Journal of Computer Science and Electrical Engineering
ISSN: 2663-1946
DOI: https://doi.org/10.61784/jcsee3033

POWER LOAD FORECASTING BASED ON THE PARTICLE SWARM OPTIMIZATION WITH BIDIRECTIONAL GATED RECURRENT UNIT NETWORKS

WenLi Tang[1*], YiZhou Fang[2]
[1]State Grid Xinyuan Anhui Xiangshuijian Pumped Storage Co.Ltd., WuHu, 241070, China.
[2]Xiaoxiang College of Hunan University of science and technology, Xiangtan, 411201, China.
Corresponding author: WenLi Tang, Email: 97163733@qq.com

Abstract: Accurate power load forecasting is pivotal in modern power systems, as it underpins efficient energy management, resource allocation, and grid stability. This paper introduces a novel hybrid forecasting model that integrates Particle Swarm Optimization (PSO) with Bidirectional Gated Recurrent Unit (Bi-GRU) networks to enhance predictive performance. The PSO algorithm is employed to systematically optimize the hyperparameters of the Bi-LSTM model, addressing challenges such as overfitting, convergence speed, and model complexity. By leveraging Bi-GRU's ability to capture bidirectional temporal dependencies and PSO's strength in global optimization, the proposed approach achieves significant improvements in forecasting accuracy. Experimental evaluations conducted on real-world power load datasets demonstrate the model's robustness and superior performance compared to standalone Bi-LSTM, PSO, and other traditional algorithms. The results highlight the potential of the PSO- Bi-GRU framework as a reliable and efficient tool for power load forecasting in complex and dynamic energy systems.
Keywords: Bidirectional gated recurrent unit; Particle Swarm Optimization; Load forecasting; Parameter optimization

1 INTRODUCTION

The widespread integration of renewable energy sources (RES) and stochastic, uncertain resources such as electric vehicles (EVs) into active distribution networks (ADNs) [1] has introduced significant challenges for power system operation and planning. These resources exhibit high variability and uncertainty [2], making it increasingly difficult to maintain grid stability and operational efficiency. Accurate power load forecasting has become an indispensable tool for addressing these challenges, as it enables proactive energy management, optimal resource allocation, and the reliable integration of distributed energy resources (DERs).
Traditional load forecasting techniques primarily rely on model-driven approaches, such as AutoRegressive Integrated Moving Average (ARIMA) [3], exponential smoothing [4], and regression-based models [5]. These methods leverage domain knowledge and mathematical formulations to predict future loads based on historical data and predefined assumptions. While effective in capturing linear trends and patterns, these techniques often struggle to account for the nonlinear, dynamic, and stochastic characteristics of modern power systems, especially in the presence of RES and EVs [5].
To overcome the limitations of traditional approaches, data-driven methods leveraging machine learning and deep learning techniques have gained traction [6]. Among these, Gated Recurrent Units (GRUs) and their bidirectional variant, Bidirectional GRU (Bi-GRU) [7], are highly effective in capturing temporal dependencies in sequential data. Bi-GRU improves on traditional GRUs by processing input sequences in both forward and backward directions, offering a more comprehensive representation of temporal features. Despite their advantages, Bi-GRU models require careful hyperparameter tuning, including learning rate, hidden layer size, and dropout rate [8]. Suboptimal hyperparameter settings can lead to poor performance, overfitting, or slow convergence, limiting the model's effectiveness.
Particle Swarm Optimization (PSO) offers a robust and efficient solution for hyperparameter optimization in Bi-GRU models. Inspired by the social behavior of particle swarms, PSO provides a global optimization mechanism that can efficiently explore and exploit the search space to identify optimal hyperparameters [9]. By integrating PSO with Bi-GRU, the combined framework can not only enhance the predictive accuracy of load forecasting but also improve model robustness and training efficiency. The feasibility of PSO lies in its simplicity, ease of implementation, and proven ability to handle complex, multidimensional optimization problems, making it an ideal choice for this task [10].
This paper proposes a hybrid PSO- Bi-GRU model for accurate power load forecasting in active distribution networks with a high penetration of stochastic resources. By leveraging Particle Swarm Optimization (PSO) to fine-tune the hyperparameters of the Bi-GRU network, the proposed approach effectively addresses the challenges of nonlinearity, temporal dependencies, and parameter optimization inherent in existing methods. Experimental results demonstrate the superiority of the PSO- Bi-GRU model in terms of forecasting accuracy and robustness, as validated through comparative analysis against conventional Bi-LSTM, PSO-based linear models, and other traditional algorithms.

2 METHODOLOGY

This section introduces the methodological framework of the proposed PSO- Bi-GRU model for power load forecasting. The primary objective is to enhance forecasting accuracy and robustness by leveraging the temporal modeling capabilities of Bi-GRU and the global optimization strengths of PSO. Section 2.1 explains the fundamentals of Bi-GRU, including its architecture and key hyperparameters that influence forecasting performance. Section 2.2 provides an overview of the Particle Swarm Optimization (PSO) algorithm, highlighting its suitability for hyperparameter optimization. Finally, Section 2.3 presents the integration of PSO and Bi-GRU, detailing the hybrid model's structure, optimization process, and workflow. Together, these components form a comprehensive solution for addressing the challenges of nonlinear and dynamic power load data forecasting.

2.1 Bidirectional Gated Recurrent Unit

Bidirectional Gated Recurrent Unit (Bi-GRU) is an advanced neural network architecture specifically designed to process sequential data and capture long-term dependencies in both forward and backward directions. Unlike traditional GRU networks, Bi- GRU consists of two separate GRU layers, one processing the input sequence in a forward direction and the other in reverse. The outputs of both layers are combined to provide a comprehensive representation of the sequence, making Bi-LSTM particularly suitable for tasks like power load forecasting, where bidirectional temporal dependencies are prevalent.
The structure is as shown in Figure 1[11].

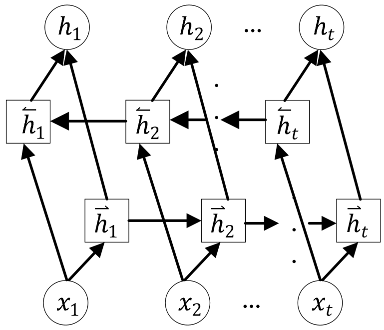

Figure 1 Structure of Bi-GRU

As shown in Figure 1, GRU is an improved version of recurrent neural network (RNN) [12], which controls the transmission of information through "reset gate" and "update gate", which helps solve the gradient disappearance problem of traditional RNN.
The "reset gate" determines the influence of the previous time step's state on the current state, calculated as follows:

$$\boldsymbol{r}_t = \sigma(\mathbf{W}_t \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \mathbf{b}_t) \tag{1}$$

The "update gate" computes the update $\mathbf{z}_t$ for the current time step *t*:

$$\mathbf{z}_t = \sigma(\mathbf{W}_z \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \mathbf{b}_z) \tag{2}$$

The candidate hidden state $\widetilde{\boldsymbol{h}}_t$ is given by:

$$\widetilde{\boldsymbol{h}}_t = \tanh\left(\mathbf{W} \cdot [\boldsymbol{r}_t * \boldsymbol{h}_{t-1}, \boldsymbol{x}_t] + \mathbf{b}\right) \tag{3}$$

The hide state $\boldsymbol{h}_t$ ris updated as:

$$\boldsymbol{h}_t = \boldsymbol{z}_t * \boldsymbol{h}_{t-1} + (1 - \boldsymbol{z}_t) * \widetilde{\boldsymbol{h}}_t \tag{4}$$

Key hyperparameters influencing Bi-GRU’s performance include:

- Hidden Layer Size: Determines the model's capacity to learn complex temporal features.
- Learning Rate: Controls the step size for weight updates, affecting convergence speed and stability.
- Dropout Rate: Reduces overfitting by randomly dropping connections during training.
- Batch Size: Impacts the computational efficiency and gradient updates during training.
- Manually tuning these hyperparameters is challenging and time-consuming, motivating the need for an efficient optimization algorithm such as PSO.

2.2 Particle Swarm Optimization

Particle Swarm Optimization (PSO) [13] is a population-based optimization algorithm inspired by the collective behavior of bird flocks and fish schools. In PSO, a swarm of particles explores the search space, where each particle represents a candidate solution (e.g., a set of Bi-GRU hyperparameters).

The particle's position and velocity are updated iteratively using the following equations:

$$v_i(t+1) = \omega \cdot v_i(t) + c_1 r_1 \cdot \left(p_i^{best} - x_i(t)\right) + c_2 r_2 \cdot \left(g^{best} - x_i(t)\right) \tag{5}$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{6}$$

Where, $x_i(t)$ and $v_i(t)$ are the position and velocity of particle i at iteration t; p_i^{best} is the particle's best-known position; g^{best} is the global best position among the swarm; ω is the inertia weight, balancing exploration and exploitation; c_1 and c_2 are acceleration coefficients; r_1 and r_2 are random values in [0,1].

PSO is well-suited for optimizing Bi-GRU hyperparameters, offering a systematic and efficient approach to explore the search space.

2.3 PSO-Bi-GRU Structure and Workflow

The PSO-Bi-GRU framework integrates PSO with Bi-GRU for power load forecasting. The workflow involves the following steps:

2.3.1 Initialization

Define the hyperparameter search space for Bi-GRU (hidden layer size, learning rate, dropout rate, and batch size). Initialize a swarm of particles, each representing a set of candidate hyperparameters.

2.3.2 Bi-GRU training

Train a Bi-GRU model for each particle's hyperparameters on the power load dataset and evaluate its performance using a fitness function (e.g., Mean Absolute Error).

2.3.3 PSO optimization

Update each particle's position and velocity based on its performance and the swarm's global best solution. Replace suboptimal hyperparameters with improved values.

2.3.4 Iteration

Repeat training and optimization until a convergence criterion is met (e.g., a maximum number of iterations or a target fitness value).

2.3.5 Final model selection

Select the Bi-GRU model with the best hyperparameter configuration from the PSO optimization process.

The PSO-Bi-GRU framework combines the temporal modeling strength of Bi-GRU with the optimization efficiency of PSO, resulting in a robust and accurate solution for power load forecasting.

3 SIMULATION

The simulations were performed on a standard desktop computer equipped with the following specifications: Processor: Intel Core i7-12700K (12 cores, 20 threads, base clock 3.6 GHz, max boost clock 5.0 GHz); RAM: 16 GB DDR4 3200 MHz; Storage: 1 TB NVMe SSD for fast read/write speeds; Operating System: Windows 11 64-bit; Software: MATLAB 2024b.

PSO-Bi-GRU model in forecasting power loads in a region of Hunan Province, which includes contributions from distributed wind turbines. The dataset used spans from January 1, 2020, to December 31, 2020, providing a comprehensive view of daily and seasonal variations influenced by renewable energy integration. The model was tasked with forecasting the load for December 20, 2020, using historical data.

The raw data underwent several preprocessing steps to ensure its suitability for model training and testing. First, normalization was applied using Min-Max scaling to transform the load values into the range [0,1], which helps improve the stability and convergence of the model. Additional temporal features, such as the day of the week and weather conditions, were engineered to capture seasonal and environmental influences on load patterns. The dataset was then divided into training and testing subsets, with data from January 1, 2020, to December 10, 2020, used for training, and data from December 11, 2020, to December 20, 2020, reserved for testing.

The performance of the PSO-Bi-GRU model was assessed using three widely adopted metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). RMSE quantified the model's ability to minimize large errors by penalizing them quadratically, MAE provided the average magnitude of prediction errors, and MAPE measured the model's accuracy as a percentage of the actual values. These metrics collectively offered a comprehensive evaluation of the model's forecasting precision and reliability.

To validate the superiority of the PSO-Bi-GRU model, its results were compared with baseline models, the results are shown in Figure 2 and Table 1.

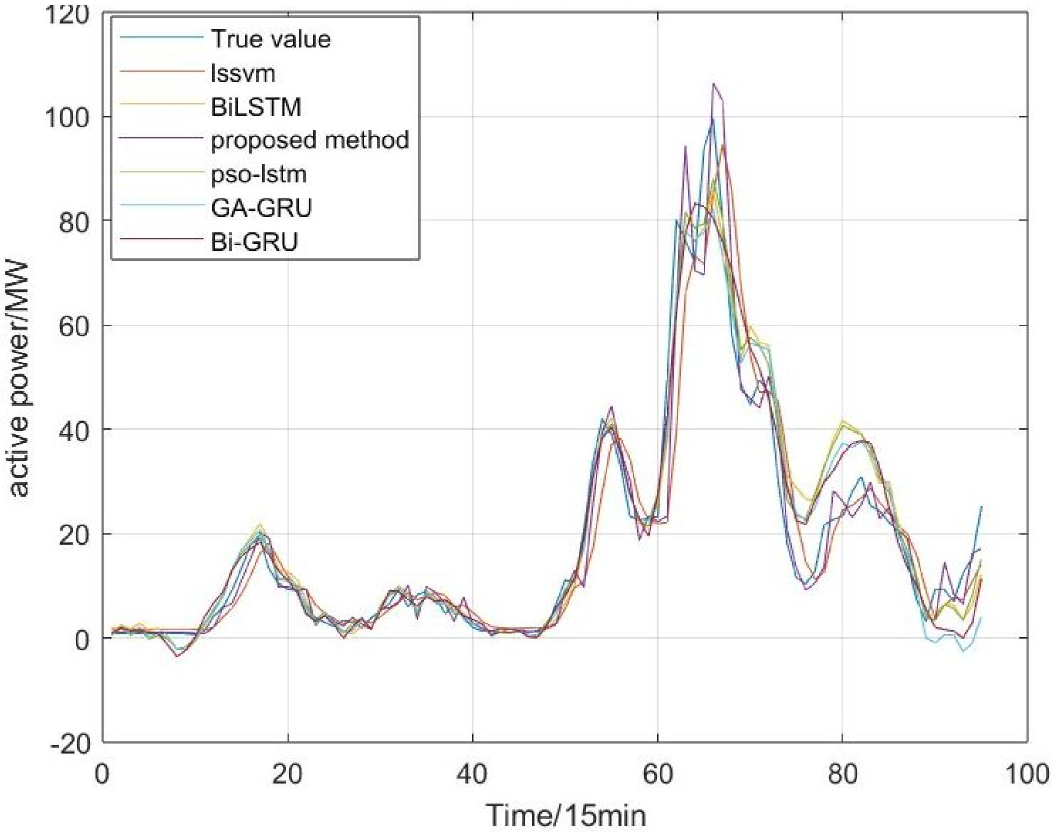

Figure 2 The Forecast Results in Different Methods

Table 1 Statistics in Different Methods

	RMSE	MAE	MAPE
LSSVM [14]	8.4282	4.9502	46.261%
Bi-LSTM [15]	6.5995	4.4555	49.127%
Proposed method	6.19	3.5991	27.4996%
Bi-GRU	5.8887	3.976	43.5609%
PSO-LSTM [16]	6.5512	4.3652	47.4303%
GA-GRU	6.1974	4.1818	48.2878%

The results presented in Table 1 demonstrate the effectiveness of different forecasting methods across three evaluation metrics: RMSE, MAE, and MAPE. The proposed PSO-Bi-GRU method shows a clear advantage in terms of overall performance, achieving the lowest MAE and MAPE values among all methods, with an RMSE that is competitive with the best-performing models. Specifically, the RMSE of the proposed method is 6.19, slightly higher than that of Bi-GRU (5.8887), which achieves the lowest RMSE. However, the difference is marginal and does not overshadow the significant improvements observed in the other two metrics. The MAE of the proposed method is the lowest at 3.5991, indicating that it provides more precise and consistent predictions compared to Bi-GRU (3.976) and all other methods. This demonstrates the model's ability to reduce the average magnitude of prediction errors effectively.

The MAPE results further highlight the superiority of the proposed method, with a value of 27.4996%, which is significantly lower than those of other models. For instance, Bi-GRU achieves a MAPE of 43.5609%, and traditional methods such as LSSVM and GA-GRU show MAPE values of 46.261% and 48.2878%, respectively. This stark contrast suggests that the proposed model performs exceptionally well in scenarios with large relative deviations, a common characteristic of power load data influenced by renewable energy and other stochastic factors. Although Bi-GRU outperforms the proposed method slightly in RMSE, its much higher MAPE indicates that it struggles with maintaining accuracy in percentage terms, particularly for smaller loads.

Overall, the proposed PSO-Bi-GRU model strikes a superior balance between all three metrics, delivering reliable and accurate predictions. While Bi-GRU shows a minor advantage in RMSE, the proposed method's lower MAE and MAPE emphasize its consistent and practical forecasting performance, making it a more robust choice for power load forecasting applications. These results validate the effectiveness of combining PSO for hyperparameter optimization with the bidirectional architecture of GRU, showcasing its ability to handle complex and dynamic energy system data efficiently.

4 CONCLUSION

In this study, a PSO-Bi-GRU model was proposed for accurate power load forecasting in active distribution networks. By integrating Particle Swarm Optimization for hyperparameter tuning with the bidirectional GRU architecture, the model effectively captured nonlinear temporal dependencies and optimized forecasting performance. Experimental results demonstrated the superiority of the proposed method, achieving the lowest MAE and MAPE while maintaining competitive RMSE compared to traditional and optimization-based models. These findings validate the PSO-Bi-GRU model as a robust and reliable tool for dynamic and complex energy system forecasting.

CONFLICT OF INTEREST

The authors have no relevant financial or non-financial interests to disclose.

REFERENCE

[1] Wang, C, Liu, C, Chen, J, et al. Cooperative planning of renewable energy generation and multi-timescale flexible resources in active distribution networks. Applied Energy, 2024, 356, 122429.
[2] Aeggegn, D B, Nyakoe, G N, Wekesa, C. A state of the art review on energy management techniques and optimal sizing of DERs in grid-connected multi-microgrids. Cogent Engineering, 2024, 11(1): 2340306.
[3] Zhong, W, Zhai, D, Xu, W, et al. Accurate and efficient daily carbon emission forecasting based on improved ARIMA. Applied Energy, 2024, 376, 124232.
[4] Sapnken, F E, Tazehkandgheshlagh, A K, Diboma, B S, et al. A whale optimization algorithm-based multivariate exponential smoothing grey-holt model for electricity price forecasting. Expert Systems with Applications, 2024, 255, 124663.
[5] Kashiri, S, Siahbalaee, J, Koochaki, A. Stochastic management of electric vehicles in an intelligent parking lot in the presence of hydrogen storage system and renewable resources. International Journal of Hydrogen Energy, 2024, 50, 1581-1597.
[6] Eren, Y, Küçükdemiral, İ. A comprehensive review on deep learning approaches for short-term load forecasting. Renewable and Sustainable Energy Reviews, 2024, 189, 114031.
[7] Alruqimi, M, Di Persio, L. Enhancing multi-step brent oil price forecasting with ensemble multi-scenario Bi-GRU networks. International Journal of Computational Intelligence Systems, 2024, 17(1): 225.
[8] Michael, N E, Bansal, R C, Ismail, A A A, et al. A cohesive structure of Bi-directional long-short-term memory (BiLSTM)-GRU for predicting hourly solar radiation. Renewable Energy, 2024, 222, 119943.
[9] Daviran, M, Maghsoudi, A, Ghezelbash, R. Optimized AI-MPM: Application of PSO for tuning the hyperparameters of SVM and RF algorithms. Computers & Geosciences, 2025, 195, 105785.
[10] Ullah, J, Li, H, Soupios, P, et al. Optimizing geothermal reservoir modeling: A unified bayesian PSO and BiGRU approach for precise history matching under uncertainty. Geothermics, 2024, 119, 102958.
[11] Y, Su, M, Tan, J, Teh, Short-Term Transmission Capacity Prediction of Hybrid Renewable Energy Systems Considering Dynamic Line Rating Based on Data-Driven Model. IEEE Transactions on Industry Applications, (early access). 2025. DOI: 10.1109/TIA.2025.3529824.
[12] Weerakody, P B, Wong, K W, Wang, G, et al. A review of irregular time series data handling with gated recurrent neural networks. Neurocomputing, 2021, 441, 161-178.
[13] Nayak, J, Swapnarekha, H, Naik, B, et al. 25 years of particle swarm optimization: Flourishing voyage of two decades. Archives of Computational Methods in Engineering, 2023, 30(3): 1663-1725.
[14] Liu, Y, Cao, Y, Wang, L, et al. Prediction of the durability of high-performance concrete using an integrated RF-LSSVM model. Construction and Building Materials, 2022, 356, 129232.
[15] Liu, F, Liang, C. Short-term power load forecasting based on AC-BiLSTM model. Energy Reports, 2024, 11, 1570-1579.
[16] Du, B, Huang, S, Guo, J, et al. Interval forecasting for urban water demand using PSO optimized KDE distribution and LSTM neural networks. Applied Soft Computing, 2022, 122, 108875.