

Day-Ahead Photovoltaic Power Forecasting Based on Dual NWP Correction with Transitional Weather Classification

Wei Zhang, Shengyi Li, Jiyuan Gao, Zhongsheng Liu, and Hannan Zhang

Abstract—Numerical weather prediction (NWP) global horizontal irradiance (GHI) suffers from inherent biases that are drastically amplified under transitional weather, severely compromising day-ahead photovoltaic (PV) power forecasting accuracy. To address this issue, this study proposes a day-ahead PV forecasting method based on dual-stage NWP GHI correction with transitional weather classification. First, a dual-sliding-window approach identifies GHI turning points and categorizes weather-transition trends. Second, trend-specific attention-bidirectional gated recurrent unit (BiGRU) models perform the first-stage correction of NWP GHI. The residual errors are then decomposed via the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) into multiscale components, each corrected by a second-stage attention-BiGRU to compensate for frequency-band errors. Finally, the dual-stage-corrected GHI serves as a key input for accurate day-ahead PV power forecasting. Case studies using operational data from Hebei, China show that the proposed method reduces mean absolute error (MAE) by up to 39% and achieves an approximately 10% relative improvement in R^2 compared with raw NWP data at station 1 for day-ahead forecasting. These improvements are particularly pronounced under transitional weather, demonstrating the method's effectiveness in enhancing forecast reliability and supporting grid dispatch and renewable energy integration.

Index Terms—NWP correction, day-ahead photovoltaic power forecasting, transitional weather classification, attention-BiGRU.

ABBREVIATION

NWP	numerical weather prediction
PV	photovoltaic
GHI	global horizontal irradiance

CNN	convolutional neural network
RNN	recurrent neural network
LSTM	long short-term memory
GRU	gated recurrent unit
DA	decaying average
BiGRU	bidirectional gated recurrent unit
GNN	graph neural network
ARMA	autoregressive moving average
SVR	support vector regression
RF	random forest
XGBoost	extreme gradient boosting
MLR	multiple linear regression
SCINet	sample convolution and interaction network
DLinear	decomposition linear model
SegRNN	segment-based RNN
iTransformer	inverted Transformer
MAE	mean absolute error
RMSE	root mean square error
R^2	coefficient of determination
CEEMDAN	complete ensemble empirical mode decomposition with adaptive noise
WRF	weather research and forecasting model
IMF	intrinsic mode function

I. INTRODUCTION

Solar energy, as one of the most promising renewable energy sources, has gained widespread application and significant global attention in recent years due to its abundant availability, low acquisition costs, and minimal environmental impact [1]. However, photovoltaic (PV) power generation is strongly influenced by meteorological factors such as global horizontal irradiance (GHI), exhibiting pronounced intermittency and uncertainty, which in turn poses challenges for grid dispatch, reserve-capacity planning, and other operational processes [2], [3]. In electricity-market environments, insuffi-

Manuscript received 24 July, 2026; revised 27 August, 2026; accepted 4 September, 2026. (Corresponding author: Wei Zhang, e-mail: zhangwei@neepu.edu.cn)

W. Zhang, S. Li, and J. Gao are with the Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology, Ministry of Education, Northeast Electric Power University, Jilin 132012, China.

Z. Liu is with Liu R&D Consulting Ltd., 8490 Montcalm St., Vancouver, BC, V6P4P7, Canada.

H. Zhang is with the China Three Gorges Materials and Tendering Management Co., Ltd., Chengdu 610041, China.

Color versions of one or more of the figures in this paper are available online at <https://doi.org/10.26599/PEF.2026.9650021>.

Digital Object Identifier 10.26599/PEF.2026.9650021

© The author(s) 2026. Published by Tsinghua University Press.

This is an open access article under the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>).

cient PV-generation forecasting accuracy can directly lead to deviation settlements and economic losses when power plants participate in spot-market trading [4]. Therefore, improving PV-generation forecasting accuracy is essential for ensuring the safe and stable operation of power systems, enhancing the economic efficiency of market transactions, and advancing the intelligence of system regulation [5], [6].

Currently, photovoltaic power forecasting methods are primarily categorized into physical modeling approaches and data-driven methods [7]. Physical modeling approaches establish theoretical frameworks based on radiation transfer mechanisms and equipment parameters, but their performance is highly dependent on the accuracy of the input data. In contrast, data-driven approaches construct forecasting models by extracting latent patterns from historical data. Depending on the specific techniques employed, data-driven approaches can be further subdivided into traditional statistical methods and deep learning methods [8]. Traditional statistical learning methods primarily rely on linear assumptions and fixed-parameter models for time series modeling. Among these, autoregressive moving average (ARMA) [9], multiple linear regression (MLR) [10], and their extended forms (ARMAX, ARX, etc.) were widely applied in early photovoltaic power forecasting [11], [12]. However, this approach exhibits limited prediction accuracy when handling the nonlinearity and high volatility inherent in photovoltaic data.

In recent years, deep learning methods have demonstrated superior performance in photovoltaic power forecasting due to their robust feature extraction and nonlinear modeling capabilities [13], [14]. Long short-term memory (LSTM) and gated recurrent unit (GRU) can effectively capture the temporal variations of PV power [15], [16], while convolutional neural network (CNN) is effective in extracting local meteorological features from numerical weather prediction (NWP) data [17]. Attention mechanisms and Transformer further enhance the representation of long-term dependencies and complex patterns [18]. Building on these advances, graph neural network (GNN), TCN, and state-space models have been introduced to capture spatial correlations and long-term temporal dependencies in PV power generation [19], [20], [21]. Hybrid models such as CNN-LSTM and Transformer-TCN, as well as meta-learning and transfer learning approaches, have further improved feature extraction and model generalization [22], [23], [24].

However, while the aforementioned advanced models significantly enhance the mapping capability from meteorological data to PV output, their effectiveness is predicated on a common prerequisite that the input meteorological data is accurate and reliable. However, NWP data are forecasts and inherently contain forecast errors. These errors can increase dramatically during sudden weather changes, directly affecting the modeling of nonlinear relationships between PV output and meteorological factors and leading to significant deviations in subsequent power predictions. To address this,

researchers have proposed various methods for correcting NWP errors. Reference [25] established a day-ahead PV forecasting framework by integrating measured generation data from multiple cities in Taiwan with WRF-Solar irradiance. It employed the decaying average (DA) method for bias correction, thereby enhancing prediction accuracy. Reference [26] utilized the bias between geostationary satellite observations and NWP to construct an error correction factor, thereby enhancing irradiance prediction accuracy. Reference [27] identified GHI correction regions through error analysis. It combines multivariate feature selection with an expert hybrid model to extract representative features and performs corrections based on information aggregation, effectively enhancing the model's robustness and prediction accuracy.

Although the aforementioned bias correction methods effectively reduce the overall error of NWP, subsequent studies indicate that NWP errors exhibit significant variations across different seasons and weather conditions. Since existing unified correction methods struggle to accurately capture the complex dynamics of error evolution influenced by factors such as seasonality and weather types, researchers have further proposed conditionally classified NWP bias correction strategies. Reference [28] utilized temporal encoding to extract seasonal variations, constructed an error balance function, and enhanced the model's generalization capability for different weather characteristics through regional segmentation. Reference [29] employed deep clustering to categorize daily data with similar patterns and developed corresponding NWP GHI correction models for each training category, effectively improving prediction accuracy. The CoST-MoELSTM model proposed in [30] achieves classification-based correction for NWP GHI forecasts across different seasons and patterns through representation learning, sparse feature activation, and encoder-decoder calibration. It demonstrates significant accuracy improvements across multiple datasets and exhibits excellent seasonal robustness.

However, existing studies still face two major limitations. First, existing condition-dependent correction methods mainly focus on longer time scales, such as seasonal and daily variations, while paying limited attention to short-term abrupt irradiance changes under transitional weather conditions. The error characteristics of NWP forecasts during rapid irradiance increases or decreases may differ substantially from those during stable periods, making it difficult for a unified correction model to adapt effectively. Kreuvel et al. [31] employed machine learning to transform coarse-resolution NWP outputs into probabilistic distributions of minute-level irradiance variability, offering new insights into short-term meteorological fluctuations. However, their study also reported that nearly half of the errors originated from irradiance forecast inaccuracies in the input NWP data, highlighting the necessity of bias correction at the source. Second, most existing studies adopt a single-stage correction strategy and do not fully exploit the residual information after the initial correction. Although

Quan et al. [32] proposed a two-stage post-processing method, and its second-stage mainly performs nonlinear fusion and does not sufficiently exploit the temporal and multiscale characteristics of the residuals, limiting further error correction.

To address these issues, this paper proposes a dual-stage NWP GHI correction method for short-term meteorological variations, consisting of trend correction and residual compensation. In the first-stage, a dual-sliding-window turning-point detection method combined with an irradiance threshold is used to identify short-term GHI variation trends, and separate attention-bidirectional gated recurrent unit (BiGRU) models are established for different trends to improve adaptability to rapid irradiance changes. In the second-stage, the residuals from the first-stage correction are decomposed to extract multiscale residual features, which are used to obtain a residual compensation term for further correcting NWP GHI. Through this progressive two-stage correction, both trend-dependent biases under short-term variations and residual errors after the initial correction are reduced, thereby improving the accuracy of NWP GHI and subsequent PV power forecasting.

(1) By employing a dual sliding window turning point detection method combined with an irradiance magnitude threshold criterion, the GHI series is partitioned into subsequences corresponding to different variation trends. The deviation characteristics of NWP GHI within each trend-based subsequence are then analyzed.

(2) Training separate correction models for each type of data to perform individual corrections, thereby enhancing the model's predictive accuracy for data during periods of rapid short-term meteorological changes.

(3) From the perspective of error compensation, a secondary correction based on the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN)-attention-BiGRU hybrid modeling approach is introduced to further optimize the accuracy of GHI series.

The structure of this paper is as follows: Section II analyzes NWP GHI errors during transitional weather to establish the basis for correction work, Section III introduces the correction method and evaluation metrics, Section IV presents case studies and analyzes the results, and Section V concludes the study.

II. ANALYSIS OF NWP GHI ERRORS DURING TRANSITIONAL WEATHER

A. Photovoltaic data analysis

This paper analyzes operational data from two PV stations in a region of Hebei Province, China, with a temporal resolution of 15 min. Fig. 1 illustrates the GHI and power curves of two representative days at station 1. The curves in the figure indicate that, in photovoltaic power prediction, the predicted power profile can, to some extent, inherit the characteristics of the historical sequence and capture the temporal patterns of

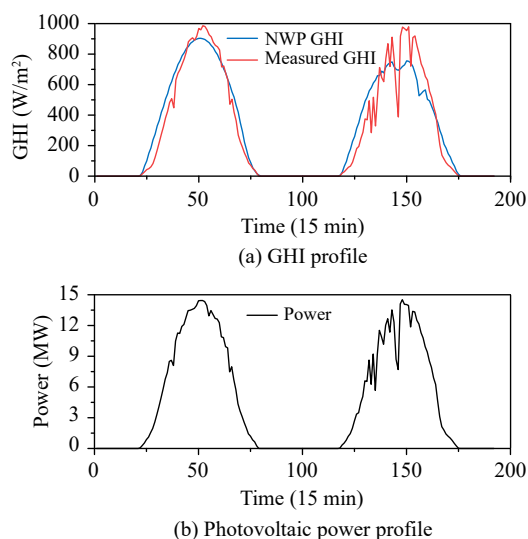


Fig. 1. Typical daily data profiles of a photovoltaic power station.

photovoltaic output. Moreover, the fluctuation amplitude of power is highly correlated with the variation in measured irradiance, further confirming that irradiance intensity is a key meteorological factor influencing photovoltaic power output. However, in actual prediction scenarios, real irradiance data are unavailable, and only irradiance data provided by NWP can be used as model inputs. Due to limitations in the spatial resolution of NWP data, its ability to depict cloud evolution and abrupt irradiance changes is insufficient, leading to significantly increased irradiance deviations during periods of sharp rise or fall in irradiance. Consequently, this adversely affects the power prediction results. In contrast, during periods of stable weather conditions such as clear skies or scattered clouds, NWP errors are relatively small, and the power prediction results are more stable. Therefore, to improve the reliability of photovoltaic power prediction, it is necessary to develop scenario-specific modeling and correction strategies tailored to different irradiance variation trends.

B. Transitional weather period identification

The variation process of photovoltaic irradiance is influenced by a combination of meteorological factors such as cloud movement and temperature changes, often exhibiting pronounced abrupt fluctuation characteristics in its temporal sequence. Traditional mutation detection methods based on a single temporal feature are highly sensitive to local noise, which can easily lead to false positives or missed detections. To address this, this paper proposes an identification method for periods of intense fluctuation based on sliding-window feature fluctuation analysis. By comprehensively modeling multi-feature temporal fluctuations, this method identifies potential abrupt-change windows within the irradiance sequence.

Assuming that the photovoltaic data comprise Q features and have a total time length of T , the entire sample can be constructed as a data matrix

$$X = [x_{ij}]_{Q \times T}, \quad i = 1, \dots, Q, \quad j = 1, \dots, T \quad (1)$$

where x_{ij} represents the value of the i^{th} feature at the j^{th} time step. A fixed sliding window of length $L = 4$ is applied to locally segment the data, forming window submatrices $X^{(k)} \in \mathbb{R}^{Q \times L}$. For each window submatrix, the feature mean fluctuation and standard deviation fluctuation are calculated to characterize the degree of change in the local feature distribution. The equation are as follows

$$\Delta\mu_{(i,k)} = \frac{\mu_{(i,k+1)} - \mu_{(i,k)}}{\max(|\mu_{(i,k)}|, |\mu_{(i,k+1)}|)} \quad (2)$$

$$\Delta\sigma_{(i,k)} = \frac{\sigma_{(i,k+1)} - \sigma_{(i,k)}}{\max(|\sigma_{(i,k+1)}|, |\sigma_{(i,k)}|)} \quad (3)$$

where $\mu_{(i,k)}$ is the mean of the i^{th} feature in the k^{th} window, and $\sigma_{(i,k)}$ is the standard deviation of the i^{th} feature in the k^{th} window. The relative fluctuation index of the window is defined to measure data variability

$$F_k = \frac{1}{Q} \sum_{i=1}^Q (w_\mu |\Delta\mu_{(i,k)}| + w_\sigma |\Delta\sigma_{(i,k)}|) \quad (4)$$

where w_μ and w_σ are the weighting coefficients, and both are set to 0.5 in this study to assign equal importance to changes in the local mean and standard deviation. For datasets with different statistical characteristics, these weights can be recalibrated accordingly. When F_k exceeds the threshold ε , the data distributions of the two adjacent windows are considered significantly different. In this study, $\varepsilon = 0.4$, which is used as a relatively loose screening threshold to identify potential transition windows rather than to directly determine the final trend category. To accurately identify the turning points of GHI within such fluctuations, thereby preparing for the extraction of transitional periods, this paper proposes a dual sliding window method for detecting GHI turning points within potential abrupt-change windows. This method divides the original window into adjacent sliding sub-windows of length $L/2$. Linear fitting is performed on the GHI sequence within each sub-window, and GHI turning points are identified by calculating the trend slopes and relative mean differences between sub-windows. The procedure is as follows

$$V = \frac{\sum_{i=1}^{L/2} |v_i^1 - v_i^2|}{\max\left(\sum_{i=1}^{L/2} |v_i^1|, \sum_{i=1}^{L/2} |v_i^2|\right)} \quad (5)$$

$$b_k = \frac{\sum_{i=1}^{L/2} (i - \bar{i})(v_i - \bar{v})}{\sum_{i=1}^{L/2} (i - \bar{i})^2} \quad (6)$$

where v_i^1 and v_i^2 represent the i^{th} GHI values in the preceding and succeeding windows, respectively. V indicates the relative change in the mean GHI between the two windows, b_k denotes the trend slope within the k^{th} window, and S represents the difference in trend slopes between adjacent windows. A turning point in GHI is marked at the junction of the windows when the condition $(V > \tau_V)$ or $(S > \tau_S)$ is satisfied. Through continuous scanning via the sliding window approach and analysis of feature fluctuations, transitional periods characterized by significant trend changes in the photovoltaic irradiance time series can be identified. Additionally, a turning point dataset $\mathcal{T}_{cp}$ is extracted for subsequent analysis.

Given that intraday photovoltaic irradiance typically follows a bell-shaped distribution, many adjacent turning intervals delineated by inflection points exhibit only minor fluctuations. For such intervals with slight variations, power predictions based on historical sequences generally remain highly accurate. To concentrate model training and evaluation on the intense fluctuations that substantially impact prediction outcomes, this paper introduces an irradiance-variation magnitude criterion based on the extracted set of inflection points. A segment is classified as a sharp-variation sample only when it simultaneously satisfies both the magnitude threshold and the rate-of-change threshold. Segments meeting these threshold conditions are further categorized into sharp increase and sharp decrease based on the direction of change, while segments not exceeding the thresholds are classified as stable periods. The specific discrimination equation are as follows

$$\left\{ \begin{array}{l} \left(|G_{j+1}^{\mathcal{T}_{cp}} - G_j^{\mathcal{T}_{cp}}| > \bar{G} \right), \quad \bar{G} = \frac{G_d}{D} \\ \left(\frac{|G_{j+1}^{\mathcal{T}_{cp}} - G_j^{\mathcal{T}_{cp}}|}{t_{j+1}^{\mathcal{T}_{cp}} - t_j^{\mathcal{T}_{cp}}} > \beta \right) \end{array} \right. \quad (8)$$

where $G_j^{\mathcal{T}_{cp}}$ denotes the GHI value at the j^{th} point in the turning point set $\mathcal{T}_{cp}$, $\bar{G}$ is the daily average GHI, G_d represents the total daily GHI, D stands for the sunshine duration, and β is the mutation rate threshold for the transitional period. In this study, $\beta = 10 \text{ W}/(\text{m}^2 \cdot \text{min})$, determined according to the statistical characteristics of short-term GHI variations in the training dataset.

During the training stage, measured GHI is available and is used to identify the actual irradiance variation trends through the proposed dual-sliding-window method. The historical samples are subsequently classified into sharp-increase, sharp-decrease, and stable periods to construct the corresponding trend-specific correction models. During actual day-ahead forecasting, however, measured GHI within the future forecasting horizon is unavailable. Since NWP provides irradiance forecasts for the subsequent day in advance, the NWP GHI forecast sequence is used for transitional-weather identification at the forecasting stage. The same dual-sliding-win-

down criteria are applied to the NWP GHI sequence to determine the corresponding trend category, after which the correction model associated with the identified category is selected.

C. NWP GHI error analysis

Based on the extracted categories of transitional weather periods, mean absolute error (MAE) and root mean square error (RMSE) are selected as evaluation metrics to analyze the deviation between NWP GHI and measured values under different variation trends.

Taking station 1 as an example, as shown in Table 1, the deviation between NWP GHI and measured values increases significantly during transitional weather. Both MAE and RMSE in sharp variation periods are considerably higher than those in stable periods. Therefore, it is necessary to apply category specific corrections to NWP data for transitional weather conditions.

TABLE 1
STATISTICAL ANALYSIS OF NWP GHI ERRORS UNDER DIFFERENT TREND TYPES (W/m²)

Trend type	MAE	RMSE
Sharp increase	145.18	186.07
Sharp decrease	148.60	191.58
Stable period	39.35	83.29

III. NWP GHI CORRECTION METHODS

A. NWP GHI correction framework

To address the problem of significant error of NWP GHI under transitional weather conditions, this paper proposes a dual-stage NWP GHI correction framework, as illustrated in Fig. 2. The framework follows a core workflow of “turning point identification–sharp variation segment extraction–stage specific correction”, enabling refined GHI correction under transitional weather.

First, a dual sliding-window model is established based on photovoltaic irradiance data characteristics. By calculating the mean fluctuation coefficients, standard deviation fluctuation coefficients, and fitted slopes between adjacent windows, turning points in the irradiance series are identified. Combined with a variation-magnitude threshold, the data are classified into three trend types: sharp increase, sharp decrease, and stable. Subsequently, independent attention-BiGRU models are constructed for each trend type to perform the first-stage correction of NWP GHI. Furthermore, CEEMDAN is introduced to decompose the first-stage correction residuals into multi-scale components, and an attention-BiGRU-based approach is applied to compensate for errors across different frequency bands, thereby achieving a multi-level and refined correction of NWP GHI.

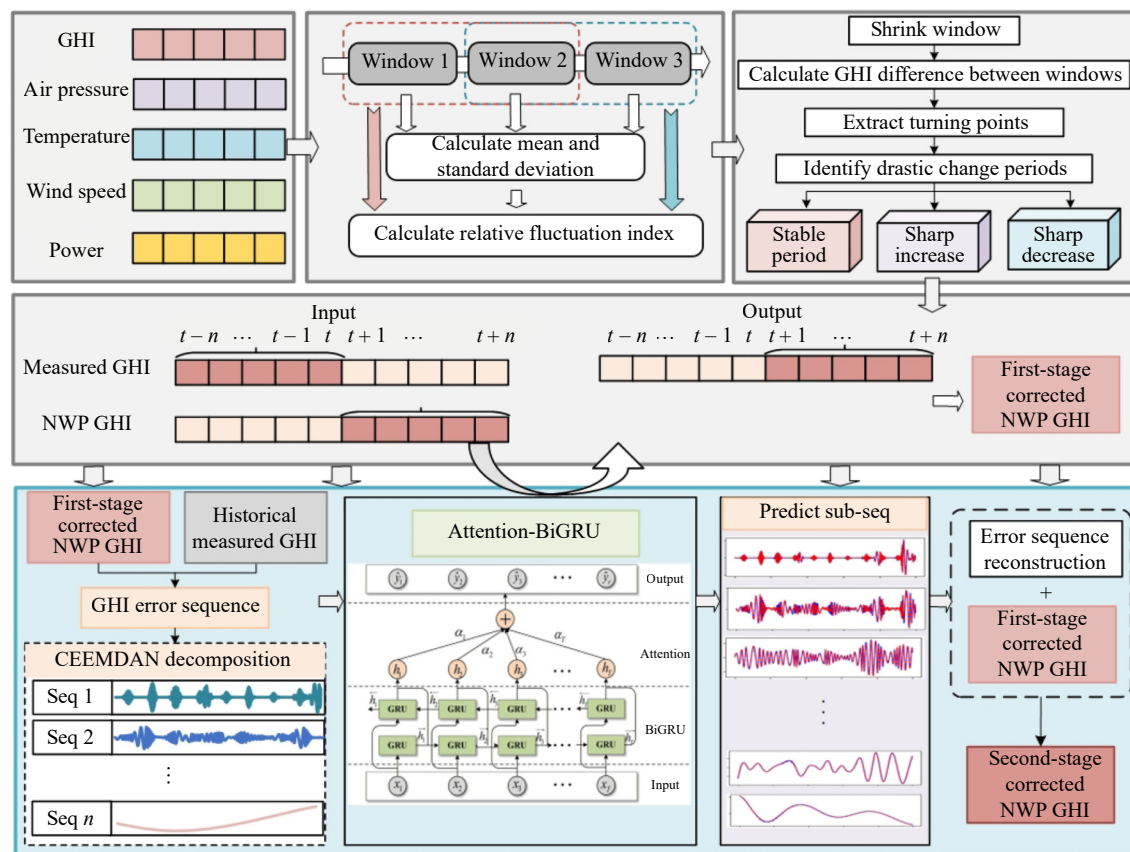


Fig. 2. Overall framework diagram.

B. Correction strategy

(1) First-stage correction: The first-stage correction strategy in this study integrates historically measured GHI sequences with NWP forecasts for future timesteps to perform GHI correction. As shown in Fig. 3, for each forecasting timestep $t+n$, the input features consist of two temporal components: a 24-h historical measured GHI sequence $\{G_{t-n}, \dots, G_t\}$ used to capture past variation trends, and the corresponding NWP GHI forecast G_{t+n}^{NWP} at the target timestep, which serves as a preliminary reference for future irradiance evolution. Since NWP provides irradiance forecasts for the subsequent forecasting period in advance, the NWP GHI corresponding to the target forecasting time is available when the forecast is issued. No measured GHI data from the future forecasting horizon are used as input, thereby avoiding look-ahead bias. This design does not rely on pre-constructed error terms; instead, it learns the temporal mapping between historical observations and NWP forecasts during training, and dynamically adjusts the forecast during inference, thereby achieving adaptive correction.

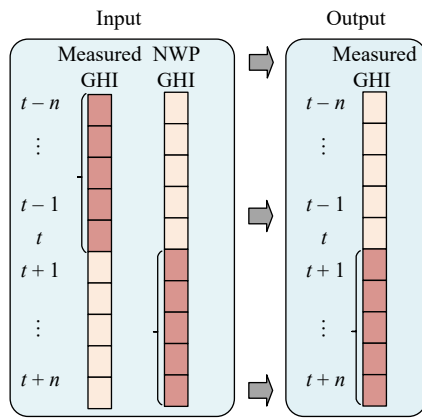


Fig. 3. First-stage correction strategy.

(2) Second-stage correction: Although the first-stage correction strategy can preliminarily adjust the NWP GHI and bring the predictions closer to historical observations, residuals inevitably remain in the forecast. These residuals contain not only high-frequency disturbance components that are difficult for conventional models to capture but also medium to long-term slow-varying trends, along with random noise. To address this, this paper proposes a second-stage correction strategy, which aims to deeply mine the underlying information within the residuals based on the first-stage corrected results, thereby achieving a more refined GHI correction.

The second-stage correction strategy employs the CEEMDAN algorithm to perform multi-scale decomposition on the residual sequence between the model output after the first-stage correction and the actual observations. CEEMDAN decomposes the complex residual series into several intrinsic mode functions (IMFs) with physical significance, which helps to reveal multi-scale error characteristics caused by fac-

tors such as abrupt meteorological changes and model limitations. By conducting in-depth analysis and modeling of each IMF component, the model can capture variation patterns of the residuals across different frequency ranges, thereby achieving effective correction of errors in various frequency bands.

For each forecasting time t , the second-stage correction is performed using only information available before t . First, the historical GHI sequence is processed by the trained first-stage correction model to obtain the corresponding historical first-stage corrected GHI. The residual sequence between the historical measured GHI and the first-stage corrected GHI is then constructed. CEEMDAN is subsequently applied to this historical residual sequence within a causal rolling window, and no residual information after the forecasting time is involved in the decomposition. The resulting IMF components are modeled independently using their historical sequences to predict the corresponding residual components over the future forecasting horizon. Finally, the predicted IMF components are reconstructed to obtain the second-stage residual correction. Therefore, the future residual sequence is not required during forecasting, and the CEEMDAN-based second-stage correction does not introduce look-ahead bias.

(3) CEEMDAN decomposition: CEEMDAN is an improved empirical mode decomposition method that progressively decomposes a non-stationary signal into a finite number of intrinsic mode functions with local characteristic scales and a residual trend term, by introducing auxiliary white noise ensembles and a recursive residual modeling mechanism. Its decomposition result takes the following form

$$x(t) = \sum_{\ell=1}^M \text{IMF}_{\ell}(t) + r_M(t) \quad (9)$$

where $x(t)$ is the original signal; $\text{IMF}_{\ell}(t)$ is the ℓ^{th} intrinsic mode function; $r_M(t)$ denotes the residual component after the M^{th} decomposition step; and M is the number of decomposition layers, which depends on the signal complexity and the decomposition termination criterion. The specific steps are as follows:

(1) Multiple independent white-noise sequences $w_i(t)$ are introduced and added to the original signal $x(t)$ to construct a set of augmented signals

$$x_i^{(1)}(t) = x(t) + \epsilon \cdot w_i(t), \quad i = 1, 2, \dots, N \quad (10)$$

where ϵ is the noise intensity, and $x_i^{(1)}(t)$ denotes the i^{th} noise-augmented signal. For each $x_i^{(1)}(t)$, EMD is applied to extract the first-order mode IMF. The first complete intrinsic mode function IMF is then obtained by averaging these modes

$$\text{IMF}_1(t) = \frac{1}{N} \sum_{i=1}^N \text{EMD}_1(x_i^{(1)}(t)) \quad (11)$$

(2) Constructing the residual signal by removing the first IMF from the original signal, yielding the first-stage residual signal

$$r_1(t) = x(t) - \text{IMF}_1(t) \quad (12)$$

(3) Recursive decomposition: For the residual $r_{\ell-1}(t)$, repeating step 1 to generate a new set of augmented signals. For each augmented residual signal $x_i^{(\ell)}(t)$, extracting the first IMF and then computing their average to obtain the ℓ^{th} IMF, while simultaneously updating the residual $r_\ell(t)$. This process is repeated until the residual satisfies the monotonicity condition or its energy becomes negligible, ultimately yielding the complete decomposition represented by Eq. (9).

C. Prediction model introduction

In photovoltaic power generation forecasting, accurately capturing the temporal features and dynamic variations of time series is crucial. While traditional recurrent neural networks (RNNs) possess the ability to process sequential data, they often suffer from gradient vanishing or gradient explosion when handling long sequences, which limits their capacity to learn long range dependencies. To address this issue, gated recurrent unit (GRU) was proposed [33]. Its structure is illustrated in Fig. 4. By introducing update and reset gate mechanisms, GRU effectively mitigates the gradient vanishing problem. Compared to LSTM, GRU has a simpler structure, fewer parameters, and higher computational efficiency, making it more suitable for photovoltaic forecasting applications that require high real time performance.

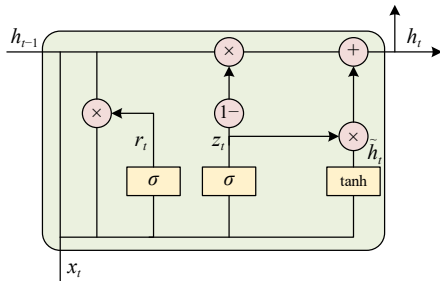


Fig. 4. Structure of the GRU.

The key computational formulas are as follows

$$z_t = \sigma(W_z[h_{t-1}, x_t] + b_z) \quad (13)$$

$$r_t = \sigma(W_r[h_{t-1}, x_t] + b_r) \quad (14)$$

$$\tilde{h}_t = \tanh(W_h[r_t \odot h_{t-1}, x_t] + b_h) \quad (15)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (16)$$

where z_t and r_t are the update gate and reset gate, respectively, which control the retention and forgetting of information; $\tilde{h}_t$ is the candidate hidden state; h_t is the final hidden state; σ denotes the sigmoid activation function; and W and b are the corresponding weight matrices and bias terms, respectively.

However, a unidirectional GRU can only utilize historical

information. In photovoltaic power generation, which is highly sensitive to weather variations, features at adjacent time steps often exhibit strong temporal correlations. To comprehensively capture bidirectional temporal dependencies, the bidirectional GRU (BiGRU) is introduced [34]. BiGRU consists of a forward and a backward GRU layer operating in parallel, with its output given by

$$h_t = [\vec{h}_t; \overleftarrow{h}_t] \quad (17)$$

Although BiGRU is effective in extracting bidirectional features, in long time series, the contributions of different time steps to the prediction outcome are not uniform. To address this, this paper introduces an attention mechanism on top of the BiGRU framework, allowing the model to adaptively assign importance weights to different time steps, thereby amplifying the influence of key-moment features on the prediction results. The attention mechanism enables the model to automatically focus on the critical time steps that are most influential for prediction, by computing the attention weight α_t at each time step. The structure of the Attention-BiGRU model is illustrated in Fig. 5.

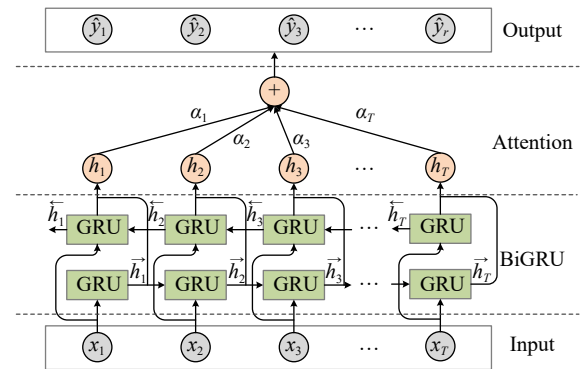


Fig. 5. Structure of the attention-BiGRU model.

It should be noted that the prediction framework proposed in this study is designed with generality; its core prediction module is not restricted to a specific network architecture. A variety of deep learning models can operate within this framework. The selection of attention-BiGRU as the exemplary model in this paper aims to validate the effectiveness of the proposed dual-correction strategy rather than to prescribe the prediction model itself.

D. Evaluation metrics

To validate the accuracy of the prediction model, this study selects MAE, RMSE, and the coefficient of determination (R^2) to evaluate the model's performance in GHI correction and power prediction. MAE reflects the average magnitude of deviation between predicted and actual values, and its calculation equation can be expressed as

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (18)$$

RMSE represents the square root of the mean squared error between predicted and actual values. Its calculation equation can be expressed as

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (19)$$

R^2 indicates how well the model's predictions fit the actual observed data. Its calculation equation can be expressed as

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (20)$$

IV. CASE STUDY

This study conducts a case analysis using data from two photovoltaic power stations in a certain region of Hebei Province, China, covering the period from July 1, 2018 to June 13, 2019. Station 1 has a rated capacity of 20 MW, while station 2 has a rated capacity of 15 MW. The dataset includes both measured meteorological data and NWP data, encompassing variables such as GHI, temperature, wind speed, and other meteorological parameters, with a temporal resolution of 15 min.

The NWP data utilized in this study were generated by the advanced research WRF (ARW) model, version 3.9.1. As a fully compressible, non-hydrostatic modeling system, ARW employs terrain-following hydrostatic pressure vertical coordinates and Arakawa C-grid spatial discretization. To ensure high-precision forecasting for the target PV stations, the model was configured with a horizontal grid resolution of 4 km and 45 terrain-following vertical layers (Eta levels), extending from the surface to a model top of 70 hPa. The model is initialized daily at 12:00 UTC using global forecast data from the European Centre for Medium-Range Weather Forecasts (ECMWF) at a resolution of $0.125^\circ \times 0.125^\circ$. In this study, the extracted NWP variables feature a 15-min temporal resolution, matching the sampling frequency of local ground-based measurements to ensure consistency during model training and evaluation.

A. First-stage correction based on transitional weather classification

The first-stage correction uses data from July to September 2018 as the training set, and data from October 2018 to June 2019 as the test set. As shown in Fig. 6, the dual sliding window turning point identification method is applied to locate significant transition points in the GHI series. Subsequently, combined with threshold criteria, the series is classified into three trend periods: “sharp increase”, “sharp decrease”, and “stable period”.

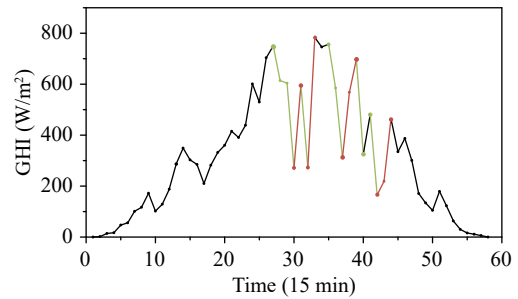


Fig. 6. Extraction of sharp-variation periods.

For the three distinct weather trend data subsets, three independent attention-BiGRU correction models are trained separately. The input to each sub model consists of a 24-h historical measured GHI sequence along with the NWP forecast value corresponding to the target time step. During the testing phase, the trend of each prediction sample is first determined and assigned to its corresponding category; then, the respective sub model is invoked for prediction. Finally, the predictions from all categories are reassembled into a complete corrected NWP GHI sequence, which serves as the input for subsequent CEEMDAN-based residual modeling and power forecasting.

To validate the effectiveness of the trend-based classification correction strategy, this study designs a control experiment. Under the condition of no trend classification, models including support vector regression (SVR), extreme gradient boosting (XGBoost), GRU, BiGRU, and attention-BiGRU are separately trained and tested on the original dataset. The parameter settings of the comparison models are shown in Table 2.

TABLE 2
MODEL PARAMETERS

Model	Parameters
XGBoost	Number of trees: 300, Max_depth: 8, Learning_rate: 0.01
SVR	Kernel: RBF, C: 100, Epsilon: 0.01
GRU	Hidden units: 64, 32, Learning_rate: 0.001, Batch_size: 64
BiGRU	Hidden units: 64, 32, Learning_rate: 0.001, Batch_size: 64
ABiGRU	Hidden units: 64, 32, Learning_rate: 0.001, Batch_size: 64

As shown in Table 3, after the first-stage correction based on trend classification, the NWP GHI forecasting accuracy is significantly improved for both stations. Taking station 1 as an example, the GHI MAE decreases to 21.942 W/m², RMSE improves from 107.146 to 56.191 W/m², and R^2 rises to 0.960. In particular, to further distinguish the contribution of trend classification from that of the underlying model architecture, the single attention-BiGRU model and the proposed first-stage

TABLE 3
FIRST-STAGE CORRECTION ERROR RESULTS OF NWP GHI FOR DIFFERENT METHODS

Station	Method	Evaluation metrics		
		MAE	RMSE	R^2
Station 1	Raw NWP GHI	53.597	107.146	0.853
	SVR	37.492	80.504	0.917
	XGBoost	36.039	68.585	0.939
	GRU	28.958	64.006	0.947
	BiGRU	26.914	59.577	0.954
	ABiGRU	25.496	58.657	0.956
	First-stage correction	21.942	56.191	0.960
Station 2	Raw NWP GHI	48.245	99.897	0.852
	SVR	30.489	69.776	0.928
	XGBoost	27.356	60.431	0.945
	GRU	24.265	56.201	0.953
	BiGRU	23.523	54.869	0.955
	ABiGRU	22.314	52.297	0.959
	First-stage correction	21.807	50.519	0.962

correction strategy employ the same attention-BiGRU architecture, while the latter further introduces trend-specific classification and constructs separate correction models for different trend types. Compared with the single attention-BiGRU model, the classification-based correction strategy achieves relative reductions of approximately 13.9% in MAE and 4.2% in RMSE, while increasing R^2 from 0.956 to 0.960. These results indicate that the observed improvement is not solely attributable to the fitting capability of the Attention-BiGRU architecture, but is also associated with the trend-based classification strategy, which enables the correction models to learn the distinct NWP GHI error characteristics under different irradiance variation trends.

B. Second-stage correction based on CEEMDAN decomposition

The second-stage correction uses data from October to

December 2018 as the training set and data from January to June 2019 as the test set. First, the residual sequence between the first-stage corrected output and the measured GHI values is extracted to form the original residual signal. Since the residual may contain trend drift, periodic disturbances, and noise interference, directly modeling it could lead to feature coupling and increased learning difficulty. Therefore, this study employs the CEEMDAN algorithm to decompose the residual sequence. Through this process, 14 intrinsic mode functions and one residual component are obtained, as shown in Fig. 7.

To reduce modeling cost and improve computational efficiency, this study introduces permutation entropy to quantify the complexity of each IMF subsequence. As can be seen from Fig. 7, some modal functions exhibit similar entropy values, indicating analogous dynamic characteristics. Accordingly, when the entropy difference between two IMF components is less than 0.05, they are considered to have similar information characteristics and are merged. IMF10 through the residual term is integrated into one composite subsequence, while IMF8 and IMF9 are combined into another subsequence. This merging strategy reduces the number of sub-models while preserving the original information as much as possible, thereby reducing the computational resources required for subsequent model training.

Subsequently, each IMF subsequence is individually input into an attention-BiGRU model for training and prediction. Finally, the outputs from all IMF sub-models are summed and reconstructed to obtain the complete predicted residual sequence, which is then added to the first-stage corrected results to form the final second-stage corrected prediction. To further evaluate the effectiveness of CEEMDAN decomposition, this study compares two configurations:

(1) Without CEEMDAN: The first-stage correction residual sequence is directly fed into an attention-BiGRU model;

(2) With CEEMDAN: The residual sequence is first decomposed by CEEMDAN, then each IMF subsequence is trained with an individual attention-BiGRU model, and the outputs are summed to reconstruct the final result.

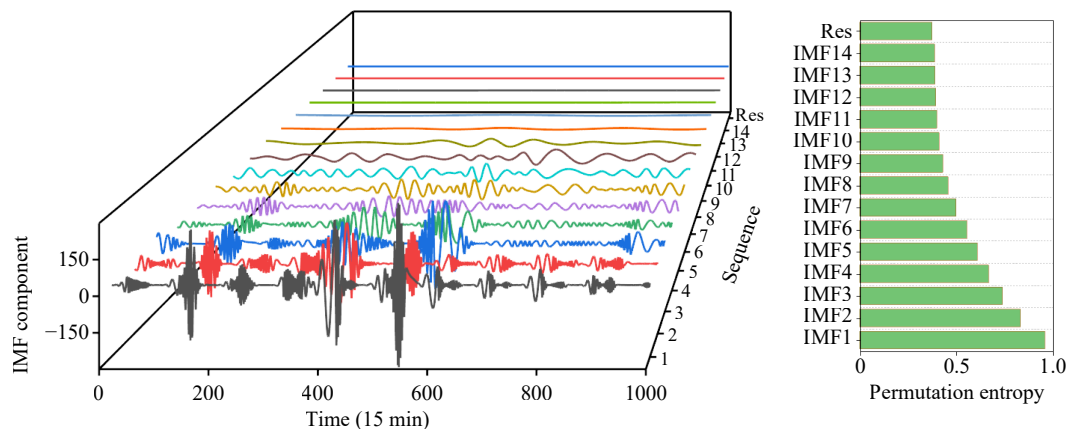


Fig. 7. CEEMDAN decomposition sequences and their permutation entropy.

The second-stage correction results are illustrated in Fig. 8.

As shown in Table 4, the model with second-stage correction achieves the best prediction performance among all tested methods, with an MAE of 17.373 W/m², RMSE of 34.065 W/m², and R^2 of 0.986. Compared with the raw NWP irradiance, the second-stage correction achieves a 16.3% relative improvement in R^2 . Even without CEEMDAN decomposition, the second-stage correction still outperforms the first-stage correction, reducing MAE and RMSE to 24.295 and 60.885 W/m², respectively, and increasing R^2 to 0.957. This

corresponds to relative reductions of approximately 1.7% in MAE and 1.1% in RMSE, and a relative improvement of approximately 0.2% in R^2 . In contrast, the second-stage correction model incorporating CEEMDAN decomposition achieves a more substantial improvement in accuracy, with a relative improvement of 3.2% in R^2 compared with the first-stage correction.

The results indicate that the second-stage correction mechanism, especially when combined with CEEMDAN decomposition, can effectively reduce systematic bias and residual errors

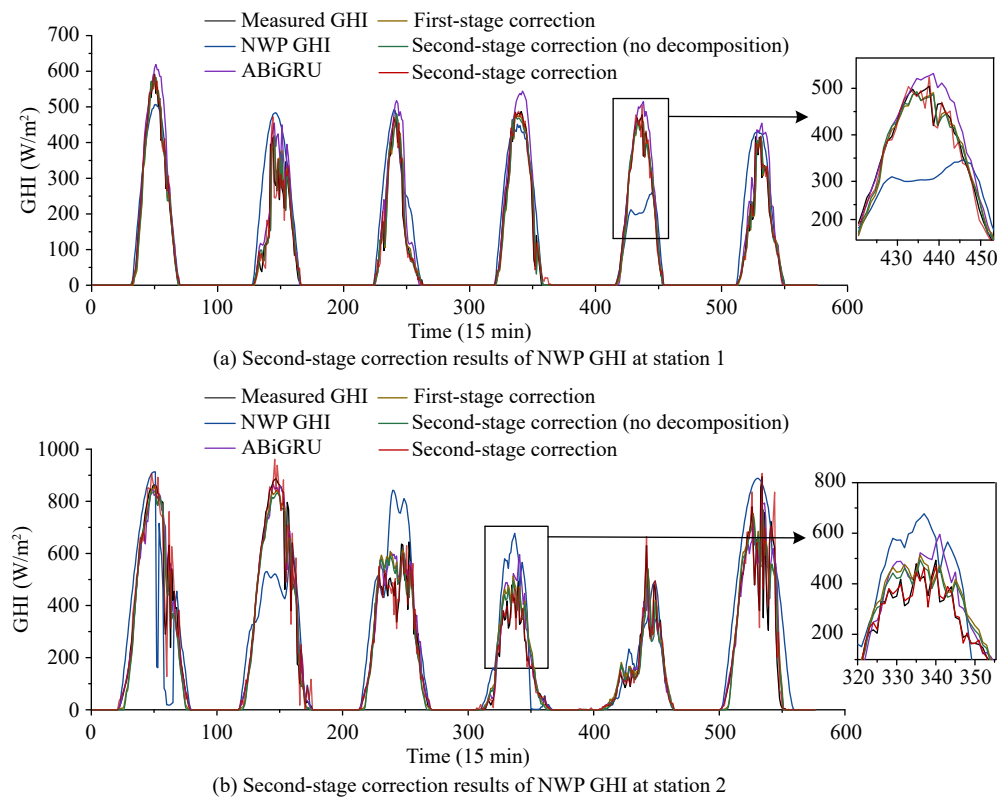


Fig. 8. Results of the second-stage correction for NWP GHI.

TABLE 4
SECOND-STAGE CORRECTION ERROR RESULTS OF NWP GHI FOR DIFFERENT METHODS

Station	Method	Evaluation metrics		
		MAE	RMSE	R^2
Station 1	Raw NWP GHI	57.304	113.774	0.848
	ABiGRU	28.976	63.759	0.952
	First-stage correction	24.704	61.584	0.955
	Second-stage correction (no decomposition)	24.295	60.885	0.957
	Second-stage correction	17.373	34.065	0.986
Station 2	Raw NWP GHI	49.607	103.271	0.847
	ABiGRU	24.269	54.283	0.956
	First-stage correction	23.826	53.169	0.959
	Second-stage correction (no decomposition)	23.169	52.828	0.960
	Second-stage correction	12.473	27.920	0.987

under complex meteorological conditions such as rapid cloud movement and abrupt weather changes. Furthermore, it dynamically compensates for omitted errors, thereby further enhancing the reliability of NWP irradiance forecasts.

The two correction stages adopt different chronological data partitions because they constitute a cascaded correction process. In the first-stage, data from July to September 2018 are used to train the trend-specific correction models, which are subsequently applied to the period from October 2018 to June 2019 to obtain the first-stage corrected GHI and corresponding residual sequences. The first-stage outputs from October to December 2018 are then used to construct the training samples for the second-stage residual correction model, while January–June 2019 is reserved as an independent test period. Therefore, although the second-stage training period overlaps temporally with the first-stage application period, no data from the second-stage test period are used during second-stage training.

To evaluate the computational complexity of the proposed model, this study recorded the time consumption of each model during both the training and testing phases. As shown in Table 5, the experimental results include the training time on the station 1 dataset and the forecasting time for a 24-h lookahead period. The results show that the dual-stage correction framework requires 840.10 ms for forecasting and 766.08 s for training. The increased computational complexity primarily stems from the secondary correction stage: The CEEMDAN algorithm decomposes the residual sequence into multiple intrinsic mode functions, requiring the framework to perform independent prediction compensation for each component. This multi sequence processing mode leads to a cumulative increase in time consumption. Despite the higher computational demands compared to the base model, considering that the training process is completed offline prior to online forecasting and that a prediction speed is well within the 15-min dispatching window of power systems, the proposed model achieves an effective balance between computational cost and forecasting accuracy.

TABLE 5
COMPARISON OF TRAINING AND FORECASTING TIME AMONG DIFFERENT METHODS

Method	Training time (s)	Forecasting time (ms)
SVR	72.98	69.88
XGBoost	141.11	20.03
GRU	78.18	46.99
BiGRU	93.29	61.79
ABiGRU	99.77	71.68
First-stage correction	112.31	94.34
Second-stage correction (no decomposition)	119.58	145.14
Second-stage correction	766.08	840.10

C. Photovoltaic power forecasting

(1) Ablation experiment: To validate the effectiveness of the proposed dual-stage NWP GHI correction method in practical power forecasting tasks, this study conducts a case analysis using data from two photovoltaic power stations in Hebei, China. Different input variables and correction strategies are progressively introduced to compare the impact of various configurations on day-ahead photovoltaic power forecasting performance. Data from January to April 2019 are used as the training set, and data from May to June 2019 serve as the test set. The specific configurations of the compared algorithms are as follows. Method I uses only historical power data without any NWP GHI information. Method II incorporates raw NWP GHI data into the input. Methods III–V progressively introduce first-stage correction using GRU, BiGRU, and ABiGRU, respectively, all without trend classification. Method VI adds trend classification to the first-stage correction with ABiGRU. Method VII further introduces second-stage residual correction without CEEMDAN decomposition, directly modeling the residuals using ABiGRU. Method VIII incorporates the complete proposed framework, which combines trend classified first-stage correction with CEEMDAN-based residual decomposition for second-stage correction.

The visual comparison of day-ahead photovoltaic power forecasting curves under different ablation schemes for the two power stations is shown in Fig. 9. It can be observed that the forecasting curve obtained by the complete framework (method VIII) achieves the best fitting performance against the real power output. Based on the data presented in Table 6, a clear improvement in forecasting performance is observed as input variables and correction strategies are sequentially enhanced. Taking station 1 as an example, method I, which utilizes only historical power data without NWP GHI correction or dedicated correction models, delivers a baseline performance with MAE, RMSE, and R^2 of 1.234 MW, 2.210 MW, and 0.810, respectively. With the introduction of NWP GHI data in method II, MAE and RMSE decrease to 1.109 and 1.953MW, respectively, indicating that the incorporation of NWP GHI data effectively enhances prediction performance.

TABLE 6
DAY-AHEAD POWER FORECASTING ERRORS FOR DIFFERENT METHODS

Method	Evaluation metrics		
	MAE	RMSE	R^2
Method I	1.234	2.210	0.810
Method II	1.109	1.953	0.852
Method III	0.878	1.516	0.911
Method IV	0.854	1.456	0.918
Method V	0.800	1.450	0.918
Method VI	0.779	1.413	0.922
Method VII	0.740	1.407	0.923
Method VIII	0.675	1.229	0.941

Further inclusion of first-stage correction along with various GRU-based architectures in methods III to V results in a continued decline in error metrics, with MAE reaching a minimum of 0.800 and R^2 exceeding 0.918. Method VI, which applies trend-based classification in addition to first-stage correction, further reduces forecasting error, with MAE and RMSE showing relative reductions of 2.6%, respectively, compared with method V, demonstrating that the trend-classification-based first-stage correction strategy effectively lowers prediction errors and enhances model fitting.

Methods VII and VIII further incorporate second-stage correction and CEEMDAN decomposition on this basis, with method VIII, combining trend-classified first-stage correction, CEEMDAN decomposition, and ABiGRU-based correction, achieving the best performance. Its MAE drops to 0.675, representing a relative reduction of nearly 39% compared with method II using raw NWP GHI, RMSE decreases to 1.229, while R^2 reaches 0.941, representing a relative improvement of approximately 10% compared with method II. These results indicate that the dual-stage correction and signal decomposition techniques effectively extract and utilize detailed information in the GHI, significantly enhancing both forecast accuracy and stability.

To further investigate the contribution of trend classification, the test samples are divided into sharp-increase, stable, and sharp-decrease periods. In particular, methods V and VI adopt the same forecasting architecture, while method VI additionally incorporates the proposed trend-classification strategy. Therefore, their performance difference can be used to evaluate the contribution of trend classification independently.

As shown in Table 7, introducing trend classification produces more substantial improvements under transitional-weather conditions than during stable periods. Compared with method V, method VI reduces MAE by 15.7% and 20.2% for sharp-increase and sharp-decrease periods, respectively, while the corresponding RMSE reductions are 8.1% and 13.1%. In contrast, the reductions during stable periods are only 7.6% in MAE and 4.1% in RMSE. The larger improvements observed for both sharp-increase and sharp-decrease periods indicate that trend classification is particularly effective in improving forecasting errors under transitional weather conditions.

In summary, combined with the analysis of the model configurations, the introduction of NWP GHI data, staged correction strategies, and CEEMDAN decomposition technology substantially improves the performance of day-ahead photovoltaic power forecasting, thereby validating the effectiveness of the proposed GHI correction methodology.

(2) Model applicability validation: To evaluate the adaptability of the proposed “dual-stage correction” strategy across different forecasting frameworks, this study selects three representative traditional machine learning models, random forest (RF), SVR, and XGBoost, as well as several recently proposed deep learning models, including DLinear, iTransformer,

TABLE 7
DAY-AHEAD POWER FORECASTING ERRORS UNDER TRANSITIONAL WEATHER CONDITIONS

Method	Trend type	Evaluation metrics	
		MAE	RMSE
Method I	Sharp increase	3.484	4.368
	Stable period	1.109	2.038
	Sharp decrease	3.307	4.105
Method II	Sharp increase	3.097	3.720
	Stable period	0.987	1.761
	Sharp decrease	3.020	3.524
Method III	Sharp increase	2.783	3.599
	Stable period	0.742	1.204
	Sharp decrease	2.742	3.528
Method IV	Sharp increase	2.654	3.441
	Stable period	0.724	1.160
	Sharp decrease	2.605	3.402
Method V	Sharp increase	2.590	3.358
	Stable period	0.667	1.147
	Sharp decrease	2.538	3.350
Method VI	Sharp increase	2.183	3.087
	Stable period	0.616	1.100
	Sharp decrease	2.025	2.912
Method VII	Sharp increase	2.067	2.972
	Stable period	0.606	1.099
	Sharp decrease	1.927	2.822
Method VIII	Sharp increase	1.818	2.794
	Stable period	0.558	0.936
	Sharp decrease	1.768	2.623

Informer, SCINet and SegRNN, as baseline models. Day-ahead power forecasting is then performed and compared under three input conditions: raw NWP GHI, first-stage corrected NWP GHI, and second-stage corrected NWP GHI.

Taking station 1 as an example, Fig. 10. and Table 8 present the comparison of forecasting performance metrics for various models under three input conditions. Specifically, input I consists of historical power plus raw NWP GHI, input II combines historical power with first-stage corrected NWP GHI, and input III includes historical power together with second-stage corrected NWP GHI. To distinguish the contribution of input quality from that of the forecasting architecture, the forecasting model is kept unchanged when comparing the three input conditions within each model. Therefore, the performance differences among inputs I, II, and III mainly reflect the impact of NWP GHI quality, whereas the differences among different models under the same input condition reflect the influence of the forecasting architecture. As shown in Table 8, the dual-stage correction strategy delivers significant and consistent performance improvements across all baseline models. In terms of MAE, the first-stage correction reduces the average error by 0.391 MW, while the second-stage correction further extends this reduction to 0.495 MW. A similar trend is

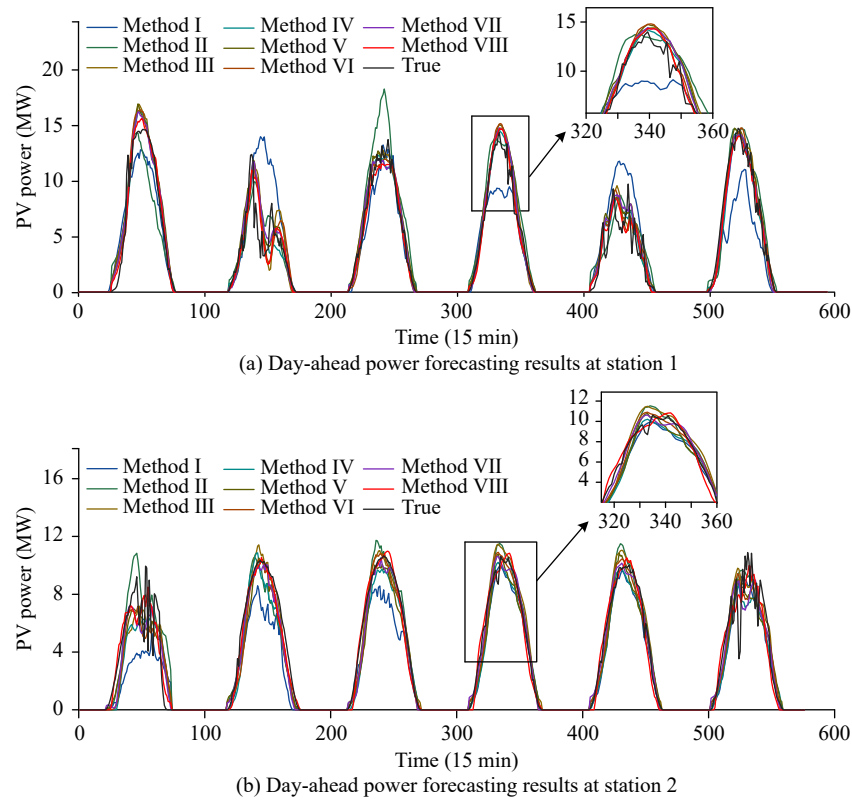


Fig. 9. Day-ahead power forecasting results.

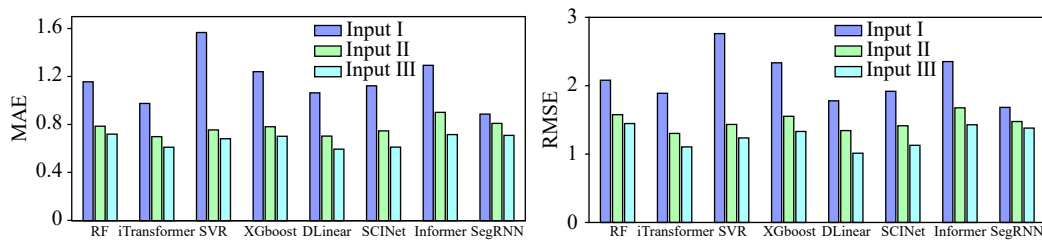


Fig. 10. Comparison of MAE and RMSE across different models.

observed for RMSE, with the first-stage correction decreasing it by an average of 0.628 MW and the second-stage correction achieving an average reduction of 0.841 MW. Furthermore, as shown in Fig. 11, the R^2 values of all models exhibit positive improvement: The first-stage correction achieves an average relative improvement of 11.22%, while the second-stage correction further increases the relative improvement to an average of 14.28%, reflecting enhanced fitting capability and better trend representation across the models. The consistent improvement obtained by replacing raw NWP GHI with corrected GHI across all forecasting models further indicates that the performance gains are primarily associated with the improved quality of the meteorological input rather than being dependent on a specific forecasting architecture.

In conclusion, the dual-stage correction framework based on “trend classification + CEEMDAN multi-scale residual compensation” not only reduces irradiance bias itself but also functions as a highly compatible preprocessing module that

can be readily integrated into various forecasting models. It significantly enhances the stability and reliability of photovoltaic power forecasting under complex meteorological conditions.

V. CONCLUSION

This paper proposes a dual-stage NWP GHI correction framework for day-ahead photovoltaic power forecasting under transitional weather conditions. The main conclusions are summarized as follows:

- (1) The proposed transitional weather classification strategy improves the adaptability of NWP GHI correction to different irradiance variation patterns. By identifying turning points and distinguishing sharp-increase, sharp-decrease, and stable periods, the correction models can learn more homogeneous error characteristics, particularly under abrupt irradiance variations.
- (2) The proposed second-stage correction further exploits

TABLE 8
DAY-AHEAD POWER FORECASTING ERRORS FOR DIFFERENT MODELS

Model	Metric	Model input		
		Input I	Input II	Input III
RF	MAE	1.158	0.788	0.722
	RMSE	2.081	1.576	1.447
	R^2	0.832	0.904	0.920
SVR	MAE	1.570	0.757	0.684
	RMSE	2.762	1.434	1.236
	R^2	0.703	0.919	0.940
XGboost	MAE	1.243	0.784	0.704
	RMSE	2.335	1.553	1.331
	R^2	0.789	0.906	0.931
DLinear	MAE	1.066	0.706	0.597
	RMSE	1.779	1.344	1.013
	R^2	0.880	0.928	0.962
SCINet	MAE	1.125	0.749	0.614
	RMSE	1.918	1.414	1.128
	R^2	0.860	0.922	0.954
Informer	MAE	1.296	0.904	0.718
	RMSE	2.353	1.677	1.428
	R^2	0.789	0.893	0.923
SegRNN	MAE	0.889	0.812	0.712
	RMSE	1.683	1.476	1.381
	R^2	0.892	0.917	0.928
iTransformer	MAE	0.978	0.701	0.613
	RMSE	1.889	1.302	1.105
	R^2	0.864	0.924	0.956

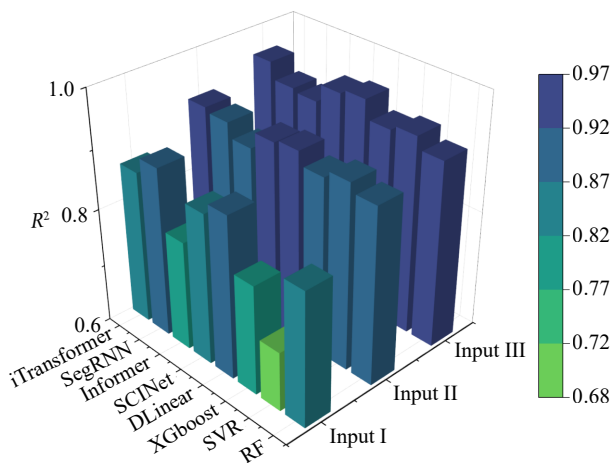


Fig. 11. Comparison of R^2 across models under different input conditions.

the residual information after the first-stage correction. CEEMDAN decomposes the residual sequence into components with different characteristic scales, while Attention-BiGRU models the decomposed components to compensate

for residual errors. This provides a complementary mechanism for reducing errors that remain after the initial correction.

(3) The corrected NWP GHI consistently improves PV power forecasting performance across different forecasting architectures. This demonstrates that the proposed dual-stage correction framework can serve as an effective preprocessing module for improving the reliability of meteorological inputs rather than being restricted to a specific forecasting model.

Overall, the proposed method provides a scenario-adaptive and multi-scale approach for reducing NWP irradiance errors and improving day-ahead PV forecasting under complex weather conditions. Future work will investigate the generalization of the method across different geographical regions and PV stations and further improve computational efficiency for large-scale operational applications.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by Smart-Grid National Science and Technology Major Project (Grant No. 2025ZD0805500).

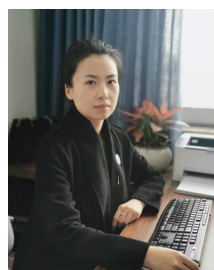
Use of AI statement

During the preparation of this manuscript, the authors used AI-assisted tools for language polishing only. The authors have reviewed and edited the content and take full responsibility for the final version of the manuscript.

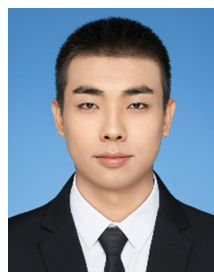
REFERENCES

- [1] G. Wang, Z. Zhang, and J. Q. Lin, "Multi-energy complementary power systems based on solar energy: A review," *Renew. Sustain. Energy Rev.*, vol. 199, pp. 114464, Jul. 2024.
- [2] L. W. Zheng, R. Su, X. Y. Sun, and S. Q. Guo, "Historical PV-output characteristic extraction based weather-type classification strategy and its forecasting method for the day-ahead prediction of PV output," *Energy*, vol. 271, pp. 127009, May 2023.
- [3] T. N. Nguyen and F. Müsgens, "What drives the accuracy of PV output forecasts?" *Appl. Energy*, vol. 323, pp. 119603, Oct. 2022.
- [4] Y. Li, R. N. Wang, and Z. Yang, "Optimal scheduling of isolated microgrids using automated reinforcement learning-based multi-period forecasting," *IEEE Trans. Sustain. Energy*, vol. 13, no. 1, pp. 159–169, Jan. 2022.
- [5] Q. Q. Li, Y. Xu, B. S. H. Chew, H. Y. Ding, and G. P. Zhao, "An integrated missing-data tolerant model for probabilistic PV power generation forecasting," *IEEE Trans. Power Syst.*, vol. 37, no. 6, pp. 4447–4459, Nov. 2022.
- [6] L. F. Yin, X. H. Cao, and D. D. Liu, "Weighted fully-connected regression networks for one-day-ahead hourly photovoltaic power forecasting," *Appl. Energy*, vol. 332, pp. 120527, Feb. 2023.
- [7] S. Pereira, P. Canhoto, T. Oozeki, and R. Salgado, "Comprehensive approach to photovoltaic power forecasting using numerical weather prediction data and physics-based models and data-driven techniques," *Renew. Energy*, vol. 251, pp. 123495, Oct. 2025.

- [8] M. J. Mayer, "Benefits of physical and machine learning hybridization for photovoltaic power forecasting," *Renew. Sustain. Energy Rev.*, vol. 168, pp. 112772, Oct. 2022.
- [9] J. Ballestrín, J. Polo, N. Martín-Chivelet, J. Barbero, E. Carra, J. Alonso-Montesinos, and A. Marzo, "Soiling forecasting of solar plants: A combined heuristic approach and autoregressive model," *Energy*, vol. 239, pp. 122442, Jan. 2022.
- [10] A. Keddouda, R. Ihaddadene, A. Boukhari, A. Atia, M. Arıcı, N. Lebbihiat, and N. Ihaddadene, "Solar photovoltaic power prediction using artificial neural network and multiple regression considering ambient and operating conditions," *Energy Convers. Manage.*, vol. 288, pp. 117186, Jul. 2023.
- [11] A. Abdellatif, H. Mubarak, S. Ahmad, T. Ahmed, G. M. Shafiullah, A. Hammoudeh, H. Abdellatif, M. M. Rahman, and H. M. Gheni, "Forecasting photovoltaic power generation with a stacking ensemble model," *Sustainability*, vol. 14, no. 17, pp. 11083, Sep. 2022.
- [12] K. Bellinguer, R. Girard, G. Bontron, and G. Kariniotakis, "A generic methodology to efficiently integrate weather information in short-term photovoltaic generation forecasting models," *Sol. Energy*, vol. 244, pp. 401–413, Sep. 2022.
- [13] P. T. Li, K. L. Zhou, X. H. Lu, and S. L. Yang, "A hybrid deep learning model for short-term PV power forecasting," *Appl. Energy*, vol. 259, pp. 114216, Feb. 2020.
- [14] Z. Y. Zhang, Y. Z. Li, Y. Li, Z. H. Qie, C. Y. Chung, and Z. G. Zeng, "Enhancing multi-step photovoltaic forecasting with large-scale meteorological data," *IEEE Trans. Smart Grid*, vol. 17, no. 1, pp. 715–732, Jan. 2026.
- [15] M. Yang, X. X. Shen, D. W. Huang, and X. Su, "Fluctuation classification and feature factor extraction to forecast very short-term photovoltaic output powers," *CSEE J. Power Energy Syst.*, vol. 11, no. 2, pp. 661–670, Mar. 2025.
- [16] X. M. Xu, L. Y. Guan, Z. Y. Wang, R. K. Yao, and X. Guan, "A double-layer forecasting model for PV power forecasting based on GRU-Informer-SVR and blending ensemble learning framework," *Appl. Soft Comput.*, vol. 172, pp. 112768, Mar. 2025.
- [17] M. Jurado, M. Samper, and R. Rosés, "An improved encoder-decoder-based CNN model for probabilistic short-term load and PV forecasting," *Electr. Power Syst. Res.*, vol. 217, pp. 109153, Apr. 2023.
- [18] S. J. Xu, H. Ma, C. Ekanayake, and Y. Cui, "Swin transformer-based transferable PV forecasting for new PV sites with insufficient PV generation data," *Renew. Energy*, vol. 246, pp. 122824, Jun. 2025.
- [19] M. Yang, Y. Jiang, Y. F. Guo, X. Su, Y. Li, and T. Huang, "Ultra-short-term prediction of photovoltaic cluster power based on spatiotemporal convergence effect and spatiotemporal dynamic graph attention network," *Renew. Energy*, vol. 255, pp. 123843, Dec. 2025.
- [20] Y. Y. Li, L. D. Song, S. Zhang, L. Kraus, T. Adcox, R. Willardson, A. Komandur, and N. Lu, "A TCN-based hybrid forecasting framework for hours-ahead utility-scale PV forecasting," *IEEE Trans. Smart Grid*, vol. 14, no. 5, pp. 4073–4085, Sep. 2023.
- [21] Q. Y. Zhang, Z. Zeng, Q. Zhou, X. Z. Ding, Z. C. Dong, M. Tian, D. H. Liu, K. Mei, and M. L. Li, "Multimodal bidirectional MAMBA framework for ultra-short-term photovoltaic power forecasting using sky images," *Renew. Energy*, vol. 256, pp. 124361, Jan. 2026.
- [22] S. C. Wang and Y. Huang, "Spatio-temporal photovoltaic prediction via a convolutional based hybrid network," *Comput. Electr. Eng.*, vol. 123, pp. 110021, Apr. 2025.
- [23] E. Sarmas, E. Spiliotis, E. Stamatopoulos, V. Marinakis, and H. Doukas, "Short-term photovoltaic power forecasting using meta-learning and numerical weather prediction independent long short-term memory models," *Renew. Energy*, vol. 216, pp. 118997, Nov. 2023.
- [24] S. Bak, S. Choi, D. Yang, D. Kim, H. Rho, and K. Lee, "Transfer learning for photovoltaic power forecasting across regions using large-scale datasets," *IEEE Access*, vol. 13, pp. 136175–136190, Jan. 2025.
- [25] C. L. Huang, Y. K. Wu, C. C. Tsai, J. S. Hong, and Y. Y. Li, "Revolutionizing solar power forecasts by correcting the outputs of the WRF-SOLAR model," *Energies*, vol. 17, no. 1, pp. 88, Dec. 2023.
- [26] T. Watanabe, H. Takenaka, and D. Nohara, "Post-processing correction method for surface solar irradiance forecast data from the numerical weather model using geostationary satellite observation data," *Sol. Energy*, vol. 223, pp. 202–216, Jul. 2021.
- [27] W. J. Dou, K. Wang, S. Shan, C. X. Li, K. J. Zhang, H. K. Wei, and V. Sreeram, "A correction framework for day-ahead NWP solar irradiance forecast based on sparsely activated multivariate-shapelets information aggregation," *Renew. Energy*, vol. 244, pp. 122638, May 2025.
- [28] M. L. Liu, Z. F. Lai, Y. Fang, and Q. Ling, "Day-ahead photovoltaic power forecasting based on corrected numeric weather prediction and domain generalization," *Energy Build.*, vol. 329, pp. 115212, Feb. 2025.
- [29] W. J. Dou, K. Wang, S. Shan, C. X. Li, Y. Y. Wang, K. J. Zhang, H. K. Wei, and V. Sreeram, "Day-ahead Numerical Weather Prediction solar irradiance correction using a clustering method based on weather conditions," *Appl. Energy*, vol. 365, pp. 123239, Jul. 2024.
- [30] W. J. Dou, K. Wang, S. Shan, K. J. Zhang, H. K. Wei, and V. Sreeram, "A hybrid correction framework using disentangled seasonal-trend representations and MoE for NWP solar irradiance forecast," *Appl. Energy*, vol. 397, pp. 126295, Nov. 2025.
- [31] F. P. M. Kreuwel, W. Knap, M. Schmeits, J. Vilà-Guerau De Arellano, and C. C. Van Heerwaarden, "Forecasting day-ahead 1-minute irradiance variability from numerical weather predictions," *Sol. Energy*, vol. 258, pp. 57–71, Jul. 2023.
- [32] H. Quan, Y. W. Ge, B. Liu, W. J. Zhang, and D. Srinivasan, "A deep-learning algorithm with two-stage training for solar forecast post-processing," *Sol. Energy*, vol. 273, pp. 112504, May 2024.
- [33] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," arXiv: 1412.3555, 2014.
- [34] Y. M. Dai, W. J. Yu, and M. M. Leng, "A hybrid ensemble optimized BiGRU method for short-term photovoltaic generation forecasting," *Energy*, vol. 299, pp. 131458, Jul. 2024.



Wei Zhang received her M.S. degree in signal and information processing from the University of Electronic Science and Technology of China, Chengdu, China, in 2013, and Ph.D. degree in electrical engineering from Northeast Electric Power University, Jilin, China, in 2020. Since 2023, she has been an Associate Professor with the School of Electrical Engineering, Northeast Electric Power University, Jilin, China. Her research interests include renewable energy generation forecasting and integrated energy system optimal scheduling.



Shengyi Li received his B.Eng. degree in communication engineering from the Northeast Electric Power University, Jilin, China, in 2024. He is currently working toward the M.Eng. degree in electrical engineering with the Northeast Electric Power University, Jilin, China. His research focuses on renewable energy generation.



Jiyuan Gao received his B.Eng. degree in electrical engineering from Shandong University of Technology, Zibo, China, in 2022. He is currently working toward the M.Eng. degree in electrical engineering with the Northeast Electric Power University, Jilin, China. His research focuses on renewable energy generation.



Zhongsheng Liu is Principal Scientist of Liu R&D Consulting Ltd. He retired in April 2026 as a Principal Research Officer from the National Research Council of Canada (NRC) after 28 years of service. He holds the Alexander von Humboldt Research Fellowship and the Queen Elizabeth II Diamond Jubilee Medal. He is a recipient of the State Council Special Government Allowance (China) and the Third Prize of the National Natural Science Award conferred by the State Science and Technology Commission of China. His technical specializations

span geochemical reaction modelling, electrochemistry, and computational mechanics. He served as Adjunct Professor at the University of Waterloo, University of Windsor, and University of Victoria, co-supervising doctoral students on joint research projects. Prior to joining the NRC, he was a Full

Professor at Jilin University. He conducted research fellowships at University of Wales Swansea UK, Technical University of Munich Germany, University of California at Berkeley USA, and Peking University.



Hannan Zhang received his B.S. degree in electrical engineering from Northeast Electric Power University, Jilin, China, and M.S. degree in engineering. He is currently with China Three Gorges Materials and Tendering Management Co., Ltd., Chengdu, China. His research interests include power-equipment condition monitoring, new-energy power generation, and power supply-chain management.