Article

Short-Term Photovoltaic Power Generation Prediction Model Based on Improved Data Decomposition and Time Convolution Network

Ranran Cao 1,2, He Tian 1,2,*; Dahua Li 1,2,*; Mingwen Feng 1,2 and Huaicong Fan 1,2

Abstract: In response to the volatility of photovoltaic power generation, this paper proposes a short-term photovoltaic power generation prediction model (HWOA-MVMD-TPA-TCN) based on a Hybrid Whale Optimization Algorithm (HWOA), multivariate variational mode decomposition (MVMD), temporal pattern attention mechanism (TPA), and temporal convolutional network (TCN). In order to improve the accuracy of photovoltaic power generation forecasting, HWOA-MVMD is used for data decomposition, the Minimum Mode Overlap Component (MMOC) is used as the objective function, the photovoltaic power generation sequence is decomposed into finite Intrinsic Mode Functions (IMFs) according to the optimal solution, and the training set is formed with key meteorological variable data such as total radiation (unit: W/m²), ambient temperature, and humidity. Then, the TPA-TCN model is trained to obtain the final predicted values after superimposing the reconstruction of the prediction results, and finally the prediction error of the photovoltaic power generation data is studied. The proposed method is applied to real photovoltaic power generation data from a commercial center in Tianjin and compared with HWOA-MVMD-BiLSTM, GWO-MVMD-TPA-TCN, and TPA-TCN prediction models. The simulation results demonstrate that the MAE value of the forecast method proposed in this paper is 1.95 MW and the RMSE value is 2.55 MW, which can be reduced by up to 33.74% and 38.85%, respectively. The HWOA-MVMD-TPA-TCN-based short-term photovoltaic power generation prediction model presented in this paper achieves higher prediction accuracy and superior performance, serving as a valuable reference for related research.

Keywords: photovoltaic power generation forecast; hybrid whale algorithm; multivariate variational mode decomposition; time convolutional network; time pattern attention mechanism

1. Introduction

With the rapid development of photovoltaic power generation, the impact of large-scale photovoltaic grid connections on the power system is becoming more and more obvious, and accurate short-term photovoltaic power generation prediction can effectively alleviate the pressure caused by photovoltaic grid connections on the power system, which is of great significance to ensure the stable operation of the power grid and the reasonable allocation of resources [1,2]. Therefore, obtaining reliable data for the power generation forecast of photovoltaic power plants has also become an important issue [3].

In recent years, research in photovoltaic power generation forecasting has predominantly focused on artificial intelligence technologies, utilizing machine learning and
neural networks to capture the randomness of photovoltaic power generation sequences [4]. Table 1 shows the current research status.

Table 1. Research status.

<table>
<thead>
<tr>
<th>File Number</th>
<th>Main Content</th>
<th>Areas of Shortcoming</th>
</tr>
</thead>
<tbody>
<tr>
<td>[5]</td>
<td>Using ant element data packets to probe the network environment, selecting information transmission links, constructing roaming paths, and enhancing predictive accuracy through iterative computations to search for optimal solutions.</td>
<td></td>
</tr>
<tr>
<td>[6]</td>
<td>Optimizing the photovoltaic power output prediction algorithm based on the Whale Optimization Algorithm for Support Vector Machines. Optimizing time delay and embedding dimensions in the kernel function to improve generalization ability and convergence speed, resulting in better adaptability.</td>
<td>Ignored the problem of poor population diversity and the tendency for the initialization to fall into local optima.</td>
</tr>
<tr>
<td>[7]</td>
<td>Optimizing the photovoltaic power prediction algorithm using the Grey Wolf Algorithm to optimize the weights of Long Short-Term Memory (LSTM) neural networks. Predicting power based on optimal weights, overcoming the drawbacks of backpropagation, and improving prediction accuracy.</td>
<td></td>
</tr>
<tr>
<td>[8–12]</td>
<td>Establishing optimization models using the Honey Badger Algorithm, Sparrow Algorithm, and PSO (Particle Swarm Optimization) Algorithm for predicting the power generation of photovoltaic power stations. Enhancing prediction accuracy.</td>
<td>Required a large amount of historical data support.</td>
</tr>
<tr>
<td>[13,14]</td>
<td>Taking into account the impact of solar irradiance on photovoltaic power generation. Integrating variational quantum circuits with Long Short-Term Memory (LSTM) neural networks, forming a Quantum LSTM neural network applied in predictive research. Accelerating the prediction algorithm through an FPGA hardware platform to reduce computational complexity.</td>
<td></td>
</tr>
<tr>
<td>[15–20]</td>
<td>Pointing out the random fluctuations due to meteorological factors in ultra-short-term photovoltaic power generation forecasting.</td>
<td>Low prediction accuracy and did not comprehensively consider all meteorological factors.</td>
</tr>
<tr>
<td>[21]</td>
<td>Utilizing the Artificial Bee Colony Optimization Support Vector Machine (ABC-SVM) classification model, combined with the Particle Swarm Optimization Random Forest (PSO-RF) model, for classification training based on meteorological data.</td>
<td>Suitable for small sample analysis, but the time complexity increased with the growth of the sample size.</td>
</tr>
<tr>
<td>[22]</td>
<td>Adopting the Complex Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) method to decompose the photovoltaic power sequence, reducing the impact of non-stationary features on the prediction.</td>
<td>There were challenges related to mode mixing and difficulty in determining the stopping conditions.</td>
</tr>
<tr>
<td>[23–26]</td>
<td>Establishing a combined forecasting model to enhance the accuracy and stability of ultra-short-term load forecasting.</td>
<td>For nonlinear time series, it did not perform well in capturing nonlinearity.</td>
</tr>
</tbody>
</table>

While there are many forecasting methods as mentioned above, they often struggle to effectively capture nonlinearity in time series data that exhibit dynamic and complex behaviors with multiple variables. Many data points within these sequences may not be adequately represented, which can directly impact the final prediction results, potentially
leading to forecast errors exceeding acceptable limits. Therefore, this paper proposes a short-term photovoltaic power generation prediction model based on the HWOA-MVMD-TPA-TCN. This approach utilizes HWOA-MVMD to decompose the photovoltaic power generation sequence comprehensively, reducing the sequence complexity while deeply mining data features. The obtained subsequences and key meteorological factors are used as inputs for iterative training in the TPA-TCN. The prediction results for each subsequence are then aggregated and reconstructed to obtain the final forecast value. When compared to HWOA-MVMD-BiLSTM, GWO-MVMD-TPA-TCN, and TPA-TCN models, the proposed forecasting model demonstrates a higher level of prediction accuracy. This approach not only provides valuable insights for energy forecasting but also has the potential to be applied in a broader context of energy prediction.

The overall organization of this article is as follows: The first chapter is an introduction. The second chapter introduces data decomposition, using a multivariational mode to decompose photovoltaic power generation data and using the improved Hybrid Whale Algorithm to optimize the hyperparameters of multivariational modal decomposition. The third chapter presents iterative training using the results and key meteorological factors obtained in the second chapter as the input of the TPA-TCN. The fourth chapter presents the prediction process of the short-term photovoltaic power generation prediction model based on the HWOA-MVMD-TPA-TCN, and the fifth chapter presents the analysis of an example.

2. Optimization of Multivariate Variational Modal Decomposition

2.1. Common Data Decomposition Methods

From the current research status, data decomposition methods generally fall into several categories, as shown in Table 2. These methods were used in many aspects, such as wind power prediction [27], bearing fault diagnosis [28], and the low-frequency oscillation mode identification of power systems [29]. These methods often exhibit certain issues during the data decomposition process.

<table>
<thead>
<tr>
<th>Data Decomposition Method</th>
<th>Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelet decomposition</td>
<td>Wavelet transform is limited by the need for the manual determination of wavelet bases and the Heisenberg uncertainty principle, which ultimately affects the accuracy of the prediction results.</td>
</tr>
<tr>
<td>Empirical mode decomposition (EMD)</td>
<td>There are problems such as modal aliasing, endpoint effects, and difficulty in determining the stopping conditions.</td>
</tr>
<tr>
<td>Variational mode decomposition (VMD)</td>
<td>VMD requires predefined modal numbers. Inaccurate modal numbers can lead to insufficient or excessive modal decomposition, and if the signal is long, the bandwidths may overlap.</td>
</tr>
<tr>
<td>Multivariate variational modal decomposition (MVMD)</td>
<td>The influencing parameters in its decomposition process are related to the number of intrinsic modes and the quadratic penalty factor, and these parameters must be preset.</td>
</tr>
</tbody>
</table>

According to Table 2, there are currently two main issues with data decomposition methods: (1) In signal decomposition, the selection of primary parameters is often based on past empirical experience, which limits the results of the decomposition. (2) Problems exist such as modal aliasing, endpoint effects, and difficulty in determining stopping conditions. To address these issues, this paper employs an improved Hybrid Whale
Optimization Algorithm. It defines the modal overlap component as the fitness function, and when this indicator reaches its minimum value, the independence and correlation between different modal components are maximized. This helps avoid modal aliasing. At this point, the corresponding optimal combination of intrinsic mode numbers and penalty factors is achieved, preventing under-decomposition or over-decomposition phenomena.

2.2. Hybrid Whale Optimization Algorithm

The Whale Optimization Algorithm (WOA) has the advantages of simple operation and few setting parameters and has been widely used in the optimization of target problems. And the introduction of the WOA in the prediction model can effectively search the solution space and find the global optimal solution or the approximate optimal solution. However, the traditional WOA is prone to fall into local optima, and the global search capability and local development capability are unbalanced [30]. Therefore, this article improves on the traditional WOA.

2.2.1. Tent Mapping

The Whale Optimization Algorithm (WOA), like most swarm intelligence optimization algorithms, initially starts with a random distribution of individuals, which can lead to population clustering and poor diversity. The population may struggle to achieve a uniform distribution throughout the search space. To enhance the performance of the Whale Optimization Algorithm, Tent Mapping is applied in the optimization process of the WOA [30]. The Tent Mapping formula is as follows:

If \( 0 < x_j < 0.5 \):

\[
x_{j+1} = \delta \times x_j
\]

(1)

If \( 0.5 \leq x_j \leq 1 \):

\[
x_{j+1} = \delta(1 - x_j)
\]

(2)

Assuming a population size of \( N \), we obtain a population \( \chi = \{x_j, j = 1, 2, \cdots N\} \), where \( \delta = (0, 2] \) is the chaos parameter, and the larger the value of this parameter, the better the chaotic effect. \( x_j \) is the position of each whale in the search space. Therefore, when \( \delta = 2 \), the population’s exploration capability and the algorithm’s solving speed are expected to be superior to other mapping methods.

2.2.2. Elite Reverse Learning Strategy

First, by employing a group selection strategy, after generating reverse solutions from elite individuals, the individuals with lower fitness values are selected as the next generation’s whale individuals [31].

In a D-dimensional space, if there is a feasible solution \( \chi = (x_1, x_2, \cdots, x_n) \), then its reverse solution \( \bar{\chi} = (\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n) \), which is an elite individual in the population, corresponds to the extremum point \( X_e^{c,j} \) in the whale population itself.

\[
X_e^{c,j}(X_e^{c,1}, X_e^{c,2}, \cdots, X_e^{c,n})\]

\( i = 1, 2, \cdots, s; j = 1, 2, \cdots, n \)

is set as its reverse solution \( \bar{X}_e^{c,j}((\bar{x}_e^{c,1}, \bar{x}_e^{c,2}, \cdots, \bar{x}_e^{c,n})) \), which can be defined as \( \bar{X}_e^{c,j}((\bar{x}_e^{c,1}, \bar{x}_e^{c,2}, \cdots, \bar{x}_e^{c,n})) \), where \( \gamma \) is a random number between 0 and 1, and \( \alpha_j \) and \( \beta_j \) are dynamic boundaries.

The elite reverse learning strategy can effectively enhance the diversity of the population and, when there are many selectable solution variables, using fitness-based sorting significantly improves the search efficiency of the algorithm. Additionally, for each
generation of the whale population, the use of the elite reverse learning strategy can generate reverse solutions that are not near extremum points. This can help the algorithm escape local optima, strengthen its global optimization capabilities, and enhance algorithm stability. Lastly, incorporating dynamic boundaries can preserve the algorithm’s search experience, aiding convergence, and improving both local exploitation and global exploration capabilities. This speeds up the algorithm’s convergence rate.

2.2.3. Nonlinear Adaptive Weight Strategy

Taking into consideration the impact of the target prey, which is the optimal location, on the whale population during the hunting process in the original WOA (Whale Optimization Algorithm), a weight is introduced before the optimal location. This weight represents the degree to which the current whale individual inherits the previous generation’s optimal location [32]. The whale position updating formula is as follows:

\[ X_i^{T+1} = \omega \cdot X_i^T - A \times D_i, \quad p < 0.5 \]  
\[ X_i^{T+1} = \omega \cdot X_i^T - D_2 \cdot e^{2\omega \cdot \cos(2\pi M)}, \quad p \geq 0.5 \]  

The weight \( \omega \), which controls the variation in the whale individual positions, is adjusted by using the following calculation formula:

\[ \omega = \omega_1 + (\omega_2 - \omega_1) \cdot \frac{2}{\pi} \cdot \arccos \left( \frac{T}{T_{\text{MAX}}} \right) \]  

where \( \omega_1 \) is the initial weight value, \( \omega_2 \) is the final weight value, \( T \) represents the current iteration count, and \( T_{\text{MAX}} \) represents the maximum iteration count.

The Hybrid Whale Optimization Algorithm (HWOA) is built upon the foundation of the WOA but incorporates the use of Tent chaotic mapping to initialize positions within the search space. It employs MMOC (Multi-Modal Optimization Competition) as the objective function and introduces both elite reverse learning and nonlinear adaptive weight strategies. This not only accelerates the algorithm’s convergence speed but also enhances its capability to track global minimum modes. By using the Hybrid Whale Algorithm to search and update the positions of the whales, it aims to find the optimal solution, ultimately improving the global search capability and the computational accuracy of the algorithm.

2.3. HWOA-MVMD Algorithm

MVMD is a typical data decomposition algorithm, which has good generalization ability when dealing with sample data decomposition problems. In practice, the selection of eigenmode numbers \( k \) and penalty parameters \( a \) is often subjective, which will directly affect the generalization ability and decomposition accuracy of the algorithm, and the core idea of the algorithm is to construct a variational problem for solving the modality [33], so the Hybrid Whale Optimization Algorithm is used to optimize these two parameters to achieve the optimal decomposition effect.

The improved multivariate variational modal decomposition applies the HWOA to the parameter optimization of MVMD, allowing the algorithm to set parameters based on the characteristics of the signal itself and achieve the best decomposition effect. The flowchart of the proposed HWOA-MVMD algorithm is shown in Figure 1, and the detailed implementation steps are as follows:

(1) Set the whale population size, maximum number of iterations, optimization parameter optimization space, and initialization population position;
(2) Take the minimum modal overlap component as the fitness function, calculate the fitness value according to the position of the population, and save the current optimal value;

(3) Search for the optimal individual update search area, and update the position of the whale’s next iteration according to the fitness value level;

(4) Determine whether the termination conditions are met: if it is met, jump out of the loop to execute step (5); otherwise, re-execute steps (2)–(3);

(5) Save the result of the final global optimal solution \([k_0, \alpha_0]\); HWOA-MVMD will decompose the original data according to the parameter optimal solution to obtain different modal components.

**Figure 1.** Flow chart of HWOA-MVMD algorithm.

3. TPA-TCN Prediction Model

3.1. Analysis of Photovoltaic Power Generation Characteristics

Solar energy has the advantages of being renewable, pollution-free, and low-cost. There are significant differences in climate types across regions, which has a significant impact on the characteristics of photovoltaic power generation. Various factors have led to the uncontrollability of natural solar energy, and photovoltaic power generation has significant randomness and volatility. In order to improve the accuracy of photovoltaic power generation prediction, according to the correlation analysis between each influencing factor and photovoltaic power generation [34], the main influencing factors considered in this paper include total radiation (unit: W/m²), direct radiation, scattered radiation, ambient temperature, wind speed, humidity, and atmospheric pressure.

3.2. Time Convolutional Network

Time convolutional networks are effective models for predicting time series. Its core structure is shown in Figure 2, where it gradually increases the receptive field directly through a series of dilated convolutions, allowing the output to contain rich information. In causal dilated convolutions, capturing the overall characteristics of long-time sequences can be achieved by adjusting parameters such as the convolutional kernel, convolutional layers, and dilation factor ‘d’, thereby further deepening the impact of the sequence on the deep network. In the causal dilation network, the part connected by the arc is the residual module, which can solve the problem of information loss that may occur when the number of TCN layers deepens.
The amount of electricity generated in time \( t \in \{1,2,3,\ldots,k\} \) is recorded as \( \mathbf{X}_t \),
\[
\mathbf{X}(t) = [x_1(t), x_2(t), x_3(t), \ldots, x_l(t)]
\]

For an input one-dimensional sequence \( x_0(t) \), the specific output expression for feature extraction through the TCN is
\[
F(s) = (x^* f)(s) = \sum_{i=0}^{k-1} x_{s-dsi}
\]
where \( d \) is the expansion coefficient, \( * \) is a convolution operation for extracting feature information, \( f : \{0, \cdots, k-1\} \) is the convolutional kernel, and \( k \) is the size of the convolutional kernel.

By combining the multivariate feature \( (x_1(t), x_2(t), x_3(t)) \) with the subcomponent sequence \( x_0(t) \) to form a multivariate time series \( \mathbf{X}(t) \) and extracting features from the multidimensional time series, the following can be concluded:
\[
h^n_t = F(h^n_{t-1}, \mathbf{X}_t)
\]
where \( n = \{1,2,\ldots,m\} \) and \( m \) is the dimension of the multivariate time series dataset.

3.3. Temporal Convolutional Network Model Based on Time Pattern Attention

In the actual prediction of photovoltaic power generation, the important data not only include the historical total radiant intensity of the sun but also include the temperature, humidity, atmospheric pressure, and other multivariate variables. Each variable has a different impact on photovoltaic power generation, and the degree of influence varies as well. Therefore, in order to improve the accuracy of photovoltaic power generation prediction and solve the complex, dynamic, and interdependent relationship between variables, in this study, multivariate variables and time series are combined to form a multivariate time series, and the time pattern attention (TPA) is introduced into the TCN model [35], which can capture the impact of each variable on the prediction sequence in the prediction model and effectively improve the prediction accuracy. Figure 3 shows the TPA-TCN prediction model.
Figure 3. TPA-TCN prediction model.

(1) TCN layer
The output of the TCN layer is the hidden state of each time step, and the hidden information $h_t \in \mathbb{R}^m$ at time $t$ is output. It can be represented by Equation (7).

(2) Time mode capture layer
Using the TCN’s convolutional kernel on the row vector of feature matrix $H = \{H_1, H_2, \cdots, H_{i-1}\}$ for feature extraction, the output is

$$H_{i,j}^C = \sum_{l=1}^{W} H_{i,l-1} \cdot C_{j,l}$$

where $H_{i,j}^C$ is the convolution output value of the $i$-th row vector and the $j$-th convolution kernel, $C$ represents all convolutional kernels, and $l$ is the length of the time series.

(3) Time Mode Attention Layer
Let $f$ be the evaluation correlation function, and select the activation function output by the weight of the sigmoid function; then,

$$f(H_t^C, h_t) = (H_t^C)^T W_h h_t$$

Then, the weight $\alpha_i$ is

$$\alpha_i = \text{sigmoid}(f(H_t^C, h_t))$$

Calculate $v_t$ by weighting $H^C$:

$$v_t = \sum_{i=1}^{l} H_t^C \alpha_i$$

Integrate $h_t$ and $v_t$, with weight coefficient matrix $W$, to obtain

$$h_t' = W_h h_t + W_v v_t$$

$$y_{t+1} = W_h' h_t'$$
4. Short-Term Photovoltaic Power Generation Prediction Model Based on HWOA-MVMD-TPA-TCN

4.1. Prediction Model Process

Based on the modelling process mentioned above, the prediction process of the short-term photovoltaic power generation prediction model based on the HWOA-MVMD-TPA-TCN can be obtained, as shown in Figure 4. The specific prediction steps of the combined model are as follows:

(1) Utilize the HWOA to optimize the modal number $k$ and penalty factor $\alpha$ in MVMD, with the MMOC as the objective function, to obtain and save the optimal solution.

(2) According to the optimal solution in (1), the original photovoltaic power generation sequence is decomposed using the HWOA-optimized MAMD to obtain $k$ different IMF components and a residual component $r_1$.

(3) Each component of the original data after decomposition, $k$ different IMF components, and the key influencing factor variable data are combined into a training set and input into the TPA-TCN model to obtain $k$ different $\text{IMF}'$ components.

(4) The superposition and reconstruction of the predicted values of $k$ different $\text{IMF}'$ components are performed to obtain the final predicted value of photovoltaic power generation.

(5) The predicted photovoltaic power generation value is compared with the actual photovoltaic power generation data, and the error is analyzed.

![Figure 4](image_url) Specific prediction principle diagram of the combination prediction model.

4.2. Data Preprocessing

4.2.1. Normalization

First, all the historical data in this paper are normalized, that is, all data are mapped to $[-1, 1]$: \[
x = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{14}
\]

where $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the input data, respectively.
4.2.2. Selection of Evaluation Indicators for Prediction Error

To ensure that the accuracy of the overall model can be effectively evaluated, two evaluation indicators are selected to evaluate the prediction model of the time series, namely, the mean absolute error (MAE) and root-mean-square error (RMSE) [22]. The error calculation formulas are

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |Y_i - \tilde{Y}_i| \tag{15}
\]

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Y_i - \tilde{Y}_i)^2} \tag{16}
\]

where \( n \) represents the total predicted amount, while \( Y_i \) and \( \tilde{Y}_i \) represent the true and predicted photovoltaic power generation values of the input data, respectively.

It can be seen from the above that MAE and RMSE are inversely proportional to the final evaluation effect, and the smaller the number of MAE and RMSE, the smaller the gap between it and the actual data, and the better the evaluation effect.

4.2.3. Preprocessing of the Photovoltaic Power Generation Sequence

To improve the final prediction accuracy of photovoltaic power generation and weaken the influence of nonstationary features on the prediction, an improved MVMD algorithm is used to decompose and preprocess the photovoltaic power generation sequence obtained from data feature extraction. During the process of optimizing the parameters of MVMD using the HWOA, when the algorithm performs the 38th iteration, the fitness value is the smallest, that is, the minimum modal overlap component is 2.111, as shown in Figure 5. At this time, the optimal solution is obtained by the modal number and penalty factor \( \alpha \), where \( k \) is 5 and \( \alpha \) is 3249. Therefore, the original data of photovoltaic power generation will be decomposed into five subsequences, and the IMF components obtained from the decomposition are shown in Figure 6.

![Figure 5. Convergence iteration diagram.](image_url)
5. Example Analysis

We select, as an example, the actual photovoltaic power generation data of a business center in Tianjin that is located at east longitude 117°2’ and north latitude 39°13’ from 23 June to 2 August 2021. Tianjin is located in northern China, and the cloud cover during the summer months (June to July) is typically influenced by seasonal weather patterns. The climate in Tianjin is generally hot with high humidity, leading to frequent cloud formation, especially in the afternoon and evening. During this season, various types of clouds may appear, including cumulus clouds and stratocumulus clouds. Afternoon convective heating often results in the formation of cumulonimbus clouds, leading to possible increased rainfall and thunderstorm activity with dense cloud cover. The cloud cover throughout the day can vary significantly. For instance, the morning may be relatively clear, but as temperatures rise in the afternoon, cloud cover may increase. For a more precise description, local meteorological data, including cloud cover records, sunshine hours, and precipitation, can be referenced. Among them, meteorological data include total radiation (unit: W/m²), direct radiation, scattered radiation, ambient temperature, humidity, wind speed, and atmospheric pressure characteristics; the sampling interval between every two groups of data is 1 h; and the number of data samples obtained for each feature is 432. The first 90% of the data are used as the training set, and the last 10% of the data are used as the test set [27].

5.1. Parameter Settings for the Prediction Models

After decomposing and normalizing the original sequence, for different reconstructed components, the data batch size b is adjusted based on their different characteristics. Repeated experiments are conducted on hyperparameters such as the size, learning rate lr, random inactivation ratio dropout, and residual block expansion coefficient d, as shown in Table 3. The settings of other training parameters (such as the convolution kernel
size $k_{size}$, optimizer, activation function, and training epoch) are shown in Table 4. The TPA-TCN model is built using the Python language, and a deep learning network model is constructed using the TensorFlow framework and Keras. The GridSearchCV method in the sklearn library is called to perform a grid search on the hyperparameters and find the optimal parameter settings.

Table 3. Training parameter setting 1.

<table>
<thead>
<tr>
<th>Weight</th>
<th>b_Size</th>
<th>Ir</th>
<th>Dropout</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMF$_1$</td>
<td>32, 64, 128</td>
<td>0.001</td>
<td>0.01</td>
<td>3, 4, 5, 6, 7</td>
</tr>
<tr>
<td>IMF$_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| IMF$_3$  | 32, 64, 128 | 0.001 | 0.00001 | 3, 4, 5, 6, 7 |
| IMF$_4$  |        |      |         |           |
| IMF$_5$  |        |      |         |           |
| $r_i$    |        |      |         |           |

Table 4. Training parameter setting 2.

<table>
<thead>
<tr>
<th>Training Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>k_size</td>
<td>3</td>
</tr>
<tr>
<td>Optimizer</td>
<td>Adam</td>
</tr>
<tr>
<td>Activation</td>
<td>Sigmoid</td>
</tr>
<tr>
<td>Epochs</td>
<td>80</td>
</tr>
</tbody>
</table>

Through repeated experiments, it was found that a higher dilation factor can yield better predictive results for high-frequency reconstruction components. Consistent with the analysis in the previous section, a higher dilation factor allows for a larger convolutional receptive field, thus more accurately capturing the overall sequence characteristics of high-frequency components while ignoring local trends. In other words, a larger dilation factor ‘d’ is more suitable for predicting high-frequency components.

5.2. Photovoltaic Power Generation Prediction Based on TAP-TCN

The IMF components obtained through HWOA-MVMD, along with the seven key influencing factors, serve as inputs to the TPA-TCN model. In order to more comprehensively uncover the patterns of photovoltaic power generation and effectively enhance the prediction accuracy, predictions were separately conducted for five sub-sequences. The predictive results for each component are shown in Figure 7.

Through the prediction results of each component in Figure 7, it can be seen that the five subseries are basically consistent with their corresponding prediction results, which proves that the photovoltaic power generation prediction performance based on the TAP-TCN is better.
5.3. Comparative Analysis of Prediction Methods

To validate the effectiveness of the HWOA-MVMD data decomposition method, this paper compares the proposed method with HWOA-MVMD-BiLSTM (Method One). Additionally, to demonstrate the effectiveness of the HWOA, the proposed method is simulated in comparison with the GWO-MVMD-TPA-TCN (Method Two) and TPA-TCN (Method Three). To increase the persuasiveness of the prediction effect, the photovoltaic power generation for five consecutive days from 29 July to 2 August is predicted. The line chart and bar chart are used to display the forecast results more intuitively, and the prediction results of each method are shown in Figure 8, of which Figure 8a,b show the corresponding photovoltaic power generation prediction results of each method on 29 July; Figure 8c,d show the prediction results of each method when the photovoltaic power generation is changed from a large fluctuation series (30 July) to a more stable series (31 July); and Figure 8e,f show the prediction results of each method when there are two successive large fluctuation sequences of photovoltaic power generation. The detailed prediction error results for each method and the average error results over the past five days are shown in Table 5.
The method proposed in this article

HWOA-MVMD-BiLSTM (Method 1)
GWO-MVMD-TPA-TCN (Method 2)
TPA-TCN (Method 3)
Figure 8. Prediction result curves of various models from 29 July to 2 August. (a) Curve of predicted results by various methods on 29 July; (b) curve of predicted results by various methods on
Analyzing the predictive results in Figure 8, it can be observed that Method One, Method Two, and Method Three to some extent capture the overall trend of power generation. However, the fitting of their prediction curves to the actual curve is poor, with significant local deviations. Notably, Method Three exhibits the poorest predictive performance, indicating that data decomposition can enhance forecasting accuracy. In contrast, the proposed method shows a high degree of alignment with the actual values, minimal volatility, and a high degree of curve similarity. This suggests that the HWOA-MVMD-TPA-TCN model outperforms the others in terms of performance.

The predictive results of Method One, Method Two, and Method Three all exhibit noticeable local deviations. In contrast, the results of the method proposed in this paper show only brief, slight fluctuations while still effectively capturing the trend in photovoltaic power generation. This indicates superior forecasting performance.

Table 5. Prediction errors of various methods.

<table>
<thead>
<tr>
<th>Time</th>
<th>Standard/MW</th>
<th>Proposed Method</th>
<th>HWOA-MVMD-BiLSTM (Method 1)</th>
<th>GWO-MVMD-TPA-TCN (Method 2)</th>
<th>TPA-TCN (Method 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>29 July</td>
<td>MAE</td>
<td>3.14</td>
<td>4.12</td>
<td>5.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>4.21</td>
<td>5.83</td>
<td>6.83</td>
<td></td>
</tr>
<tr>
<td>30 July</td>
<td>MAE</td>
<td>4.75</td>
<td>6.01</td>
<td>11.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>6.76</td>
<td>7.18</td>
<td>10.58</td>
<td></td>
</tr>
<tr>
<td>31 July</td>
<td>MAE</td>
<td>1.69</td>
<td>3.68</td>
<td>5.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>2.09</td>
<td>4.53</td>
<td>7.12</td>
<td></td>
</tr>
<tr>
<td>1 August</td>
<td>MAE</td>
<td>3.32</td>
<td>5.08</td>
<td>6.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>4.50</td>
<td>6.18</td>
<td>7.62</td>
<td></td>
</tr>
<tr>
<td>2 August</td>
<td>MAE</td>
<td>2.58</td>
<td>4.13</td>
<td>5.51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>3.56</td>
<td>5.67</td>
<td>6.79</td>
<td></td>
</tr>
<tr>
<td>average error</td>
<td>MAE</td>
<td>3.10</td>
<td>4.60</td>
<td>6.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>4.17</td>
<td>5.88</td>
<td>7.79</td>
<td></td>
</tr>
</tbody>
</table>

Based on Figure 8 and Table 5, Method One and Method Two proposed in this paper decompose the data first, while Method Three does not decompose the data, and the prediction error of Method Three is the largest, indicating that decomposition of the original data is conducive to power generation prediction. In addition, compared with Methods Two and Three, the method proposed in this paper shows that after HWOA-MVMD, the change trend of short-term photovoltaic power generation can be better predicted as a whole, and the prediction error is significantly reduced, among which the average MAE value is 1.95 MW and the RMSE value is 2.55 MW, which can be reduced by 33.74% and 38.85%, respectively. In addition, under the condition that the actual data of photovoltaic power generation fluctuate greatly (29 July–2 August), the prediction error of the method in this paper remains basically stable, which verifies the effectiveness of the method proposed in this paper.

6. Conclusions

This article proposes a short-term photovoltaic power generation prediction model based on the HWOA-MVMD-TPA-TCN model, which improves the prediction accuracy for photovoltaic power generation and draws the following conclusions:

(1) Using the actual data example analysis of a business center in Tianjin, the HWOA-MVMD-TPA-TCN model effectively reduces the prediction error of photovoltaic power generation, in which the MAE value is 1.95 MW and the RMSE value is 2.55
MW, which can be reduced by up to 33.74% and 38.85%, respectively. The effectiveness of the forecasting model proposed in this paper is proved, and the model has certain reference value in the field of time series forecasting.

(2) The IMF components resulting from the decomposition of photovoltaic power generation possess distinct characteristic changes. Exploring more efficient prediction methods for these components requires further research. Additionally, analyzing error components and improving prediction accuracy through error correction should be considered. Future research can focus on these aspects to advance photovoltaic power generation prediction.

**Author Contributions:** Conceptualization, R.C., H.T. and D.L.; methodology, R.C.; software, R.C.; validation, R.C.; formal analysis, R.C.; investigation, R.C.; resources, M.F.; data curation, H.F.; writing—original draft preparation, R.C.; writing—review and editing, R.C., H.T. and D.L.; visualization, R.C.; supervision, H.T. and D.L.; project administration, R.C., H.T. and D.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by State Grid Tianjin Electric Power Company Science and Technology Project, grant number KJ21-1-21; Tianjin Postgraduate Scientific Research Innovation Project, grant number 2022SKYZ070; and Tianjin University of Technology 2022 school-level postgraduate scientific research innovation practice project, grant number YJ2209.

**Data Availability Statement:** The data involve project confidentiality and cannot be disclosed.

**Conflicts of Interest:** The corresponding author represents all authors and declares that there are no conflict of interest.

**References**


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.