▪ Journal of Competitiveness

Toward Stable Power Systems in Complex Environments: A Hybrid Forecasting Model Validated Across Two Typical Regions in China

Jinfeng Zhao, Changqi Tao*, Jiahao Fu, Yong Wang

Abstract

This study explores the application of artificial-intelligence technology in power prediction, demonstrating that deep learning-based high-precision forecasting can provide critical technical support for power system stability in complex environments. By integrating deep-learning methods, this study proposes a hybrid approach combining complete ensemble empirical mode decomposition with adaptive noise and variational mode decomposition to uncover intrinsic patterns in photovoltaic power sequences. A convolutional neural network – bidirectional long short-term memory – Transformer prediction model is constructed, and the slime mold algorithm is introduced to optimize its hyperparameters. The model is validated using heterogeneous datasets from two geographically distinct regions—the East China Plain and Central China Hills—facilitating cross-climate and cross-terrain experimentation. The impact of improved prediction accuracy on power-system stability is illustrated through theoretical analyses and practical cases from the two regions. Results show that the integrated signal decomposition and deep-learning model maintains high prediction accuracy under complex climatic and topographic conditions, contributing to power system stability. A key innovation of this research lies in the construction of an intelligent forecasting framework that provides a reliable decision-support tool for power enterprises in the two regions, mitigating the operational risks caused by dispatch deviations and enhancing the system stability and management efficacy during the energy-structure transition.

Keywords: *Power prediction, Deep learning, Stability*

JEL Classification: C53, C45, Q40

Article history: Received: May 2025; Accepted: September 2026; Published: September 2026

1 Introduction

Amid the global endeavors to tackle climate change and push forward energy transition, developing and applying clean energy is no longer a choice but a necessity of our time. Renewable energy, solar energy being a typical example, is characterized by safety, cleanliness, and ample resources, and as such, occupies an important position in long-term energy strategies (Singh, 2013; Li et al., 2024). China's photovoltaic (PV) power-installation capacity had grown to exceeding 609 GW as of the end of 2023, maintaining its position as the global leader for the eighth year running. Despite the continuous growth in PV capacity, the variability and unpredictability of PV power output remain critical concerns, jeopardizing the secure and stable functioning of electrical grids (Zhang et al., 2020).

Artificial intelligence (AI) can potentially provide a rational resolution to challenges concerning safe and stable power system operations. A key requirement of modern energy systems is the highly accurate short-term prediction of PV power generation. This ensures the reliable integration of PV power generation in optimizing the whole-life-cycle efficiency of power

https://doi.org/10.7441/joc.2026.03.11

plants, facilitating intelligent operation and maintenance decisions, supporting power-grid dispatching departments in formulating day-ahead, intraday, and real-time coordinated dispatching strategies, and enhancing the frequency and voltage stability of power systems. Starting from the vantage point of the performance advantages of deep learning, which is a part of AI, in high-precision forecasting, and focusing on the core demand for the stable operation of power systems, this study explores the key application value of AI technologies in improving the forecasting accuracy. It further applies such technologies to the PV power forecasting scenarios in two major regions, namely, Central China Hills and East China Plain, and conducts multiscenario verification studies. The deployment of this method offers effective technical assistance to maintain power system stability and efficiency.

1.1 Challenges Faced by Modern Power Systems

With the energy-structure evolution, the integration of new energy sources introduces considerable complexities that strain the reliable operation of power grids. The dispatch, operation, and management of modern power systems are undergoing profound changes, encountering traditional and emerging issues being deeply intertwined. This section elaborates on these challenges concerning three aspects of the power system: generating side, grid network, and load demand.

Revolutionary challenges on the generation side

The high penetration of renewable-energy integration has introduced significant uncertainties. In particular, solar and wind-power generation—being highly weather-dependent—pose considerable difficulties in output prediction, posing a major challenge for maintaining real-time power-system balance. The grid must efficiently deploy extensive resources to “shave peaks and fill valleys.” Deficiencies in prediction accuracy plague both short-term and ultrashort-term horizons, resulting in considerable deviations and compelling system operators to reserve additional backup capacity, which in turn increases the operational costs.

Increased complexity on the grid side

With the widespread adoption of distributed PV systems (e.g., rooftop PV panels), the traditional, linear power flow—from generation to transmission, distribution, and consumption—has been fundamentally upended. Moreover, reverse power flows in distribution networks have complicated the power-flow patterns, making them more difficult to monitor and control. In addition, renewable power plants are often located in remote areas, such as solar and wind farms in Western China, which are geographically mismatched with the load centers in the central and eastern regions. Therefore, transmission-channel capacity constraints frequently lead to grid congestion, hindering the effective integration of clean energy.

Growing uncertainty on the load side

The rapid growth in new high-power and stochastic loads, such as electric-vehicle charging, large-scale data centers, and fifth-generation base stations, has considerably increased the uncertainty on the demand side. In particular, the simultaneous charging of electric vehicles may create significant “peak stacking,” greatly challenging the system's peak-load–management capabilities. Furthermore, the increased frequency of extreme climate events, notably, the extreme high temperatures experienced in recent years, exacerbates the contradiction between the increased residential-electricity demand in the evening and decline in solar-power generation after sunset. This results in a sharply steep “duck curve” load profile,

https://doi.org/10.7441/joc.2026.03.11

placing enormous pressure on the peak regulation during evening hours. As a result, the system faces immense pressure during evening-peak regulation.

Although pricing signals and incentive mechanisms have been employed as part of national policy to encourage users to reduce peak demand (Siano, 2014), challenges such as low user participation, uncertain responsiveness, and imperfect market mechanisms remain. Addressing these issues requires not only coordination between climate and economic policies but also the use of advanced meteorological technologies to achieve higher forecasting accuracies. Advances in meteorological technologies are essential for better aligning electricity consumption with generation, facilitating coordinated and sustainable development; specifically, digital transformation technology can enhance competitiveness (Qazi, 2025). This study focuses primarily on challenges at the generation level, and proposes an deep-learning–based application model to provide robust technical support for managing uncertainties associated with renewable-energy integration.

1.2 Practical Predicaments in the Application of Traditional Forecasting Technologies

Conventional approaches to PV-power forecasting typically can be grouped into two main classes: physical and statistical models. Physical models, grounded in the physical properties of PV cells, predict power outputs through numerical simulations. These methods offer high accuracy, but require detailed meteorological and component parameters and involve complex computations (De Soto et al., 2006; Villalva et al., 2009). Physical models are built upon mathematical representations of the governing physical processes in solar PV systems; these intricate equations necessitate comprehensive knowledge of the system's core physical and operational traits (Dobos, 2014). Moreover, the efficacies of these models are contingent upon precise meteorological predictions, as they employ weather conditions as the primary input variables (Yang et al., 2018). Utilizing statistical methodologies, these models operate by fitting historical data to delineate the interrelationships among various factors, such as weather conditions, solar irradiance, and the subsequent energy production (Gaboitaolelwe et al., 2023). They also include time-series models, probabilistic forecasting models, and regression-analysis models. The Bayesian model averaging post-processing method, proposed by Doubleday et al. (2020), which incorporated a novel beta-kernel parameterization, significantly enhanced the forecasting accuracy. The approach put forward by Thaker et al. (2022), which clustered datasets using an irradiance-classification method to derive forecasting equations, substantially reduced forecasting errors. Nevertheless, statistical modeling methods exhibit certain limitations in short-term PV-power forecasting. On the one hand, they struggle to handle large-scale datasets (Singh et al., 2021); on the other hand, their insufficient flexibility in handling complex nonlinearities can result in incomplete representations of the dynamic interdependencies within PV power-generation systems (Li et al., 2021; Tahir et al., 2024).

Traditional forecasting techniques primarily focus on centralized PV systems, and heavily rely on meteorological data. For instance, physical-model–based PV forecasting predominantly employs numerical weather prediction models (Sweeney et al., 2020), which simulate mathematical representations of atmospheric dynamics governed by physical equations. However, as the resolution of these models increases, their computational efficiency declines significantly. Statistical approaches falling under the category of the autoregressive moving average (ARMA) model are computationally simpler but impose stringent requirements on data preprocessing and time-series stationarity, thereby struggling to capture nonlinear influences. In their comprehensive work on time-series analysis, Box et al. (2015) noted that traditional

https://doi.org/10.7441/joc.2026.03.11

statistical methods such as ARMA are effective for linear processes but face limitations when applied to highly complex, nonlinear systems. Moreover, many methods lack in-depth exploration of salient features in raw data, resulting in limited model generalizability (Xie et al., 2021). With the rapid expansion of distributed PV systems, the increasing integration of distributed generation into power grids poses serious challenges to grid scheduling. Conventional forecasting models now fall short of meeting the practical requirements for accuracy in power-system management and dispatch.

1.3 Emergence of AI Technology Applications

The fast-paced advancement of distributed PV installations has made them a core part of the global energy transition. However, the scarcity of high-quality meteorological data severely restricts the traditional forecasting methods in cross-regional and heterogeneous scenarios (Chen et al., 2023), complicating power-system management and supply–demand balance. Against this backdrop, there is an urgent need for efficient short-term cross-regional PV power-forecasting models, and deep-learning techniques provide a promising solution by enhancing the predictability of renewable energy (Ren et al., 2022).

To address these challenges, existing studies have proposed various hybrid deep-learning models. Li et al. (2025) developed the white-shark-optimizer–optimized neural-prophet long short-term memory (LSTM) informer model, which performs well in medium-term forecasting. Niu et al. (2022) proposed the CNN - bidirectional gated recurrent unit (BiGRU) -attention multitask learning model to strengthen feature fusion. Zhou et al. (2025) verified the effectiveness of complete-ensemble empirical-mode decomposition with adaptive noise (CEEMDAN) single decomposition in ultrashort-term forecasting. Other studies also confirmed the value of decomposition and AI technologies in renewable-energy prediction (RajasundrapandiyanLeebanon et al., 2025; Kim & Kim, 2023). Nevertheless, critical limitations remain. First, single-decomposition techniques such as CEEMDAN cannot fully solve the mode aliasing and end effects (Zhou et al., 2025). Second, existing hybrid models fail to deeply integrate local spatial-feature extraction, bidirectional temporal modeling, and global correlation capture (Niu et al., 2022; Li et al., 2025). Third, most studies focus on single-region forecasting, and lack adaptation to cross-regional topographic and climatic heterogeneity, leading to poor generalization. To fill these gaps, this study proposes a short-term cross-regional PV power-forecasting model based on CEEMDAN–variable mode decomposition (VMD) dual decomposition and slime mold algorithm (SMA)-optimized CNN–BiLSTM–Transformer. The dual decomposition compensates for the deficiencies of single decomposition. The hybrid architecture integrates CNN, BiLSTM, and Transformer to characterize spatiotemporal features comprehensively, while SMA optimizes the hyperparameters to improve the prediction accuracy and generalization.

Under the dual-carbon goals, power systems are evolving toward a new “double-high” featuring a high share of renewable energy and high penetration of power-electronic devices. This “double-high” feature poses major difficulties for maintaining stability and making control decisions within the power grid. Although the increasing penetration rate of renewable energy reduces CO_2 emissions, its intermittency poses a threat to the stable running of conventional thermal units (Suri et al., 2025). Therefore, precise forecasting of output from renewable energy systems is of particular importance. This paper commences with a summary of the constraints inherent in the traditional forecasting methods, followed by an overview of the present landscape of deep learning in predictive applications. Regarding the enhancement of power-system the stability, this paper first theoretically analyzes how the accuracy of PV power

https://doi.org/10.7441/joc.2026.03.11

forecasting affects the power-system stability, and then describes the construction of a PV power-forecasting accuracy-improvement model, and verifies the model with multiterrain datasets. Finally, the practical impacts of forecasting accuracy on the stable operation of power systems are illustrated by introducing actual engineering cases.

2 Connotation and Impact

2.1 Connotation of Power-System Stability

Institute of Electrical and Electronics Engineers (IEEE) and Conseil International des Grands Réseaux Électriques (CIGRE) define power-system stability as the ability of an electric power system, under a given initial operating condition, to regain a state of operating equilibrium after being subjected to a physical disturbance, with most system state variables remaining within limits so that system integrity is maintained (Kundur, 2004). China's Power System Security and Stability Guideline (DL 755-2001) describes power-system stability as the ability of an electric power system to keep functioning in a stable state after being impacted by a fault disturbance. Both definitions emphasize the system's capability to restore equilibrium post-disturbance, and classify it into rotor-angle stability, voltage stability, and frequency stability, along with numerous subcategories. The primary distinction between the two definitions lies in the detailed classification of these subcategories (Sun et al., 2006). Rotor-angle stability is defined as the capability of synchronous generators within an interconnected system to stay in synchronism after experiencing a disturbance. Voltage stability, under a given initial operating condition, describes a power system's ability to maintain steady voltages at all buses following a disturbance. Both forms of stability depend on the power system's capacity to maintain or restore a balance between load demand and power supply. Frequency stability denotes a power system's capability to keep its frequency steady after a major disturbance causes a substantial mismatch between power supply and demand.

Based on the aforementioned definitions of power-system stability, this study defines it as the ability of a power system to maintain secure and stable synchronous operation, voltage, and frequency, under the dual challenges of physical disturbances and stochastic fluctuations from renewable energy sources, by leveraging accurate forecasting and rapid control. It is likewise categorized into frequency stability, voltage stability, and rotor-angle stability.

2.2 Impact of Prediction Accuracy on Power-System Stability

Enhancing PV-prediction accuracy improves the system's “predictability,” which directly strengthens the “proactive nature” and “initiative” of the power grid in responding to power fluctuations. This enhancement serves as a cornerstone for ensuring stable operation of modern power systems, especially those with a high penetration of renewable energy. Supporting this, Malinkovich et al. (2024) found that high-accuracy ultrashort-term prediction is crucial for maintaining power continuity, managing ramp rates, and ensuring grid stability. They also concluded that combining such predictions with energy-storage systems presents an effective solution for managing grid stability.

Prediction accuracy has a direct impact on both frequency stability and voltage stability. Frequency stability directly determines the accuracy of generation/load-balancing schedules. Reducing power deficits and minimizing frequency deviations effectively enhances frequency stability. Conversely, inaccurate predictions can lead to frequency imbalance in microgrids, while accurate forecasting enables the grid to precisely reserve capacity to balance fluctuations. In an islanded microgrid, employing a more accurate prediction model, combined with an

https://doi.org/10.7441/joc.2026.03.11

optimized controller, can significantly reduce the standard deviation of frequency (Guptah & Singh, 2024). Accurate prediction facilitates proactive anticipation of the voltage changes induced by PV fluctuations, allowing active adjustment of reactive-power-compensation devices, thereby influencing voltage stability. Studies employing a predictive control strategy and simulations based on a modified IEEE 123-node test-feeder system revealed that, compared to traditional real-time control strategies, the predictive strategy can proactively adjust the tap positions of voltage regulators based on PV power forecast data. This approach reduces unnecessary and frequent switching. Consequently, it not only decreases the number of tap-changing steps—mitigating equipment wear—but also optimizes the system's voltage profile, significantly reducing the voltage deviation. During periods of high load fluctuation, the predictive-control strategy's proactive tap adjustments avoid the impacts associated with abrupt changes, resulting in a more stable system voltage (Cui et al., 2025). An improvement in the prediction accuracy does not directly affect rotor-angle stability. The value of accurate prediction lies in providing dispatchers with more accurate forecasts of renewable-energy output, thereby facilitating a more rational arrangement of conventional generator-unit operating schedules and reducing the risk of large disturbances (Song et al., 2025).

The impact of prediction accuracy on the power-system stability has been analyzed above, from a theoretical perspective. However, the power-system stability is influenced not only by prediction accuracy, but also highly correlated with factors such as power-source structure (Li et al., 2024), power-grid scale (Smith et al., 2024), load characteristics (Ozan, 2025), and operational constraints (Ahmed et al., 2020), in different application scenarios.

This study mainly focuses on the model's performance in terms of prediction. To verify the universality of the prediction model, the East China Plain and Central China Hills regions were selected. These two regions exhibited significant differences in terms of topography, climatic characteristics, PV output characteristics, and forecasting challenges, which could fully verify the universality and robustness of the prediction model in complex environments. The East China Plain features a flat and open terrain with no obvious topographic shading. It presents a subtropical monsoon climate characterized by frequent cloud-cover changes and high humidity. This region is dominated by the distributed PV systems, and their output exhibited synchronous cliff-like fluctuations affected by cloud movement. The Central China Hills feature undulating terrain with diverse slope orientations and significant topographic shading. This region also has a subtropical monsoon climate, but the terrain introduces local microclimates, leading to strong spatial heterogeneity of meteorological elements. PV systems in this region consisted of both centralized and distributed mountainous ones, and their outputs presented significant asynchrony owing to the different slope orientations.

3 Research Methodology

This study explores the technical support solutions for promoting sustainable high-quality development of the power grid by constructing a combination prediction model based on deep learning.

3.1 Benchmark Model Method

CEEMDAN

The Empirical Mode Decomposition (EMD), first proposed by Huang et al. (1998), provides a foundational adaptive approach for analyzing nonlinear and non-stationary signals. To address its inherent mode-mixing problem, CEEMDAN was proposed by Torres et al. (2011).

https://doi.org/10.7441/joc.2026.03.11

CEEMDAN reduces computational complexity while enhancing decomposition accuracy. Its core innovation lies in progressively adding white noise during each iteration to the intrinsic mode functions (IMFs) extracted by EMD. The stepwise addition of noise aims to counteract the adverse effects of white noise on the outcomes, leading to a significant enhancement of both the precision and reliability of the signal decomposition.The specific decomposition steps are given in Appendix A.

VMD

VMD is an adaptive time-frequency domain signal processing method (Dragomiretskiy et al., 2013). It assumes that a signal consists of limited-bandwidth IMFs, each of which has its own center frequency and bandwidth. PV power and its associated features are initially decomposed using CEEMDAN to address the issue of mode mixing. Subsequently, VMD is applied to perform a secondary frequency-domain optimization decomposition on the key IMF components. VMD formulates this as a constrained variational problem and employs the Alternating Direction Method of Multipliers (ADMM) algorithm to iteratively update the modal components, center frequencies, and Lagrange multipliers until convergence conditions are satisfied, thereby achieving optimal signal separation in both the time and frequency domains. This approach constructs a composite feature matrix through multi-scale feature extraction, encompassing high-frequency noise, medium-frequency periodic components, and low-frequency trends. Finally, a model is developed by BiLSTM network with a Transformer mechanism. The details are given in Appendix B.

Sample Entropy (SE)

CEEMDAN decomposes a time series into multiple IMF components, each containing valid signals and random noise. However, an excessive number of IMF components will increase the computational complexity of the prediction model. SE has a dual role (Richman & Moorman, 2000): it can not only suppress noise interference within IMF components but also identify the similarity characteristics between different components, thereby effectively alleviating the problem of excessive computation caused by component redundancy in the model. The details are given in Appendix C.

SMA

The SMA, proposed by Li et al. (2020), is a novel metaheuristic optimization algorithm. It simulates the foraging behaviors and morphological adaptations of slime mould, a type of eukaryote living in humid and cold environments. When slime mould perceives a food source, its biological oscillators generate propagation waves within the veins, leading to rhythmic contractions. In the presence of multiple food sources, slime mould constructs vein networks of varying thicknesses. Higher food concentrations strengthen propagation waves, accelerate cytoplasmic flow, and enhance adaptive exploration capabilities. The core processes of SMA are given in Appendix D.

CNN

CNNs, developed by LeCun et al. (1989), exhibit outstanding classification and regression capabilities. From a structural perspective, a CNN is organized into three core components: an input layer, a series of hidden layers for feature extraction and transformation (convolutional, pooling), and a final, fully connected output layer. The convolutional layer serves to extract localized features via sliding kernels, the pooling layer functions to reduce dimensionality

https://doi.org/10.7441/joc.2026.03.11

through down-sampling while retaining essential information, and the fully connected layer is tasked with integrating these features for final classification or regression tasks.
The core of CNN is the convolutional layer, where the convolution kernel a^1acts as a filter to extract local features from the data, with the specific calculation as follows: $a^1 = \sigma(b + w * a^0)$, σ is the activation function, a^0 is the input of the convolution layer, $*$ is the convolution operation, w is the weight matrix, b is the bias matrix.

BiLSTM

LSTM networks, initially proposed by Hochreiter & Schmidhuber (1997), were designed to overcome the long-term dependency issues and gradient problems of recurrent neural networks (RNNs). By introducing gated structures, LSTM dynamically regulates information flow through three main gates: the forget gate regulates the retention of historical data, while the input gate controls the assimilation of new information, and the output gate controls the transmission of memory content. Within this paper, BiLSTM is adopted to conduct bidirectional time-series modeling. The forward LSTM learns the power variation trends from historical data, while the backward LSTM explores potential correlations from a future perspective. The concatenation of these two components forms global time-series features, enhancing the ability to parse complex fluctuation patterns. By leveraging bidirectional temporal context, the model compensates for the limitations of unidirectional approaches, which often overlook implicit future dependencies. The details are given in Appendix E.

Transformer

At the heart of the Transformer model (Vaswani et al., 2017) lies its self-attention mechanism, which constitutes the foundation of its deep-learning architecture. Unlike traditional sequence models such as recurrent neural networks and LSTMs, which rely on recursive computations, the Transformer dispenses with recurrent structures entirely, relying instead on attention mechanisms, to capture sequence dependencies. This facilitates parallel computation, global-dependency modeling, and greater scalability, which have demonstrated significant advantages in predicting and analyzing the stability of PV power generation. In the context of PV power systems, the self-attention mechanism can effectively model the complex, nonlinear relationships between high-dimensional, time-varying meteorological inputs (e.g., solar irradiance, temperature, cloud cover) and the resultant PV output-power fluctuations. By capturing long-range dependencies within historical and real-time sequential data, Transformer-based models enhance the accuracy of short-term and ultrashort-term PV power forecasting, which is critical for assessing and mitigating the instability risks that variable PV generation poses to grid operation. Furthermore, the architecture's capability to integrate heterogeneous data sources and perform multivariate time-series analyses provides a robust framework for evaluating the stability margin of PV-integrated power systems under diverse and evolving operational scenarios. The architecture follows an encoder–decoder structure, in which self-attention serves as the core component, providing enhanced capability in modeling long-range sequential relationships essential for understanding and predicting the stability dynamics of PV generation. The details are provided in Appendix F.

3.2 Construction of the Forecasting Model

A systematic examination of the existing relevant research reveals that numerous studies have adopted hybrid forecasting models for enhancing PV generation forecasting precision (Liu et al., 2024; Zhou et al., 2025). However, no study has simultaneously addressed all the core challenges in PV power forecasting—including the adaptation to spatiotemporal dual

https://doi.org/10.7441/joc.2026.03.11

characteristics, capture of long and short-term temporal dependencies in power sequences, and modeling of nonlinear fluctuations under extreme weather conditions. Decomposing PV power signals with CEEMDAN alone suffers from high-frequency mode aliasing and residual end effects. VMD can perform targeted adaptive band-limited modal decomposition on its decomposition residuals, thereby precisely mitigating the aforementioned defects (Li et al., 2020). Therefore, this study draws on the mainstream application approach of hybrid forecasting models in existing research and proposes a short-term PV power forecasting method for the CNN–BiLSTM–Transformer architecture, which is based on secondary decomposition and SMA optimization. The implementation process (Fig.1) and model framework (Fig. 2) are described below:

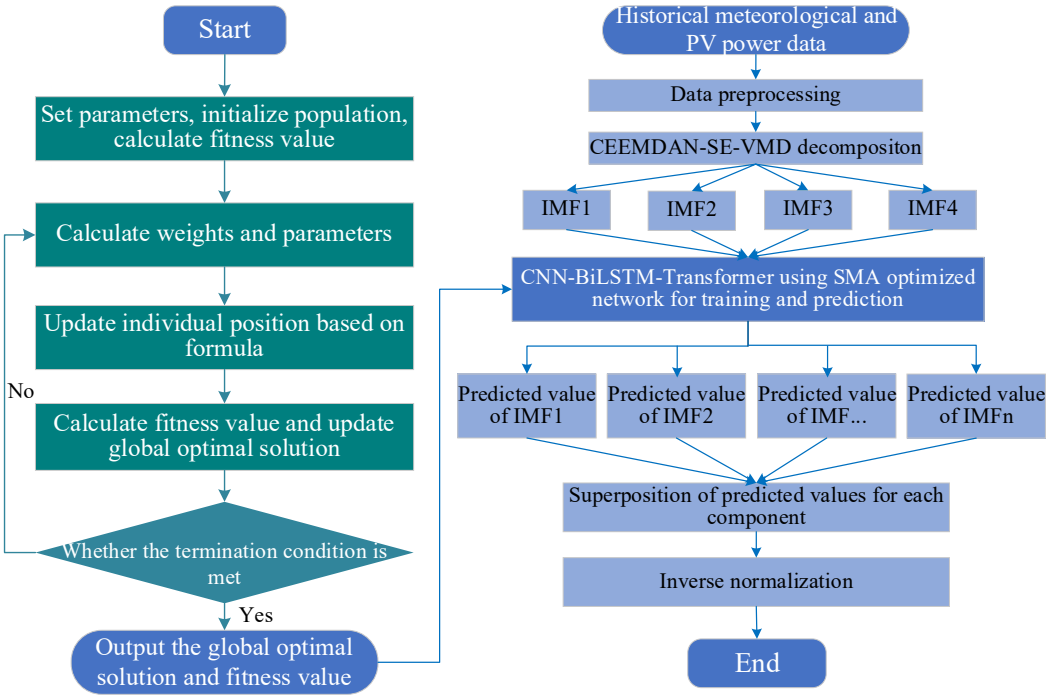


Fig. 1 – Methodological workflow. Source: own research

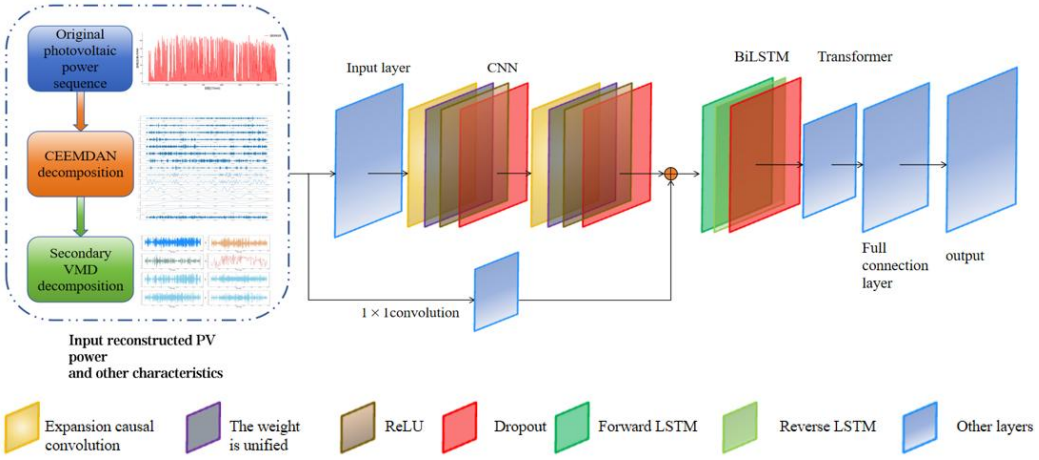


Fig. 2 – CNN-BiLSTM-Transformer model framework diagram. Source: own research

https://doi.org/10.7441/joc.2026.03.11

First, the raw PV power data are normalized to eliminate scale differences across variables. To suppress nonstationarity and noise, the CEEMDAN algorithm is applied to adaptively decompose the original series into intrinsic mode functions (IMFs) and a residual. Considering that excessive IMF components may increase complexity, sample entropy (SE) is employed to identify significant high-frequency components. These selected components are refined through a secondary decomposition stage employing VMD, producing the final modal inputs. This secondary decomposition ensures that the dataset is denoised, stabilized, and information-rich, while reducing redundancy.

Second, the decomposed modal components are combined with meteorological features and passed into a one-dimensional CNN. In this stage, convolutional layers perform nonlinear operations to extract localized spatial dependencies, while max-pooling layers down-sample and compress the outputs, retaining critical features. The product of this preprocessing is a refined feature map that encodes the essential structural patterns present in the data.

Third, the CNN-extracted feature representations are modeled using BiLSTM layers. The BiLSTM architecture models temporal dependencies in both forward and backward directions, enabling the model to represent both past and future sequences. Through iterative hidden-state updates, BiLSTM progressively constructs more abstract and global temporal features, better reflecting the inherent sequential dynamics of PV power generation.

Fourth, the features generated by BiLSTM are fed into a Transformer module. Leveraging self-attention and multihead attention mechanisms, the Transformer adaptively reweights the different time steps, distinguishing the relative importance of inputs and highlighting critical historical periods, while reducing redundancy. Together, these mechanisms prevent the loss of important contextual information and improve the model's capability to model long-range temporal dependencies. To mitigate overfitting risks, dropout regularization is applied after each BiLSTM layer, which improves generalization and shortens training time.

Fifth, to ensure optimal model configuration, the SMA is adopted to search for and update hyperparameters efficiently. An early-stopping strategy is also implemented to avoid overfitting during training, which results in improved model robustness and generalization capability.

Finally, the trained CNN–BiLSTM–Transformer model is evaluated using independent test sets. Prediction results are quantitatively assessed with standard error metrics, and residual patterns are analyzed to identify potential deficiencies. Insights gained from this evaluation are expected to guide iterative improvements, ensuring that the forecasting framework remains adaptive and effective in practical applications.

3.3 Model Configuration

3.3.1 Data Sources and Preprocessing

The data for this study were sourced from operational PV power plants located in the coastal region of East China Plain and hilly region of Central China Hills, and comprised data for the entire year of 2019. The datasets included measured PV power outputs and meteorological features (such as atmospheric pressure, solar radiation, temperature, humidity, etc.), with a temporal resolution of 15 min. Given that solar panels produced no power at night, the analysis utilized data from the 7:00 AM to 7:30 PM time frame alone. A small number of abnormal data points were addressed using linear interpolation, and the final dataset comprised 17,155 valid data points. Given the significant differences in climate and topography between the two regions, data from both locations were selected for model training and prediction. This approach

https://doi.org/10.7441/joc.2026.03.11

enhanced the model's generalizability, validated its interpretability, and strengthened its potential for transfer learning.

3.3.2 Parameter Settings

Inspired by natural phenomena, the SMA is a metaheuristic optimizer that emulates the foraging behavior of slime molds, which dynamically adjust their paths by secreting chemical substances. This mechanism allows an effective balance between global exploration and local exploitation. Building upon the research of Bhatt et al. (2022), Mohanasundaram & Rangaswamy (2025), in this study, SMA was employed to optimize the hyperparameters of a deep--learning model, including the number of convolutional kernels in the CNN, number of hidden units in the LSTM, head cardinality of the multihead attention mechanism, and learning rate. A population of hyperparameter combinations was randomly initialized, and the R^2 score on the validation set was defined as the fitness function to automatically search for the optimal parameter set. This approach enhanced the model's performance in terms of prediction, while avoiding the inefficiency and subjectivity associated with manual tuning. The training batch size was set to 1000, and an early-stopping mechanism is implemented—training halted if the loss function fails to decrease for 20 consecutive steps—to ensure adequate learning of data characteristics. The ranges of the optimized hyperparameters used in the experiments are presented in Tab.1. To ensure rigorous evaluation and generalization capability, a rolling-window strategy was adopted along with a strictly isolated test set. Chronologically, the most recent six months of data were reserved as an independent test set, which remained entirely unused throughout all stages of model development, SMA-based hyperparameter tuning, and feature engineering. Training and validation within the rolling windows were confined to the data prior to the test period, and all feature processing was performed based solely on statistics derived from the training data within each window, thereby preventing data leakage. The model was evaluated only once on the final held-out test set, and all reported performance metrics, including R^2, were derived from this set to ensure unbiased and reliable results.

Tab. 1 – Key hyperparameter settings of the model. Source: own research

Hyperparameter	**Setting**
cnn_filters	[2，16]
cnn_kernel_size	[1，5]
lstm_units LSTM	[1，32]
num_heads LSTM	[1，4]
num-heads Transformer	[1，4]
droupt_rate	[0.01，0.5]
Learning rate α	[0.0001，0.01]
Optimizer	Adma
Sigmod	Relu
Running enviroment	Python 3.9.21
Regularization	Dropout

3.3.3 Model Evaluation

To evaluate the model's performance, the approach adopted by Zhou et al. (2025) was followed, and three metrics were employed: root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2). The predictive accuracy is quantified using three indices. The MAE and RMSE were inversely proportional to the error between the actual measurements

https://doi.org/10.7441/joc.2026.03.11

and the predicted ones. The R^2 statistic, bounded between 0 and 1, was interpreted as a measure of the goodness-of-fit, with values nearing 1 indicating superior model performance. The formulas were as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (1)$$

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

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (3)$$

where n represents the number of predicted sample points, y_i represents the true value at the ith moment, $\bar{y}$ represents the average value of the output power, and $\hat{y}_i$ represents the predicted value.

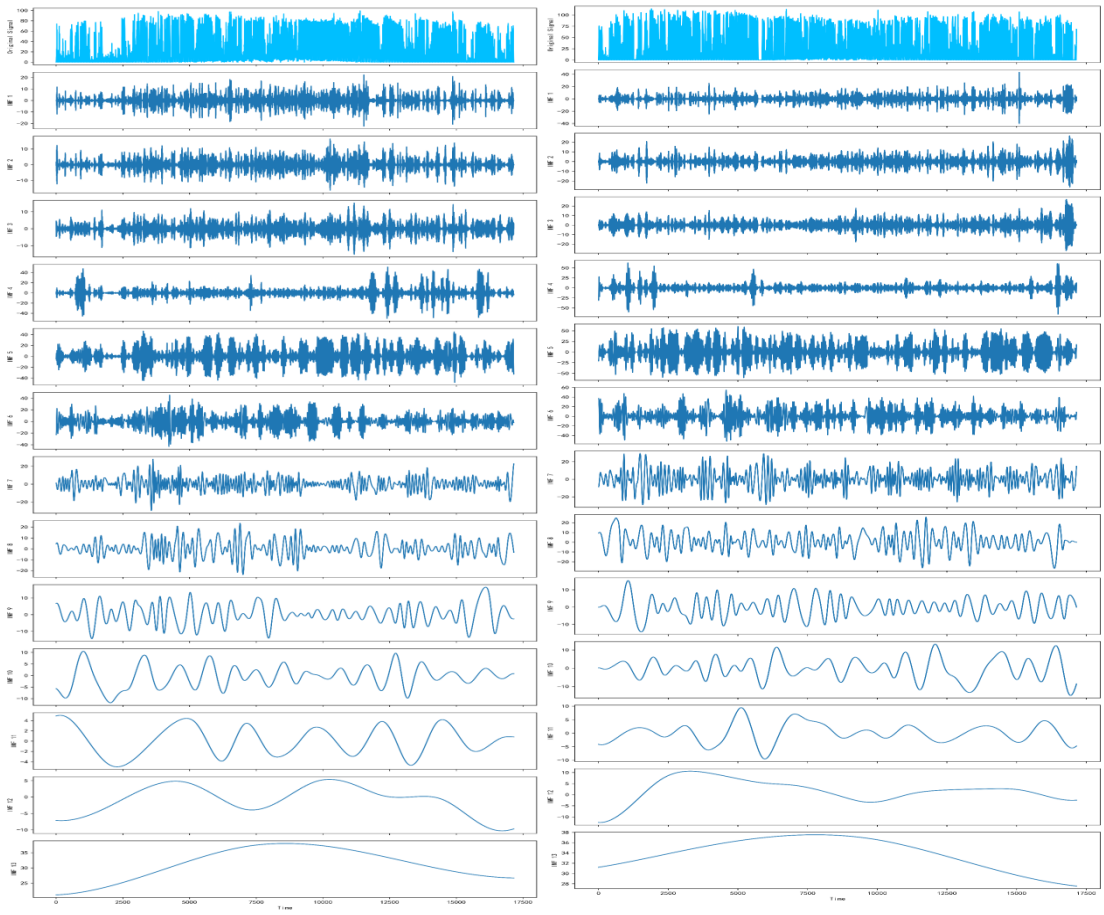
4 Validation of Model Application Scenarios

This investigation utilized a hybrid model leveraging deep learning to assess the predictive performances in two regions: PV power-generation forecasting in Central China Hills and East China Plain.

4.1 Application Results of the CEEMDAN–SE–VMD Dual-Decomposition Algorithm

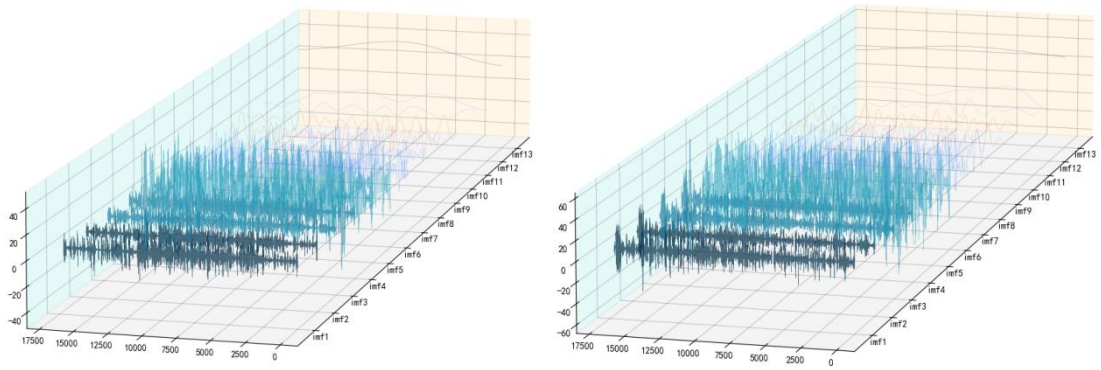
Decomposition was performed on the raw PV power data from both datasets, via the CEEMDAN method. This process generated 13 IMFs for each individual dataset. The derived IMFs were systematically ordered in a vertical sequence, yielding the component-decomposition diagram (Fig.3) and a three-dimensional diagram (Fig.4). In the component diagram, the first column corresponds to the actual power time series, with its data points plotted along the horizontal axis. The vertical axis, in turn, indicates the information value contributed by each component from the original sequence. As shown in Fig. 3 and 4, both the frequency and the degree of complexity of the components gradually decrease, while the variation trend of power generation becomes more apparent.

https://doi.org/10.7441/joc.2026.03.11



a. Central China Hills b. East China Plain

Fig. 3 – CEEMDAN decomposition diagram. Source: own research



a. Central China Hills b. East China Plain

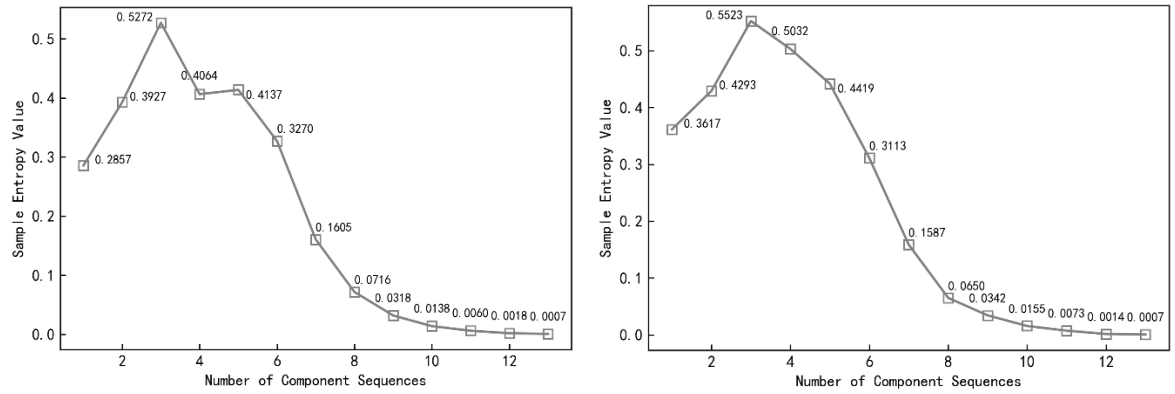
Fig.4 – Decomposition of CEEMDAN components visualization diagram. Source: own research

4.1.1 SE Reconstruction

SE provides a measure of the complexity inherent in each IMF component. By grouping components that exhibited analogous SEs, the workload of the forecasting model could be

https://doi.org/10.7441/joc.2026.03.11

reduced, thereby lowering the time cost. The entropy values for all components in the two datasets were calculated using the SE method, as shown in Fig.5. The analysis revealed that the SEs of the components in the two scenarios exhibit comparable trends and distributions. To form four new modal sequences, the components were categorized into four groups using K-means clustering, according to their SE values. Owing to the similarities in SE values of the two datasets, the clustering results were identical. The reconstruction results of the components for both datasets are presented in Tab. 2.



a. Central China Hills b. East China Plain

Fig.5 – SE line chart. Source: own research

Tab.2 – Combined results of the IMF components. Source: own research

HF-IMF	MF-IMF	LF-IMF	TF-IMF
IMF2, IMF3, IMF4, IMF5	IMF1, IMF6	IMF7	IMF8, IMF9, IMF10, IMF11, IMF12, IMF13

4.1.2 VMD Denoising

Taking the Central China Hills as an example, the IMF components were reconstructed according to the method presented in Tab.2, resulting in reconstructed sequences for the two datasets, as shown in the left half of Fig.5. As can be observed in Fig.6, the reconstructed sequences HF–IMF and MF–IMF exhibit high frequency and randomness, and can be categorized as high-frequency and relatively high-frequency sequences, respectively. In contrast, the LF–IMF sequence shows significantly lower frequency and retains certain randomness compared to the high-frequency components, making it suitable as a low-frequency sequence. The TF–IMF sequence captures the overall trend of power generation and is thus identified as the trend sequence. After decomposition and reconstruction, the original PV power generation data were transformed into four more stable and smoother reconstructed datasets, which enableld more effective feature extraction from the generation data, bolstering the model's forecasting accuracy and training robustness.

High-frequency sequences are often accompanied by strong stochastic volatility. To augment predictive accuracy in a manner that retains the full original data fidelity, we applied VMD to the first highest-frequency component obtained after merging and reconstruction. VMD can decompose a signal into multiple IMFs, meticulously extracting components of different frequencies, thereby effectively removing high-frequency noise while preserving useful signal characteristics. As a result, the forecasting potential of the data is increased, alongside the

https://doi.org/10.7441/joc.2026.03.11

model's predictive accuracy. Envelope entropy (EE) is typically used to measure the complexity or disorder of a signal. Minimum EE usually indicates a high concentration of signal energy, implying lower signal fluctuation. This study leveraged the particle swarm optimization (PSO) to identify the optimal parameter set for the VMD model that yielded the minimum EE. By minimizing the EE, the impact of high-frequency noise could be reduced and important signal components could be retained—especially within high-frequency modes—aiding in improving the signal's predictability or enabling more accurate modeling. Consequently, the high-frequency sequence HF–IMF was decomposed. PSO optimized the VMD process to minimize the EE, decomposing the high-frequency sequence HF–IMF into four components. The final seven input components obtained were HF-IMF1, HF–IMF2, HF–IMF3, HF–IMF4, MF–IMF, LF–IMF, and TF–IMF, as shown in Fig.6.

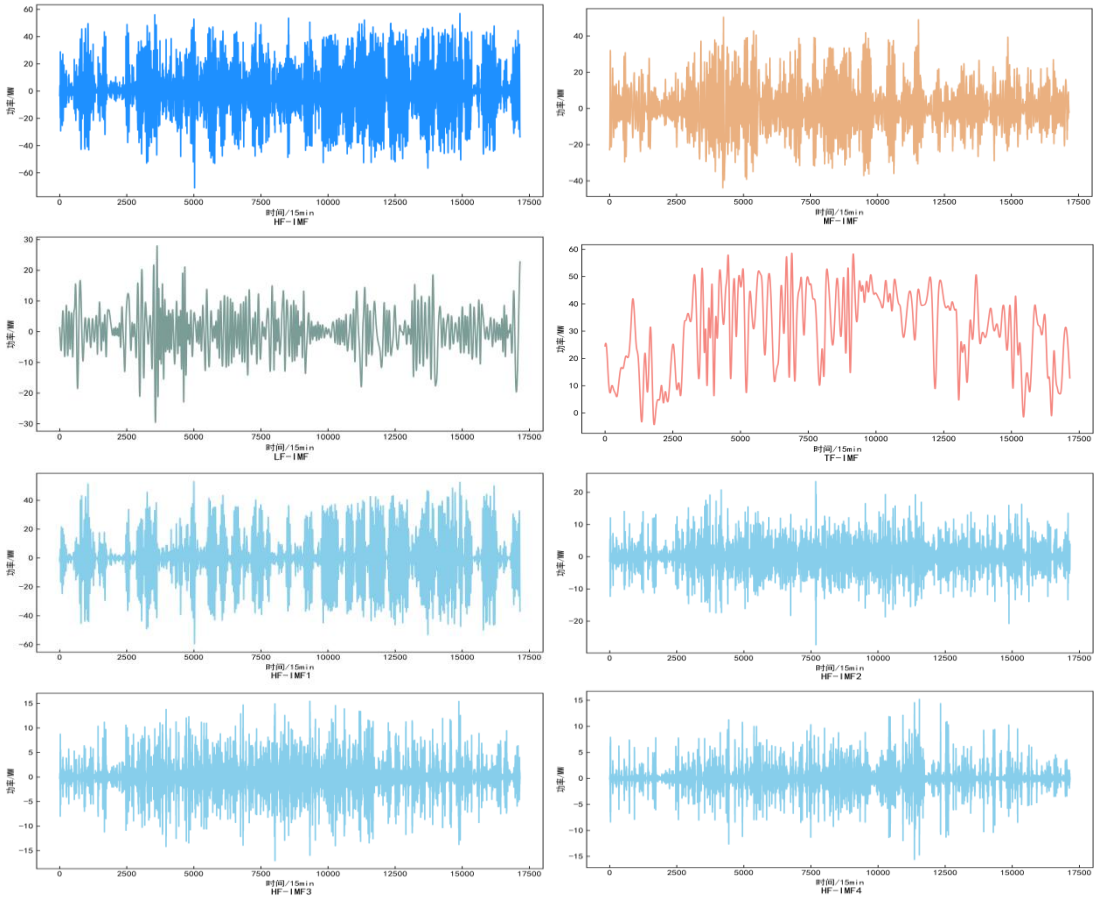
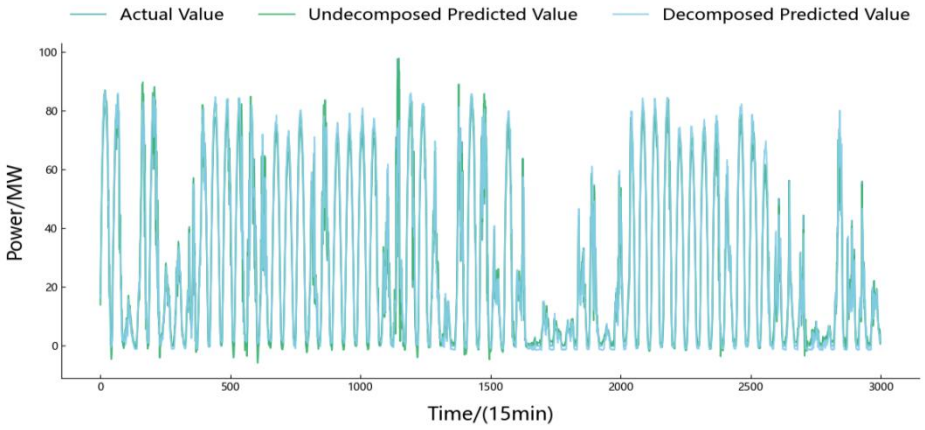


Fig. 6 – The process of PV power generation sequence decomposition and VMD Denoising in Central China Hills. Source: own research

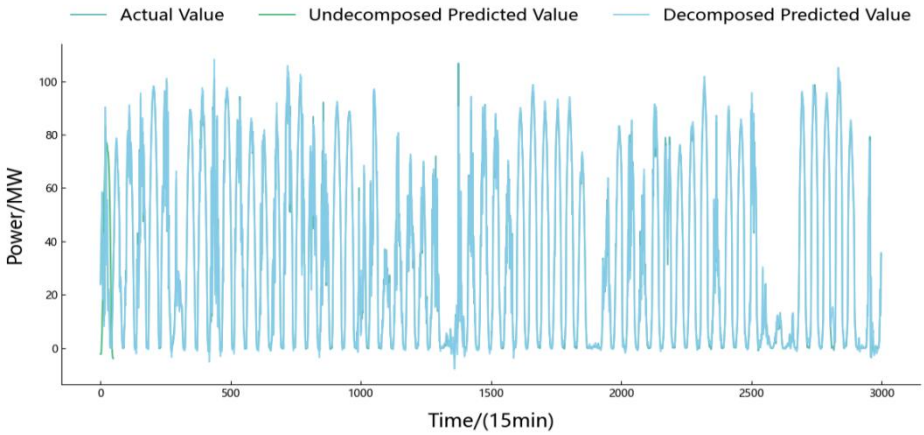
This study adopted a combined approach integrating the CEEMDAN–SE–VMD dual-decomposition algorithm with the CNN–BiLSTM–Transformer model to predict PV power generation. A comparative analysis was performed to evaluate the performance of the present method against that of several benchmark models, including CNN, MLP, GRU, BiLSTM, Attention, CNN–BiLSTM, BiGRU–Attention, and CNN–BiLSTM–Attention. A rigorous chronological split was applied to divide the dataset into training and validation sets, preventing data leakage in the time series. All models adhered to the same data-partitioning protocol, maintaining identical training and testing sets for an unbiased evaluation of their PV-generation forecasting capabilities.

https://doi.org/10.7441/joc.2026.03.11

The initial step demonstrated that the CEEMDAN–SE–VMD dual-decomposition algorithm effectively enhanced the prediction accuracy. The CNN–BiLSTM–Transformer model was used to predict both the original dataset without dual decomposition and reconstructed dataset processed through dual decomposition. The prediction results were compared and analyzed. A time-series plot of the predicted values versus actual values is presented in Fig.7, and the evaluation metrics are presented in Tab.3.



a. Central China Hills



b. East China Plain

Fig.7 – Effect of CEEMDAN-SE-VMD decomposition on prediction effect. Source: own research

Tab.3 – Effects of CEEMDAN-SE-VMD decomposition on evaluation indexes. Source: own research

Region	Model	MAE/MW	RMSE/MW	R^2
Central China Hills	No	4.3974	5.5839	0.9592
	VMD	1.8569	2.5775	0.9898
	CEEMDAN	3.2710	4.8802	0.9634
	CEEMDAN-VMD	1.8018	2.4653	0.9902
East China Plain	No	5.2897	9.8046	0.9084
	VMD	2.6608	3.6237	0.9874
	CEEMDAN	5.1719	7.8254,	0.9417
	CEEMDAN-VMD	1.6890	3.1440	0.9906

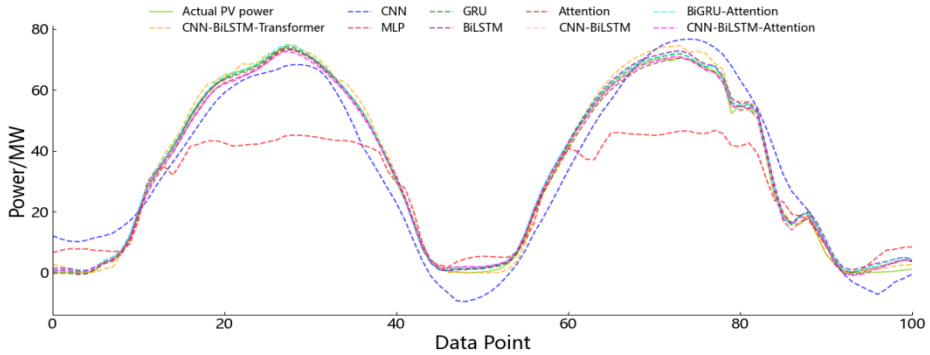
https://doi.org/10.7441/joc.2026.03.11

The evaluation metrics in Tab.3 reveal that different decomposition methods exert significant impacts on the model's forecasting performance. Under the same CNN–BiLSTM–Transformer forecasting model, the forecasting error of raw undecomposed data was relatively large, indicating that noise and nonstationary components in the original PV power sequence exerted a certain interference on the model's forecasting accuracy. After VMD, the model's forecasting performance was significantly improved: both MAE and RMSE were lower than those of the undecomposed scenario, and the correlation coefficient was higher, which demonstrated that VMD could, to a certain extent, extract the main features of the signal and mitigate the impact of noise. In comparison, CEEMDAN decomposition also played a positive role in noise reduction; however, the improvement in its error metrics was limited, with a less significant increase in the forecasting accuracy than that produced by VMD. Furthermore, CEEMDAN–VMD combined technique achieves more sufficient signal separation and noise reduction, on the basis of the above two single-decomposition methods, and its forecasting results exhibited optimal performance in terms of MAE, RMSE, and correlation coefficient. Taking the Central China Hills as an example, the forecasting results after CEEMDAN–VMD achieved an MAE of 1.8018, RMSE of 2.4653, and correlation coefficient of 0.9902, which were significantly superior to those of the undecomposed, single-VMD, and single-CEEMDAN-decomposed scenarios.

From the results listed in Tab.3 and Fig.7, the CEEMDAN–SE–VMD dual decomposition method could effectively reduce the noise and nonstationarity in PV power data, and enhance the model's ability to learn key features. Thus, it outperformed other decomposition methods in terms of forecasting accuracy, stability, and correlation, making the PV power-generation forecasting results closer to the actual values. This demonstrated that the CEEMDAN–SE–VMD dual-decomposition effectively reduced prediction errors and deviations, enhanced prediction accuracy and relevance, and enabled more precise forecasting of PV power generation.

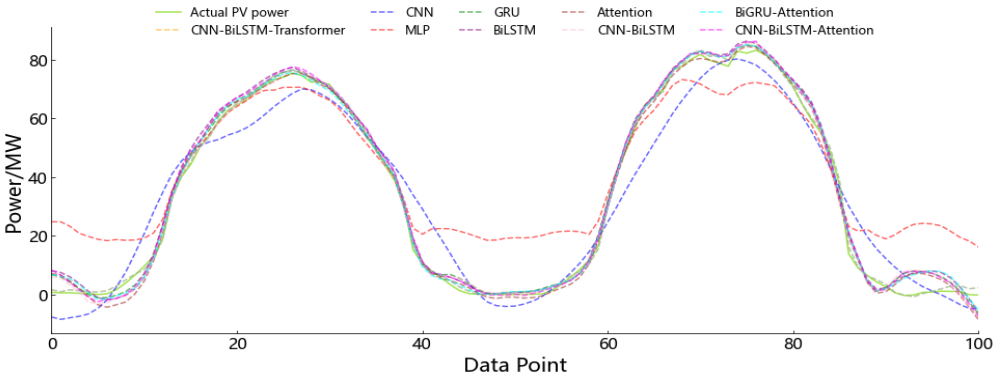
4.2 Model Application under Different Weather Conditions

Considering that PV system output is inherently contingent on atmospheric variables, short-term forecasting serves a key function in facilitating grid-dispatch departments on adjust plans and design appropriate control strategies. To quantify the technical proficiency of the newly developed hybrid deep-learning architecture, the CNN-BiLSTM-Transformer architecture was applied under both stable and unstable weather conditions. The forecasting outcomes are illustrated in Fig.8 and 9.



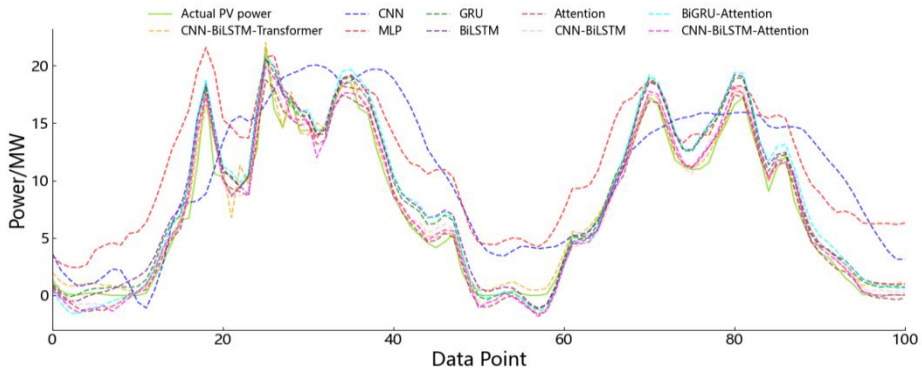
a. Central China Hills

https://doi.org/10.7441/joc.2026.03.11

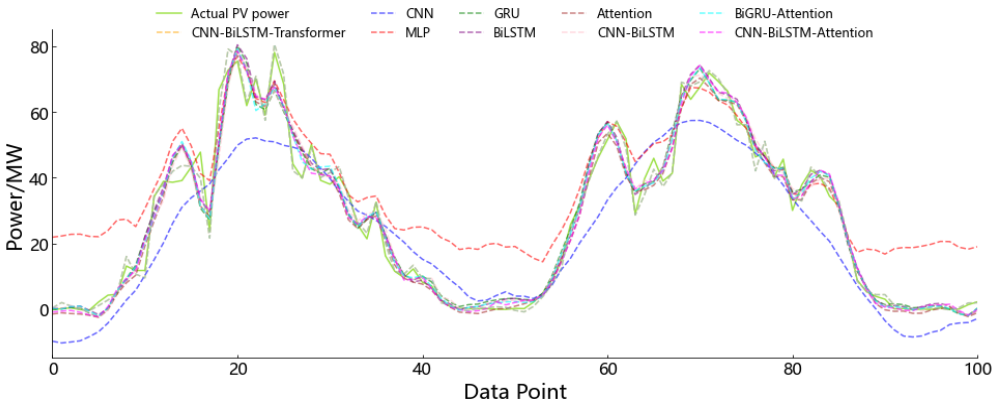


b. East China Plain

Fig.8 – Forecast results in stable weather. Source: own research



a. Central China Hills



b. East China Plain

Fig. 9 – Forecast results in non-stable weather. Source: own research

As shown in the figures, benchmark time-series models such as CNN and GRU were able to capture the periodicity and fluctuation trends of PV output but tended to fall into local optima and were prone to overfitting, resulting in relatively large forecast errors. The CNN–BiLSTM model provides moderate improvements, producing predictions closer to the actual values, though the enhancement remained limited. In contrast, under both stable and unstable weather scenarios, the CNN–BiLSTM–Transformer consistently delivered forecasts with higher accuracy and curves that were more closely aligned with the observed values, compared to the other eight models.

https://doi.org/10.7441/joc.2026.03.11

4.3 Model-Performance Evaluation

4.3.1 Prediction Performance under Different Weather Conditions of the Two Regions

Tab.4 and 5 present the error-evaluation metrics for both the Central China Hills and East China Plain. All models in this study, including the proposed model and all benchmark comparison models, were trained and evaluated on the same preprocessed dataset. The results provide intuitive and quantitative validation of the proposed model. Although all models demonstrated strong predictive performances under stable weather conditions, the CNN–BiLSTM–Transformer achieved the smallest MAE and RMSE, while also obtaining the highest correlation coefficient. This indicates the superior precision and fitting ability of the CNN–BiLSTM–Transformer in explaining data variations. Remarkably, across all four performance indicators, this model yielded the most favorable results, confirming its efficiency and accuracy in forecasting PV power generation for the immediate horizon.

Under unstable weather conditions also, the CNN–BiLSTM–Transformer delivered the best results. For example, in the Central China Hills dataset, the MAE and RMSE reached 0.6479 MW and 0.7698 MW, respectively—representing reductions of 82.99% and 83.76% compared to that of CNN 43.70% and 47.23% compared to that of GRU, 52.49% and 54.02% compared to that of BiLSTM, 25.40% and 28.03% compared to that of CNN–BiLSTM, and 25.95% and 32.03% compared to that CNN–BiLSTM–Attention. Furthermore, the correlation coefficient attained 0.9848, representing improvements of 56.90% over that of CNN, 3.99% over that of GRU, 5.75% over that of BiLSTM, 1.43% over that of CNN–BiLSTM, and 1.80% over that of CNN–BiLSTM–Attention.

Tab. 4 – Error evaluation of prediction by model under stable weather. Source: own research

Model	MAE		RMSE		R^2	
	CCHills	ECPlain	CCHills	ECPlain	CCHills	ECPlain
CNN	5.7600	6.6080	6.8229	8.1805	0.9262	0.9337
MLP	8.1549	10.4823	11.2505	12.9684	0.7993	0.8335
GRU	1.6684	2.1158	2.3123	2.8178	0.9915	0.9921
BiLSTM	2.1083	2.5127	2.6608	3.1897	0.9888	0.9899
Attention	1.6784	2.0851	2.4434	2.8478	0.9905	0.9920
CNN-BiLSTM	1.6919	2.0632	2.3151	2.8517	0.9915	0.9919
BiGRU-Attention	1.9184	2.0362	2.4228	2.7073	0.9907	0.9927
CNN-BiLSTM-Attention	1.5688	2.3822	2.3239	3.0476	0.9914	0.9908
CNN-BiLSTM-Transformer	1.5077	0.8909	2.1738	1.0767	0.9925	0.9989

CCHills represents Central China Hills, ECPlain represents East China Plain.

Overall, the application results highlighted that the proposed data-decomposition strategy effectively reduced noise and stabilized PV power series, enabling the model to learn more representative features and thereby improve the prediction accuracy. In stable weather, all models performed relatively well; yet, the hybrid approach remained the most accurate. In unstable conditions, the hybrid model significantly reduced the deviations between forecasts and actual values, while achieving higher correlation, thus demonstrating superior ability to capture latent patterns in complex forecasting tasks. This confirmed the model's robustness, precision, and applicability in diverse operational scenarios.

https://doi.org/10.7441/joc.2026.03.11

Tab. 5 – Error evaluation of prediction by model under non-stable weather. Source: own research

Model	MAE		RMSE		R^2	
	CCHills	ECPlain	CCHills	ECPlain	CCHills	ECPlain
CNN	3.8095	8.7813	4.7403	12.0594	0.4244	0.7741
MLP	4.2366	12.4450	4.6768	14.0460	0.4397	0.6936
GRU	1.1508	5.0966	1.4589	6.6687	0.9455	0.9309
BiLSTM	1.3637	4.8906	1.6742	6.3579	0.9282	0.9372
Attention	0.7239	4.6407	0.9388	6.2532	0.9774	0.9393
CNN-BiLSTM	0.8685	4.8927	1.0696	6.6326	0.9707	0.9317
BiGRU-Attention	1.4682	4.8460	1.7766	6.4298	0.9191	0.9358
CNN-BiLSTM-Attention	0.8749	4.7863	1.1326	6.1672	0.9671	0.9409
CNN-BiLSTM-Transformer	0.6479	1.7216	0.7698	2.4115	0.9848	0.9910

CCHills represents Central China Hills, ECPlain represents East China Plain.

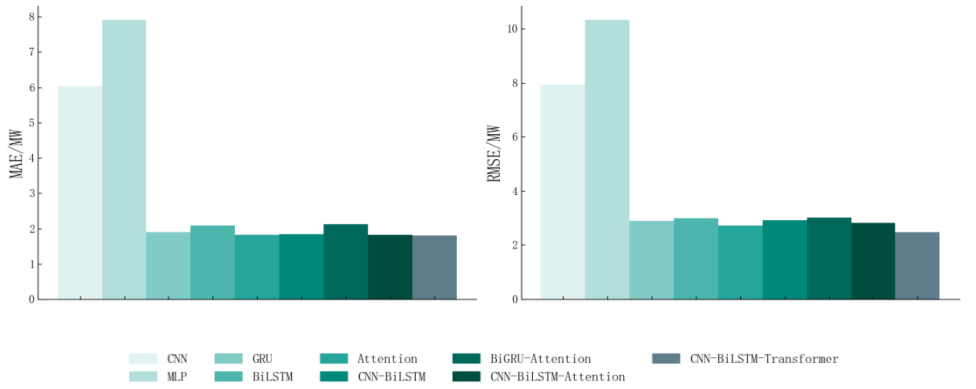
4.3.2 Overall Prediction Performance under Different Scenarios

To further assess the predictive performance of the CNN–BiLSTM–Transformer model against alternative approaches across the entire dataset, this study continued to employ the same three evaluation metrics (RMSE, MAE, and R^2). The comparative error assessment table and histogram of prediction errors are presented in Tab. 6 and Fig.10, respectively.

Tab.6 – Overall prediction error evaluation of each model. Source: own research

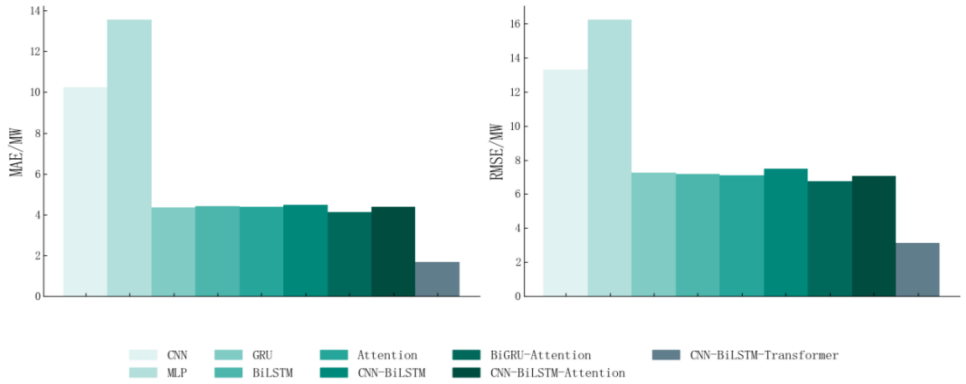
Model	MAE		RMSE		R^2	
	CCHills	ECPlain	CCHills	ECPlain	CCHills	ECPlain
CNN	6.0379	10.2446	7.9170	13.3155	0.9026	0.8309
MLP	7.9074	13.5535	10.3229	16.2397	0.8345	0.7485
GRU	1.9058	4.3626	2.8883	7.2520	0.9870	0.9499
BiLSTM	2.0883	4.4065	2.9994	7.1827	0.9860	0.9508
Attention	1.8304	4.3844	2.7283	7.1129	0.9884	0.9518
CNN-BiLSTM	1.8360	4.4610	2.9176	7.4690	0.9868	0.9468
BiGRU-Attention	2.1240	4.1056	3.0236	6.7463	0.9858	0.9566
CNN-BiLSTM-Attention	1.8138	4.3691	2.8307	7.0661	0.9876	0.9524
CNN-BiLSTM-Transformer	1.8018	1.6890	2.4653	3.1440	0.9902	0.9906

CCHills represents Central China Hills, ECPlain represents East China Plain.



a. Central China Hills

https://doi.org/10.7441/joc.2026.03.11

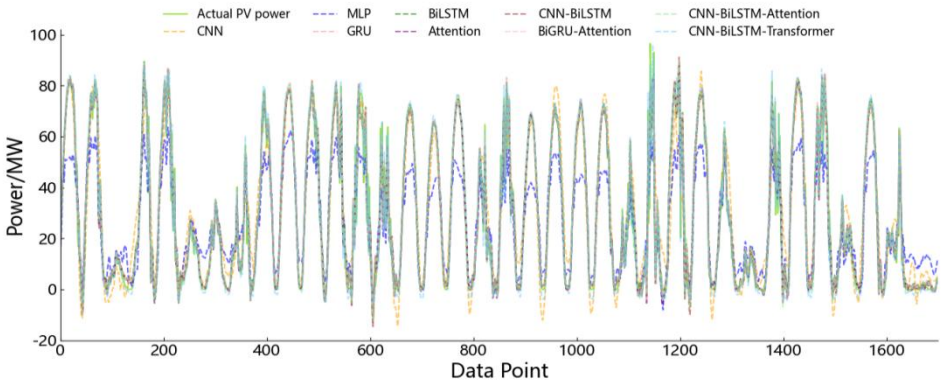


b. East China Plain

Fig. 10 – Histogram of the overall prediction error of each model. Source: own research

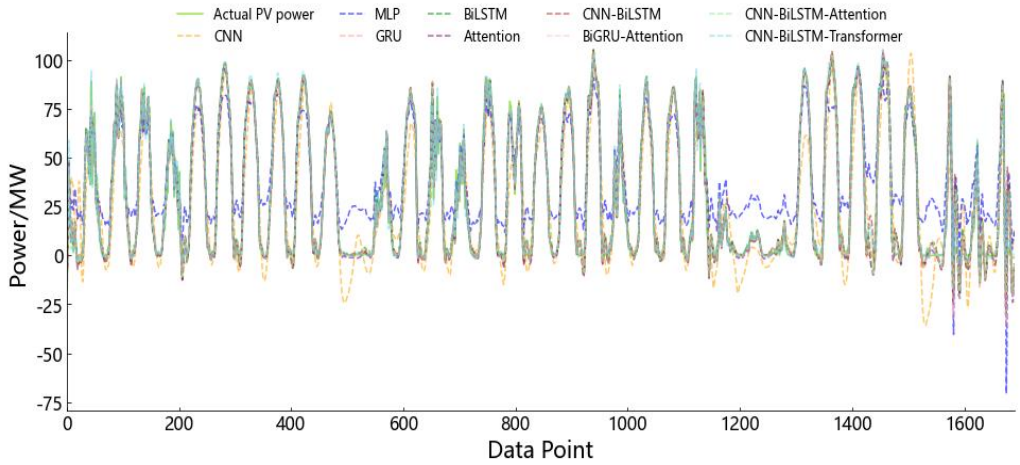
As shown in Tab.6 and Fig.10, the CNN–BiLSTM–Transformer model demonstrated superior performance, achieving the best goodness-of-fit, smallest prediction deviations, and lowest MAE and RMSE values, which collectively attested to its superior accuracy and robustness. The overall predictive performance curves of all models can be found in Fig.11.

Synthesizing the results from Tab.6, Fig.10, and Fig.11 revealed that the CNN–BiLSTM–Transformer model exhibits clear advantages relative to GRU, LSTM, and BiLSTM. Specifically, its prediction curves demonstrated the smallest error and highest degree of fit to the observed values, thereby enhancing both the accuracy and reliability. A closer agreement between predictions and actual measurements demonstrated the hybrid-model's enhanced capability, achieved through architectural fusion, to characterize the underlying complex dynamics of PV power data. Consequently, the model exhibited distinct superiority in both precision and stability.



a. Central China Hills

https://doi.org/10.7441/joc.2026.03.11



b. East China Plain

Fig. 11 – Overall prediction curve of each model. Source: own research

Moreover, across different datasets and weather conditions, the CNN–BiLSTM–Transformer approach consistently maintains accurate and stable performance, underscoring its adaptability and applicability in diverse forecasting scenarios. This outcome validated the model's practical utility and confirmed its viability for deployment within intricate real-world energy-management infrastructures.

4.3.3 Diebold Mariano test

Tab.7 – Diebold Mariano test. Source: own research

Model	DM statistic		P-value	
	CCHills	ECPlain	CCHills	ECPlain
CNN	-29.887	-33.803	0.000	0.000
MLP	-36.316	-38.349	0.000	0.000
GRU	-14.988	-13.345	0.000	0.000
BiLSTM	-15.624	-13.122	0.000	0.000
Attention	-14.142	-13.201	0.000	0.000
CNN-BiLSTM	-14.940	-12.867	0.000	0.000
BiGRU-Attention	-15.999	-12.718	0.000	0.000
CNN-BiLSTM-Attention	-14.184	-12.828	0.000	0.000

Following Diebold & Mariano (1995), The Diebold-Mariano test results, as presented in Tab. 7, reveal that the CNN-BiLSTM-Transformer model achieves significantly superior predictive accuracy compared to the benchmark models across both regions. The DM test statistics are consistently negative, with corresponding p-values approaching zero. This indicates that the null hypothesis of equal predictive accuracy is rejected in all pairwise comparisons. Consequently, the forecasting errors generated by the CNN-BiLSTM-Transformer model are statistically significantly lower than those of the baseline models. These findings robustly demonstrate the efficacy of this hybrid approach.

https://doi.org/10.7441/joc.2026.03.11

5 Case Study on the Practice of Enhancing Power-System Stability

5.1 Practical Cases from the East China Plain and Central China Hills

This study primarily focused on the enhancement of PV-forecasting accuracy and examined its impact on power-system stability through the analysis of existing practical cases. In terms of case selection, examples from the State Grid Xuzhou Power Supply Company and State Grid Zhejiang Electric Power Company were used to represent the East China Plain region. For Central China Hills, two representative cases were selected: first, the "one-site-one-strategy" meteorological service model adopted by the Hubei Meteorological Service Center, and second, the coordinated hydro-wind-solar dispatch practice implemented by the Hubei Qingjiang Company.

State Grid Xuzhou Power Supply Company:

In July 2023, the State Grid Xuzhou Power Supply Company, in collaboration with the Xuzhou Meteorological Observatory, developed a new numerical wind–solar forecasting technology based on local geographical and meteorological conditions, and applied it to PV power prediction. This facilitated accurate forecasting with a 15 min resolution for periods exceeding 10 days, effectively supporting flexible grid dispatch and power-flow regulation. After implementation, the PV forecasting accuracy improved by approximately 2.85%, and through measures such as optimized maintenance, it helped increase the monthly green-power generation from PV stations by approximately 400,000 kWh.

State Grid Zhejiang Electric Power Company:

Relying on the AI-based PV forecasting system named "RuiCe" (Intelligent Forecast) developed by its Electric Power Research Institute, the State Grid Zhejiang Electric Power Company increased the accuracy of PV power prediction to 96.5% during typhoon periods. Based on these forecasts, the dispatch center formulated daily generation plans and accurately reserved standby capacity, effectively mitigating fluctuations from renewable energy. The forecast error during morning peak hours was reduced by 4,000 MW, equivalent to saving the regulating capacity of at least four million-kilowatt-level generating units.

Hubei Qingjiang Company:

In hilly areas, the PV output fluctuates significantly owing to terrain-influenced cloud systems. Accurate forecasting helps prepare flexible hydropower in advance and quickly fill power shortages when the solar output drops abruptly. For example, the Hubei Qingjiang Company used wind–solar power-forecast curves to formulate daily hydropower output plans. During an autumn flood event in October 2025, forecast-based optimized dispatch achieved both safety and economic benefits.

Hubei Meteorological Service Center:

For PV stations scattered across hilly terrain, forecasting helps dispatch departments anticipate local voltage variations. The Hubei Meteorological Service Center provides power forecasts for 47 renewable-energy stations in the Huanggang area, supporting dispatch decision-making.

These practical application cases demonstrate that improving the forecasting accuracy directly reduces the power shortage caused by sudden drops in PV output, which is crucial for maintaining the frequency stability. More accurate forecasts also facilitate advance detection of local voltage changes, allowing proactive adjustment of reactive-power equipment, thereby, maintaining voltage levels.

https://doi.org/10.7441/joc.2026.03.11

5.2 Key Factors Influencing the Adoption of AI-based PV Power Forecasting Systems

This paper next presents the construction of a technology–organization–environment framework analysis of the factors influencing the application of AI prediction systems, drawing on the aforementioned cases. The adoption of AI prediction systems in power systems results from a combination of technological feasibility, organizational compatibility, and environmental necessity (Tornatzky et al., 1990). Drawing on specific cases from Central and Eastern China, the following sections provide a detailed elaboration from these three dimensions.

The adoption of AI prediction systems hinges primarily on the persuasiveness and feasibility of their technical attributes. Relative advantages and performances are the most direct drivers, requiring systems to demonstrate marked superiority over conventional methods. For example, Zhejiang's “RuiCe” System achieved 96.5% forecast accuracy during typhoons, and research in Hubei reduced the prediction volatility (standard deviation) by 50%. Such quantifiable gains are central to adoption decisions. Secondly, compatibility and integrability determine successful deployment. Systems must integrate seamlessly with the grid's existing dispatch platforms, data interfaces, and business processes. Instances include Jiangsu's province-wide system facilitating minute-level data access and Zhejiang's forecasts directly informing day-ahead generation plans, underscoring the value of deep integration. Finally, complexity and adaptability impact scalability and implementation scope. Systems need sufficient flexibility for diverse environments: Hubei employed customized “one-site-one-policy” models for hilly terrain, while Jiangsu's system had to handle massive distributed PV integration. These demanded technologies with high configurability and environmental adaptability, to reduce deployment and operational complexity.

The realization of technological advantages is highly dependent on internal organizational conditions. Organizational scale and resources form the foundational prerequisite. Large grid enterprises possess the capital, data, and talent reserves necessary for deploying province-wide systems, representing a key advantage that smaller organizations struggle to match. Management style and collaborative culture act as the “lubricant” for integrating technology into workflows. The effective operation of prediction systems requires departments such as dispatch, operations, maintenance, and meteorology to break down data silos and reengineer processes. The deep collaborative service mechanism established between the Hubei Meteorological Service and power company served as a successful model of cross-organizational, cross-departmental cooperation, ensuring smooth data and workflow integration. Furthermore, top-level support and strategic alignment provided sustained momentum and legitimacy for projects. Only when technology adoption explicitly served the organization's highest strategic goals, such as “ensuring grid security and stability” and “enhancing renewable energy integration capacity”, could it secure enduring resource allocation and policy support. This explains why related projects received high priority and continuous advancement.

External environments shape the urgency and direction of technology adoption. Policy and regulatory pressure constitute the strongest external push. Under China's “Dual Carbon” goals and stringent renewable energy-integration assessment policies, improving forecast accuracy is not merely a technical enhancement, but rather, a mandatory measure to meet regulatory requirements. This policy environment creates a compulsory market demand for prediction systems. Market and competitive pressure provide the economic pull, justifying the investment. For instance, Hubei Qingjiang Company, through coordinated hydro–wind–PV dispatch integrated with forecasting, achieved water savings exceeding 100 million cubic meters and

https://doi.org/10.7441/joc.2026.03.11

direct economic benefits of over ten million yuan during a single flood event. This clearly demonstrated the return on investment, strengthening organizational adoption willingness. Finally, regional endowment and partner networks determine the differentiation of technological pathways. Different regions face distinct core challenges: Zhejiang must cope with typhoon weather, Hubei needs to address output volatility in hilly terrain, and Jiangsu must manage high-penetration distributed generation. These unique environmental constraints drive collaborations with different universities and research institutes, fostering localized innovative solutions.

6 Conclusions, Implications, Limitations, and Future

6.1 Conclusions

As a technological integration that collects and processes data by simulating human intelligence (Biloslavo et al., 2024), AI can support human decision-making. This paper examined the application of AI technology in the power sector, for PV power forecasting. High-accuracy PV power prediction could effectively assist power companies in making dispatch decisions (Peng et al., 2024), thereby facilitating the transition of the energy structure and contributing to sustainable development goals (Ferilli et al., 2021). However, the introduction of more advanced deep-learning technologies in business not only entailed technical upgrades but also triggered business-model restructuring (Jobstreibizer et al., 2025). Through accurate forecasting, intelligent dispatch, and ecological synergy, PV companies could establish an advantage in the evolving electricity market, while providing critical support for the global carbon-neutrality goals. This technology-enabled commercial value extended from the generation side to the entire energy-value chain, becoming a core driver of the new energy revolution.

Due to the variability and unpredictability of PV output, its large-scale integration placed significant stress on the operation of power systems, thereby affecting overall system stability (Rahman, 2022). This study established a unified framework that incorporated both power-system stability and PV power forecasting. It theoretically explored the mechanism through which forecast accuracy impacted system stability and conducted empirical studies using real-world cases from Central China Hills and East China Plain. This approach addressed a gap in the existing literature, which predominantly focuses on technical forecasting methods, while often neglecting practical application analyses (Wang & Xiang, 2025). Building on this foundation, we further developed a high-accuracy PV power forecasting model based on machine learning, aimed at providing technical support for enhancing power-system stability through improved forecast precision. The research findings indicated that data processed through the CEEMDAN–VMD dual-decomposition technique exhibits enhanced stationarity, effectively boosting the model's forecasting accuracy (Xue & Li, 2026). The proposed CNN–BiLSTM-–Transformer model demonstrates superior performance in short-term PV power-forecasting tasks, with its predictions closest to the actual values, compared to other models, showcasing its excellent forecasting capability (Dai et al., 2025). This provided robust technical support for power dispatch decision-making in the two regions. Furthermore, the study validated the model using PV plants in two distinct climatic and topographic regions—the Central China Hills and East China Plain. The results showed that the model effectively integrated multisource meteorological and geographical features, demonstrating strong adaptability in both scenarios. To some extent, this model overcame the challenges related to insufficient cross-regional data utilization and forecasting difficulties under complex weather conditions, consequently increasing the model's generalization capability to some degree.

https://doi.org/10.7441/joc.2026.03.11

6.2 Implications

This study yields the following policy implications from its findings:

Promoting the standardized application of high-precision forecasting technologies. Research indicated that employing CEEMDAN–VMD dual decomposition combined with deep-learning models such as CNN–BiLSTM–Transformer could dramatically refine the correctness of PV power forecasts. It was recommended that relevant authorities promote the standardization and modularization of such advanced forecasting algorithms, incorporating them into the assessment and subsidy mechanisms for new energy stations. This would encourage power-generation enterprises to adopt high-precision forecasting models, thereby mitigating the grid impact of PV output uncertainty at its source.

Establishing differentiated dispatching mechanisms based on meteorological and geographical characteristics. This study validated the adaptability of the proposed model in two distinct scenarios: Central China Hills and East China Plain. Policymakers should fully consider the influence of regional climatic and topographical characteristics on PV output, avoiding a "one-size-fits-all" approach to dispatching. It is recommended that grid companies develop differentiated forecasting-accuracy standards and reserve capacity-allocation plans based on the local prevailing weather types (e.g., cloudy, rainy, sunny) and terrain conditions, thereby facilitating refined dispatching.

Improving ancillary services and compensation mechanisms in the electricity market. Enhancements in PV power forecasting accuracy are directly related to the demand for grid reserve capacity. Policymakers should further refine the rules of the electricity ancillary services market, linking forecasting accuracy with the revenues of power-generation enterprises. Incentives should be provided to enterprises with high forecasting accuracies that effectively reduce system operational risks, while those with significant forecasting errors that increase grid-balancing pressure should be subject to assessments. This market-based leverage can guide the industry toward improving forecasting standards.

Increasing investment in the interdisciplinary field of AI-driven energy meteorology. The research outcomes demonstrate the potential of AI in addressing the volatility of new energy sources. It is recommended that governments and scientific-research-management agencies enhance support for the interdisciplinary field integrating "meteorological science, artificial intelligence, and power systems." Encouragement should be given to the development of ultrashort-term downscaling forecasting techniques tailored for complex terrains, providing robust scientific and technological support for the construction of power systems featuring a high share of renewable energy.

6.3 Limitations and Future

This paper has only established relevant technical models to explore the application value of artificial intelligence technologies in the power system field, and from the perspective of improving forecasting accuracy, it elucidates the promotional effect of such technologies on the stable operation of power systems. However, this study has merely analyzed the mechanism underlying the impact of forecasting accuracy on power system stability at the theoretical level, and has not yet conducted quantitative analysis work targeted at power system stability. In future research, quantitative research on the impact of forecasting accuracy on power system stability can be carried out based on the model established in this paper by adopting numerical

https://doi.org/10.7441/joc.2026.03.11

simulation methods; furthermore, from the perspective of organizational management, the dispatching optimization problems of power systems can be further investigated in depth by combining with cases that are more in line with actual engineering practices. Furthermore, the model constructed in this study was only validated in Central China Hills and East China Plain, and its generalization capability remains to be further tested. Future research could conduct model validation under a wider range of climatic and topographic conditions to enhance its applicability and broader value.

Acknowledgements

This work was supported by the National Natural Science Foundation of China [Grant No. 72473058,72163008,72563012], Jiangxi Provincial Social Sciences Fund [Grant No. 24YJ12], the research project of Jiangxi University of Water Resources and Electric Power [Grant No. 2022kyqd046], Zhejiang Provincial Philosophy and Social Sciences Planning Project [Grant No. 26YJRC015ZD], Zhejiang International Studies University, Institute for Advanced Studies on New International Relations (IASNIR) Annual Major Project under Grant number [Grant No. 2025GYYZD08].

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. https://doi.org/10.1016/j.rser. 2020.109792
2. Bhatt, A., Ongsakul, W., & Singh, J. G. (2022). Sliding window approach with first-order differencing for very short-term solar irradiance forecasting using deep learning models. *Sustainable Energy Technologies and Assessments*, 50, 101864. https://doi.org/10.1016/j.seta.2021.101864
3. Biloslavo, R., Edgar, D., Aydin, E., & Bulut, C. (2025). Artificial intelligence (AI) and strategic planning process within VUCA environments: a research agenda and guidelines. *Management Decision,* 63(10), 3599-3624. https://doi.org/10.1108/MD-10-2023-1944
4. Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time Series Analysis: *Forecasting and Control* (5th ed.). Wiley.
5. Chen, F., Su, J., Li, Q., Duan, M., Liu, Z., & Zheng, H. (2023). Short-term power prediction of distributed photovoltaic based on CNN-LSTM and meteorological interpolation. In Fourth International Conference on Artificial Intelligence and Electromechanical Automation. *AIEA*, 12709, 1331-1337. https://doi.org/10.1117/12. 2685707
6. Cui, L., Li, C. H., et al. (2025). Distributed cooperative voltage control method based on PV power generation prediction. *Renewable Energy Resources*, 43(08), 1106-1113. doi:10.13941/j.cnki.21-1469/tk.2025.08.010
7. Dai, Y., Wang, Y., Chen, Y., Wu, J., Chao, J. (2025). Combining meteorological and power information of station-measurement and model-prediction with the hybrid CNN-Transformer and CNN-BiLSTM for ultra-short-term photovoltaic power forecasting, *International Journal of Electrical Power & Energy Systems*, 171, 111009. https://doi.org/10.1016/j.ijepes.2025.111009
8. De Soto, W., Klein, S. A., & Beckman, W. A. (2006). Improvement and validation of a model for photovoltaic array performance. *Solar energy*, 80(1), 78-88. https://doi.org/10.1016/j.solener.2005.06.010

https://doi.org/10.7441/joc.2026.03.11

9. Diebold, FX., Mariano, RS. (1995). Comparing Predictive Accuracy. J*ournal of Business and Economic Statistics*, 13(3), 134-144. doi:10.1080/07350015.1995.10524599
10. Dragomiretskiy, K., & Zosso, D. (2013). Variational mode decomposition. *IEEE transactions on signal processing*, 62(3), 531-544. https://doi.org/10.1109/TSP.2013.2288675
11. Dobos, A. P. (2014). PVWatts version 5 manual (No. NREL/TP-6A20-62641). *National Renewable Energy Lab.(NREL), Golden, CO (United States)*.
12. Doubleday, K., Jascourt, S., Kleiber, W., & Hodge, B. M. (2020). Probabilistic solar power forecasting using bayesian model averaging. *IEEE Transactions on Sustainable Energy*, 12(1), 325-337. https://doi.org/10.1109/TSTE.2020.2993524
13. Gaboitaolelwe, J., Zungeru, A. M., Yahya, A., Lebekwe, C. K., Vinod, D. N., & Salau, A. O. (2023). Machine learning based solar photovoltaic power forecasting: A review and comparison. *IEEE Access*,11,40820-40845. https://doi.org/10.1109/ACCESS.2023.3270041
14. Gupta, P., Singh, R. (2024). Effect of PV power forecast error on the frequency of a standalone microgrid system. *Journal of Renewable and Sustainable Energy*,16(4), 046101. doi:10.1063/5.0209826
15. Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971), 903-995. https://doi.org/10.1098/rspa.1998.0193
16. Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735
17. Kundur, P., Paserba, J., Ajjarapu, V., Andersson, G., Bose, A., Canizares, C., ... & Vittal, V. (2004). IEEE/CIGRE joint task force on stability terms and definitions, definition and classification of power system stability. *IEEE Trans on Power Systems*, 19(2),1387-1401.
18. Kim, H. J., & Kim, M. K. (2023). A novel deep learning-based forecasting model optimized by heuristic algorithm for energy management of microgrid. *Applied Energy*, 332, 120525. https://doi.org/10.1016/j.apenergy.2022.120525
19. Li, S., Chen, H., Wang, M., Heidari, A. A., & Mirjalili, S. (2020). Slime mould algorithm: A new method for stochastic optimization. *Future generation computer systems*, 111, 300-323. https://doi.org/10.1016/j.future.2020.03.055
20. Li, W., Shi, Q., Sibtain, M., Li, D., & Mbanze, D. E. (2020). A hybrid forecasting model for short-term power load based on sample entropy, two-phase decomposition and whale algorithm optimized support vector regression. *IEEE access*, 8, 166907-166921. doi: 10.1109/ACCESS.2020.3023143
21. LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551. https://doi.org/10.1162/neco.1989.1.4.541
22. Li, Y., Ye, F., Liu, Z., Wang, Z., & Mao, Y. (2021). A short-term photovoltaic power generation forecast method based on LSTM. *Mathematical Problems in Engineering*, 1, 6613123. https://doi.org/10.1155/2021/6613123
23. Li, Y., Liu, G., Cao, Y., Chen, J., Gang, X., & Tang, J. (2025). WNPS-LSTM-Informer: A Hybrid Stacking model for medium-term photovoltaic power forecasting with ranked feature selection. *Renewable Energy*, 244, 122687. https://doi.org/10.1016/j.renene.2025.122687

https://doi.org/10.7441/joc.2026.03.11

24. Li, Y.K., Huang, W., Lou, K.Y., Zhang, X.Z. (2024). Short-term PV power prediction based on meteorological similarity days and SSA-BiLSTM. *Systems and Soft Computing*, 6, 200084. https://doi.org/10.1016/j.sasc.2024.200084
25. Liu, W., Liu, Q., & Li, Y. (2024). Ultra-short-term photovoltaic power prediction based on modal reconstruction and BiLSTM-CNN-Attention model. *Earth Science Informatics*, 17(3), 2711-2725. https://doi.org/10.1007/s12145-024-01308-4
26. Malinkovich, Y., Sitbon, M., Lineykin, S., Dagan, K. J., Baimel, D. (2024). A combined persistence and physical approach for Ultra-Shorterm photovoltaic power forecasting using distributed sensors. *Sensors (Basel)*, 24(9), 2866. doi: 10.3390/s24092866
27. Mohanasundaram, V., & Rangaswamy, B. (2025). Photovoltaic solar energy prediction using the seasonal-trend decomposition layer and ASOA optimized LSTM neural network model. *Scientific Reports*, 15(1), 4032. https://doi.org/10.1038/s41598-025-876 25-0
28. Niu, D., Yu, M., Sun, L., Gao, T., & Wang, K. (2022). Short-term multi-energy load forecasting for integrated energy systems based on CNN-BiGRU optimized by attention mechanism. *Applied Energy*, 313, 118801. https://doi.org/10.1016/j.apenergy.2022. 118801
29. Ozan, G. Dynamic load flow in modern power systems: renewables, crypto mining, and electric vehicles. *Sustainability*,2025,17(6), 2515. https://doi.org/10.3390/su17062 515
30. Peng, D., Liu, Y., Wang, D., Luo, L., Zhao, H., and Qu, B. (2024). Short-term PV-Wind forecasting of large-scale regional site clusters based on FCM clustering and hybrid Inception-ResNet embedded with Informer. *Energy Conversion and Management*, 320, 118992. https://doi.org/10.1016/j.enconman.2024.118992
31. Qazi, A. (2025). Exploring digital competitiveness through Bayesian belief networks. *Journal of Competitiveness*, 17(2), 57-87. https://doi.org/10.7441/joc.2025.02.03
32. RajasundrapandiyanLeebanon, T., Murugan, N. S., Kumaresan, K., & Jeyabose, A. (2025). Long-term solar radiation forecasting in India using EMD, EEMD, and advanced machine learning algorithms. *Environmental Monitoring and Assessment*, 197(3),1-36. https://doi.org/10.1007/s10661-025-13738-8
33. Rahman, S., Saha, S., Haque, M.E., et al. (2022). A framework to assess voltage stability of power grids with high penetration of solar PV systems. *International Journal of Electrical Power & Energy Systems*, 139,107815. https://doi.org/10.1016/j.ijepes.2021. 107815
34. Ren, J., J., Wei, H., H., Zou, Z., L., Hou, T., T., Yuan, Y., L., Shen, J., Q., & Wang, X., M. (2022). W-BiLSTM based Ultra-short-term generation power prediction method of renewable energy. *Power System Protection and Control*, 50(08),108-116. https://doi. org/10.19783/j.cnki.pspc.211187
35. Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American journal of physiology-heart and circulatory physiology*, 278(6), 2039-2049. https://doi.org/10.1152/ajpheart.2000.278.6. H2039
36. Siano, P. (2014). Demand response and smart grids—A survey. *Renewable and sustainable energy reviews*, 30, 461-478. https://doi.org/10.1016/j.rser.2013.10.022
37. Singh, U., Rizwan, M., Alaraj, M., & Alsaidan, I. (2021). A machine learning-based gradient boosting regression approach for wind power production forecasting: A step towards smart grid environments. *Energies*, 14(16), 5196. https://doi.org/10.3390/en141 65196
38. Singh, G. K. (2013). Solar power generation by PV (photovoltaic) technology: A review. *Energy*, 53, 1-13. https://doi.org/10.1016/j.energy.2013.02.057

https://doi.org/10.7441/joc.2026.03.11

39. Smith, O., Coombes, S., & O'Dea, R. D. (2024). Stability analysis of electrical microgrids and their control systems. *PRX Energy*, 2024, 3(1), 013011. https://doi.org/10. 1103/PRXEnergy.3.013011
40. Song, S., Min, Sw. & Jung, S. (2025). Steady-state data-driven dynamic stability assessment in the Korean power system. *Scientific Reports,* 15, 7756. https://doi.org/10.1038/s41598-025- 90798-3
41. Sun, H. D.,Tang, Y., Ma, S. Y. (2006). A commentary on definition and classification of power system stability. *Power System Technology*, 30(17), 31-35. doi:10.13335/j.1000-3673.pst.2006.17.006
42. Suri, D., de Chalendar, J., & Azevedo, I. M. L. (2025). Assessing the real implications for CO2 as generation from renewables increases. *Nature Communications*, 16(1), 7124. https://doi.org/10.1038/s41467-025-59800-4
43. Sweeney, C., Bessa, R. J., Browell, J., & Pinson, P. (2020). The future of forecasting for renewable energy. *Wiley Interdisciplinary Reviews: Energy and Environment*, 9(2), e365. https://doi.org/10.1002/wene.365
44. Thaker, J., & Hoeller, R. (2022). A comparative study of time series forecasting of solar energy based on irradiance classification. *Energies*, 15(8), 2837. https://doi.org/10.3390 /en 15082837
45. Tahir, M. F., Yousaf, M. Z., Tzes, A., El Moursi, M. S., & El-Fouly, T. H. (2024). Enhanced solar photovoltaic power prediction using diverse machine learning algorithms with hyperparameter optimization. *Renewable and Sustainable Energy Reviews*, 200, 114581. https://doi.org/10.1016/j.rser.2024.114581
46. Torres, M. E., Colominas, M. A., Schlotthauer, G., & Flandrin, P. (2011). A complete ensemble empirical mode decomposition with adaptive noise. *In 2011 IEEE international conference on acoustics, speech and signal processing(ICASSP)*, IEEE, 4144-4147. https://doi.org/10.1109/ICASSP.2011.5947265
47. Tornatzky, L.G., and Fleischer, M. (1990). The Processes of Technological Innovation, Lexington Books.
48. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30, 5999-6009. http://arxiv.org/abs/1706.03762
49. Villalva, M. G., Gazoli, J. R., & Ruppert Filho, E. (2009). Comprehensive approach to modeling and simulation of photovoltaic arrays. *IEEE Transactions on power electronics*, 24(5)5, 1198-1208. https://doi.org/10.1109/TPEL.2009.2013862
50. Wang, M., Xiang, C. (2025). Machine learning for power forecasting in BIPV and BAPV Systems: A review. *Journal of Building Engineering*, 113, 114117. https://doi.org/10.10 16/j. jobe. 2025.114117
51. Xie, X. Y., Zhou, J., H., Zhang, Y., J., Wang, J. & Su, J., Y. (2021). W-BiLSTM Based Ultra-short-term Generation Power Prediction Method of Renewable Energy. *Automation of Electric Power Systems*, 45(08), 175-184.
52. Xue, S., Li, L. (2026). Photovoltaic power forecasting based on secondary decomposition strategy and hybrid model. *Scientific Reports*. https://doi.org/10.1038/s41598-026-42896-z
53. Yang, D., Kleissl, J., Gueymard, C. A., Pedro, H. T., & Coimbra, C. F. (2018). History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Solar Energy*, 168, 60-101. https://doi.org/10.1016/j.solener. 2017.11. 023

https://doi.org/10.7441/joc.2026.03.11

54. Zhang Y, Shi S, Yang Q, et al. (2020). Design of output fluctuation smoothing strategy in photovoltaic power station [C]//*E3S Web of Conferences*. EDP Sciences, 194: 03021
55. Zhou, D., Liu, Y., Wang, X., Wang, F., & Jia, Y. (2025). Combined ultra-short-term photovoltaic power prediction based on CEEMDAN decomposition and RIME optimized AM-TCN-BiLSTM. *Energy*, 318, 134847. https://doi.org/10.1016/j.energy.2025.134847

Appendix A The specific decomposition steps of CEEMDAN

The specific decomposition steps are as follows:

Step 1: Initialize

Set the noise amplitude coefficient ε_0, integration times I, and maximum decomposition layers K.

Initializing the residual signal: $r_0(t) = x(t)$

Initialize IMF set:

$$IMF_k(t) = 0, k = 1,2,\cdots,K$$

Step 2: Extraction of the First Stage IMF

Generate the noise signal in Group I:

$$X^{(i)}(t) = r_0(t) + \varepsilon_0\omega^{(i)}(t), t = 1,2,\cdots,I$$

Perform EMD decomposition on each signal group and extract the first-order IMF:

$$IMF_1^{(i)}(t) = E_1(x^{(i)}(t))$$

The final first-order IMF derivative is obtained by integration of the average.

$$IMF_1(t) = \frac{1}{I}\sum_{i=1}^{I} IMF_1^{(i)}(t)$$

Including original PV power sequence $x(t), t = 1,2,\cdots,N$; Gaussian white noise added in the i-th iteration $\omega^{(i)}(t)$; k stage noise amplitude coefficient ε_k; Operator for extracting the i-th eigenmode function (IMF) through EMD decomposition $E_k(\cdot)$; The residual signal of k stage $r_k(t)$; The k intrinsic modal function $IMF_k(t)$.

Step 3: Iteratively extract subsequent IMFs ($k \geq 2$) and update residuals r_k:

$$IMF_{k+1}(t) = \frac{1}{I}\sum_{i=1}^{I} E_1[r_k(t) + \varepsilon_k E_k(\omega_i(t))]$$

$$r_k(t) = r_{k-1}(t) - IMF_k(t)$$

Step 4: Termination Condition

When the residual $r_k(t)$ is a monotone function or the number of extreme points is less than 2, the algorithm terminates;

The final residual is denoted as $R(t) = r_k(t)$.

Step 5: Signal reconstruction, where the original photovoltaic power is represented as follows:

$$x(t) = \sum_{i=1}^{I} IMF_k(t) + R(t)$$

Appendix B The details of VMD

Step 1: The unilateral spectrum of each mode is obtained by analyzing K modal functions via the Hilbert transform.

$$\left(\delta(t) + \frac{j}{\pi x}\right) * Lu_k(t)$$

https://doi.org/10.7441/joc.2026.03.11

Step 2: The spectrum is multiplied by the exponential term $e^{-jw_k t}$ corresponding to the center frequency w_k, there by modulating the signal's central frequency band down to the corresponding baseband.

$$\left[\left(\delta(t)+\frac{j}{\pi x}\right)*Lu_k(t)\right]*e^{-jw_k t}$$

Step 3: The squared sum of the $L2$ norms of the time-domain gradients of the above expression is calculated, and the sum of all bandwidths is estimated. The corresponding constrained variational problem is formulated as follows:

$$\begin{cases} \boldsymbol{min}_{\{u_k\},\{w_k\}}\left\{\sum_{k=1}^{K}\left\|\frac{\boldsymbol{\partial}}{\boldsymbol{\partial t}}\left[\left(\delta(t)+\frac{j}{\pi t}\right)*Lu_k(t)\right]*e^{-jw_k t}\right\|_{\mathbf{2}}^{\mathbf{2}}\right\} \\ s.t.\ \sum_{k=1}^{K}Lu_k(t)=f(t) \end{cases}$$

where $Lu_k(t)$ denotes the corresponding set of K IMFs, min represents the minimization constraint, $w_k(t)$ are the center frequencies of the K IMFs, j is the imaginary unit, $\delta(t)$ denotes the unit impulse function, and $*$ indicates the convolution operation.

Step 4: A penalty factor α and an augmented Lagrange multiplier λ are introduced into the above constrained variational model, thereby transforming the constrained variational problem into an unconstrained variational problem:

$$L(\{Lu_k\},\{w_k\},\lambda)=\alpha\sum_{k=1}^{K}\left\|\frac{\boldsymbol{\partial}}{\boldsymbol{\partial t}}\left[\left(\delta(t)+\frac{j}{\pi t}\right)*Lu_k(t)\right]*e^{-jw_k t}\right\|_{\mathbf{2}}^{\mathbf{2}}+$$
$$\left\|f(t)-\sum_{k=1}^{K}Lu_k(t)\right\|_{\mathbf{2}}^{\mathbf{2}}+\langle\lambda(t),f(t)-\sum_{k=1}^{K}Lu_k(t)\rangle$$

Step 5: Solve the equations above. The Alternating Direction Method of Multipliers (ADMM) is employed to obtain the iteratively updated variables Lu^{n+1}, w_v^{n+1} and λ_v^{n+1}, thereby finding the optimal solution to the unconstrained variational problem:

$$\hat{u}_k^{n+1}(w)=\frac{\hat{f}(w)-\sum_{i\neq k}L\hat{u}_i(w)+\hat{\lambda}(w)/2}{1+2\alpha(w-w_k)^2}$$

$$w_k^{n+1}=\frac{\int_0^\infty w|L\hat{u}_k^{n+1}|^2\,dw}{\int_0^\infty\left|L\hat{u}_k^{n+1}\right|^2\,dw}$$

$$\hat{\lambda}^{n+1}(w)=\hat{\lambda}^n(w)+\tau\left[\hat{f}(w)-\sum_{k=1}^{K}Lu_k^{n+1}(w)\right]$$

Step 6: Execute the loop to calculate the equation until the iteration termination condition is met:

$$\sum_{k=1}^{K}\frac{\|u_k^{n+1}-u_k^n\|_2^2}{\|u_k^n\|_2^2}<\varepsilon$$

$$\begin{cases} \boldsymbol{min}_{\{u_k\},\{w_k\}}\left\{\sum_{k=1}^{K}\left\|\frac{\boldsymbol{\partial}}{\boldsymbol{\partial t}}\left[\left(\delta(t)+\frac{j}{\pi t}\right)*Lu_k(t)\right]*e^{-jw_k t}\right\|_{\mathbf{2}}^{\mathbf{2}}\right\} \\ s.t.\ \sum_{k=1}^{K}Lu_k(t)=f(t) \end{cases}$$

$$L(\{Lu_k\},\{w_k\},\lambda)=\alpha\sum_{k=1}^{K}\left\|\frac{\boldsymbol{\partial}}{\boldsymbol{\partial t}}\left[\left(\delta(t)+\frac{j}{\pi t}\right)*Lu_k(t)\right]*e^{-jw_k t}\right\|_{\mathbf{2}}^{\mathbf{2}}+$$

https://doi.org/10.7441/joc.2026.03.11

$$\left\| f(t) - \sum_{k=1}^{K} Lu_k(t) \right\|_2^2 + \langle \lambda(t), f(t) - \sum_{k=1}^{K} Lu_k(t) \rangle$$

Appendix C The calculation formula of Sample Entropy

For time series, the sample entropy is calculated as follows.

Reconstruct the time series X into an m-dimensional vector $X_m(i)$:

$$X_m(i) = \{x(i), x(i+1), \cdots, x(i+m-1)\}, 1 \le i \le N-m+1$$

The distance between vectors $X_m(i)$ and $X_m(j)$ can be calculated using the following formula: $d[X_m(i), X_m(j)] = max_{0 \le k \le m-1} |x(i+k) - x(j+k)|$

Given threshold r, Calculate the number B_i and ratio of $B_i^m(r)$ satisfied items

$d[X_m(i), X_m(j)] < \mathrm{r}: B_i^m(r) = \frac{1}{N-m} B_i$

Calculate the $B_i^m(r)$ mean: $B^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} B_i^m(r)$

Continuously expand the dimensionality and repeat the above formula to calculate the sample entropy of the theoretical time series:

$$SE(m,r) = \lim_{N \to \infty} \left\{ -ln\left(\frac{B^{m+1}(r)}{B^m(r)} \right) \right\}$$

When the population N is finite, the sample entropy can be estimated as:

$$SE(m,r,N) = -ln\left(\frac{B^{m+1}(r)}{B^m(r)} \right)$$

Appendix D The core processes of SMA

Step 1: Food approaching phase

During the food approaching phase, the slime mould approaches the food source by sensing odor in the air. Its approaching behavior is represented by the following formula:

$$\boldsymbol{X(t+1)} = \begin{cases} \boldsymbol{X_b(t) + vb\left(W \times X_A(t) - X_B(t)\right)}, & \boldsymbol{r < p} \\ \boldsymbol{vc \times X(t)}, & \boldsymbol{r \ge p} \end{cases}$$

where, $X(t+1)$ and $X(t)$ represent the positions of the slime mould at the $(t+1)$-th and t-th iterations, respectively. $X_b(t)$ denotes the position with the highest food odor concentration found up to the t-th iteration, while $X_A(t)$ and $X_B(t)$ represent two randomly selected individual positions from the current slime mould population at the t-th iteration. vb ranges [-a,a], $\boldsymbol{a = \arctan h\left(1 - {}^{t}/_{T}\right)}$, t is the current iteration number, T is the maximum number of iterations. vc decreases linearly from 1 to 0; r is a random number ranged within [0,1]; $\boldsymbol{\mathrm{p} = \tanh(|S(i) - DF|), i = 1, 2, \cdots, N}$. $S(i)$ represents the fitness value of the i-th slime mould individual, DF indicates the best fitness value obtained in all iterations so far, and N is the population size of the slime mould. W represents the slime mould's weight, calculated by the formula:

$$\boldsymbol{W(SortIndex(i))} = \begin{cases} \boldsymbol{1 + r \times \log\left(\frac{bF - S(i)}{bF - wF} + 1\right)}, & \boldsymbol{condition} \\ \boldsymbol{1 - r \times \log\left(\frac{bF - S(i)}{bF - wF} + 1\right)}, & \boldsymbol{others} \end{cases}$$

$$\boldsymbol{SortIndex = sort(S)}$$

https://doi.org/10.7441/joc.2026.03.11

where, *condition* refers to the individuals whose fitness ranks in the top half of the population. *bF* and *wF* represent the best and worst fitness values found in the current iteration, respectively. *SortIndex* denotes the sorted sequence of fitness values.

Step 2: Food wrapping phase

This phase simulates the contraction pattern of the slime mould's venous tissue. The higher the food concentration contacted by the vein, the stronger the propagation wave generated by the biological oscillator and the faster the cytoplasmic flow. The formula for updating the slime mould's position in this phase is as follows:

$$\boldsymbol{SortIndex}(\boldsymbol{x}) = \begin{cases} \boldsymbol{1 + r \times \log\left(\frac{bF - S(i)}{bF - wF} + 1\right)}, & \boldsymbol{condition} \\ \boldsymbol{1 - r \times \log\left(\frac{bF - S(i)}{bF - wF} + 1\right)}, & \boldsymbol{others} \end{cases}$$

$$\boldsymbol{SortIndex = sort(S)}$$

where *rand* and *r* denote random values generated within [0,1]. *ub* and *lb* represent the upper and lower bounds of the search space, respectively. *z* refers to the proportion of randomly distributed slime mould individuals relative to the total population, which is used to switch the algorithm between the global search phase and the local search phase.

Step 3: Food grabbing phase

In this phase, the slime mould adjusts the cytoplasmic flow rate within its veins using the propagation wave generated by the biological oscillator. The algorithm simulates the variations in venous width and the oscillation frequency of the biological oscillator through the parameters *vb*, *vc*, and *W*. This enables the slime mould to approach food slowly when the concentration is low and move toward it more rapidly upon locating high-quality food. This stage fundamentally achieves adaptive behavior through parameter updates.

Appendix E The details of BiLSTM

This study employs a bidirectional LSTM architecture for time-series modeling. The forward LSTM learns power fluctuation trends from historical data, while the backward LSTM identifies latent correlations from a future perspective. By integrating these two components, the model generates global temporal features that enhance the ability to analyze complex fluctuation patterns. The bidirectional structure effectively combines forward and backward temporal information, thereby reducing prediction biases caused by one-way models that overlook future implicit patterns. The theoretical formulas related to LSTM are as follows:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$
$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$
$$Q_t = \sigma(W_O[h_{t-1}, x_t] + b_O)$$
$$h_t = O_t \tanh(C_t)$$
$$C_t = f_t \odot C_{t+1} + i_t \odot \tanh(W_C.[x_t, h_{t-1}] + b_C)$$

W_i, W_O and W_f represent the weights of the input gate, output gate, and forget gate, respectively. b_i, b_O, b_f represent the biases of the input gate, output gate, and forget gate, respectively. C_t, C_{t-1} represent memory unit states at different time periods, while W_C and b_C denote the weights and biases of internal memory units. $\odot$ is element-by-element multiplication.

BiLSTM enhances the model's ability to model time series by integrating two independent LSTM units—one for processing forward sequences and the other for backward sequences. The

https://doi.org/10.7441/joc.2026.03.11

BiLSTM architecture is illustrated in Fig.1. x_t denotes the input data at time t, A_t and B_t represent the forward and backward LSTM hidden states respectively, Y_t is the output data at time t, and ω indicates the weights of each layer.

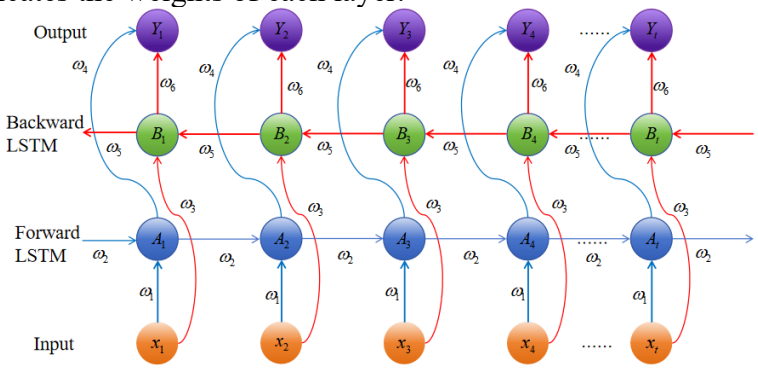


Fig. 1 – BiLSTM assumption diagram. Source: own research

Appendix F The computational approach of the Transformer

The computational approach of the Transformer is as follows **(**Vaswani et al., 2017**)** :

Calculate the *Q*, *K*, *V* matrices:

The input *X* is obtained via linear transformation:

$$Q = XW^Q, K = XW^K, V = XW^V$$

where, *Q* denotes the query vector, indicating which words the current word needs to focus on; *K* denotes the key vector, representing the identity markers of other words; *V* denotes the value vector, which contains the actual information.
Calculate the attention scores:

$$\boldsymbol{Attention(Q,K,V) = soft\,max(\frac{QK^T}{\sqrt{d_k}})V}$$

where $\sqrt{d_k}$ is the scaling factor to prevent *soft* max vanishing gradients caused by excessively large dot products; *soft* max is the weight obtained by normalization, which represents the importance of each word.
Multi-head attention:
Multiple groups of *Q*, *K*, *V* are used and computed in parallel to enhance the expressive capacity of the model:

$$Multihead(Q,K,V) = Concat(head_1,\cdots,head_h)W^O$$

where $\boldsymbol{head_i = Attention(QW_i^Q, KW_i^K, VW_i^V)}$

https://doi.org/10.7441/joc.2026.03.11

Contact information

Jinfeng Zhao
College of Science, Key Laboratory of Engineering Mathematics and Advanced Computing
Jiangxi University of Water Resources and Electric Power
Nanchang, China
E-mail: zhaojinfeng009@163.com

Changqi Tao* (Corresponding author)
Institute for Advanced Studies,
Zhejiang International Studies University
Hangzhou, China
E-mail: tcq_822@163.com

Jiahao Fu
College of Science
Jiangxi University of Water Resources and Electric Power
Nanchang, China
E-mail: 2596895971@qq.com

Yong Wang
College of Science
Jiangxi University of Water Resources and Electric Power
Nanchang, China
E-mail: 1023175310@qq.com

https://doi.org/10.7441/joc.2026.03.11