x represents the original feature value, and x_{min} and x_{max} denote the minimum and maximum values of that feature, respectively.

2.2 Parameter Settings of the Photovoltaic System

Calculation of Total Installed Capacity:

$$P_{\mathrm{total}} = S_{\mathrm{total}} \times \eta_{\mathrm{effective}} \times \rho_{\mathrm{power}} \tag{2}$$

The reference values for the parameters are as follows: $S_{\mathrm{total}} = 5000\ \mathrm{m}^2$; $\eta_{\mathrm{effective}} = 70\%$; $\rho_{\mathrm{power}} = 331.3\ \mathrm{Wp/m}^2$.

$$P_{\mathrm{total}} = 5000 \times 70\% \times \frac{331.3}{1000} = 1159.55\ \mathrm{kW}$$

Equivalent Full Load Hours:

$$H_{\mathrm{equivalent}} = H_{\mathrm{sunshine}} \times 0.85$$

0.85 is an empirical statistical value representing irradiance utilization efficiency

Comprehensive Correction Coefficient:

$$K_{\mathrm{comprehensive}} = K_{\mathrm{cloud}} \times K_{\mathrm{temperature}} \times K_{\mathrm{UV}} \times K_{\mathrm{weather}} \tag{3}$$

$$K_{\mathrm{temperature}} = 1 - 0.0034 \times (T_{\mathrm{average\ daily}} - 25)(\text{temperature coefficient for monocrystalline silicon modules})$$

The values of K_{cloud}, K_{weather}, and K_{UV} were established based on measured data from distributed PV projects in the Guangdong region.

2.3 Calculation of Power Generation

The formula for calculating photovoltaic power generation is as follows:

$$E = P_{\mathrm{total}} \times H_{\mathrm{equivalent}} \times K_{\mathrm{comprehensive}} \tag{4}$$

III. PHOTOVOLTAIC POWER FORECASTING BASED ON TRADITIONAL TIME SERIES

Traditional time series algorithms are among the earliest forecasting techniques applied in the field. This approach builds mathematical models to extrapolate future power demand by identifying trends, seasonality, and periodic patterns in historical power generation data based on their statistical properties. Traditional time series methods mainly include moving average, exponential smoothing, and autoregressive integrated moving average (ARIMA) models.

The moving average (MA) method uses the arithmetic mean of the data from the most recent 7 days and 30 days, respectively, to forecast the power generation for the next day:

$$\hat{y}_{t+1} = \frac{1}{7}\sum_{i=0}^{6} y_{t-i} \tag{5}$$

$$\hat{y}_{t+1} = \frac{1}{30}\sum_{i=0}^{29} y_{t-i} \tag{6}$$

The exponential smoothing (ES) method employs simple exponential smoothing, assigning exponentially decreasing weights to historical data:

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t \quad (7)$$

Here, $\alpha = 0.3$ is the smoothing coefficient, which controls the weight assigned to recent data.

Mean squared error (MSE) is used to evaluate the performance of the fitted functions:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (8)$$

The forecast curves for power generation were fitted using the above methods, as shown in the figures.

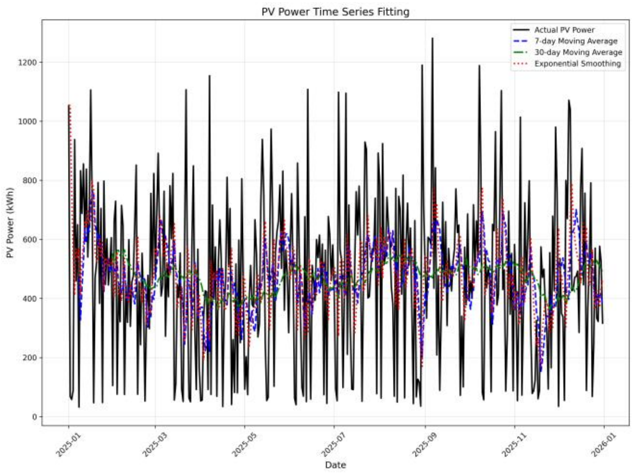

Fig. 1: PV Power Time Series Fitting

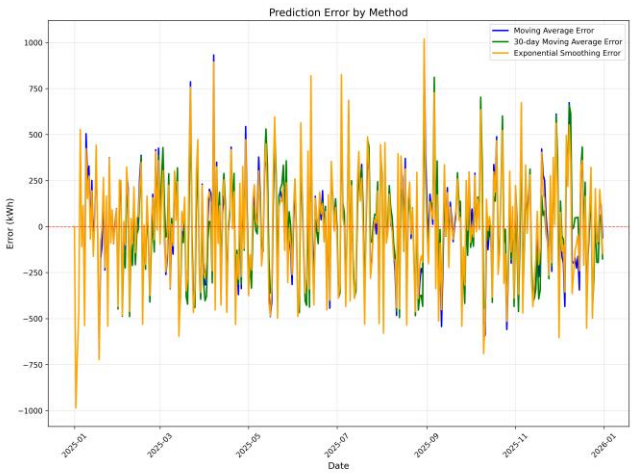

Fig. 2: Prediction Error by Method

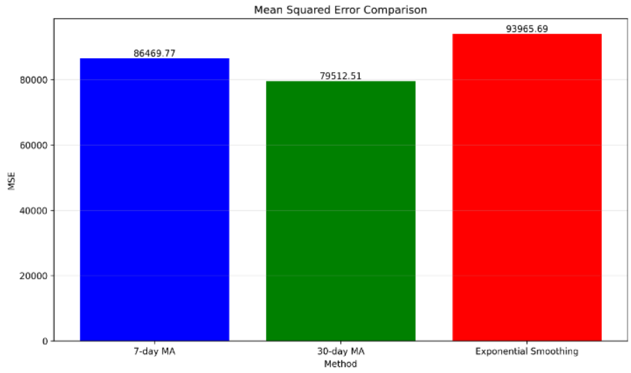

Fig. 3: Mean Squared Error Comparison

Based on the analysis of the figures, the time series fitting results for PV power in Fig. 1 show that the actual PV power fluctuates dramatically, with a wide range from 0 to 1300 kWh, exhibiting numerous peaks and valleys while also displaying a certain degree of periodicity. This reflects the generation pattern of distributed PV systems, which is centered around the daily sunshine cycle, with high power output during the day and almost no generation at night. Fig. 1 visually demonstrates the strong randomness and intermittency of PV power influenced by factors such as weather, seasons, and day-night cycles. The 7-day moving average method provides short-term smoothing of the actual power, preserving weekly trend variations while significantly narrowing the fluctuation range, thereby clearly reflecting short-term power fluctuations. The 30-day moving average method achieves the highest degree of smoothing, retaining only the long-term monthly trend, remaining relatively stable within the range of 400 to 600 kWh, and nearly eliminating short-term noise. The smoothing effect of exponential smoothing falls between that of the 7-day and 30-day moving averages; by assigning higher weights to recent data, it responds more quickly to trend changes, but its smoothing effect is weaker than that of the 30-day moving average. From the comparison of MSE values in the bar chart of Fig. 3, a smaller MSE indicates a smaller deviation between the fitted and actual values and thus better fitting performance. Among the three methods, the 30-day moving average achieved the best performance with an MSE of 79,512.51, followed by the 7-day moving average with an MSE of 86,469.77, while exponential smoothing exhibited the poorest fitting performance with an MSE of 93,965.69. However, the MSE values for all three fitting methods exceeded 70,000, and they all suffered from insufficient capability to capture peaks and prediction lag. The three forecast curves in Fig. 2 show numerous instances of extreme errors in the prediction error plot, indicating that these traditional time series forecasting methods are inadequate for achieving accurate predictions on this dataset, with poor adaptability.

The aforementioned traditional methods have significant drawbacks in forecasting power generation. The algorithms only consider the single variable of historical power generation during the forecasting process and fail to incorporate exogenous variables such as temperature and weather conditions, making them ill-suited for microgrid scenarios with substantial fluctuations. Consequently, this study attempts to introduce multivariate algorithms to address the limitations of traditional time series methods in terms of adaptability.

IV. PHOTOVOLTAIC POWER FORECASTING MODELS AND METHODS

4.1 Long Short-Term Memory (LSTM)

LSTM is a variant of Recurrent Neural Networks (RNNs), proposed by Hochreiter and Schmidhuber in 1997, primarily designed to address the vanishing/exploding gradient problem that conventional RNNs encounter when processing long sequences, enabling the model to effectively capture long-range dependencies. Its core concept lies in the cell state and the gating mechanism, characterized by the ability to selectively remember or forget irrelevant information, resulting in superior performance on longer time series compared to standard RNNs. Theoretically, information that needs to be retained can be transmitted through the cell state of the model. Consequently, the impact of short-term memory on model training is mitigated, allowing information from earlier data to be conveyed to subsequent cell states. Through the gating structure, the LSTM is trained to understand which information needs to be retained or forgotten, and which information should be added or removed.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{9}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{10}$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{11}$$

where i_t is the output vector of the input gate; W_i is the weight matrix of the input gate; b_i is the bias term of the input gate; $\tilde{C}_t$ is the candidate cell state; tanh is the hyperbolic tangent activation function; W_c is the weight matrix for the candidate state; and b_c is the bias term for the candidate state.

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{12}$$

where C_t is the cell state at the current time step; C_{t-1} is the cell state at the previous time step; and $\odot$ denotes element-wise multiplication.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{13}$$

$$h_t = o_t \odot \tanh(C_t) \tag{14}$$

where o_t is the output vector of the output gate; W_o is the weight matrix of the output gate; b_o is the bias term of the output gate; and h_t is the hidden state at the current time step.

The algorithmic flow of LSTM is illustrated in Fig. 4. At time step t, the algorithm receives the cell state C_{t-1} and hidden state h_{t-1} from the previous time step, along with the current input x_t. First, h_{t-1} and x_t are passed through a sigmoid activation to generate the forget gate f_t, which is element-wise multiplied by C_{t-1} to control the retention of historical cell state information. Second, h_{t-1} and x_t are passed through sigmoid and tanh activations, respectively, to generate the input gate i_t and the candidate cell state $\tilde{C}_t$; i_t is element-wise multiplied by $\tilde{C}_t$ to filter the current new information, and the result is added to the forget-modified C_{t-1} to obtain the updated current cell state C_t. Finally, h_{t-1} and x_t are passed through a sigmoid activation to generate the output gate o_t; C_t is activated by tanh and then element-wise multiplied by o_t to produce the current hidden state h_t, which serves both as the current output and as the input for the next time step in the iterative process.

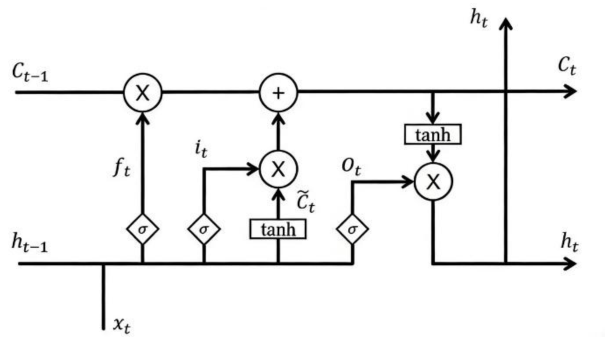

Fig. 4: LSTM Flowchart

4.2 Grey Wolf Optimizer (GWO)

The Grey Wolf Optimizer (GWO) is a swarm intelligence optimization algorithm proposed by Mirjalili et al. in 2014, based on the hunting behavior of grey wolf packs. It iteratively searches for the optimal solution by simulating the division of labor and collaboration among four hierarchical ranks: α, β, δ, and ω.Table 1 presents the core hierarchical ranks and role definitions of the Grey Wolf Optimizer.

Table.1: Core Hierarchical Ranks and Role Definitions of the Grey Wolf Optimizer

Rank	Symbol	Role Definition
Optimal solution	α	Global best position, dominates the search direction
Suboptimal solution	β	Assists decision-making, collaborates

		with α to guide the population
Third-optimal solution	δ	Follows α and β, participates in position updates
Other solutions	ω	Follows the top three ranks, performs position iteration

The convergence factor a decreases linearly with the number of iterations, achieving a smooth transition from global exploration to local exploitation:

$$a = 2 - t \cdot \frac{2}{T_{\max}} \tag{15}$$

where t is the current iteration number and $T_{\max}$ is the maximum number of iterations.

$$A = 2a \cdot r_1 - a, C = 2 \cdot r_2 \tag{16}$$

where A and C are coefficient vectors, and $r_1, r_2 \sim U(0,1)$ are random vectors.

Distance calculation:

$$D_\alpha = \mid C_1 \cdot X_\alpha - X \mid, D_\beta = \mid C_2 \cdot X_\beta - X \mid, D_\delta = \mid C_3 \cdot X_\delta - X \mid \tag{17}$$

Position update:

$$\begin{aligned} X_1 &= X_\alpha - A_1 \cdot D_\alpha \\ X_2 &= X_\beta - A_2 \cdot D_\beta \\ X_3 &= X_\delta - A_3 \cdot D_\delta \\ X(t+1) &= \frac{X_1 + X_2 + X_3}{3} \end{aligned} \tag{18}$$

V. DISTRIBUTED PHOTOVOLTAIC POWER FORECASTING MODEL BASED ON GREY WOLF OPTIMIZER OPTIMIZED LSTM (GWO-LSTM)

The training performance of LSTM neural networks is significantly influenced by the choice of hyperparameters. Selecting appropriate hyperparameters can enhance the prediction accuracy of the model. Currently, most hyperparameter selection processes rely on empirical tuning. By employing a swarm intelligence optimization algorithm to transform the selection of various hyperparameters into a multidimensional space optimization problem, the randomness associated with hyperparameter selection can be reduced. Therefore, this study proposes a distributed photovoltaic power forecasting model based on LSTM optimized by the Grey Wolf Optimizer (GWO-LSTM).

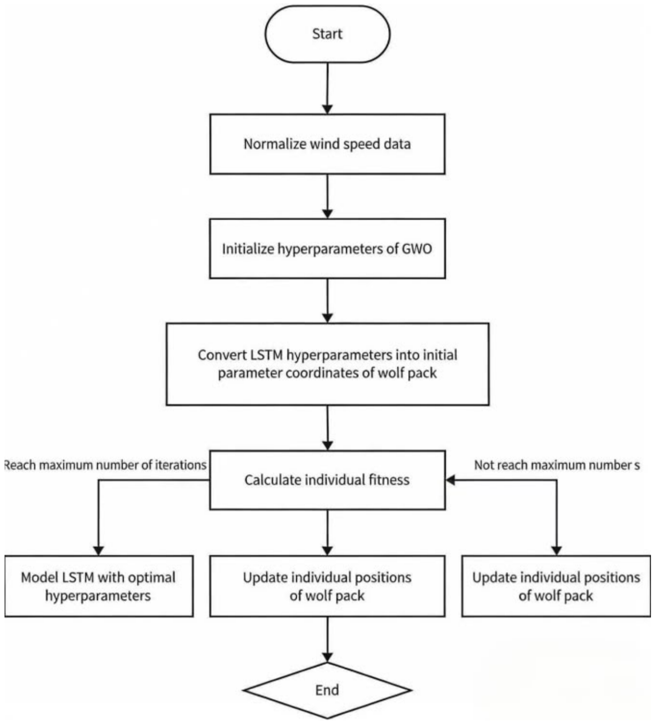

Fig. 5: GWO-LSTM Flowchart

The PV power and meteorological feature data are normalized and then divided into training, validation, and test sets according to a specified ratio.

The hyperparameters of the Grey Wolf Optimizer, such as the population size N and the maximum number of iterations, are set. The hyperparameters of the LSTM other than the number of neurons, learning rate, and batch size, such as the number of training epochs and dropout rate, are also configured.

The number of neurons, learning rate, and batch size of the LSTM are encoded as the position of a grey wolf. N grey wolf positions are randomly initialized, and N LSTM models are trained using the training dataset for a specified number of steps or until the optimal solution remains unchanged for 30 iterations.

The fitness of each grey wolf individual, corresponding to its LSTM model, is calculated on the validation set. In this study, the sum of squared errors between the predicted and actual values is used as the fitness function.

Based on the fitness values, the positions of the grey wolf individuals are updated using the GWO position update formulas. When the number of iterations reaches the maximum or the global optimal position meets the requirements, the optimal grey wolf position is output from the GWO algorithm, yielding the optimal combination of hyperparameters for the LSTM.

The test set is used to evaluate the GWO-LSTM model with the optimal hyperparameters, producing the PV power forecasting results.

The curve fitted by the GWO-LSTM model for the power generation of the distributed PV system is shown in Fig. 6, and the corresponding error curve is presented in Fig. 7.

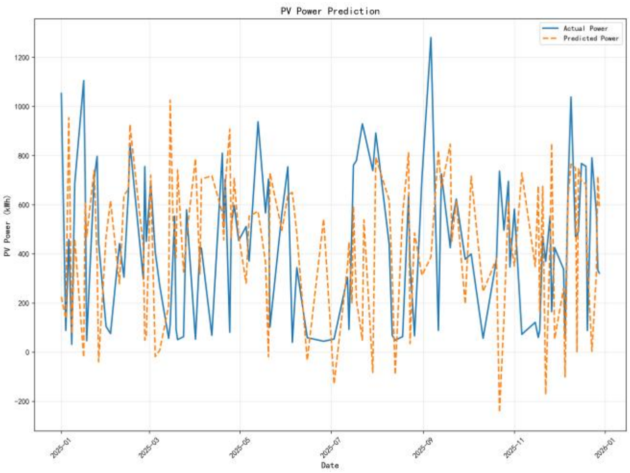

Fig. 6: PV Power Prediction

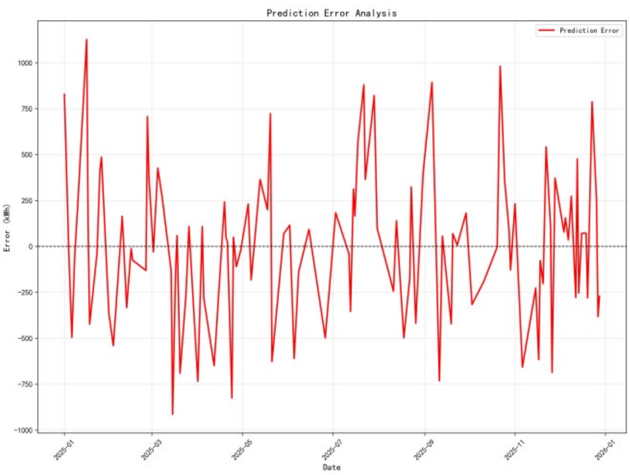

Fig. 7: Prediction Error Analysis

Analysis of the figures: Fig. 6, PV Power Prediction, illustrates the model's capability to fit the temporal sequence of PV power. The solid blue line represents the actual PV power, while the orange dashed line represents the model's predicted power. The predicted power generally follows the upward and downward trends of the actual power, indicating that the model effectively captures the periodic patterns of PV power. Based on this, it can be concluded that GWO-LSTM possesses a certain degree of effectiveness in learning long-term temporal dependencies. Fig. 7, Prediction Error Analysis, further quantifies the model's prediction deviation. The red curve represents the prediction error (defined as the difference between actual power and predicted power), and the black dashed line serves as the zero-error baseline. The error ranges between 1000 kWh and 1100 kWh, exhibiting substantial volatility, with errors at some time points approaching 100% of the actual power, suggesting lower prediction reliability under extreme scenarios. The errors are randomly distributed around the zero-error line without persistent positive or negative bias, indicating that the model's predictions are unbiased but exhibit high variance, resulting in insufficient overall stability.

VI. EVALUATION METRICS

To evaluate the prediction results of the models, three error analysis metrics are employed in this study to verify forecasting accuracy: mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE). Their calculation formulas are as follows:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{19}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{20}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{21}$$

where n is the total number of samples; y_i is the actual value of the i-th sample; $\hat{y}_i$ is the predicted value of the i-th sample; MSE is the mean squared error; RMSE is the root mean squared error; and MAE is the mean absolute error.

VII. DATA COMPARISON

Table 2 presents a comparison of prediction accuracy between different models and the proposed GWO-LSTM model.

Table.2: Comparison of Prediction Accuracy Between Different Models and GWO-LSTM

Model	MSE	RMSE	MAE
7-Day Moving Average	44,304.68	210.49	170.96
Exponential Smoothing	47,209.73	217.28	173.65
30-Day Moving Average	45,074.55	212.31	174.62
GWO-LSTM	6,546.39	80.91	56.82

From the comparative analysis of prediction accuracy among different models in Table 2, it can be observed that among the traditional time series forecasting algorithms, the 7-day moving average model achieves the highest prediction accuracy, followed by the 30-day moving average model, while exponential smoothing exhibits the weakest performance. In terms of the MSE metric, the error of the GWO-LSTM model is only approximately 14.8% to 15.4% of that of the traditional models (approximately one-seventh of the traditional models), indicating a substantial reduction in the sum of squared errors. For the RMSE metric, the error of the GWO-LSTM model is only about 38% to 39% of that of the traditional models, representing a reduction of over 60% in the average deviation between predicted and actual values. Regarding the MAE metric, the error of the GWO-LSTM model is only about 32% to 33% of that of the traditional models, roughly one-third of the traditional models. In summary, the GWO-LSTM model outperforms traditional time series methods in terms of prediction accuracy. This result fully demonstrates that the model combining the Grey Wolf Optimizer (GWO) with Long Short-Term Memory (LSTM) networks possesses stronger fitting capability and lower prediction error when handling such time series forecasting tasks, exhibiting superior performance advantages.

VIII. CONCLUSION

This study addresses the core characteristics of distributed photovoltaic (PV) power generation systems—namely, the significant influence of meteorological factors and the presence of strong randomness and volatility—by proposing a hybrid forecasting model that integrates the Grey Wolf Optimizer (GWO) with Long Short-Term Memory (LSTM) networks, referred to as GWO-LSTM. A systematic empirical analysis is conducted using a distributed PV system in a region of Guangdong Province as the research object.

In terms of research contributions, this study constructs a comprehensive PV power generation dataset based on daily meteorological data for a location in Guangdong Province in 2025, obtained from the China Meteorological Administration Meteorological Data Center, combined with the parameters of the distributed PV system itself. The Min-Max normalization method is employed for data preprocessing, effectively eliminating dimensional differences among the various feature variables. Additionally, the study systematically compares the predictive performance of three traditional time series forecasting methods against the proposed GWO-LSTM hybrid model, using mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE) as unified model evaluation metrics. Furthermore, by leveraging the Grey Wolf Optimizer to automatically optimize the key hyperparameters of the LSTM model, the hyperparameter selection problem is transformed into a global optimization problem in a multidimensional space, effectively overcoming the limitations of traditional empirical tuning and grid search methods, such as low efficiency and susceptibility to local optima.

Regarding the main findings, the experimental results demonstrate that the GWO-LSTM model significantly outperforms the traditional forecasting methods across all evaluation metrics. Specifically, the MSE is reduced by approximately 85% compared to the traditional methods, the RMSE is reduced to about 38% of that of the traditional methods, and the MAE is only one-third of that of the traditional methods. The average deviation between predicted and actual values is reduced by over 60%. These results fully validate the effectiveness and superiority of the proposed hybrid model in the task of PV power forecasting.

In terms of model advantages, the GWO-LSTM model possesses four core strengths. First, through its unique gating mechanism, LSTM effectively captures long-term and short-term dependencies in time series, enabling accurate learning of the complex nonlinear mapping relationship between PV power and meteorological factors. Second, the GWO algorithm, by simulating the hunting behavior of grey wolf packs, achieves a smooth transition from global exploration to local exploitation, characterized by fast convergence and the ability to effectively avoid local optima. Third, in contrast to traditional time series methods that rely solely on historical power data, the proposed model can incorporate multidimensional meteorological features such as temperature, weather conditions, and sunshine duration, significantly enhancing its adaptability in microgrid scenarios with substantial fluctuations. Fourth, the error analysis indicates that the prediction errors are randomly distributed around the zero-error line without persistent positive or negative bias, demonstrating the model's favorable unbiasedness.

Regarding limitations and shortcomings, although the GWO-LSTM model achieves significantly higher prediction accuracy than traditional methods, this study still has certain limitations. The prediction errors at some time points approach 100% of the actual power, indicating that there remains room for improvement in the model's prediction stability under extreme weather conditions. This study only utilizes meteorological data from a specific region in Guangdong Province for empirical analysis, and the model's generalizability across different regions and diverse climatic conditions has yet to be fully validated. The current model only incorporates basic meteorological features and does not account for more detailed meteorological variables such as cloud movement trajectories or aerosol concentrations, leaving room for further expansion of the feature dimensions. During the GWO optimization process, the LSTM model must be trained multiple times, and each fitness evaluation requires a complete training process, resulting in relatively high computational time.

In terms of future research directions, subsequent studies can be further explored in the following areas. First, explore the integration of GWO with other optimization algorithms, or introduce attention mechanisms and Transformer architectures to further enhance model prediction performance. Second, combine signal decomposition methods such as empirical mode decomposition (EMD) or variational mode decomposition (VMD) to decompose the original power sequence into multiple subsequences for separate prediction, thereby reducing the impact of sequence non-stationarity on model prediction effectiveness. Third, establish specialized prediction sub-models for extreme weather conditions such as typhoons, heavy rainfall, and prolonged cloudy weather, or introduce uncertainty quantification metrics to evaluate prediction reliability. Fourth, collect PV data from different climatic regions to validate the model's cross-regional generalizability and adaptability. Fifth, deploy the model proposed in this study in actual grid dispatch systems to achieve real-time online prediction updates, providing robust technical support for the accurate integration of distributed PV power and stable grid operation.

In summary, the GWO-LSTM hybrid forecasting model proposed in this study demonstrates significant performance advantages in the task of distributed PV power forecasting, effectively validating the feasibility and effectiveness of integrating swarm intelligence optimization algorithms with deep learning models. Although there remains room for further improvement, this study provides a feasible technical pathway and methodological reference for accurate forecasting of distributed PV systems and optimization of grid dispatch, holding both theoretical value and practical significance for promoting efficient renewable energy integration and ensuring stable power system operation.

ACKNOWLEDGEMENTS

This work was supported by the Research Funding of GDUPT, Research on Heat Transfer Enhancement of Heat Sink by Inverse Calculation Design Method (No. 2019rc074).

REFERENCES

[1]. Ahmed, R., Sreeram, V., Mishra, Y., & Arif, M. D. (2020). A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. Renewable and Sustainable Energy Reviews, 124, 109792.

[2]. Ozdemir, A., & Polat, K. (2025). Optimized CNN-LSTM with hybrid metaheuristic approaches for solar radiation forecasting. Heliyon, *11*(10), e42161.

[3]. Zhao, B. B., Wang, Y., Wang, B., Xuan, W. B., Lei, Z., Ge, L. J., & Xu, X. M. (2019). Research on output power prediction method of distributed photovoltaic system based on ARIMA time series. Renewable Energy, *37*(6), 820-823.

[4]. Kumari, S., & Singh, S. K. (2025). Exploring deep learning methods for solar photovoltaic power output forecasting: A review. Renewable Energy Focus, *53*, 100682.

[5]. Liu, Y., & Zhang, H. (2025). Short-term photovoltaic power prediction based on RF-SGMD-GWO-BiLSTM hybrid models. Energy, *316*, 134545.

[6]. Short-term photovoltaic power prediction based on GWO-CNN-LSTM-MATT hybrid model. (2025). In 2025 3rd International Conference on Power Electronics and Electrical Engineering (PEEE) (pp. 1-6). IEEE.

[7]. Wang, D., Zhang, H., Lin, X., Yu, Y., Zheng, J., & Zhu, J. (2025). A GWO-LSTM based approach for photovoltaic power generation prediction under extreme climate conditions. Journal of Physics: Conference Series, *3163*, 012009.

[8]. Research on photovoltaic power prediction method based on variational mode decomposition and LSTM optimization. (2025). In 2024 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA) (pp. 1-5). IEEE.

[9]. Yang, Y., Ding, S., & Wu, X. (2025). Optimization of solar power generation prediction based on the XGBoost-LSTM model and the grey wolf optimization algorithm. In 2025 4th International Conference on Power System and Energy Technology (ICPSET) (pp. 1–6). IEEE.

[10]. Dong, M., Li, X. F., Yang, Z., Chang, Y., Ren, M., Zhang, C. X., & Jiao, Z. B. (2024). Research progress of data-driven distributed photovoltaic power prediction methods. Power System and Clean Energy, *40*(1), 8-17+28.