processes

MDPI

Article

Precision Prediction Strategy for Renewable Energy Power in Power Systems—A Physical-Knowledge Integrated Model

Shenbing Hua [1], Shuanglong Jin [1], Zongpeng Song [1], Xiaolin Liu [1], Shu Wang [1] and Hengrui Ma [2,3,*]

[1] State Key Laboratory of Renewable Energy Grid-Integration, China Electric Power Research Institute, Haidian District, Beijing 100192, China; ceprihua@163.com (S.H.); ceprijinsl@163.com (S.J.); szp_cepri@163.com (Z.S.); liuxiaolin02@163.com (X.L.); wangshu2011@pku.edu.cn (S.W.)
[2] Hubei Engineering and Technology Research Center for AC/DC Intelligent Distribution Network, School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China
[3] School of Electrical and Automation, Wuhan University, Wuhan 430072, China
* Correspondence: henry3764@foxmail.com

Abstract: To promote energy transformation and upgrading and achieve the dual carbon strategic goals, the proportion of renewable energy generation in China has been increasing annually. Among them, wind and photovoltaic power generation play a significant role in renewable energy generation due to their widely distributed energy sources, and their development has been particularly rapid. However, renewable energy generation exhibits strong volatility, and the integration of a large amount of renewable energy into the power grid poses a threat to the stable operation of the power system. Therefore, renewable energy power prediction is of great significance for the safe and stable operation of the power system. This paper proposes a physical-knowledge integrated renewable energy power prediction model. Firstly, the Fuzzy C-Means (FCM) method is used to handle missing data. Then, the variational mode decomposition (VMD) algorithm is applied to decompose the renewable energy power into high-frequency and low-frequency components. The high-frequency components are input into a Convolutional Neural Network (CNN) model, while the low-frequency components are input into a Long Short-Term Memory (LSTM) neural network for training and prediction. Finally, the effectiveness and feasibility of the proposed method in this paper are verified through real data from a certain actual power grid. After analysis, the proposed method demonstrates a prediction accuracy improvement of up to 5.71% compared to conventional approaches. Simultaneously, in high-noise interference environments, the prediction error can be reduced by up to 11.16%.

Keywords: renewable energy; Fuzzy C-Means method; VMD algorithm; CNN model; Long Short-Term Memory neural network

check for updates

Academic Editor: Takuya Oda

Received: 13 February 2025
Revised: 25 March 2025
Accepted: 26 March 2025
Published: 1 April 2025

Citation: Hua, S.; Jin, S.; Song, Z.; Liu, X.; Wang, S.; Ma, H. Precision Prediction Strategy for Renewable Energy Power in Power Systems—A Physical-Knowledge Integrated Model. *Processes* **2025**, *13*, 1049. https://doi.org/10.3390/pr13041049

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

As global energy demand continues to rise, China is actively exploring the development and application of new energy sources. Vigorously developing renewable energy not only alleviates the energy crisis caused by the shortage of oil and gas resources but also constitutes an important strategy for the construction of China's modern energy system [1]. Wind and solar power generation technologies, as mature and technologically advanced renewable energy generation methods in China's clean energy development, exhibit vast development prospects. However, the inherent strong randomness of wind and solar energy poses significant challenges to existing modern power systems with a high proportion of new energy sources [2,3]. Therefore, accurate prediction of new energy output power is

crucial for the stable operation of large-scale new energy integration into the grid. Precise prediction of output power is a vital task in enhancing the stability of new energy output [4]. This technology requires modeling using relevant parameters of new energy generation based on numerical weather forecasting, historical output power data from new energy stations, and weather observation data. The prediction model forecasts the trend of output changes at specific future times, aiding in the arrangement of maintenance schedules, grid dispatching, safety and stability analysis, and the improvement of new energy consumption rates. With the future development of electricity-related technologies, the advancement of distributed power grids, and the maturation of electricity trading markets, predictions of new energy generation power will provide support for emerging grid businesses, such as integrated energy services, clean energy trading, and demand response [5,6].

There is a vast amount of research on new energy power prediction models, both domestically and internationally. Commonly used methods for new energy power prediction include physical methods and statistical methods [7]. Statistical methods are further divided into traditional statistical methods and new statistical methods based on artificial intelligence algorithms. Physical methods are based on the principles of photovoltaic power generation and directly model the physical characteristics of the photovoltaic power generation process. They require detailed geographic information, module parameters, and environmental meteorological data of the photovoltaic power station and have a high degree of dependence on meteorological data and hardware information [8]. Although physical methods do not rely on a large amount of historical data, the modeling process is complex, the model robustness is not strong, and the prediction accuracy is relatively low [9]. Statistical methods do not consider physical processes such as changes in solar radiation intensity but rely on a large amount of historical data to identify statistical laws between input and output variables and establish mapping relationships for prediction. Compared with physical methods, statistical methods are simpler to model, and the information is easier to obtain, so current research methods for photovoltaic power prediction mainly focus on statistical methods. Traditional statistical methods include time series methods [10], fuzzy logic methods [11], regression analysis methods [12], and Markov chain methods [13]. Using time series as the theoretical basis and modeling with historical photovoltaic power data provides a complete theoretical system and strong model interpretability, but the prediction accuracy is not high. In contrast, artificial intelligence models can fully explore the internal characteristics and hidden changing laws of the data, extract high-dimensional complex nonlinear features, and make predictions. Compared with traditional statistical models, they have greatly improved prediction accuracy, stability, and versatility.

With the development of the big data industry on the Internet, machine learning methods have gradually replaced physical and statistical models to achieve better prediction results in the field of wind power and photovoltaic power output prediction. Neural networks and support vector machines are the two most representative machine learning methods, which describe the randomness of new energy generation by establishing a mapping relationship between input and output data. In neural network prediction models, there are further distinctions between traditional neural networks and deep learning networks. In Reference [14], to improve the accuracy of traditional neural networks, an optimization method using the Sparrow Search Algorithm (SSA) to optimize the thresholds of neural networks is proposed for effective short-term power prediction of photovoltaic power stations in various environments such as sunny, cloudy, and abnormal weather conditions. The results show that the neural network optimized by the sparrow algorithm has better accuracy than the negative gradient method and particle swarm algorithm and can search for the optimal threshold value in a short time. In Reference [15], to address

the issues of abnormal historical data and unstable numerical weather forecast data, a neural network wind power prediction model based on association rules is adopted, using the Apriori algorithm to associate wind power with meteorological data. Experiments demonstrate that the maximum and minimum relative errors are 5.76% and 0.01%, respectively, proving the effectiveness of this method. In Reference [16], a deep deterministic and recurrent deterministic policy model under the attention mechanism is proposed, unifying historical output data and numerical weather forecast data, adjusting the weights of different component learning to highlight important information, and obtaining the optimal power prediction value. In Reference [17], a support vector machine optimized by the Particle Swarm Optimization algorithm (PSO-LSSVM) is proposed for predicting day-ahead photovoltaic power. The penalty function and kernel function width in the Least Squares Support Vector Machine can both be obtained through PSO's global search. Experiments show that PSO-LSSVM has higher accuracy than the Particle Swarm Optimization-Back Propagation (PSO-BP) neural network algorithm in tests with data samples under different meteorological conditions. Reference [18] combines the Pearson correlation coefficient with the genetic algorithm to optimize and train data in the ELM hybrid model. In the prediction results for typical days of the four seasons, the estimated deviation rate decreased by 19% compared to a single model. In Reference [19], an ELM prediction model is proposed that decomposes ELM multiple times and reconstructs it for optimization. This method fully leverages the characteristics of fast adaptive learning and effectively fits historical wind power, but due to the limitations of ELM's single-layer neural network, the learning model cannot delve deeper and cannot extract deep-level features of wind power signals. In Reference [20], the extreme learning machine is optimized using dynamic safety and elite opposition-based SSA. This method is affected by meteorological factors, resulting in low prediction accuracy and poor generalization ability for photovoltaic power under rainy weather. In Reference [21], PSO and ELM are combined to predict photovoltaic power output, with PSO dynamically optimizing parameters at different stages.

Despite significant progress in existing research, renewable energy output power prediction faces the following issues:

The prediction error is relatively large, and the accuracy cannot meet requirements. In China's research models, the prediction accuracy of new energy output power is relatively low, lagging behind other prediction systems, such as transportation. Therefore, meeting the needs of grid dispatch, reducing the impact of new energy volatility on the grid, and improving power prediction accuracy have become difficult research challenges today.

There are numerous factors affecting new energy generation, resulting in a huge amount of related data samples. An excessive sample size will directly interfere with prediction efficiency and reduce prediction accuracy. If outliers in the data samples can be screened before establishing the prediction model, this will significantly improve prediction accuracy and speed. Selecting the best data screening and classification methods for different new energy generation becomes a key issue in new energy prediction research.

To address this challenge, this paper proposes a physical-knowledge integrated model for renewable energy power forecasting. The core innovations of the proposed framework can be summarized as follows:

(1) A hybrid physics-informed prediction architecture is developed by integrating the FCM clustering algorithm for handling missing data and the VMD technique to decompose renewable power generation into high-frequency and low-frequency components. This dual-processing mechanism enables the precise capture of the dynamic fluctuation characteristics inherent in renewable energy outputs.

(2) A collaborative deep learning paradigm is implemented where high-frequency components are fed into a CNN to extract transient features from rapid fluctuations, and

low-frequency components are processed through an LSTM network to leverage its long-term memory capacity for capturing temporal dependencies in sequential patterns. This division-of-labor strategy significantly enhances prediction accuracy by synergizing the strengths of both architectures.

2. Measures for Improving the Quality of Renewable Energy Output Data

2.1. Data Preprocessing

In renewable energy power forecasting, the identification of certain data points as outliers is based on conventional statistical criteria, where these values significantly deviate from the distribution pattern of the majority dataset. The output power of renewable energy sources is subject to multiple external factors, including weather variability, equipment malfunction, and grid dispatch adjustments. These factors can induce abrupt changes in system output within short timeframes, generating data points that may appear anomalous. In some scenarios, measurement equipment may experience temporary malfunction or deviation due to environmental conditions (e.g., extreme weather events), yet such deviating data may still contain valuable information regarding system status.

Crucially, renewable energy sources inherently exhibit intermittent and volatile characteristics, leading to substantial temporal variations in their output power. This natural behavior may produce extreme values in datasets that are statistically classified as outliers but actually reflect the authentic operational characteristics of renewable energy systems. This paper employs mean and median imputation for missing data, which partially mitigates data incompleteness. Concurrently, a probabilistic threshold is applied to filter data, representing a common outlier management strategy.

To address missing data in the original dataset, we use the mean to fill in missing values for feature variables with little difference between their median and mean and the median for those with a significant difference. To eliminate the impact of outliers in the dataset, we only retain values within the range of $[\mu - 3\sigma, \mu + 3\sigma]$, ensuring the rationality of the dataset and improving prediction accuracy by excluding outliers. To eliminate the dimensionality effect among feature variables and make them comparable, we need to standardize them so that each feature variable indicator is on the same scale. In this paper, we adopt the z-score normalization method to standardize the data, with the formula being

$$\begin{cases} Z = \frac{x-\mu}{\sigma}, \\ \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)}. \end{cases} \tag{1}$$

In the formula, Z represents the processed data; μ represents the mean of the overall data; σ represents the standard deviation of the overall data. In the dataset used in this paper, there exists a certain linear relationship between different feature variables and the actual power output of renewable energy. Taking photovoltaic power generation as an example, the Pearson correlation coefficient can be used to measure the correlation and degree of correlation between different feature variables and the actual power output of photovoltaics. The calculation formula is as follows:

$$\rho = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2)E^2(X)}\sqrt{E(Y^2)E^2(Y)}}. \tag{2}$$

Here, X and Y represent two different variables.

2.2. Dimensionality Reduction and Clustering of Renewable Energy Output Data

Principal Component Analysis (PCA) is a commonly used data analysis method whose core idea is to achieve dimensionality reduction by extracting main feature components. Through linear transformation, PCA can convert original feature variables into linearly independent data representations. When analyzing the actual power output of photovoltaics and related feature variables, it was found that wind speed, temperature, humidity, and irradiance have significant correlations with the actual power output of photovoltaics, while wind direction and rainfall have less impact on the actual power output of renewable energy. In order to reduce feature variables that have little influence on the prediction of actual renewable energy power output while maintaining information integrity and to accelerate calculation and convergence speed, this paper adopts the PCA method to convert multiple feature variables into a small number of comprehensive indicators, thereby achieving information compression and dimensionality reduction. The criteria for retaining principal components in PCA have two points: (1) Cumulative variance contribution rate: The variance of principal components reflects their ability to carry information from the original data. To preserve sufficient information, we retained principal components whose cumulative contribution rate reached 85–95% of the total variance; (2) Kaiser criterion: Components with eigenvalues greater than one were retained, as eigenvalues below this threshold indicate insufficient explanatory power relative to a single standardized original variable. This dual-criterion approach balances information retention and dimensionality reduction efficiency, aligning with best practices in renewable energy prediction studies. The specific steps are as follows:

(1) Assume that the original data has m samples, $X_1, X_2, \ldots, X_m$, and each sample has n-dimensional features $X_i = (x_{i1}, x_{i2}, \ldots, x_{in})^T$, where each feature x_j has its own characteristic value. After standardizing the data, the covariance matrix is obtained. In the case of three-dimensional data, the covariance matrix is as follows:

$$C = \begin{bmatrix} \mathrm{cov}(x_1, x_1) & \mathrm{cov}(x_1, x_2) & \mathrm{cov}(x_1, x_3) \\ \mathrm{cov}(x_2, x_1) & \mathrm{cov}(x_2, x_2) & \mathrm{cov}(x_2, x_3) \\ \mathrm{cov}(x_3, x_1) & \mathrm{cov}(x_3, x_2) & \mathrm{cov}(x_3, x_3) \end{bmatrix}. \quad (3)$$

(2) Calculate the eigenvalues λ and eigenvectors u of the covariance matrix C.

(3) Project the original features onto the selected eigenvectors to obtain the new k-dimensional features after dimensionality reduction. The calculation formula is as follows:

$$y_k^i = u_k^T(x_1^i, x_1^i, \cdots, x_1^i). \quad (4)$$

In summary, for each sample X_i, it will be transformed from the original $X_i = (x_{i1}, x_{i2}, \ldots, x_{in})^T$ to $X_i = y_{ik}$, achieving the purpose of dimensionality reduction. The core objective of data clustering is to optimize the clustering effect by reducing the internal differences within the same cluster and increasing the differences between different clusters. The FCM clustering algorithm is a fuzzy clustering algorithm based on an objective function. Compared to hard clustering algorithms, FCM focuses more on the membership degree of samples rather than a simple binary classification standard, thus possessing higher flexibility and accuracy. As a soft clustering technique, FCM allows data points to belong to multiple clusters with membership degrees ranging between 0 and 1. The mathematical formulation minimizes the weighted sum of squared distances between data points and cluster centers, subject to the constraint that membership degrees for each data point sum to one. Key steps include formulating the objective function with Lagrangian multipliers and iteratively updating cluster centers and membership matrices.

FCM clustering introduces the concept of membership degree from fuzzy mathematics, differing from the traditional hard clustering approach where a sample belongs either to a cluster (1) or not (0). It is an unsupervised learning "soft clustering" algorithm. The FCM method obtains the membership degree of each sample point to all cluster centers (i.e., the probability that a sample point belongs to each cluster) by optimizing the objective function and determines the cluster membership of the sample points based on the membership degree values. The optimization objective function for FCM clustering is

$$J = \sum_{i=1}^{n} \sum_{k=1}^{K} r_{ik}^{m} \left\| x_i - \mu_k^F \right\|^2. \tag{5}$$

In the formula, K represents the number of cluster centers; μ_k^F denotes the k-th cluster center in FCM clustering; x_i represents the i-th sample; $\|*\|$ indicates a measure of data similarity, most commonly the Euclidean distance. The membership degree r_{ik} represents the probability that the i-th sample point belongs to the k-th cluster, with values closer to one indicating a higher probability of the sample point belonging to that cluster. For a single sample x_i, the sum of its membership degrees to all clusters is one. The fuzzification parameter m controls the degree of fuzziness of the membership degree r_{ik}, and in this paper, m is set to two. When performing FCM clustering, it is often necessary to determine the optimal number of clusters. Generally, the Bayesian Information Criterion method, which balances model fit and complexity, is used for this purpose. FCM clustering optimizes the objective function by iteratively updating r_{ik} and μ_k^F:

$$\mu_k^F = \frac{\sum_{i=1}^{n} r_{ik}^{m} x_i}{\sum_{i=1}^{n} r_{ik}^{m}}. \tag{6}$$

$$r_{ik} = \frac{1}{\sum_{j=1}^{K} \left(\frac{\left\| x_i - \mu_k^F \right\|}{\left\| x_i - \mu_j^F \right\|} \right)^{2/(m-1)}}. \tag{7}$$

3. Decomposition Technique for Renewable Energy Output Sequences Based on the VMD Algorithm

Renewable energy power data exhibit periodicity, fluctuation, and randomness. To reduce the non-stationarity of renewable energy power sequences and uncover internal data patterns, photovoltaic power data need to be decomposed. Commonly used sequence decomposition methods include wavelet decomposition, empirical mode decomposition (EMD), and VMD. The basis function for wavelet transform (WT) is difficult to select and susceptible to white noise. The basis function for EMD can be obtained from the data themselves and is convenient to use, but it is prone to modal aliasing when processing data with large fluctuations, such as photovoltaic power. When modal aliasing occurs, the resulting intrinsic mode functions (IMFs) are meaningless. Additionally, EMD can produce endpoint effects that affect decomposition results. In contrast, VMD not only does not require preset basis functions but also overcomes modal aliasing and has strong noise resistance and good adaptability. VMD also allows for the manual selection of the number of IMF components based on prediction needs. In summary, this paper uses VMD to decompose renewable energy power data and extract patterns from them. As a signal decomposition method analogous to EMD but grounded in variational principles, VMD aims to decompose a signal into multiple IMFs. Each IMF is constrained to a central frequency band. Mathematically, this involves solving a constrained variational problem

where the objective is to minimize the sum of bandwidths of all IMFs while ensuring their summation equals the original signal. The solution typically employs Lagrangian multipliers to handle constraints and alternating Direction Method of Multipliers for iterative optimization. Below is a detailed explanation of the relevant theory.

VMD is a non-recursive signal processing method well-suited for handling non-stationary data such as photovoltaic power. Through iteration, this method decomposes photovoltaic power into several IMFs with limited bandwidths, each centered around its respective central frequency. The goal of VMD is to minimize the sum of the bandwidths of all IMFs, with the constraint that the sum of all IMFs equals the original signal. The frequency spectrum of each mode is calculated using the Hilbert transform, and the frequency spectra of each mode are then input into the basis function to obtain their respective central frequencies. Finally, Gaussian smoothness is used to estimate the bandwidth. The specific expression is as follows:

$$\begin{cases} \min\left\{\sum_k \left\|\partial_t[(\delta(t)+\frac{j}{\pi t})\times u_k(t)]e^{-jw_k t}\right\|_2^2\right\}, \\ s.t.\sum_k u_k = f. \end{cases} \tag{8}$$

In the formula, $\{u_k\} = \{u_1, \cdots, u_n\}$ represents the set of IMFs obtained through decomposition; $\{w_k\} = \{w_1, \cdots, w_n\}$ represents the set of central frequencies of each IMF; and $\delta(t)$ represents the unit impulse function.

By introducing a quadratic penalty factor a and a Lagrange multiplier $\lambda(t)$, the minimization problem is transformed into an unconstrained optimization problem. The augmented Lagrangian function is as follows:

$$\begin{cases} \min\left\{\sum_k \left\|\partial_t[(\delta(t)+\frac{j}{\pi t})\times u_k(t)]e^{-jw_k t}\right\|_2^2\right\}, \\ s.t.\sum_k u_k = f. \end{cases} \tag{9}$$

Using the Alternating Direction Method of Multipliers (ADMM), we iteratively update the variables u_k^{n+1}, w_k^{n+1} and λ^{n+1} until a preset discriminant accuracy is achieved, at which point the iteration stops. Ultimately, we obtain K IMFs after decomposing the renewable energy power. The expression is as follows:

$$\hat{u}_k^{n+1}(w) = \frac{\hat{f}(w) - \sum_{i\neq k}\hat{u}(w) + \frac{1}{2}\hat{\lambda}(w)}{1+2a(w-w_k)^2}. \tag{10}$$

$$w_k^{n+1} = \frac{\int_0^\infty w|\hat{u}_k(w)|^2 dw}{\int_0^\infty |\hat{u}_k(w)|^2 dw}. \tag{11}$$

$$\hat{\lambda}^{n+1}(w) = \hat{\lambda}^n(w) + (\hat{f}(w) - \sum_k \hat{u}_k^{n+1}(w)). \tag{12}$$

$$\sum_k \left\|\hat{u}_k^{n+1} - \hat{u}_k^n\right\|_2^2 \leq \varepsilon\|\hat{u}_k^n\|_2^2. \tag{13}$$

4. Renewable Energy Forecasting Method Based on CNN-LSTM Hybrid Model

After applying the VMD algorithm in Section 3, the high-frequency and low-frequency components of renewable energy can be obtained. The high-frequency components are input into the CNN model, while the low-frequency components are input into the LSTM

neural network for training and prediction. In Section 4.1, the relevant principles of the CNN model are introduced in detail, and a detailed LSTM model is presented in Section 4.2.

4.1. CNN Prediction Model for High-Frequency Components

When dealing with high-frequency components, CNN has stronger local feature extraction capability. The instantaneous fluctuations of high-frequency components, such as wind speed and lighting, have local abrupt changes and spatial correlations. The convolutional kernel of CNN automatically captures local temporal features through sliding windows, without the need for manual feature design. Through multi-layer convolution and pooling operations, CNN can extract features at different time scales, such as minute-level fluctuations and hour-level trends, adapting to the complex high-frequency characteristics of new energy power. CNNs are effective at extracting features from two-dimensional images and high-dimensional data. CNNs have advantages such as reduced memory usage, fewer network parameters, and mitigation of overfitting issues. Therefore, a renewable energy generation power forecasting model is built based on CNNs. A CNN consists of an input layer, hidden layers, and an output layer. The hidden layers are further divided into convolutional layers, pooling layers, and fully connected layers. The structure of CNN is shown in Figure 1 [10]. The functions of each part are described below:

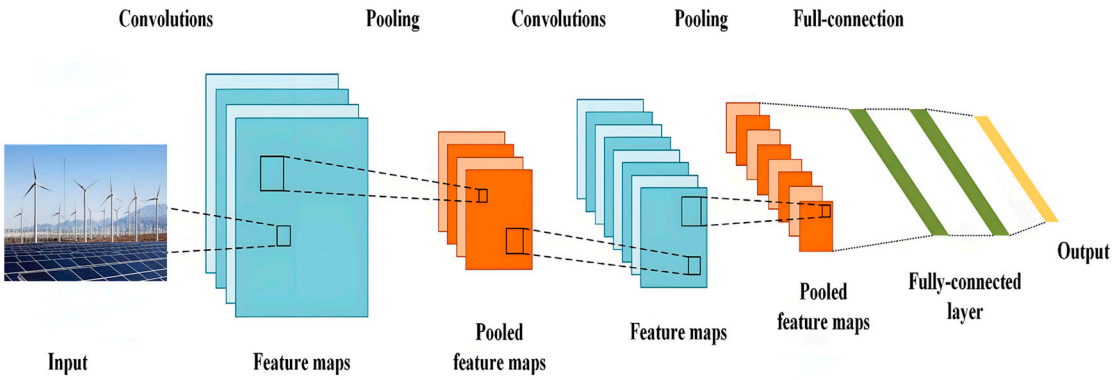


Figure 1. The structure of the CNN.

(1) Input Layer: The role of the input layer is to preprocess the input images or data. Preprocessing methods can reduce the impact of differences in data dimensions on the model and improve the learning efficiency of the model.
(2) Hidden Layers: The hidden layers, including convolutional layers, pooling layers, and fully connected layers, are responsible for feature extraction and learning.
(3) Output Layer: A Softmax layer is added, and the output values from the fully connected layer are input into the Softmax layer to obtain the final output of the neural network, which is the result of the nonlinear transformation through the neural network's nonlinear mapping.

The training process of a CNN is divided into forward propagation and backward propagation. The propagation of data from the input layer to the output layer is called forward propagation. The error is calculated through an error function and then propagated backward through the network. This process is called backward propagation. Training stops when the error reaches the expected standard or the maximum number of iterations is reached. The training process is as follows:

(1) Initialize all parameters in the network.
(2) Forward Propagation: Sample data are sequentially passed through the input layer, hidden layers, and output layer.

(3) Calculate the error between the output values and the actual values using an error function.
(4) Backward Propagation: When the error is greater than the set threshold, and the model training has not reached the maximum number of iterations, the error is propagated backward. The update of network parameters is based on the errors from the fully connected layer, pooling layer, and convolutional layer.
(5) Continue to adjust the parameters based on the error values and return to step two. This process repeats until the error is less than the threshold or the maximum number of iterations is reached, and then the training ends.

4.2. LSTM Prediction Model for Low-Frequency Components

LSTM, built upon the foundation of RNNs, effectively addresses the issues of gradient vanishing or exploding that RNNs encounter when dealing with long time series. LSTM outperforms RNNs in terms of convergence speed and performance when handling long time series problems. The overall structure of LSTM is roughly the same as RNN, with the major difference being that RNN lacks a cell state, whereas LSTM uses a cell state to remember information and employs forget gates, input gates, and output gates to maintain and control this information. LSTM typically uses the Sigmoid function as the activation function for the forget gate and input gate. The Sigmoid function maps the input information to the interval [0, 1], allowing for the selection of input information. An output of 0 indicates discarding all information, while an output of 1 indicates retaining all information. The structure of an LSTM is shown in Figure 2 [13]. LSTM relies on long-term modeling when dealing with low-frequency components. Low-frequency components need to capture long-term dependencies. The gating mechanism of LSTM can effectively establish temporal correlations ranging from a few days to several days, avoiding the problem of gradient vanishing. The power of new energy is affected by multiple factors, such as weather and power grid scheduling, and has non-stationarity. The dynamic memory unit of LSTM can adaptively adjust the temporal weights, which is superior to the static, fully connected structure of BPNN. Meanwhile, LSTM can fuse cross-modal temporal features, while SVM has difficulty efficiently handling such high-dimensional temporal correlations. Compared to traditional SVM, SVM relies on kernel functions to map high-dimensional data, but the original temporal dimension of new energy power is high, and the computational complexity increases exponentially, making it difficult to meet real-time prediction requirements. SVM is essentially a static classifier that relies on manually constructing temporal features and cannot automatically learn temporal dynamic patterns like LSTM, especially in low-frequency trend prediction where performance is limited. SVM performs well in small sample scenarios, but new energy prediction usually requires massive historical data, in which case the generalization ability of deep learning models is more advantageous. BPNN processes temporal data through fully connected layers but has a weak ability to model long-term dependencies, and deep networks are prone to gradient vanishing. In contrast, LSTM explicitly controls the information flow through gating mechanisms. Finally, CNN and LSTM have mature deployment frameworks in the industry, such as TensorFlow PyTorch, supporting GPU acceleration to meet the timeliness requirements of real-time power grid prediction, while SVM or BPNN have poor scalability in large-scale time series prediction.

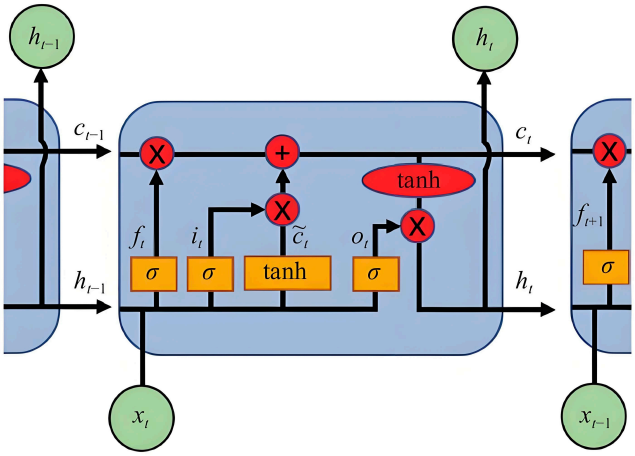


Figure 2. The structure of an LSTM.

The forget gate determines the selection and rejection of information during the process of cellular information transmission, as represented by the following formula:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f). \tag{14}$$

In the above formula, W_f and b_f denote the weight bias of the forget gate. The sigmoid function activation of the input gate yields i_t, and this is then passed through a tanh layer to create a new candidate value vector C_t. The input gate determines the addition of new information to the cell state, as represented by the following Formulas (15)–(17):

$$i_t = \sigma(W_f[h_{t-1}, x_t] + b_t). \tag{15}$$

$$\widetilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C). \tag{16}$$

$$C_t = f_t C_{t-1} + i_t \widetilde{C}_t. \tag{17}$$

Finally, the hidden state h_t is updated, and the output gate determines the final output at the current time step, as represented by the following formula.

$$o_t = \sigma(W_0[h_{t-1}, x_t] + b_o). \tag{18}$$

$$h_t = o_t \tanh(C_t). \tag{19}$$

In the CNN model, the activation functions of the output layer are the Sigmoid function and Tanh function, while the activation functions in other structures are the ReLU function. The ReLU function is widely used due to its simplicity, efficiency, fast computation speed, and ability to effectively alleviate gradient vanishing problems. In the LSTM model, the forget gate uses the Sigmoid function as the activation function to determine which information to discard from the cell state of the previous time step. The input gate also uses the Sigmoid function to determine which values will be used to update the cell state. The candidate memory unit uses the Tanh function as the activation function to generate a new candidate memory value. The output gate uses the Sigmoid function to determine which part of the cell state will be output to the current output value of the LSTM. During the training process, metrics are calculated, such as mean square error (MSE) and coefficient of determination (R2), to evaluate the performance of the model. The smaller the MSE, the closer the R^2 is to one, indicating better model performance. In this article, MSE and R^2 of the CNN model are 24.15 kW and 0.95, respectively. The MSE and R^2 of the LSTM model are 26.34 kW and 0.92, respectively.

The detailed process diagram of the proposed method is given in Figure 3 below to help the reader gain a better understanding.

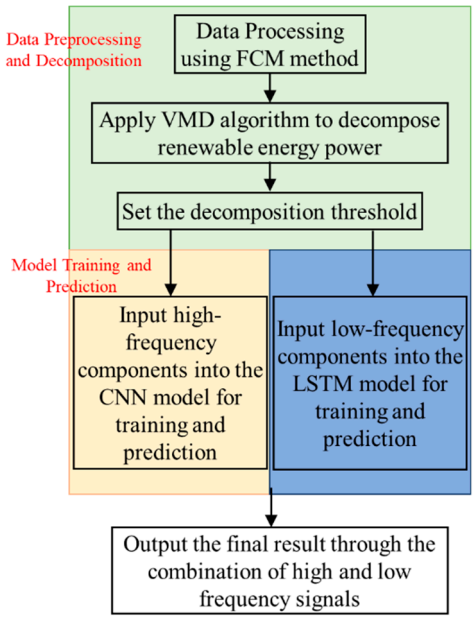


Figure 3. Process diagram of the proposed method.

It should be noted that this paper uses a grid search and cross-validation to optimize the hyperparameters of CNN and LSTM models. Grid search is an exhaustive search method that searches for the optimal hyperparameters by traversing all possible combinations of hyperparameters. It first sets a set of candidate values for each hyperparameter and then generates the Cartesian product of these candidate values to form a composite grid of hyperparameters. Next, the grid search will train and evaluate each hyperparameter combination to find the best-performing hyperparameter combination. Cross-validation is a fundamental technique used in machine learning to evaluate the performance of a model by reducing the risk of overfitting and determining the extent to which the model may generalize to unseen data. It involves dividing the dataset into multiple subsets or folds, then training the model on the training set, and finally evaluating its performance on the remaining validation set. This article sets a set of candidate values for each hyperparameter (including learning rate, convolution kernel size, LSTM layers, number of neurons, etc.) of CNN and LSTM models. Generate Cartesian products of these candidate values to form a composite grid of hyperparameters. For each hyperparameter combination, evaluate the model performance using cross-validation methods. In each cross-validation, the training set is divided into k subsets, and then $k-1$ subsets are used as training data in sequence, with the remaining subset used as validation data. Train and validate each hyperparameter combination k times, and record the performance metrics (accuracy, loss, etc.) for each validation. Calculate the average performance metric for each hyperparameter combination and select the hyperparameter combination with the best performance. Train the final model on the complete training set using the selected optimal hyperparameter combination.

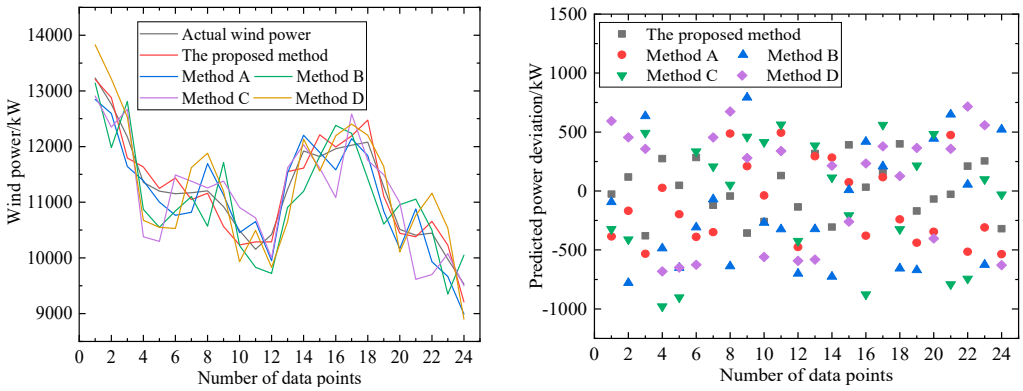
5. Case Study

In order to further demonstrate the effectiveness and feasibility of the method proposed in this paper, predictions are conducted separately for wind power and photovoltaic power. The wind power generation system is equipped with an XYZ-120k turbine made in China, which can continuously output 120 kilowatts of power at a rated wind speed of 12 m per second. Its 12.5 m diameter impeller forms a sweeping area of 122.7 square

meters, coupled with a low cut in wind speed of 3 m per second and a high cut-out wind speed of 25 m per second design, effectively extending the power generation time. The core power generation component adopts a permanent magnet direct drive synchronous generator, and the gearbox-free design significantly reduces mechanical losses. Paired with a yaw system with a response sensitivity of ± 3 degrees, it achieves high-precision wind direction tracking of ± 0.5 degrees through a dual braking mechanism of a hydraulic brake and electromagnetic lock. The photovoltaic power generation system is based on ABC-300 W high-efficiency modules and adopts PERC monocrystalline silicon technology. Under standard testing conditions (STC: irradiance 1000 watts/square meter, temperature 25 °C, atmospheric quality AM1.5), the conversion efficiency reaches 21.5%. The component has a temperature coefficient of $-0.38\%/^{\circ}\mathrm{C}$ and a weak light power generation capacity of 100 watts/square meter, demonstrating excellent environmental adaptability. The component size is $1956 \times 992 \times 40$ mm, equipped with anodized aluminum frame. The electrical energy is converted by a DEF-5 kW inverter made in China, which supports a wide voltage input of 110–500 volts with a maximum conversion efficiency of 97.5%. It is equipped with six independent MPPT tracking and islanding detection mechanisms and has dual-mode communication functions of RS485 and WiFi. The system is equipped with a 60A intelligent controller to achieve $\pm$ 1% accuracy in charge and discharge management. It adopts a four-stage charge and discharge algorithm and integrates a 3.5-inch touch display screen. The bracket system is made of 6063-T5 aluminum alloy material and supports stepless tilt angle adjustment from 5 degrees to 60 degrees. It has a wind load resistance of 2.5 kilonewtons per square meter and a grounding resistance strictly controlled below four ohms. The overall protection level has passed IP65 certification, forming a complete and efficient solution from photovoltaic modules to inverter control and bracket installation.

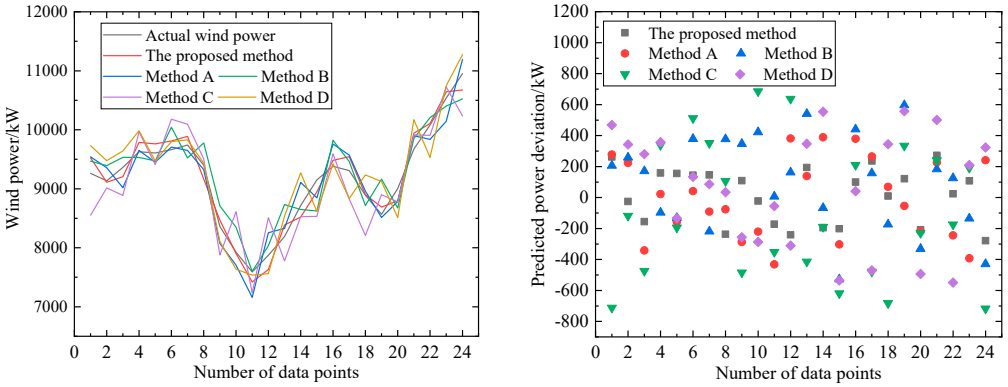
Even if the wind direction changes frequently, the system can still maintain the blades perpendicular to the wind direction, reducing energy loss. In the dataset of this article, more attention is paid to wind speed rather than wind direction. The wind power data are sourced from Spain [22], while the photovoltaic power data are sourced from Guoneng Corporation [23], and the relevant data can be downloaded from the aforementioned links. Here, four commonly used methods are selected for comparison in the case study. Specifically, Methods A, B, C, and D represent the Recurrent Neural Network (RNN) [24], Long Short-Term Memory (LSTM) network [25], Support Vector Machine (SVM) [26], and Convolutional Neural Network (CNN) model [27], respectively. The relevant prediction results are shown in Figures 4–8 below. Among them, subgraph (a) in Figures 4–8 reflects the actual data and predicted wind power using different methods, while the symbols in font (b) represent the difference between predicted data and actual data using different methods. The closer the symbol is to 0, the smaller the prediction error. The programs in this article are all written on MATLAB 2019b software, with a computer configuration of Intel (R) Core (TM) i5-7300HQ CPU @ 2.50 GHz 2.50 GHz and 12 GB of memory. It should be noted that this article did not strictly screen the computing environment. In the field of load forecasting, the focus is more on the accuracy of prediction rather than the efficiency of prediction. This article uses deep learning libraries in MATLAB to call programs. In the CNN model, the sequenceInputLayer is used to define the time series input and adapt it to the temporal characteristics of new energy power. Build convolution kernels using convolution2dLayer, and use batch Normalization Layer and reluLayer to accelerate convergence. Use average pooling 2dLayer for downsampling to preserve the main features. Overlay multiple convolutional structures and use the global Average Pooling2dLayer instead of fully connected layers to reduce parameter count and improve generalization ability. Use regressoLayer to output the prediction results of high-frequency components. In the LSTM model, the sequenceInputLayer is also used to process low-

frequency component inputs. Capture long-term dependencies through lstmLayer and set OutputMode to last to only output the prediction at the last moment. Add a dropout layer to prevent overfitting. Use Fully ConnectedLayer and Regression Layer to output low-frequency component prediction values.



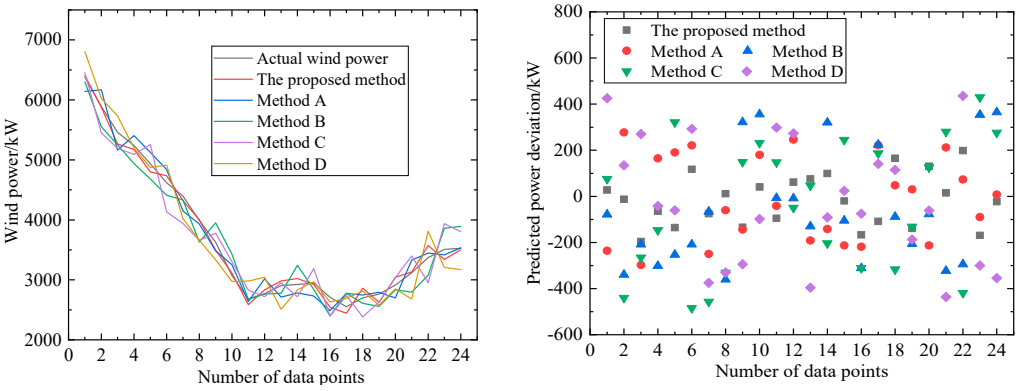
(**a**) Prediction Results (**b**) Error Analysis

Figure 4. Analysis and comparison of prediction results for typical Scenario 1 of wind power.



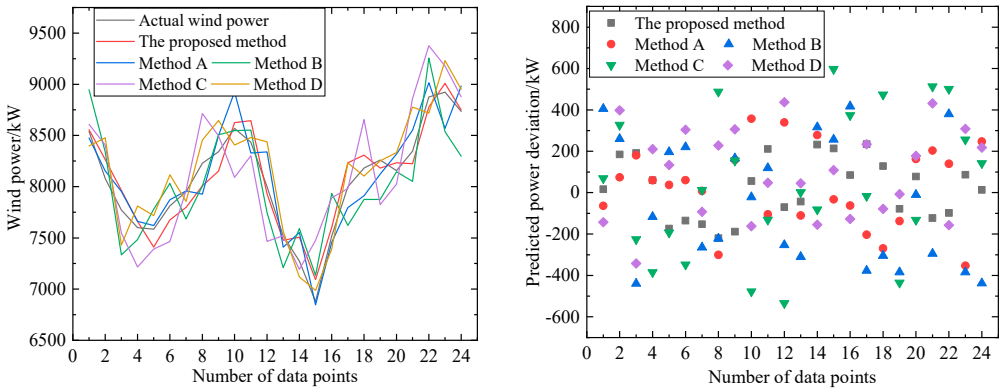
(**a**) Prediction Results (**b**) Error Analysis

Figure 5. Analysis and comparison of prediction results for typical Scenario 2 of wind power.



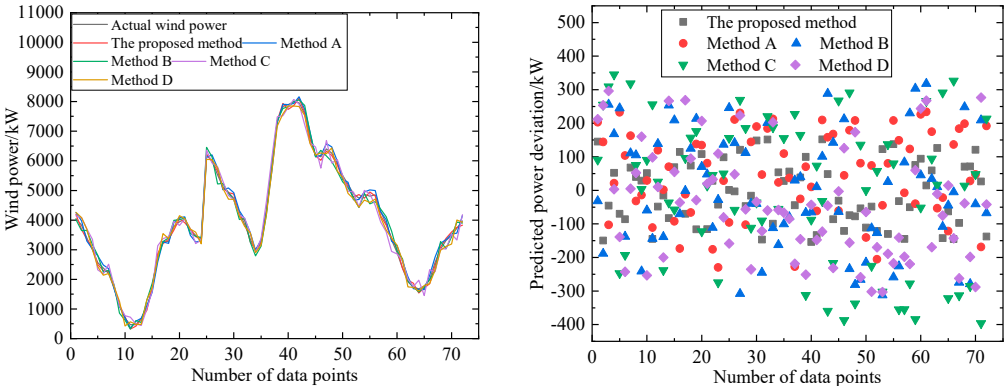
(**a**) Prediction Results (**b**) Error Analysis

Figure 6. Analysis and comparison of prediction results for typical Scenario 3 of wind power.



(**a**) Prediction Results (**b**) Error Analysis

Figure 7. Analysis and comparison of prediction results for typical Scenario 4 of wind power.



(**a**) Prediction Results (**b**) Error Analysis

Figure 8. Analysis and comparison of prediction results for long-term wind power continuous output scenario.

Observing the above figure, the hybrid model proposed in this paper demonstrates the highest prediction accuracy. The renewable energy power prediction model presented in this paper may outperform traditional models such as RNN, LSTM, SVM, and CNN in terms of prediction accuracy. Instead of relying solely on a single machine learning or deep learning algorithm, our model combines physical knowledge with data-driven methods. This integration enables a more comprehensive capture of the physical characteristics and dynamic behaviors of renewable energy generation (especially wind power), thereby enhancing prediction accuracy. The FCM clustering method is employed to handle missing data, which is generally more effective than simple interpolation or deletion of missing values in dealing with incomplete datasets, preserving more useful information. The VMD algorithm is adopted to decompose the renewable energy power into high-frequency and low-frequency components. This decomposition helps separate different features within the signal, allowing subsequent prediction models to focus more on learning from their respective features, thus improving prediction accuracy. For high-frequency components, the CNN model is selected for prediction due to its advantages in processing high-frequency features in images and time series data, which is capable of capturing local variations and detailed information in signals. For low-frequency components, the LSTM model is chosen for its proficiency in handling long-term dependencies and global features of sequence data, which is particularly important for predicting low-frequency components. By inputting high-frequency and low-frequency components into different models and fusing them at the final stage, this strategy fully leverages the strengths of different models, enhancing the overall prediction accuracy and robustness.

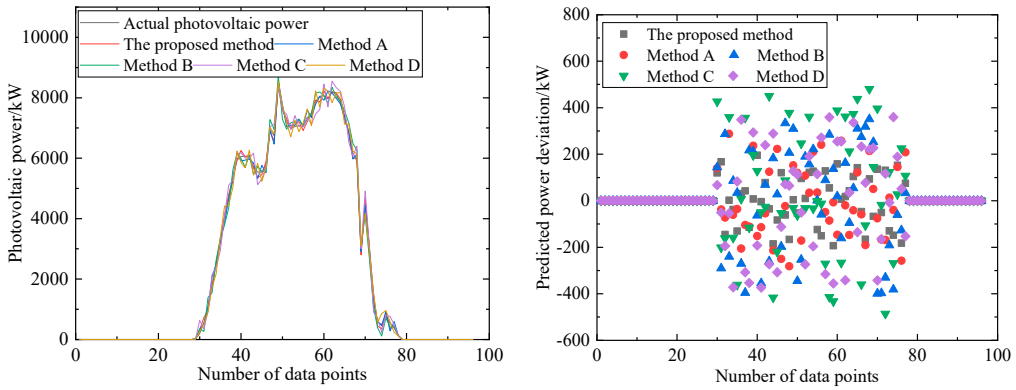
RNNs face challenges when processing long sequences due to parameter sharing and sequence recursion, which can lead to multiple multiplications of gradients during back-propagation, causing the gradient values to gradually decay to zero (gradient vanishing) or grow excessively large (gradient explosion). Gradient vanishing makes it difficult for the network to learn long-distance dependencies, while gradient explosion may result in unstable training processes and even failure to converge. The computational process of RNNs is based on time-step unfolding, with each time step requiring sequential calculation, leading to relatively low computational efficiency. Especially when dealing with long sequences, the computational complexity increases significantly, affecting training and inference speeds. The sequential computation of RNNs is inherently serial, meaning that the computation of each time step depends on the output result of the previous time step. This dependency limits the parallel computing capabilities of RNNs, making them more time-consuming when processing large-scale data.

Although LSTM alleviates the gradient vanishing problem of RNNs to some extent, it still faces challenges. LSTM introduces gating mechanisms to control the flow and forgetting of information, which increases model complexity and training difficulty. This may result in more time-consuming training processes, difficulty in tuning, and an increased risk of overfitting. The numerous parameters in LSTM models require careful adjustment to achieve optimal performance. The selection and optimization of parameters have a significant impact on the model's prediction results, increasing the complexity of parameter tuning. Training SVM models on large-scale datasets can consume considerable time and computational resources. The time complexity of the SVM algorithm is proportional to the number of training samples, so training time increases significantly when the dataset is very large. SVM involves important parameters such as the regularization parameter C and kernel function parameters, whose selection has a significant impact on model performance. The reasonable selection of parameters and kernel functions is a critical issue for SVM, but it usually requires cross-validation of different parameter combinations, increasing the complexity of parameter tuning. The SVM algorithm is sensitive to missing data, and even a small number of missing features can degrade model performance. In practical applications, many datasets contain missing values, requiring preprocessing of missing data to ensure model accuracy.

The performance of CNNs largely depends on the quantity and quality of training data. If training data are insufficient or noisy, the prediction capabilities of CNNs may be limited. CNNs excel at extracting local features of images but are relatively weaker in extracting global features of time series data. This may result in CNNs performing less well than models such as LSTM, which specialize in processing time series data in tasks such as wind power prediction. The generalization ability of CNNs is affected to some extent by their network structure and parameter settings. If the network structure is too complex or the parameters are improperly set, the model may perform poorly on unseen data.

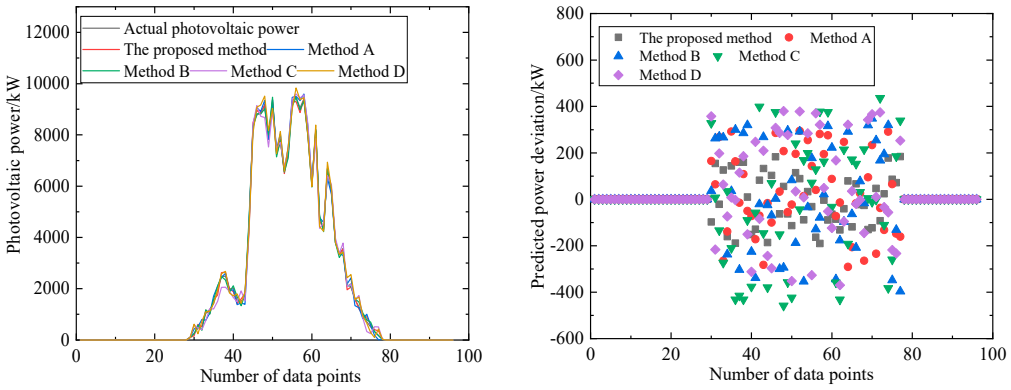
Unlike wind power forecasting, photovoltaic power forecasting presents greater challenges. Photovoltaic power generation primarily relies on solar radiation, which is influenced by multiple factors such as cloud cover and weather changes. This results in significant intermittency and uncertainty in photovoltaic power generation. In contrast, although wind speed is also affected by various factors, it exhibits relatively better continuity and smoother variations. Photovoltaic power generation is not only influenced by solar radiation but also by a multitude of meteorological factors, including temperature, humidity, and atmospheric pressure. The complex interactions between these factors make photovoltaic power forecasting even more difficult. While wind power forecasting is also impacted by meteorological factors, the types and complexity of these factors are relatively lower. Due to the multiple and intricate factors affecting photovoltaic power generation

and the difficulty in accurately describing their interactions, existing forecasting models have certain limitations in terms of adaptability. The significant variations in solar radiation and meteorological conditions across different regions limit the applicability of forecasting models in those areas. In comparison, wind power forecasting models, although affected by geographical differences, demonstrate stronger adaptability. To further illustrate the accuracy of the method proposed in this paper, four typical photovoltaic scenarios were selected for validation. The relevant results are shown in Figures 9–13 below.



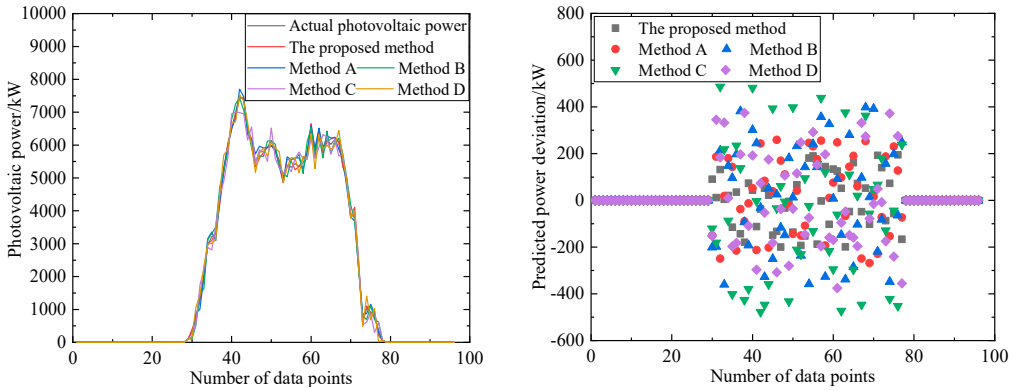
(**a**) Prediction Results (**b**) Error Analysis

Figure 9. Analysis and comparison of prediction results for typical Scenario 1 of photovoltaic power.



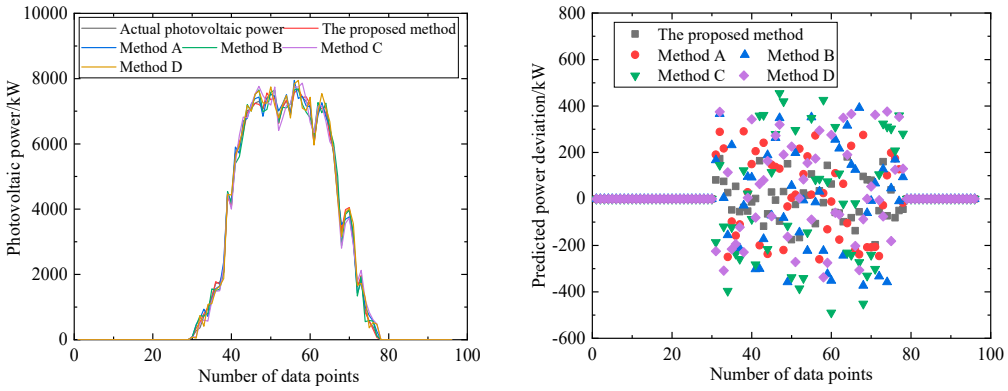
(**a**) Prediction Results (**b**) Error Analysis

Figure 10. Analysis and comparison of prediction results for typical Scenario 2 of photovoltaic power.



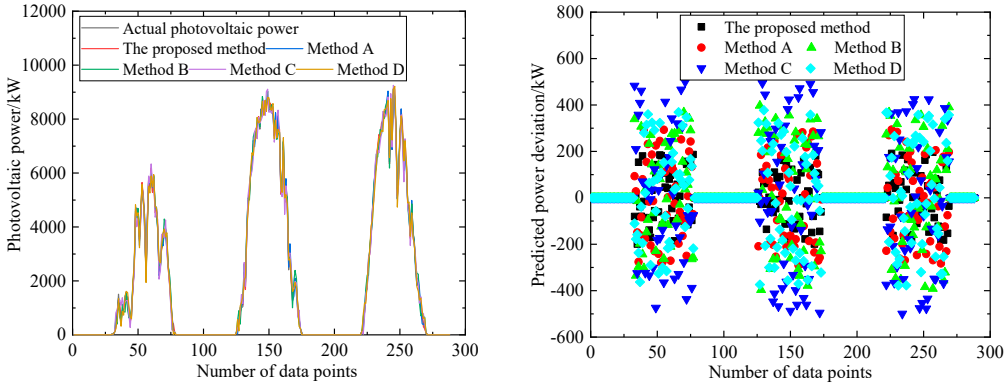
(**a**) Prediction Results (**b**) Error Analysis

Figure 11. Analysis and comparison of prediction results for typical Scenario 3 of photovoltaic power.



(**a**) Prediction Results (**b**) Error Analysis

Figure 12. Analysis and comparison of prediction results for typical Scenario 4 of photovoltaic power.



(**a**) Prediction Results (**b**) Error Analysis

Figure 13. Analysis and comparison of prediction results for long-term photovoltaic power continuous output scenario.

Furthermore, in order to further verify the robustness of the method proposed in this paper against noise interference, this paper introduces disturbed data into the original wind power and photovoltaic power data and calculates the average prediction error rate. The results are presented in Table 1 below.

Table 1. Analysis of prediction error under noise interference.

Proportion of Disturbed Data	The Proposed Method	Method A	Method B	Method C	Method D
5%	4.52%	6.28%	8.52%	7.61%	8.26%
10%	7.56%	8.97%	11.40%	12.56%	17.52%
15%	9.24%	12.40%	15.62%	16.83%	20.4%

From the table above, it is evident that as the proportion of disturbed data increases, the accuracy of each method decreases accordingly. However, the maximum error of the method proposed in this paper remains within 10%. The proposed method combines physical characteristics with machine learning models. By incorporating physical knowledge, the model can better understand the physical processes of renewable energy generation, thereby improving the accuracy and robustness of predictions. This method not only leverages the advantages of data-driven approaches but also integrates the physical mechanisms of renewable energy generation, enabling the model to more accurately capture core features when facing disturbed data. The use of the FCM method to handle missing data effectively fills in the gaps in the data, reducing the impact of incomplete data on prediction

results. The VMD algorithm decomposes renewable energy power into high-frequency and low-frequency components, helping the model better capture feature information at different frequencies. Due to the combination of physical knowledge and advanced machine learning algorithms, the proposed method can better filter out noise and extract useful information from raw data containing disturbed data. This makes the model more stable and reliable in practical applications, capable of adapting to different data environments and conditions.

To further visualize the error distributions, this study conducts a comparative analysis of prediction errors between the proposed method and the baseline approach without data preprocessing, as presented in Table 2 below.

Table 2. Comparison of forecasting errors before and after data processing.

Model	Short-Term Photovoltaic Power Prediction	Long-Term Photovoltaic Power Prediction	Short-Term Wind Power Prediction	Long-Term Wind Power Prediction
Error after data processing	3.57%	5.65%	3.17%	4.98%
Error before data processing	4.95%	8.24%	4.56%	7.88%

As demonstrated in the table above, after implementing data preprocessing procedures, both long-term and short-term forecasting errors exhibit significant reductions. When data follow a near-normal distribution, mean imputation may yield relatively accurate results. However, in the presence of skewed distributions or outliers, the mean becomes susceptible to distortion from extreme values, potentially leading to substantial discrepancies between imputed and true values. Median imputation demonstrates greater robustness against skewed data and outliers, as it remains unaffected by extreme values. Conversely, when data exhibit high symmetry without outliers, the discrepancy between median and mean may become negligible, rendering median imputation no more advantageous in terms of accuracy. Furthermore, when missingness occurs completely at random, mean or median imputation can produce reasonable results. Yet, under non-random missingness mechanisms (e.g., missingness dependent on unobserved variables or data values themselves), both imputation methods may introduce systematic biases regardless of distribution characteristics. The exclusion of outliers beyond the [$\mu - 3\sigma$, $\mu + 3\sigma$] range helps mitigate the distortive effects of extreme values on central tendency measures. Standardization procedures are implemented to eliminate dimensional discrepancies across features, ensuring the comparability of variables with different scales. The error curve is convex during the training and testing process, indicating that the model is learning steadily during the training process and has not fallen into local optima or saddle points. The error can be reduced by 1–2 orders of magnitude, and there is no overfitting phenomenon.

The error used in Table 1 reflects the size of MAPE. Choosing appropriate evaluation parameters is crucial in the field of new energy forecasting. Different parameters have different emphases. MAE focuses more on the average size of absolute error, RMSE emphasizes the degree of deviation, and MAPE focuses on the percentage form of relative error. In order to facilitate readers' better understanding, this article further compares the MAE and RMSE indicators under different prediction methods, as shown in Table 3.

Table 3. Comparison of performance indicators of different prediction methods.

Indicators	The Proposed Method	Method A	Method B	Method C	Method D
MAE/kW	10.20	15.60	20.10	17.30	16.50
RMSE/kW	15.78	21.45	27.52	25.40	24.50

This paper further selected five typical new energy output datasets provided in previous references, as shown below, to verify the feasibility of the proposed method.

1. Dataset 1: actual wind power data within 30 days from sites located in Gansu Province [28]
2. Dataset 2: actual Klim Wind Farm [29]
3. Dataset 3: actual load data in a certain region of China [30]
4. Dataset 4: real-world solar farms [31]
5. Dataset 5: solar power measurements in California [32]

Observing Table 3 and Figure 14, it can be observed that the method proposed in this article performs the best in both MAE and RMSE indicators. In terms of MAE indicators, the method proposed in this article can reduce errors by up to 49.3%; on the RMSE index, the method proposed in this article can reduce errors by up to 42.7%. The smaller the MAE value, the smaller the prediction error of the model and the stronger its predictive ability. The smaller the RMSE value, the higher the prediction accuracy of the model. In five typical scenarios, the method proposed in this article achieved the best performance in four of them, except for a decreasing trend in accuracy in load power prediction scenarios, which is due to the chaotic spatiotemporal distribution characteristics of load power and the insignificant significance of decomposing it into high- and low-frequency signals. In the other four scenarios, the method proposed in this article can improve accuracy by a maximum of 6.4% and an average improvement of 3.5%. RMSE is easy to understand and easy to calculate, sensitive to outliers, and can effectively reflect the precision of measurements. This article combines the advantages of physics knowledge and data-driven models. By combining physical characteristics with the learning ability of data-driven models, it is possible to capture more accurately the changing patterns of new energy power. The high-frequency components are input into a CNN model, which excels at processing data with local correlations and can effectively capture short-term fluctuation features in the high-frequency components. The low-frequency components are input into an LSTM neural network, which is suitable for processing time series data and can capture long-term dependencies to accurately predict long-term trends in the low-frequency components. This hybrid model, which combines CNN and LSTM, can fully utilize both advantages and improve the accuracy of prediction.

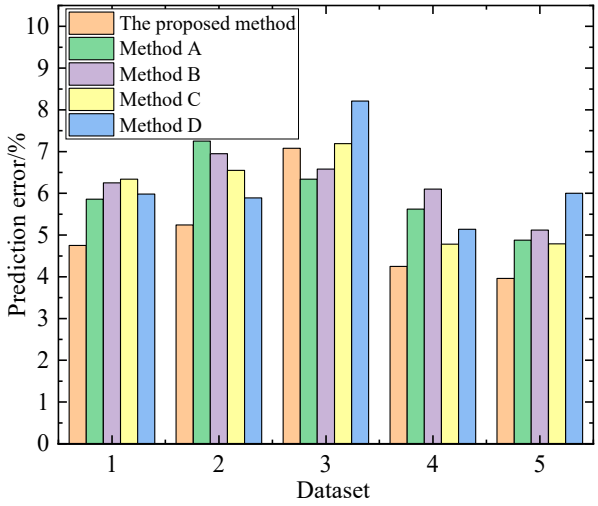


Figure 14. Effectiveness validation of the proposed methods on different datasets.

6. Conclusions

This paper addresses the challenges posed by the high proportion of renewable energy integration into the power grid for the stable operation of the power system. With the core goal of improving power prediction accuracy, a hybrid prediction model that integrates physical knowledge is proposed. By combining FCM data restoration, VMD signal decomposition, and CNN-LSTM deep learning architecture, this model can efficiently handle the volatility and noise interference of renewable energy. The measured results show that its prediction accuracy is 5.71% higher than traditional methods, and the error reduction in complex noise environments is 11.16%. This achievement not only provides key technical support for the consumption of renewable energy and the safe dispatch of the power grid but also lays an important foundation for the steady progress of energy structure transformation.

In the future, we will continue to study the generalization ability of the model, further verify its applicability in different types of renewable energy such as hydropower and tidal power and multi-energy coupling scenarios, and explore collaborative optimization mechanisms for cross-regional and multi-time scale prediction. At the same time, more refined physical constraints are introduced to construct a bidirectional feedback mechanism of "mechanism driven data correction" to enhance the robustness of the model to extreme working conditions.

Author Contributions: Conceptualization, S.H., S.J., Z.S., X.L., S.W. and H.M.; methodology, S.H., S.J., Z.S., X.L., S.W. and H.M.; software, S.H., S.J., Z.S., X.L., S.W. and H.M.; investigation, S.H., S.J., Z.S., X.L., S.W. and H.M.; writing—original draft preparation, S.H., S.J., Z.S., X.L., S.W. and H.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by State Grid Corporation of China Headquarters Technology Project: Extreme Weather Warning and Risk Assessment Affecting Power System Operation (No. 5200-202355779A-3-7-CB).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, J.; Luo, Y.; Yang, S.; Wei, S.; Huang, Q. Review of Uncertainty Prediction Methods for Renewable Energy Power. *High Volt. Eng.* **2021**, *47*, 1144–1155.
2. Rafi, S.H.; Masood, N.A.; Deeba, S.R.; Hossain, E. A Short-Term Load Forecasting Method Using an Integrated CNN and LSTM Network. *IEEE Access* **2021**, *9*, 32436–32448.
3. Tong, H.; Zeng, X.; Yu, K.; Zhou, Z. A Fault Identification Method for Animal Electric Shocks Considering Unstable Contact Situations in Low-Voltage Distribution Grids. *IEEE Trans. Ind. Inform.* **2025**, 1–12. [CrossRef]
4. Huang, Y.; Wei, Y.; Tong, D.; Wang, W. Short-Term Photovoltaic Power Forecasting Based on VMD and Improved TCN. *Electron. Sci. Technol.* **2023**, *36*, 42–49.
5. Xia, Y.; Li, Z.; Xi, Y.; Wu, G.; Peng, W.; Mu, L. Accurate Fault Location Method for Multiple Faults in Transmission Networks Using Travelling Waves. *IEEE Trans. Ind. Inform.* **2024**, *20*, 8717–8728.
6. Dong, X.; Zhao, H.; Zhao, S.; Lu, D. Ultra-Short-Term Photovoltaic Power Forecasting Based on SOM Clustering and Secondary Decomposition BiGRU. *Acta Energiae Solaris Sin.* **2022**, *43*, 85–93.
7. Wu, H.; Xu, J.; Wang, J.; Long, M. Autoformer: Transformers with Auto-correlation for Decomposition in Long-term Series Forecasting. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 22419–22430.
8. Zhou, H.; Zhang, S.; Peng, J.; Zhang, S.; Li, J.; Xiong, H.; Zhang, W. Informer: Beyond Efficient Transformer for Long Sequence Time-series Forecasting. *Proc. AAAI Conf. Artif. Intell.* **2021**, *35*, 11106–11115.
9. Liu, H.; Xiang, M.; Zhou, B.; Duan, Y.; Fu, D. Long-sequence Time Series Power Load Forecasting Based on Informer. *J. Hubei Minzu Univ. (Nat. Sci. Ed.)* **2021**, *39*, 326–331.
10. Wu, Z.; Pan, F.; Li, D.; He, H.; Zhang, T.; Yang, S. Prediction of Photovoltaic Power by the Informer Model Based on Convolutional Neural Network. *Sustainability* **2022**, *14*, 13022. [CrossRef]
11. Yu, F.; Wang, L.; Jiang, Q.; Yan, Q.; Huang, J. Short-term Load Forecasting Based on Feature Selection and VMD-MIC-SSA-Informer. *J. Shaanxi Univ. Sci. Technol.* **2022**, *40*, 191–196+203.
12. Chang, D.; Nan, X. Short-term Photovoltaic Power Forecasting Based on Improved Backpropagation Neural Network with Hybrid Sparrow Search Algorithm. *Mod. Electr. Power* **2022**, *39*, 287–298.
13. Sun, H.; Yang, F.; Gao, Z.; Hu, S.; Wang, Z.; Liu, J. MIBILSTM-based Short-term Load Forecasting Considering Feature Importance Value Fluctuation. *Autom. Electr. Power Syst.* **2022**, *46*, 95–103.
14. Lu, W.; Xiao, H.; Wu, Z.; Wang, Z.; Zhao, S. Ultra-short-term combined forecasting of photovoltaic power based on chaotic CSO-WNN-RBF. *Chin. J. Power Sources* **2021**, *45*, 485–489.
15. Li, R.; Ding, X.; Sun, F.; Han, Y.; Liu, H.; Yan, J. Ultra-short-term photovoltaic power forecasting based on the Wide & Deep-XGB2LSTM model. *Electr. Power Autom. Equip.* **2021**, *41*, 31–37.
16. Yang, G.; Zhang, K.; Wang, D.; Liu, J.; Qin, M. Multi-mode fusion ultra-short-term photovoltaic power forecasting algorithm based on envelope clustering. *Electr. Power Autom. Equip.* **2021**, *41*, 39–46.
17. Guo, F.; Liu, D.; Zhang, Z.; Tang, F. Short-term load forecasting based on GPSO-LSTM with feature correlation analysis correction. *Electr. Meas. Instrum.* **2021**, *58*, 39–48.
18. Qi, X. Research on Photovoltaic/Wind Power Forecasting Based on Deep Learning. Ph.D. Thesis, Harbin Engineering University, Harbin, China, 2020.
19. Jia, P.; Lin, K.; Guo, F. Coal spontaneous combustion temperature prediction model based on PSO-SRU deep neural network. *Ind. Mine Autom.* **2022**, *48*, 105–113.
20. Zeng, W.; Li, J.; Sun, C.; Cao, L.; Tang, X.; Shu, S.; Zheng, J. Ultra Short-Term Power Load Forecasting Based on Similar Day Clustering and Ensemble Empirical Mode Decomposition. *Energies* **2023**, *16*, 1989. [CrossRef]
21. Xu, S.; Si, C.; Wan, C.; Yu, J.; Cao, Z. Load curve ensemble spectral clustering algorithm considering dual-scale similarity. *Autom. Electr. Power Syst.* **2020**, *44*, 152–160.
22. Available online: https://download.csdn.net/download/qq_39727487/15420454?ops_request_misc=%257B%2522request%255Fid%2522%253A%2522b8d60b54817f371a66062bc04c25eced%2522%252C%2522scm%2522%253A%252220140713.130102334..%2522%257D&request_id=b8d60b54817f371a66062bc04c25eced&biz_id=1&utm_medium=distribute.pc_search_result.none-task-download-2~download~baidu_landing_v2~default-1-15420454-null-null.269%5Ev2%5Econtrol&utm_term=%E8%A5%BF%E7%8F%AD%E7%89%99%E9%A3%8E%E7%94%B5%E6%95%B0%E6%8D%AE&spm=1018.2226.3001.4451.2 (accessed on 25 March 2025).
23. Available online: https://download.csdn.net/download/liu864225597/12761340?ops_request_misc=%257B%2522request%255Fid%2522%253A%25221593a4fd0d1e1e77f902ad9f6735dbe7%2522%252C%2522scm%2522%253A%252220140713.130102334.pc%255Fdownload.%2522%257D&request_id=1593a4fd0d1e1e77f902ad9f6735dbe7&biz_id=1&utm_medium=distribute.pc_search_result.none-task-download-2~download~first_rank_ecpm_v1~rank_v31_ecpm-1-12761340-null-null.269%5Ev2%5Econtrol&utm_term=%E5%9B%BD%E8%83%BD%E6%97%A5%E6%96%B0%E5%85%89%E4%BC%8F%E5%8A%9F%E7%8E%87%E9%A2%84%E6%B5%8B&spm=1018.2226.3001.4451.2 (accessed on 25 March 2025).

24. Qiu, R.; Dai, W.; Wang, G.; Luo, Z.; Li, M. Evaluation of Different Deep Learning Methods for Meteorological Element Forecasting. *IEEE Access* **2024**, *12*, 81772–81782.
25. Lin, X.; Tan, Z.; Liu, Y. Wind Power Forecasting: LSTM-Combined Deep Reinforcement Learning Approach. In Proceedings of the 2023 IEEE 7th Conference on Energy Internet and Energy System Integration (EI2), Hangzhou, China, 15–18 December 2023; pp. 5202–5206.
26. Nikitha, M.S.; Nisha, K.C.R.; Gowda, M.S.; Aithal, P.; Mudakkayil, N.M. Solar PV Forecasting Using Machine Learning Models. In Proceedings of the 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 23–25 February 2022; pp. 109–114.
27. Thanh, S.T.; Dinh, H.D.; Minh, G.N.H.; Trong, T.N.; Duc, T.N. Spatial-Temporal Graph Hybrid Neural Network for PV Power Forecast. In Proceedings of the 2024 8th International Conference on Green Energy and Applications (ICGEA), Singapore, 14–16 March 2024; pp. 317–322.
28. Chen, N.; Wang, Q.; Yao, L.; Zhu, L.; Tang, Y.; Wu, F.; Chen, M.; Wang, N. Wind power forecasting error-based dispatch method for wind farm cluster. *J. Mod. Power Syst. Clean Energy* **2013**, *1*, 65–72. [CrossRef]
29. Sideratos, G.; Hatziargyriou, N.D. An Advanced Statistical Method for Wind Power Forecasting. *IEEE Trans. Power Syst.* **2007**, *22*, 258–265. [CrossRef]
30. Wang, Y.; Xia, Q.; Kang, C. Secondary Forecasting Based on Deviation Analysis for Short-Term Load Forecasting. *IEEE Trans. Power Syst.* **2011**, *26*, 500–507. [CrossRef]
31. Kharazi, S.; Amjady, N.; Nejati, M.; Zareipour, H. A New Closed-Loop Solar Power Forecasting Method with Sample Selection. *IEEE Trans. Sustain. Energy* **2024**, *15*, 687–698. [CrossRef]
32. Ramakrishna, R.; Scaglione, A.; Vittal, V.; Dall'Anese, E.; Bernstein, A. A Model for Joint Probabilistic Forecast of Solar Photovoltaic Power and Outdoor Temperature. *IEEE Trans. Signal Process.* **2019**, *67*, 6368–6383. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.