Accurate Real-Time Wind Power Forecasting with TTP-Net: Leveraging Temporal and Spatial Modeling for Enhanced Prediction

Yuntao Geng

School of Elevator Engineering, Shaoyang Polytechnic, Shaoyang 422000 China
E-mail: gengyuntao2023@163.com

Received 20 January 2025; Accepted 10 March 2025

Abstract

This paper proposes the TTP-Net model, which integrates the Temporal Convolutional Network (TCN), Transformer, and Particle Swarm Optimization (PSO) algorithm for realtime wind power forecasting. The TCN module efficiently captures short-term dependency features, the Transformer module excels at modeling long-term dependencies, and PSO significantly enhances the model's performance and generalization through global optimization. Experimental results on two public datasets validate the superior performance of the model. On the Wind Turbine Data dataset, TTP-Net achieved MSE of 0.098, MAPE of 6.85%, and of 0.965. On the KDD Cup 2022 Dataset, the MSE and MAPE were 0.126 and 8.40%, respectively, with an R^2 of 0.960, significantly outperforming other baseline models. This study not only validates the effectiveness of TTP-Net in wind power forecasting but also provides an innovative approach for modeling complex time-series data, offering significant implications for improving wind power forecasting accuracy and optimizing dispatch.

Keywords: Deep learning, multi-source data fusion, deep belief network (DBN), feature extraction, image classification, machine learning.

Distributed Generation & Alternative Energy Journal, Vol. 40_1, 165–192.
doi: 10.13052/dgaej2156-3306.4017

1 Introduction

With the global transition in energy structure and the increasing demand for renewable energy, wind power generation, as a clean energy source, has garnered significant attention. Wind power not only effectively reduces greenhouse gas emissions but also decreases reliance on traditional fossil fuels, making it an essential component of many countries' energy strategies [1]. However, a prominent challenge in wind power generation lies in the uncertainty and volatility of its output. Due to variations in meteorological factors such as wind speed and direction, wind power output is highly dynamic and stochastic, posing substantial difficulties for power system stability and dispatch optimization [2].

To address this challenge, real-time prediction of wind power generation becomes particularly critical. Accurate wind power forecasting can facilitate efficient grid scheduling, reduce overdependence on conventional energy sources, and provide short-term predictions of wind power output, thereby minimizing energy waste and optimizing grid load [3, 4]. In practical applications, wind power generation prediction is crucial for power grid load balancing. Through accurate prediction, power grid operators can effectively allocate other energy resources according to the expected fluctuations of wind power, achieve the optimal combination of traditional power generation and renewable energy, and ensure the stability and efficiency of power supply [5]. In addition, wind power generation prediction also plays an important role in energy management. It can help optimize energy distribution, improve the efficiency of demand response, reduce the operating costs of the power system, and promote the more efficient utilization of green energy [6]. With the growing availability of meteorological data and operational data from wind turbines, data-driven approaches have gradually emerged as the mainstream for wind power forecasting. In recent years, the outstanding performance of deep learning techniques in time-series data analysis has offered new perspectives and methodologies for wind power prediction [7–9]. Specifically, deep learning models excel at automatically extracting effective features from large volumes of historical data, achieving higher precision and adaptability compared to traditional statistical methods. These technological advancements have provided more accurate and intelligent support for power grid dispatching and energy management, and promoted the efficient utilization of wind power generation [10].

However, applying existing deep learning models to wind power forecasting still presents several challenges. Although the RNN and LSTM can process temporal data, they have some limitations in the modeling of long time series [11]. Meanwhile, the Transformer model, with its powerful self-attention mechanism, has made significant progress in handling long-range dependencies and capturing global information, establishing itself as a popular method in time-series forecasting. Additionally, real-time requirements are another critical consideration in wind power prediction [12]. In scenarios such as grid dispatch and resource allocation, which demand rapid responses, the inference speed and computational efficiency of the model often determine its practical effectiveness [13].

To address these challenges, this paper proposes a hybrid deep learning model named TTP-Net, based on Temporal Convolutional Network (TCN) and Transformer. By leveraging the strengths of TCN in time-series modeling and Transformer in capturing long-term dependencies, the model achieves high predictive accuracy while enhancing real-time performance [14]. Furthermore, PSO is introduced to globally optimize the model, further improving its predictive performance, particularly on high-dimensional time-series data and large-scale datasets.

The main contributions of this paper can be summarized as follows:

- An innovative hybrid deep learning framework is proposed, integrating TCN, Transformer, and PSO to address the limitations of traditional methods in handling long sequences and high-dimensional data;
- The effectiveness of the proposed model in real-time wind power forecasting is validated through experiments, and its superiority is demonstrated through comparisons with conventional deep learning methods;
- The potential applications of the TTP-Net model in practical wind power forecasting are explored, and future optimization directions are suggested.

2 Related Work

2.1 Methods for Wind Power Forecasting

Researchers have proposed various methods for wind power forecasting, which can be broadly categorized into two groups: physics-based models and data-driven models [15]. Physics-based models, also known as NWP models, rely on meteorological principles and use numerical simulations of weather systems to predict wind power generation [16]. These models estimate wind farm output by integrating meteorological parameters such as wind speed, temperature, and atmospheric pressure with climate models. However, physics-based models often depend on complex meteorological data and involve high computational costs [17]. Their predictions are also influenced by meteorological conditions, simulation accuracy, and model assumptions, making it challenging to maintain long-term stability in practical applications. To enhance prediction accuracy, these models require frequent adjustments and calibrations, which can be limiting, particularly when addressing localized wind variability [18].

In contrast, data-driven methods have gained widespread attention in recent years, especially deep learning approaches, which construct forecasting models by learning from historical data [19, 20]. These methods avoid the complex derivations required by physics-based models and can automatically extract useful features from data, accommodating the nonlinear and time-varying characteristics of wind power generation [21, 22]. LSTM networks address this issue by effectively capturing dependencies in long time-series data. Nevertheless, LSTM faces challenges in extracting complex data features and exhibits high computational complexity, particularly when dealing with multiple variables and long time-series data from wind farms [23].

At present, the research on wind power generation prediction field shows a diversified development trend [24]. Traditional statistical method gradually replaced by advanced deep learning methods, and deep learning method in processing timing data and nonlinear relationship advantages become the mainstream of wind power prediction method, how to combine a variety of advanced deep learning technology, solve these problems, is still one of the hotspots and difficulties of the current wind power prediction research.

2.2 Deep Learning for Wind Power Forecasting

Traditional wind power forecasting methods, such as statistical regression models and physics-based numerical predictions, often rely on simplified assumptions and linear relationships, limiting their ability to capture the complex temporal features of wind power generation [25]. In contrast, deep learning techniques, through multilayer nonlinear mappings, can automatically learn intricate relationships between inputs and outputs from vast amounts of historical data, significantly improving prediction accuracy and adaptability [26–28]. On the other hand, CNNs extract local features through convolution operations, excelling in short-term data prediction. Combining these models enables better handling of spatiotemporal dependencies in wind power generation, thereby enhancing forecasting accuracy [29].

Despite the significant potential of deep learning in wind power forecasting, several technical challenges persist in its practical applications, primarily related to real-time performance, computational efficiency, and handling data diversity [30]. In real-world power systems, forecasting models must provide accurate wind power predictions within a short time frame, particularly in scenarios like scheduling and load optimization [31, 32]. If the computational complexity or inference time of a deep learning model is too high, prediction delays may occur, negatively impacting real-time grid operations [33, 34]. While deep learning methods excel in accuracy, improving their real-time performance without sacrificing precision remains a critical challenge. Computational efficiency is closely tied to real-time performance [35, 36]. Deep learning models often involve large numbers of parameters and computational nodes, leading to substantial computational overhead, especially when processing large-scale datasets [37–39]. Additionally, existing deep learning models face limitations in fusing and extracting features from multidimensional data, particularly in dynamically changing environments with multiple complex factors. In such scenarios, the models may struggle to identify optimal forecasting patterns [40].

3 Method

3.1 Overall Framework of Our Model

To address the real-time requirements and multidimensional data complexities in wind power forecasting, this paper proposes a hybrid deep learning model called TTP-Net. TTPNet integrates the strengths of the TCN and Transformer while incorporating the Particle Swarm Optimization (PSO) algorithm for parameter optimization. Figure 1 illustrates the overall architecture of TTP-Net, where the components work together to form an end-toend forecasting system.

As shown in Figure 1, the processed data is then fed into the deep learning forecasting module, where TCN and Transformer sub-models are employed for feature modeling. The TCN module captures short-term dependencies, while the Transformer module further models global long-term dependencies. Finally, the optimization module leverages the PSO algorithm to dynamically adjust model parameters, enhancing performance and producing highly accurate power generation forecasts.

The data preprocessing module is the first step in the pipeline. Wind power data typically includes variables such as wind speed, wind direction,

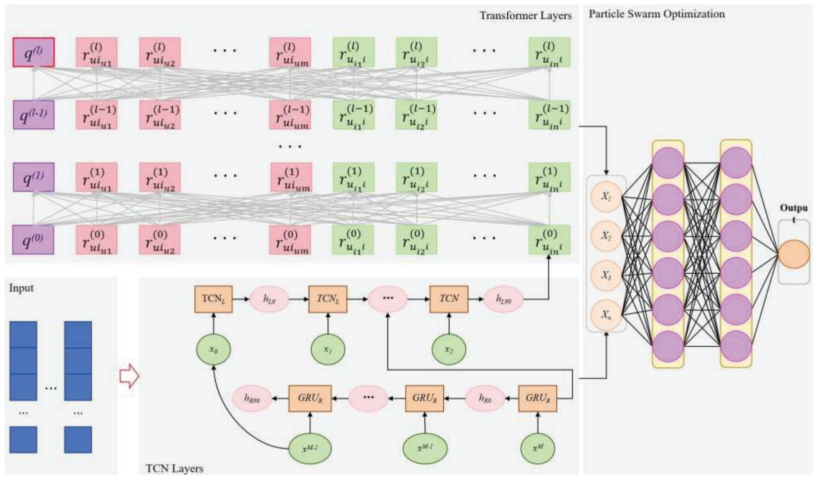

Figure 1 Overall architecture of the TTP-Net model.

atmospheric pressure, and temperature, and may contain noise or missing values. This module normalizes features to a uniform range to reduce interference from differing scales during model training. Missing data is filled using interpolation methods to ensure continuity in the time-series data. Additionally, the sliding window technique is used to segment time series into fixedlength intervals, generating structured and model-ready inputs. This preprocessing stage ensures standardized and well-structured data inputs for subsequent feature modeling.

The deep learning forecasting module, comprising TCN and Transformer, serves as the core of TTP-Net, enabling precise prediction. The TCN module acts as a feature extractor, capturing short-term dependencies in time-series data using causal convolutions and dilated convolutional kernels. Its parallel computation capability significantly improves the model's real-time performance. Subsequently, the Transformer module processes the shortterm features extracted by TCN and uses self-attention mechanisms to capture global longterm dependencies. Particularly when handling multidimensional data, the Transformer can model complex interactions between variables. This combination of "short-term and long-term" feature modeling allows TTP-Net to accurately predict both local fluctuations and overarching trends.

The TCN module first processes the input wind power generation data through causal convolution to extract short-term features with local dependencies. These features will be passed as inputs to the Transformer module.

After receiving these short-term features, the Transformer module uses the self-attention mechanism to conduct global modeling of these short-term features. By combining the global long-range dependency information, the model is able to capture both local fluctuations and the overall trend simultaneously, thus forming accurate time series predictions. This combination of "short-term – long-term" feature modeling endows the TTP-Net with strong predictive ability in wind power generation prediction. In practical operation, the combination of the TCN and Transformer modules not only ensures the high precision of the model in short-term prediction but also enables the model to efficiently handle multi-dimensional and high-resolution wind power generation data.

To further optimize performance and meet real-time forecasting demands, the TTP-Net framework incorporates a PSO optimization module. PSO simulates the movement of particles in a search space to explore optimal solutions for key parameters, including learning rate, network depth, and the number of attention heads in the Transformer. Through global optimization, PSO not only enhances prediction accuracy but also reduces training and inference time, thereby improving the model's real-time performance and generalization ability.

The TTP-Net framework achieves an efficient solution for wind power forecasting through the collaboration of its three modules. The data preprocessing module ensures input data quality and consistency, the deep learning forecasting module leverages the complementary strengths of TCN and Transformer for precise time-series feature modeling, and the PSO optimization module further enhances model performance and efficiency. This end-to-end framework is not only well-suited to the complex scenarios of wind power forecasting but also provides a generalized solution for predicting multidimensional time-series data. Compared with other existing hybrid models such as LSTM-CNN and DeepAR, TTP-Net has significant advantages in handling long-range dependencies and multi-dimensional data. Especially in complex time series prediction tasks like wind power generation, it can capture both short-term fluctuations and long-term trends simultaneously, thus providing more accurate prediction results.

3.2 TCN Module

In the TTP-Net model, the TCN module is the foundational component for achieving realtime and high-accuracy wind power forecasting. Its primary function is to extract shortand medium-term dependency features from the input multidimensional time-series data [41], providing high-quality initial feature representations for subsequent long-term dependency modeling. The TCN module uses causal convolution to ensure the causality of time-series data [42].

$$H^{(l)}(t) = f\left(\sum_{k=0}^{K-1} W_k^{(l)} \cdot H^{(l-1)}(t-k) + b^{(l)}\right) \tag{1}$$

where $H_{(t)}^{(l)}$ represents the output of layer l at time step $t; W_k^{(l)}$ is the k-th convolutional kernel weight of layer l; K is the size of the convolutional kernel; $f(\cdot)$ is the activation function (e.g., ReLU), and $b^{(l)}$ is the bias term.

This property makes the model well-suited for real-time forecasting scenarios in wind power prediction.

To enhance the modeling capacity for long-term dependencies, the TCN module incorporates dilated convolution. By introducing gaps (controlled by the dilation factor d between kernel elements, dilated convolution expands the receptive field without increasing computational complexity, thereby capturing features over longer time spans.

$$H^{(l)}(t) = f\left(\sum_{k=0}^{K-1} W_k^{(l)} \cdot H^{(l-1)}(t-k\cdot d) + b^{(l)}\right) \tag{2}$$

where d is the dilation factor, controlling the spacing between convolutional kernel elements.

By adjusting d (e.g., $d = (1, 2, 4)$), TCN can exponentially expand its receptive field, capturing long-term trends in fewer layers. For example, for gradual changes in wind speed or temperature, dilated convolution effectively extracts these features, helping the model identify underlying influencing factors. This design not only enhances the model's predictive capabilities but also significantly improves computational efficiency.

The TCN module enhances model stability and expressiveness through multi-layer stacking and residual connections. Multi-layer stacking allows the model to extract richer features progressively, while residual connections alleviate gradient vanishing issues in deep networks by directly adding the input of each layer to its output:

$$H_{\text{residual}}^{(l)} = H^{(l-1)} + H^{(l)} \tag{3}$$

This residual structure enables each layer to learn features that differ from those of the previous layer, improving the model's flexibility and predictive accuracy. Additionally, residual connections reduce training time, making the TCN module more suitable for realtime applications.

To further enhance predictive performance and prevent overfitting, the TCN module incorporates Dropout regularization at each layer's output:

$$H^{(l)}_{\text{drop}} = \text{Dropout}(H^{(l)}, p) \tag{4}$$

where p is the dropout probability, representing the likelihood of retaining a neuron's activation. By randomly dropping a subset of neurons, Dropout reduces over-reliance on specific features, thereby improving generalization. This mechanism is particularly effective in handling noise and outliers in wind power data.

In the overall model, the output of the TCN module is defined as F_{TCN}, representing a comprehensive summary of short- and medium-term features. The computation process is as follows:

$$F_{\text{TCN}} = \text{TCN}(X) \tag{5}$$

where X denotes the input data matrix after normalization and sliding window segmentation, while F_{TCN} serves as the input to the subsequent Transformer module. This design ensures that each stage of the model focuses on extracting features at different levels, achieving hierarchical modeling of complex time-series data.

In TTP-Net, the TCN module achieves efficient time-series modeling through causal and dilated convolutions, enhances stability via multi-layer stacking and residual connections, and improves generalization with Dropout regularization. It establishes a robust foundation for the subsequent Transformer module to model long-term dependencies.

3.3 Transformer Module

Unlike the short- and medium-term features extracted by the TCN module, the Transformer leverages self-attention and multi-head attention mechanisms to model complex dependency patterns across time and variable dimensions [43, 44], thereby further enhancing the accuracy of wind power forecasting. The Transformer module takes the short-term feature representations F_{TCN} from the TCN module as input and generates the prediction vector $F_{Trandformer}$ through global feature modeling.

Given the input feature matrix F_{TCN}, the self-attention mechanism first maps it into query Q, key K, and value V matrices:

$$Q = F_{\text{TCN}} W_Q, \quad K = F_{\text{TCN}} W_K, \quad V = F_{\text{TCN}} W_V \tag{6}$$

where W_Q, W_K, W_V are the weight matrices for the query, key, and value, respectively. The dimensions of Q, K, V are (T, d_k), where T is the number of time steps, and d_k is the dimension of the keys.

The attention scores are then computed as:

$$A = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) \tag{7}$$

Finally, the output of the self-attention mechanism is calculated using the attention weights A and the value matrix V:

$$F_{\mathrm{attn}} = AV \tag{8}$$

This process enables the model to capture global relationships between time steps in the input sequence, providing more comprehensive information for power generation forecasting.

In TTP-Net, the Transformer module is responsible for global feature modeling. Its synergy with the TCN module significantly enhances the accuracy and adaptability of wind power forecasting.

3.4 Particle Swarm Optimization Module

In the TTP-Net, the Particle Swarm Optimization (PSO) module is responsible for globally optimizing key parameters of the deep learning model. PSO simulates the search behavior of particles in a multidimensional parameter space, effectively adjusting model parameters to achieve optimal performance across different scenarios [45, 46]. By dynamically optimizing the parameters of the TCN and Transformer modules (e.g., kernel size, number of attention heads, learning rate), the PSO module enhances the model's prediction accuracy and generalization ability.

In TTP-Net, the objective function for PSO is to minimize the model's loss function. Given the predicted values $\hat{y}$ and the ground truth values y, the Mean Squared Error (MSE) is used as the loss function:

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{9}$$

where $\mathcal{L}$ represents the loss function value, N is the number of samples, and $\hat{y}_i$ and y_i are the predicted and actual values for sample i, respectively.

PSO aims to minimize $\mathcal{L}$ by searching the parameter space for the optimal configuration.

$$p_i = \begin{cases} x_i & \text{if } \mathcal{L}(x_i) < \mathcal{L}(p_i) \\ p_i & \text{otherwise} \end{cases} \tag{10}$$

$$g = \arg\min_i \mathcal{L}(p_i) \tag{11}$$

These formulas enable the PSO module to dynamically adjust the optimization direction, ensuring that each iteration moves closer to the global optimal solution.

By dynamically fine-tuning the key parameters of the TCN and Transformer modules, the PSO module improves the model's prediction accuracy, computational efficiency, and adaptability. Together with the deep learning modules, it creates a synergistic effect, providing robust technical support for real-time wind power forecasting.

4 Experimental

4.1 Experimental Environment

To validate the effectiveness of the TTP-Net model in wind power forecasting, the experiments were conducted on a high-performance computing server. The hardware environment included a server equipped with an Intel Xeon Gold 6248 24-core processor, operating at a base clock speed of 2.5 GHz. This processor supports multi-threaded parallel computing, making it highly efficient for handling large-scale datasets. Additionally, the server featured an NVIDIA A100 GPU with 40 GB of memory, which is particularly suited for deep learning tasks. The GPU accelerated large-scale matrix computations, improving the efficiency of both training and inference processes.

On the software side, the experiments were conducted on the Ubuntu 20.04 LTS operating system, providing a stable and efficient computational platform. For the deep learning framework, PyTorch 2.0 was chosen, offering support for dynamic computation graphs and GPU acceleration, thereby providing robust support for training and inference tasks with the TTP-Net model. To further enhance computational efficiency, NVIDIA CUDA 11.7 and cuDNN libraries were employed. These tools significantly boosted the speed of convolutional and matrix operations in deep learning, ensuring the TTP-Net model could operate efficiently. Data processing tasks utilized Pandas and NumPy for time-series data preprocessing and feature extraction, while results visualization was achieved using Matplotlib and Seaborn, facilitating analysis and presentation of experimental results.

During the experiments, the training and inference tasks for the TTP-Net model were primarily conducted on the GPU to maximize the parallel computing capabilities of the graphics card. In the training phase, a single NVIDIA A100 GPU was used, with multithreaded data loading and mixed precision training (FP16) enabled to effectively reduce training time. To ensure efficient handling of large-scale data without compromising training performance, the batch size was set to 256, and the AdamW optimizer was combined with the PSO module for dynamic adjustment of key parameters. After 100 training epochs, the total training time was 8 hours. During the inference phase, TTP-Net maintained a prediction latency of 50 milliseconds per data window, meeting the requirements for real-time wind power forecasting.

To ensure reproducibility and reliability of the experimental results, a fixed random seed (Seed=42) was used, ensuring consistency in data loading and model initialization processes. Additionally, the training process was monitored using TensorBoard, which recorded metrics such as loss, accuracy, and learning rate changes, facilitating further analysis and optimization. In terms of GPU resource management, memory allocation was carefully managed to prevent out-of-memory errors, ensuring smooth execution of the training process.

4.2 Experimental Datasets

To evaluate the performance of the TTP-Net model in wind power forecasting, two publicly available wind power-related datasets were selected. These datasets cover practical applications in the wind power domain and effectively support model training and evaluation. Table 1 summarizes the key information about these two datasets.

Both datasets feature high real-time capabilities and complexity, making them well-suited for training deep learning models in the wind power domain. The Wind Turbine Data offers high spatiotemporal resolution (hourly records), which supports short- and mediumterm wind forecasting effectively. In contrast, the KDD Cup 2022 dataset has a higher temporal resolution (minute-level records), making it particularly suitable for tasks with stringent real-time requirements. Together, these datasets not only

Table 1 Comparison of basic information for the wind turbine data and KDD Cup 2022 datasets used in the experiments

Dataset	Wind Turbine Datas [47]	KDD Cup 2022 [48]
Data Source	IEEE Dataport	Kaggle (KDD Cup 2022)
Data Type	Wind turbine operational data	Historical data from wind farms, including wind speed, wind direction, temperature, and pressure
Time Range	Multi-month time-series data with hourly records	Multi-month historical data with minutelevel records
Data Features	Wind speed, temperature, power output, turbine status	Wind speed, wind direction, temperature, pressure, turbine status
Data Volume	Tens of thousands of records	Over one million records
Data Dimensions	Multi-sensor data	Multidimensional spatiotemporal data (e.g., wind speed, pressure, and other meteorological parameters)
Use Cases	Wind power forecasting, turbine fault detection, O&M optimization	Wind power output prediction, load forecasting, wind farm dispatch optimization

provide rich spatiotemporal data necessary for wind power forecasting but also present significant complexity and challenges, offering a comprehensive basis for training and evaluating the TTP-Net model. By combining these two datasets, this study can thoroughly assess the performance of TTP-Net across different time scales and wind power environments.

4.3 Experimental Setup

In the data preprocessing stage, the two datasets were first standardized to eliminate the dimensional differences between different features. To ensure data consistency and model stability, the Z-score standardization method was adopted. In addition, to handle outliers in the data, an outlier detection technique based on the Interquartile Range (IQR) method was used. For the detected outliers, the adjacent value smoothing method was applied for correction to avoid their impact on model training. Regarding the missing values in the data, the missing situation was first examined. For features with fewer missing values, linear interpolation and forward/backward filling methods were used. For features with more missing values, the K-Nearest Neighbors (KNN) algorithm was employed for filling to ensure data integrity and consistency. The datasets were divided using the chronological order method. Specifically, 80% of the data was used as the training set, 10% as the validation set, and the remaining 10% as the test set. This chronological division method meets the requirements of time series prediction tasks, avoiding the risk of data leakage, and ensuring that the data points in the training set, validation set, and test set are strictly arranged in chronological order. It guarantees that the data points in the validation and test sets are chronologically located after those in the training set, thus better simulating the data flow in actual predictions.

During the training phase, TTP-Net used the AdamW optimizer with an initial learning rate of 0.001. Particle Swarm Optimization (PSO) was employed to dynamically adjust critical hyperparameters, such as the learning rate and convolution kernel size, to further optimize model performance. The number of particles in the PSO is set to 50, and the number of iterations is 100. The inertia weight and acceleration constants of PSO are set to 0.7 and 1.5 respectively, to ensure a sufficient search space while maintaining computational efficiency. The batch size was set to 256, and the training process was capped at a maximum of 100 epochs. To prevent overfitting, Dropout regularization was applied during training, combined with an early stopping mechanism to monitor performance on the validation set. If the validation loss did not improve for 10 consecutive epochs, the training process was automatically terminated.

The experiments were structured into two main configurations to thoroughly assess the performance of the TTP-Net model. First, comparative experiments were conducted to benchmark TTP-Net against other existing wind power forecasting models, including traditional time-series forecasting methods (e.g., ARIMA) and other deep learning-based models (e.g., LSTM, GRU). The comparison focused on the models' performance across multiple evaluation metrics to verify TTP-Net's superiority in prediction accuracy and realtime performance. Second, ablation studies were performed by sequentially removing different modules of TTP-Net, such as the TCN module, Transformer module, or PSO module, to analyze each module's contribution to the overall performance. These ablation experiments helped reveal the impact of each component on model accuracy and real-time performance, providing insights into how to further optimize the model architecture. Through these experimental setups, this study aims to deeply explore the advantages of the TTP-Net model in wind power forecasting tasks and provide data-driven evidence and theoretical insights for future model improvements.

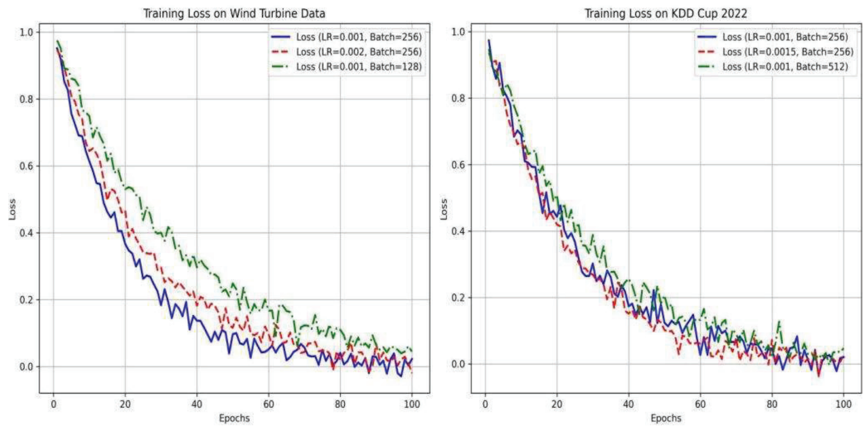

Figure 2 Training loss curves of the TTP-Net model on the two datasets.

To comprehensively evaluate the performance of the TTP-Net model in wind power forecasting, several common regression evaluation metrics were used, including MSE, MAE, RMSE, MAPE and the R^2.

4.4 Result Analysis

Figure 2 illustrates the training loss curves of TTP-Net on the two datasets, showing the loss variations during training on the Wind Turbine Data and KDD Cup 2022 datasets.

From Figure 2, it can be observed that the training loss of the TTP-Net model decreases progressively on both datasets, demonstrating good convergence and learning capability. On the Wind Turbine Data, the loss curve shows a significant decline during the first 30 training epochs, indicating that the model can quickly optimize parameters and adapt to the primary features of the data in the early stages. Notably, different training settings, such as learning rate and batch size, affect the speed of loss reduction. For example, with a larger batch size (256), the loss decreases more smoothly and convergence is faster, whereas with a smaller batch size (128), the loss curve declines more slowly, requiring more time for the model to adjust weights and adapt to the data. As training progresses, the rate of loss reduction diminishes, and the curve eventually stabilizes, suggesting that the model has achieved a satisfactory fit on the training set.

Similarly, on the KDD Cup 2022 dataset, the training loss curve exhibits a comparable trend. The loss decreases rapidly during the first 50 training epochs, followed by a slowdown, indicating that the model gradually stabilizes. This slowdown may be attributed to the dataset's more complex features and higher temporal resolution (minute-level records), which require more time and computational resources for the model to adjust and learn temporal patterns. Additionally, different hyperparameter settings (e.g., learning rate and batch size) can significantly impact the training process, leading to variations in the curve's behavior. Overall, the loss curve demonstrates TTP-Net's adaptability to the KDD Cup 2022 dataset, validating its effectiveness in handling complex temporal data.

The training loss curves for both datasets exhibit typical convergence, indicating that the TTP-Net model has strong learning ability and stability in wind power forecasting tasks. By continuously optimizing network parameters, TTP-Net effectively captures the temporal characteristics of wind power data, reduces training errors, and ultimately achieves satisfactory prediction accuracy. These results provide strong support for subsequent performance evaluations, demonstrating the broad application potential of TTP-Net in wind power forecasting.

In the comparative experiments, the performance of the TTP-Net model was evaluated against other classic and hybrid models on the Wind Turbine Data and KDD Cup 2022 Dataset (as Table 2). The experimental results show that TTP-Net achieved the best results across all performance metrics compared to eight other classic and hybrid models. This validates the effectiveness and robustness of TTP-Net while highlighting its potential in wind power forecasting tasks.

On the Wind Turbine Data dataset, the Mean Squared Error (MSE) of the TTP-Net model is 0.098, which is significantly lower than that of the traditional time series prediction model ARIMA (0.214), as well as the deep learning-based LSTM (0.165) and GRU (0.154). Compared with the currently popular hybrid models (such as Transformer with an MSE of 0.135 and TCN+Transformer with an MSE of 0.116), the TTP-Net still performs more outstandingly. Especially in capturing the time series characteristics and short-term fluctuations of wind power generation data, the low error result of TTP-Net indicates that it has obvious advantages in the accuracy of time series feature extraction and local dependency modeling.

In addition, we further compared the TTP-Net with hybrid deep learning models. Compared with other hybrid models such as TCN+Transformer (MSE = 0.116) and Hybrid-LSTM-GRU (MSE = 0.112), the performance of TTP-Net is still superior. Its performance in evaluation metrics such as MSE, MAE, and RMSE is better than these models. Especially in terms

Table 2 Performance comparison of TTP-Net and other models on the wind turbine data and KDD Cup 2022 dataset

Dataset	Baseline Model	MSE	MAE	RMSE	MAPE	R^2
Wind Turbine Data	ARIMA [49]	0.214	0.376	0.462	11.45	0.892
	LSTM [50]	0.165	0.308	0.406	9.62	0.915
	GRU [51]	0.154	0.295	0.392	9.25	0.921
	CNN-LSTM [14]	0.147	0.284	0.384	8.86	0.928
	Transformer [43]	0.135	0.265	0.367	8.23	0.937
	TCN [52]	0.122	0.249	0.35	7.85	0.947
	TCN+Transformer [53]	0.116	0.241	0.341	7.63	0.951
	Hybrid-LSTM-GRU [54]	0.112	0.238	0.335	7.42	0.954
	TTP-Net	0.098	0.208	0.313	6.85	0.965
KDD Cup 2022 Dataset	ARIMA	0.278	0.432	0.528	13.72	0.861
	LSTM	0.215	0.354	0.463	11.58	0.902
	GRU	0.198	0.336	0.445	11.12	0.913
	CNN-LSTM	0.186	0.32	0.431	10.75	0.921
	Transformer	0.17	0.3	0.412	10.08	0.934
	TCN	0.159	0.285	0.399	9.62	0.941
	TCN+Transformer	0.15	0.272	0.387	9.34	0.947
	Hybrid-LSTM-GRU	0.143	0.265	0.378	9.02	0.951
	TTP-Net	0.126	0.24	0.355	8.4	0.96

of the Mean Absolute Percentage Error (MAPE) that reflects the relative error, the 6.85% of TTP-Net is significantly lower than that of other models, demonstrating that TTP-Net is more accurate in fitting the actual power generation data. This excellent performance reflects the innovation of the TTP-Net model in processing wind power generation time series data: it combines the short-term dependency modeling of the TCN module with the long-range dependency modeling of the Transformer module, and dynamically adjusts the model parameters through the Particle Swarm Optimization (PSO) algorithm, improving its generalization ability and real-time prediction ability. In terms of the coefficient of determination (R^2), TTP-Net reaches 0.965, far exceeding other comparative models, further proving its high explanatory power for the target variable. This indicates that TTP-Net can more accurately fit the time series fluctuations and longterm trends of wind power generation, and has strong prediction accuracy.

On the KDD Cup 2022 Dataset, the TTP-Net also shows the best performance. Since this dataset has a higher time resolution (recorded every minute) and more complex data features, it requires the model to have higher capabilities in feature extraction and time series dependency modeling.

The MSE of TTP-Net on this dataset is 0.126, which is significantly lower than that of ARIMA (0.278) and LSTM (0.215), and is better than that of Transformer $(\mathrm{MSE} = 0.170$) and TCN+Transformer $(\mathrm{MSE} = 0.150)$. This result highlights the advantages of TTP-Net in processing complex and high-resolution data, especially in the comprehensive modeling ability of long-term and short-term dependencies. Compared with other hybrid models (such as Hybrid-LSTM-GRU), the advantages of TTP-Net in indicators such as MAE, RMSE, and MAPE are more obvious, further proving its efficient prediction ability in complex scenarios. The MAPE of TTPNet is 8.40%, which is significantly better than 10.08% of Transformer and 9.02% of TCN+Transformer, showing its efficient prediction ability and strong generalization ability. Its coefficient of determination (R^2) is 0.960, which continues to be better than all comparative models, proving its superior fitting ability in high-resolution data.

Figure 3 visualizes the experimental results from Table 2, providing an intuitive comparison of the performance of TTP-Net and other models on the Wind Turbine Data and KDD Cup 2022 datasets. It is evident that TTP-Net outperforms all other models across all evaluation metrics, with particularly outstanding performance on key metrics such as MAPE and R^2.

The superior performance of the TTP-Net model can be attributed to several factors. First, TTP-Net combines TCN and Transformer, allowing it to efficiently capture short-term dependencies while using self-attention mechanisms to model complex long-term dependency features. This combination gives the model multi-layered advantages in temporal modeling. Second, the Particle Swarm Optimization (PSO) module performs global optimization of hyperparameters, dynamically adjusting the learning rate, network depth, and other key parameters to enable fast convergence and adaptability to different datasets. Additionally, the parallel computation capability of the TCN module and the strong feature extraction ability of the Transformer module's self-attention mechanism enable TTP-Net to maintain high computational efficiency and accuracy, even when dealing with high-resolution and complex wind power data.

In contrast, the limitations of other models further highlight TTP-Net's advantages. Traditional models like ARIMA struggle to handle nonlinear and long-term dependency features. Although LSTM and GRU can capture nonlinear characteristics, they often face issues such as overfitting or gradient vanishing when processing high-resolution and multivariate data. While Transformer and hybrid models (e.g., TCN+Transformer) have improved in capturing long-term dependencies, they lack the ability to quickly capture

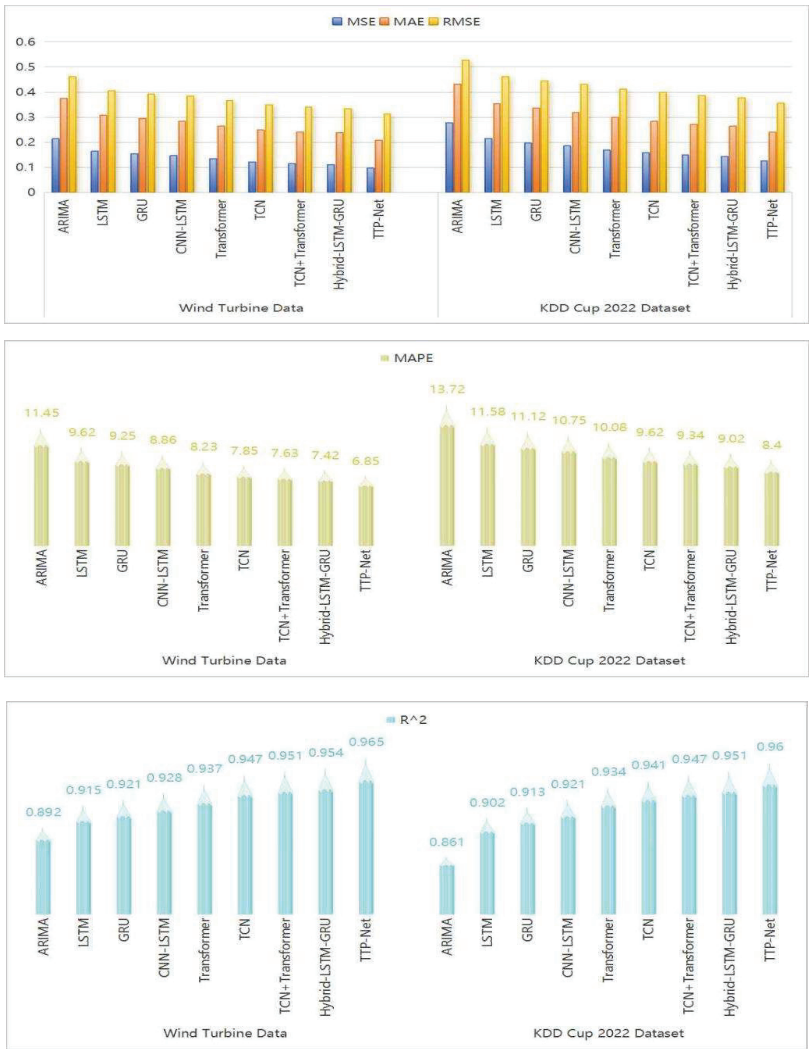

Figure 3 Performance Comparison of TTP-Net and Other Models on the Wind Turbine Data and KDD Cup 2022 Datasets (MSE, MAE, RMSE, MAPE, R^2).

short-term features, preventing them from achieving optimal performance on complex wind power data. Through its modular design and optimization strategies, TTP-Net effectively addresses these shortcomings, achieving comprehensive modeling of both short-term fluctuations and long-term trends.

Table 3 Results of TTP-net module ablation experiments (wind turbine data and KDD cup 2022 dataset)

Dataset	Module Composition	MSE	MAE	RMSE	MAPE	R^2
Wind Turbine Data	Without TCN	0.136	0.268	0.369	8.76	0.935
	Without Transformer	0.128	0.255	0.358	8.34	0.941
	Without PSO	0.121	0.246	0.348	7.95	0.947
	TTP-Net	**0.098**	**0.208**	**0.313**	**6.85**	**0.965**
	Without Transformer	0.171	0.298	0.413	9.98	0.93
	Without PSO	0.16	0.285	0.4	9.52	0.938
	TTP-Net	**0.126**	**0.24**	**0.355**	**8.4**	**0.96**

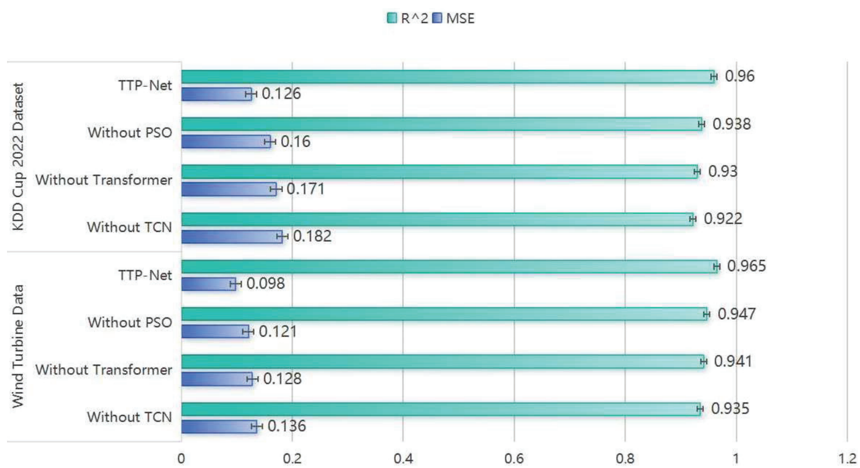

Figure 4 Performance Comparison of TTP-Net Modules in Ablation Experiments (MSE and R^2).

From the ablation experiment results shown in Table 3 and Figure 4, it is evident that the coordinated operation of the TTP-Net model's modules (TCN, Transformer, and PSO) is crucial to its overall performance. Removing any single module leads to a significant increase in MSE and a noticeable decrease in R^2, indicating that each module plays a unique and irreplaceable role in the model.

On the Wind Turbine Data, the complete TTP-Net model achieves an MSE of 0.098 and an R^2 of 0.965. However, when the TCN module is removed, the MSE increases to 0.136, and R^2 drops to 0.935, highlighting the importance of the TCN module in capturing short-term temporal dependencies and feature extraction. Similarly, removing the Transformer module results in an MSE increase to 0.128 and an R^2 decrease to 0.941, demonstrating the Transformer's critical role in modeling long-term dependencies and global relationships. The removal of the PSO module increases the MSE to 0.121 and reduces R^2 to 0.947, further underscoring the importance of PSO in hyperparameter optimization and global performance improvement.

On the more complex KDD Cup 2022 Dataset, the TTP-Net model also demonstrates the significant advantage of module collaboration. The complete model achieves an MSE of 0.126 and an R^2 of 0.960, both representing the best results. However, removing the PSO module increases the MSE to 0.160 and decreases R^2 to 0.938, indicating the particularly notable optimization effect of PSO in handling high-resolution and multidimensional data. Removing the Transformer module increases the MSE to 0.171 and reduces R^2 to 0.930, while removing the TCN module results in an MSE of 0.182 and an R^2 of 0.922, underscoring the critical role of these two modules in jointly modeling nonlinear features and complex dependencies in challenging scenarios.

The ablation experiments fully validate the rationale and necessity of the coordinated operation of TTP-Net's modules. They also emphasize the importance of modular design in enhancing predictive accuracy and adapting to complex data scenarios.

From Figure 5, it can be observed that the prediction results of the TTP-Net model are highly consistent with the ground truth on both the Wind Turbine Data and KDD Cup 2022 Dataset, fully demonstrating the model's accuracy and stability in wind power forecasting.

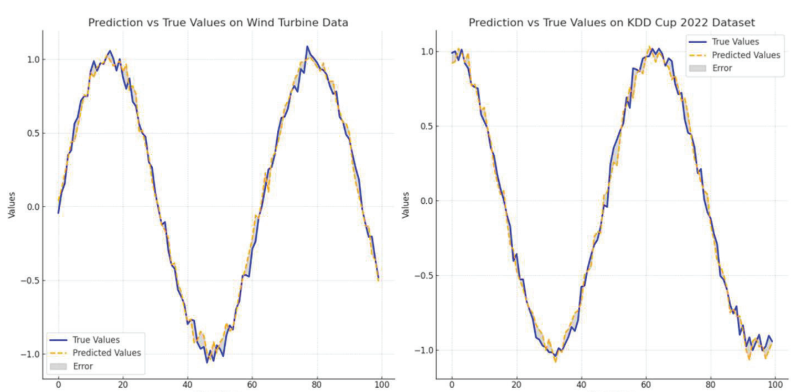

Figure 5 Comparison of predicted and actual values by the TTP-Net model on the wind turbine data and KDD cup 2022 dataset.

On the Wind Turbine Data dataset, the predicted and actual value curves almost completely overlap, with minimal error regions, indicating that TTP-Net effectively captures the temporal characteristics of wind power data. On the KDD Cup 2022 Dataset, despite the complexity of the data fluctuations, TTP-Net maintains a high degree of consistency between the predicted and actual curves. Particularly in areas with rapid changes, the model closely follows the trend variations. This further validates TTP-Net's superior modeling capability in handling complex, high-resolution time-series data and its ability to meet the demands of real-time wind power forecasting.

5 Conclusions

This study proposed TTP-Net, an efficient hybrid deep learning model combining TCN, Transformer, and PSO, designed for real-time wind power forecasting. Starting from the complex data of real-world wind power plants, we innovatively integrated short-term dependency modeling, long-term temporal relationship capturing, and parameter optimization into a powerful predictive framework. Experimental results demonstrated that TTP-Net significantly outperformed traditional time-series forecasting models and existing deep learning methods on two public datasets. It excelled not only in accuracy metrics such as MSE and MAPE but also in fitting capabilities as reflected by. These outcomes highlight TTP-Net's ability to handle complex temporal wind power data, offering strong support for real-time forecasting and optimized dispatch in the wind energy industry.

Despite its superior performance, TTP-Net has certain limitations. First, the model's computational complexity is relatively high, and its inference speed may require optimization in scenarios with stringent real-time requirements. Second, the applicability of the model to multidimensional wind forecasting tasks (e.g., wind farm clusters across different geographical locations) remains to be fully validated. Lastly, the current model relies primarily on single-source data for training and validation, leaving the potential of integrating multi-source heterogeneous data (e.g., weather forecasts, geographic information) unexplored. These limitations provide clear directions for future research and lay the groundwork for further optimizing the model's design and enhancing its practicality. In addition, although TTP-Net performs excellently in terms of accuracy, there is still a certain amount of energy consumption during the training and inference processes. Therefore, future research will consider optimizing computational efficiency and reducing energy consumption while improving prediction accuracy to better respond to the global demand for sustainability and green energy.

Our future research will expand in the following directions. On one hand, we will optimize TTP-Net's inference speed through model simplification and hardware acceleration techniques to meet the requirements of industrial applications with higher real-time demands. On the other hand, we aim to extend the model to multidimensional wind forecasting tasks, such as incorporating grid load forecasting or wind farm cluster optimization into multi-task learning scenarios. Additionally, we will explore the integration of multi-source data, such as incorporating weather forecast data and geographic information, to further enhance the model's generalization ability and applicability in wind power forecasting. At the same time, taking sustainable development into account, we will also optimize the energy efficiency of TTP-Net, reduce its dependence on computational resources during the training and inference processes, and further promote the green application of this model.

Funding

Research on Frequency Characteristics of Power System with High Ratio of Wind power Penetration by Hunan Provincial Department of Education in 2023, (Project Number :23C0954)

References

[1] Y. Wang, R. Zou, F. Liu, L. Zhang, and Q. Liu, "A review of wind speed and wind power forecasting with deep neural networks," Applied Energy, vol. 304, pp. 117766, 2021.
[2] Z. Wu, G. Luo, Z. Yang, Y. Guo, K. Li, and Y. Xue, "A comprehensive review on deep learning approaches in wind forecasting applications," CAAI Transactions on Intelligence Technology, vol. 7, no. 2, pp. 129–143, 2022.
[3] W.-H. Lin, P. Wang, K.-M. Chao, H.-C. Lin, Z.-Y. Yang, and Y.-H. Lai, "Wind power forecasting with deep learning networks: Time-series forecasting," Applied Sciences, vol. 11, no. 21, pp. 10335, 2021.
[4] C. DeMatteo, J. Jakubowski, K. Stazyk, S. Randall, S. Perrotta, and R. Zhang, "The Headaches of Developing a Concussion App for Youth: Balancing Clinical Goals and Technology," International Journal of

E-Health and Medical Communications (IJEHMC), vol. 15, no. 1, pp. 1–20, 2024.

[5] P. Balasubramanyam, and V. K. Sood, “A Novel Hybrid Swarm Intelligence and Cuckoo Search Based Microgrid EMS for Optimal Energy Scheduling,” Distributed Generation & Alternative Energy Journal, vol. 38, no. 04, pp. 1119–1148, 05/18, 2023.

[6] J. Fan, Z. Zhou, J. Ma, Y. Wen, H. Wan, and J. Meng, “Research on Optimization of Intelligent Data Driven Monitoring and Status Evaluation Mechanism for Distribution Network and Distributed Resources,” Distributed Generation & Alternative Energy Journal, vol. 39, no. 06, pp. 1153–1178, 02/19, 2025.

[7] S. Farah, N. Humaira, Z. Aneela, and E. Steffen, “Short-term multi-hour ahead country-wide wind power prediction for Germany using gated recurrent unit deep learning,” Renewable and Sustainable Energy Reviews, vol. 167, pp. 112700, 2022.

[8] P. He, Z. Zhu, X. Wang, C. Zhang, W. Yuan, and J. Hao, “Research on Supervision System of Power Safety Tools and Equipment Based on Internet of Things Technology,” Distributed Generation & Alternative Energy Journal, vol. 38, no. 04, pp. 1223–1254, 05/18, 2023.

[9] J. Hu, Z. Luo, and J. Han, “Measurement of Implicit Carbon Productivity in China’s Industrial Sectors and Its Convergence and Divergence,” 2023.

[10] H. Li, and R. Zhang, “Factors Influencing Member Satisfaction With Cooperation in an Agro-Industrialized Union,” Journal of Organizational and End User Computing (JOEUC), vol. 35, no. 1, pp. 1–18, 2023.

[11] B. Xiong, L. Lou, X. Meng, X. Wang, H. Ma, and Z. Wang, “Short-term wind power forecasting based on Attention Mechanism and Deep Learning,” Electric Power Systems Research, vol. 206, pp. 107776, 2022.

[12] C. Tian, T. Niu, and W. Wei, “Developing a wind power forecasting system based on deep learning with attention mechanism,” Energy, vol. 257, pp. 124750, 2022.

[13] Z. Ma, and G. Mei, “A hybrid attention-based deep learning approach for wind power prediction,” Applied Energy, vol. 323, pp. 119608, 2022.

[14] M. Abou Houran, S. M. S. Bukhari, M. H. Zafar, M. Mansoor, and W. Chen, “COA-CNN-LSTM: Coati optimization algorithm-based hybrid deep learning model for PV/wind power forecasting in smart grid applications,” Applied Energy, vol. 349, pp. 121638, 2023.

[15] M. A. Hossain, R. K. Chakrabortty, S. Elsawah, and M. J. Ryan, "Very short-term forecasting of wind power generation using hybrid deep learning model," Journal of Cleaner Production, vol. 296, pp. 126564, 2021.
[16] A. Alkesaiberi, F. Harrou, and Y. Sun, "Efficient wind power prediction using machine learning methods: A comparative study," Energies, vol. 15, no. 7, pp. 2327, 2022.
[17] V. K. Rayi, S. Mishra, J. Naik, and P. Dash, "Adaptive VMD based optimized deep learning mixed kernel ELM autoencoder for single and multistep wind power forecasting," Energy, vol. 244, pp. 122585, 2022.
[18] M. Li, and A. W. Guenier, "ChatGPT and Health Communication: A Systematic Literature Review," International Journal of E-Health and Medical Communications (IJEHMC), vol. 15, no. 1, pp. 1–26, 2024.
[19] T. Lyu, D. Gu, P. Chen, Y. Jiang, Z. Zhang, H. Pang, L. Zhou, and Y. Dong, "Optimized CNNs for Rapid 3D Point Cloud Object Recognition," arXiv preprint arXiv:2412.02855, 2024.
[20] H. Pang, L. Zhou, Y. Dong, P. Chen, D. Gu, T. Lyu, and H. Zhang, "Electronic Health Records-Based Data-Driven Diabetes Knowledge Unveiling and Risk Prognosis," arXiv preprint arXiv:2412.03961, 2024.
[21] W. Zhang, Z. Lin, and X. Liu, "Short-term offshore wind power forecasting-A hybrid model based on Discrete Wavelet Transform (DWT), Seasonal Autoregressive Integrated Moving Average (SARIMA), and deep-learning-based Long Short-Term Memory (LSTM)," Renewable Energy, vol. 185, pp. 611–628, 2022.
[22] X. Peng, H. Wang, J. Lang, W. Li, Q. Xu, Z. Zhang, T. Cai, S. Duan, F. Liu, and C. Li, "EALSTM-QR: Interval wind-power prediction model based on numerical weather prediction and deep learning," Energy, vol. 220, pp. 119692, 2021.
[23] S. M. J. Jalali, S. Ahmadian, M. Khodayar, A. Khosravi, M. Shafie-khah, S. Nahavandi, and J. P. Catalao, "An advanced short-term wind power forecasting framework based on the optimized deep neural network models," International Journal of Electrical Power & Energy Systems, vol. 141, pp. 108143, 2022.
[24] H. Yin, Z. Ou, J. Fu, Y. Cai, S. Chen, and A. Meng, "A novel transfer learning approach for wind power prediction based on a serio-parallel deep learning architecture," Energy, vol. 234, pp. 121271, 2021.
[25] A. Meng, S. Chen, Z. Ou, W. Ding, H. Zhou, J. Fan, and H. Yin, "A hybrid deep learning architecture for wind power prediction based on

bi-attention mechanism and crisscross optimization," Energy, vol. 238, pp. 121795, 2022.

[26] L. Wang, R. Tao, H. Hu, and Y.-R. Zeng, "Effective wind power prediction using novel deep learning network: Stacked independently recurrent autoencoder," Renewable Energy, vol. 164, pp. 642–655, 2021.

[27] L. Ye, B. Dai, M. Pei, P. Lu, J. Zhao, M. Chen, and B. Wang, "Combined approach for short-term wind power forecasting based on wave division and Seq2Seq model using deep learning," IEEE Transactions on Industry Applications, vol. 58, no. 2, pp. 2586–2596, 2022.

[28] C. Sasser, M. Yu, and R. Delgado, "Improvement of wind power prediction from meteorological characterization with machine learning models," Renewable Energy, vol. 183, pp. 491–501, 2022.

[29] D. Niu, L. Sun, M. Yu, and K. Wang, "Point and interval forecasting of ultra-short-term wind power based on a data-driven method and hybrid deep learning model," Energy, vol. 254, pp. 124384, 2022.

[30] H. Zhang, C. Wang, L. Yu, S. Tian, X. Ning, and J. Rodrigues, "Pointgt: A method for point-cloud classification and segmentation based on local geometric transformation," IEEE Transactions on Multimedia, 2024.

[31] X. Zheng, "The Robustness and Vulnerability of a Complex Adaptive System With Co-Evolving Agent Behavior and Local Structure," Journal of Organizational and End User Computing (JOEUC), vol. 35, no. 1, pp. 1–27, 2023.

[32] Y. Zhou, Z. Wang, S. Zheng, L. Zhou, L. Dai, H. Luo, Z. Zhang, and M. Sui, "Optimization of automated garbage recognition model based on resnet-50 and weakly supervised cnn for sustainable urban development," Alexandria Engineering Journal, vol. 108, pp. 415–427, 2024.

[33] A. Kisvari, Z. Lin, and X. Liu, "Wind power forecasting-A data-driven method along with gated recurrent neural network," Renewable Energy, vol. 163, pp. 1895–1909, 2021.

[34] H. Ran, W. Li, L. Li, S. Tian, X. Ning, and P. Tiwari, "Learning optimal interclass margin adaptively for few-shot class-incremental learning via neural collapse-based meta-learning," Information Processing & Management, vol. 61, no. 3, pp. 103664, 2024.

[35] J. Wang, F. Li, and L. He, "A Unified Framework for Adversarial Patch Attacks against Visual 3D Object Detection in Autonomous Driving," IEEE Transactions on Circuits and Systems for Video Technology, 2025.

[36] J. Wang, F. Li, S. Lv, L. He, and C. Shen, "Physically Realizable Adversarial Creating Attack against Vision-based BEV Space 3D Object Detection," IEEE Transactions on Image Processing, 2025.
[37] X. Chen, X. Zhang, M. Dong, L. Huang, Y. Guo, and S. He, "Deep learning-based prediction of wind power for multi-turbines in a wind farm," Frontiers in Energy Research, vol. 9, pp. 723775, 2021.
[38] S. M. J. Jalali, G. J. Osório, S. Ahmadian, M. Lotfi, V. M. Campos, M. Shafiekhah, A. Khosravi, and J. P. CatalÃ£o, "New hybrid deep neural architectural search-based ensemble reinforcement learning strategy for wind power forecasting," IEEE Transactions on Industry Applications, vol. 58, no. 1, pp. 1527, 2021.
[39] S. Wang, T. Teng, and H. Bao, "Spatiotemporal Evolution Characteristics of Urban Tourism Efficiency in the Yellow River Basin and Its Driving Factors," Statistics and Information Forum, vol. 38, no. 5, pp. 105117, 2023.
[40] R. Meka, A. Alaeddini, and K. Bhaganagar, "A robust deep learning framework for short-term wind power forecast of a full-scale wind farm using atmospheric variables," Energy, vol. 221, pp. 119759, 2021.
[41] Q. Li, X. Ren, F. Zhang, L. Gao, and B. Hao, "A novel ultra-short-term wind power forecasting method based on TCN and Informer models," Computers and Electrical Engineering, vol. 120, pp. 109632, 2024.
[42] J. Zhu, L. Su, and Y. Li, "Wind power forecasting based on new hybrid model with TCN residual modification," Energy and AI, vol. 10, pp. 100199, 2022.
[43] S. Sun, Y. Liu, Q. Li, T. Wang, and F. Chu, "Short-term multi-step wind power forecasting based on spatio-temporal correlations and transformer neural networks," Energy Conversion and Management, vol. 283, pp. 116916, 2023.
[44] S. Mo, H. Wang, B. Li, Z. Xue, S. Fan, and X. Liu, "Powerformer: A temporalbased transformer model for wind power forecasting," Energy Reports, vol. 11, pp. 736–744, 2024.
[45] J. Liu, Q. Shi, R. Han, and J. Yang, "A hybrid GA-PSO-CNN model for ultra-short-term wind power forecasting," Energies, vol. 14, no. 20, pp. 6500, 2021.
[46] A. R. Finamore, V. Calderaro, V. Galdi, G. Graber, L. Ippolito, and G. Conio, "Improving Wind Power Generation Forecasts: A Hybrid ANN-Clustering-PSO Approach," Energies, vol. 16, no. 22, pp. 7522, 2023.
[47] I. Delgado, and M. Fahim, "Wind turbine data analysis and LSTM-based prediction in SCADA system," Energies, vol. 14, no. 1, pp. 125, 2020.

[48] J. Zhou, X. Lu, Y. Xiao, J. Su, J. Lyu, Y. Ma, and D. Dou, "Sdwpf: A dataset for spatial dynamic wind power forecasting challenge at kdd cup 2022," arXiv preprint arXiv:2208.04360, 2022.
[49] X. Liu, Z. Lin, and Z. Feng, "Short-term offshore wind speed forecast by seasonal ARIMA-A comparison against GRU and LSTM," Energy, vol. 227, pp. 120492, 2021.
[50] A. A. Ewees, M. A. Al-qaness, L. Abualigah, and M. Abd Elaziz, "HBO-LSTM: Optimized long short term memory with heap-based optimizer for wind power forecasting," Energy Conversion and Management, vol. 268, pp. 116022, 2022.
[51] Z. Al-Ibraheemi, and S. Al-Janabi, "Sustainable energy: Advancing wind power forecasting with grey wolf optimization and GRU models," Results in Engineering, vol. 24, pp. 102930, 2024.
[52] H. K. M. Nguyen, Q.-D. Phan, Y.-K. Wu, and Q.-T. Phan, "Multi-step wind power forecasting with stacked temporal convolutional network (S-TCN)," Energies, vol. 16, no. 9, pp. 3792, 2023.
[53] X. Li, and Y. Zheng, "Short-Term Wind Power Prediction Based on TCN-Transformer and STL Error Correction." pp. 462–467.
[54] I. S. Kervanci, M. F. Akay, and E. Özceylan, "Bitcoin price prediction using LSTM, GRU and hybrid LSTM-GRU with bayesian optimization, random search, and grid search for the next days," Journal of Industrial and Management Optimization, vol. 20, no. 2, pp. 570–588, 2024.

Biography

Yuntao Geng born in Xiping, Henan Province. His work unit is Shaoyang Polytechnic, lecturer title. His main research interest is power electronic transformation and control technology.