Hybrid Photovoltaic Output Forecasting Model with Temporal Convolutional Network Using Maximal Information Coefficient and White Shark Optimizer

Xilong Lin [1], Yisen Niu [1], Zixuan Yan [1], Lianglin Zou [1], Ping Tang [1] and Jifeng Song [2],*

[1] School of New Energy, North China Electric Power University, Beijing 102206, China
[2] Institute of Energy Power Innovation, North China Electric Power University, Beijing 102206, China
* Correspondence: solarsjf@163.com

Abstract: Accurate forecasting of PV power not only enhances the utilization of solar energy but also assists power system operators in planning and executing efficient power management. The Temporal Convolutional Network (TCN) is utilized for feature extraction from the data, while the White Shark Optimization (WSO) algorithm optimizes the TCN parameters. Given the extensive dataset and the complex variables influencing PV output in this study, the maximal information coefficient (MIC) method is employed. Initially, mutual information values are computed for the base data, and less significant variables are eliminated. Subsequently, the refined data are fed into the TCN, which is fine-tuned using WSO. Finally, the model outputs the prediction results. For testing, one year of data from a dual-axis tracking PV system is used, and the robustness of the model is further confirmed using data from single-axis and stationary PV systems. The findings demonstrate that the MIC-WSO-TCN model outperforms several benchmark models in terms of accuracy and reliability for predicting PV power.

Keywords: photovoltaic power forecasting; maximal information coefficient; temporal convolutional network; white shark optimizer

1. Introduction

The progress of human society has increased the energy demand. But fossil fuel reserves are limited and non-renewable. Therefore, more and more countries are starting to support the growth of renewable energy technology for generating power [1]. As energy demand rises in developing nations, the installed capacity of photovoltaic (PV) systems is also increasing annually [2], and Building Integrated Photovoltaics (BIPVs) have been developed, which are important for reducing energy consumption and improving thermal quality [3]. Due to factors such as solar radiation and ambient temperature, the output power of PV systems exhibits intermittency, volatility, and randomness, which brings great uncertainty to the operation, scheduling, and planning of the power system [4]. Forecasting photovoltaic power is considered one of the most economically feasible solutions for managing solar intermittency. Precise photovoltaic power forecasting significantly enhances the efficiency of solar energy utilization, thus increasing the revenue of the power plant and reducing the economic loss caused by power limitation [5].

In current studies, photovoltaic power prediction typically falls into three main categories: physical models, statistical approaches, and machine learning methods [6]. Instead of requiring historical data, physical prediction methods rely on accurate meteorological information, power plant geographic information, and PV module information [7]. Meteorological information generally comes from the following three sources: numerical weather prediction (NWP), satellite cloud images, and sky images. The obtained meteorological parameters are combined with parameters such as module mounting angle, PV array conversion efficiency, and battery status to build a physical model, which in

turn directly calculates the power generation. Physical modeling of PV cells requires a large number of circuit parameters, which greatly affects the accuracy of the models [8]. Physical methods, while offering excellent long-term forecasting capabilities, have some shortcomings. Satellite cloud images suffer from low spatial and temporal resolution, which makes it difficult to capture meteorological information at small scales [9]. A full-sky imager is unable to provide a wide range of cloud coverage information [10]. Although the coverage can be expanded by means of arrays, it requires high hardware costs and complex communication technologies.

Statistical methods do not require much information about the PV system compared to physical models [11]. Statistical approaches offer the advantages of straightforward modeling and suitability across various regions. The application of statistical methods in PV power prediction also has certain challenges. Collection and calculation of accurate data in actual implementation remains challenging [12]. Statistical models have high requirements for accurate historical data, and their relatively low computational speed and computational volume make it difficult to meet the requirements of short-term PV power prediction [13].

In recent years, researchers have shown significant interest in machine learning, and machine learning models can extract nonlinear features from photovoltaic (PV) power generation data to improve prediction accuracy [14]. An artificial neural network is employed for short-term solar radiation prediction [15]; as a classical algorithm, it is also used in the fault detection of lines in the power grid [16–18]. In Ref. [19], the outcomes of the proposed Extreme Learning Machine (ELM) are contrasted with those of conventional models, demonstrating the capability of the proposed model to predict short-term wind speeds. In Ref. [20], support vector machine (SVM) is employed to forecast the output from a photovoltaic power station. However, these machine learning models encounter challenges due to the pronounced volatility and nonlinearity of photovoltaic power generation time series; conventional machine learning models may struggle to capture the intricate nonlinear and dynamic nature of photovoltaic power generation data [21]. In Ref. [22], various machine learning algorithms are used for PV power prediction, including ensemble of regression trees, support vector machine, Gaussian process regression, and artificial neural networks. Therefore, some scholars have begun to shift their attention to deep learning (DL) models because they have sufficient feature extraction and feature transformation capabilities [23]. Recently, common deep learning models used in photovoltaic power prediction include convolutional neural networks (CNNs) and recurrent neural networks (RNNs). The widely used long short-term memory network (LSTM) is an improvement over RNN, which solves the problem of RNN gradient disappearance. In Ref. [24], a bidirectional long short-term memory network with Bayesian optimization is used to predict solar photovoltaic power generation. In Ref. [25], an algorithm based on LSTM is proposed for predicting solar radiation, outperforming the persistence algorithm, linear least squares regression, and other algorithms in terms of prediction accuracy. In Ref. [26], a hybrid model employing CNN and support vector regression (SVR) is proposed to enhance the accuracy of solar radiation prediction. In Ref. [27], it has been demonstrated that integrating CNN and LSTM to predict PV power shows superior performance compared to using either model individually.

The Temporal Convolutional Network (TCN) structure outperforms conventional recurrent structures like LSTM and GRU across various sequence modeling tasks. Hence, we utilize TCN for addressing the PV power prediction task [28]. TCN is a new convolution architecture specially designed for sequence modeling. While maintaining the convolution operation characteristic of CNN, it incorporates dilated causal convolution and residual connections, enhancing its performance in handling time series data [29]. TCN has recently been applied to tackle various complex prediction tasks. For example, Reference [30] applies TCN to short-term wind power prediction and obtains high prediction accuracy even in the case of large fluctuations in wind power. Reference [31] uses TCN to predict the ship's motion attitude.

Constructing a deep learning model involves numerous hyperparameters, making it challenging to establish a model with both strong robustness and accurate prediction capabilities. The traditional enumeration method and grid search method have the disadvantages of low efficiency and a large amount of calculation. Therefore, some researchers began to pay attention to meta-heuristic algorithms. Many meta-heuristic algorithms have been demonstrated to enhance prediction accuracy. In Ref. [32], Particle Swarm Optimization (PSO) is used to optimize the proposed adaptive network based fuzzy inference system. Reference [33] utilizes a Genetic Algorithm (GA) to optimize the hyperparameters of LSTM for forecasting PV power generation four hours ahead. Ref. [34] proposes an SCA-BILSTM architecture for hourly solar radiation forecasting. An SSA-RNN-LSTM architecture for predicting PV power output one hour in advance is proposed in [35]. In order to better tune and optimize the model, in this study, a new meta-heuristic algorithm known as White Shark Optimizer (WSO) is used. Compared to many previous meta-heuristics, WSO performs better in global optimality and avoiding local minima [36]. In [37], the WSO algorithm is employed to optimize the design parameters of proton exchange membrane fuel cell, which improves its performance in practical applications. Compared with SSA, HHO, DBO, ASO, and other algorithms, its effect is better.

In predicting PV power, potential colinearity between explanatory variables may lead to feature redundancy, which in turn may adversely affect the performance of the prediction model. Moreover, explanatory variables that lack a strong correlation with the output power might also detrimentally influence performance. Given the intricate nature of PV output power, complex nonlinear or nonfunctional relationships between variables may exist. The maximal information coefficient (MIC) exhibits stronger robustness and fairness compared to the traditional correlation coefficient. MIC serves to detect both linear and nonlinear relationships in extensive datasets, while also revealing possible non-functional correlations [38]. Currently, MIC is successfully used in various fields [39,40].

According to the literature review, the research gaps in photovoltaic power prediction that are addressed in this study are as follows: (1) This study investigates the power prediction of three different photovoltaic systems, namely dual-axis tracking photovoltaic systems, single-axis tracking photovoltaic systems, and fixed photovoltaic systems. (2) The maximal information coefficient (MIC) is employed for feature vector selection. (3) The application of Temporal Convolutional Networks (TCNs) in photovoltaic power prediction is still rare, especially for the three different photovoltaic systems. (4) This paper is the first to use the White Shark Optimizer (WSO) algorithm to adjust the hyperparameters of the TCN to improve the accuracy of photovoltaic power output prediction.

To improve the accuracy of PV power prediction in this study, we propose a novel hybrid approach of Temporal Convolutional Networks (TCNs), maximal information coefficients (MICs), and the White Shark Optimizer (WSO). Among them, the MIC is used to deal with the complex relationships of the variables in the dataset in this study, the data are processed and then fed into the TCN, and the WSO algorithm adjusts and optimizes the structure and hyperparameters of the model during the training process to further enhance its prediction capabilities. The primary contributions of this paper are as follows:

(1) A novel hybrid model (MIC-WSO-TCN) for photovoltaic power prediction is proposed, capable of achieving accurate prediction results across different seasons.
(2) The prediction accuracy of the proposed model is compared with MIC-TCN, MIC-WSO-BP, and MIC-WSO-LSTM.
(3) The performance of the proposed model in predicting photovoltaic output is evaluated across various seasons using a dual-axis tracking photovoltaic system. In addition, the robustness of the model is verified on single-axis and fixed photovoltaic systems, and its prediction accuracy is evaluated using real power generation data.

2. Methodology

This section introduces three different photovoltaic systems, summarizes the data preprocessing steps, describes the proposed deep learning model, and explains the indicators

used to evaluate the performance of the model. In addition, Figure 1 shows a schematic diagram of the research conducted in this paper.
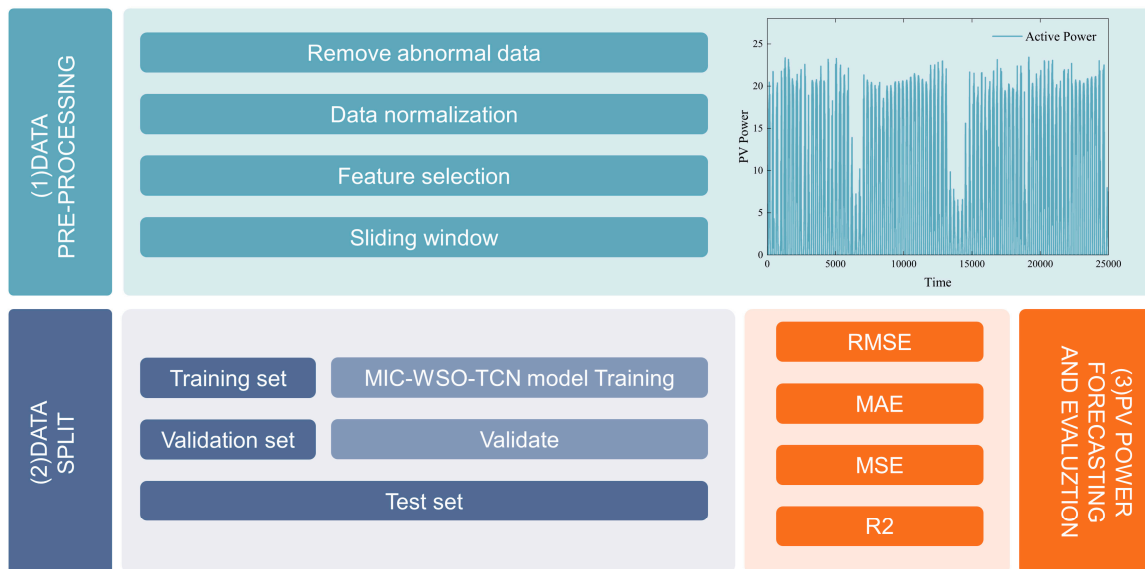

Figure 1. The diagram of this study.

2.1. Overview of PV Systems

The three photovoltaic systems are all located in Alice Springs, Australia, with a latitude of 23.7618° S and a longitude of 133.8748° E. Figure 2 shows an overview of the three systems, which are the dual-axis tracking photovoltaic system (1B), the single-axis tracking photovoltaic system (5), and the fixed photovoltaic system (11). The power generation data of the photovoltaic system can be downloaded in Ref. [41].
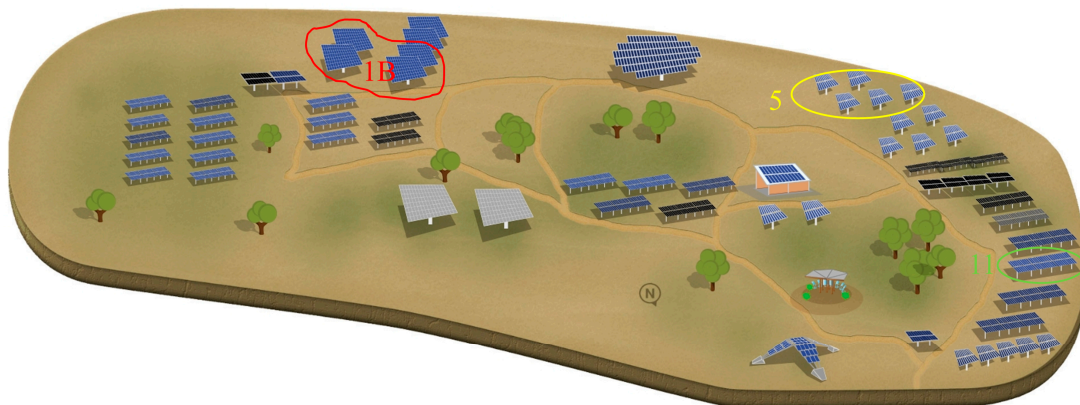

Figure 2. Location distribution of PV power plants.

2.2. Data Preprocessing and Data Split

Data preprocessing includes data segmentation, data standardization, abnormal data processing, and feature selection. Data preprocessing can improve the convergence speed of the model and remove the influence of dimension. The outliers in the dataset are removed, and thresholds are set based on the installed generating capacity and then filtered. Missing data can affect the accuracy of the model, especially when predicting PV power generation data that requires continuous measurements. To address this issue, we utilize widely adopted cubic spline interpolation for handling missing values. In this experiment, the data from December 2013 to December 2014 are selected. These data are sampled every 5 min. The data of this year are divided into four seasons for the experiment. Since the

photovoltaic system studied in this paper is located in Australia, the season is opposite to the northern hemisphere. The data for each season are partitioned into training, testing, and validation sets, comprising 80%, 10%, and 10% of the total data, respectively.

Due to the fluctuating nature of PV power, it is categorized as time series data. For forecasting using deep learning, it is crucial to reformat the dataset into a supervised regression framework. This involves structuring the dataset so that both input features and their corresponding outputs are explicitly identified [42]. The sliding window technique effectively addresses this requirement by dividing the dataset to create sequences used as model inputs, with a defined number of sequences designated as outputs. In this study, a single-step prediction model is implemented to forecast PV power generation, where data from intervals 0 to $t-1$ serve as inputs, and the data point at time t is used as the output. The dataset is continuously partitioned by shifting the window one time step forward to create new input and output pairs.

Because the data include many features, different features have different dimensions. In order to eliminate the influence of dimension, this experiment adopts data standardization. The calculation method is as follows:

$$u = \frac{1}{N} \sum_{n=1}^{N} X \tag{1}$$

$$\sigma = std(X) = \sqrt{\frac{\sum (x_i - u)^2}{N}} \tag{2}$$

$$X_{std} = (X - u)/\sigma \tag{3}$$

Here, X represents the actual data, N is the size of the data, u denotes the average value, x_i represents each value of the data, and σ is the standard deviation of the data.

Feature selection is a common feature engineering technology in deep learning and data mining. When training a model, the input variables are called features, and training a model with too many useless features can lead to longer training times and reduce the predictive power of the model and complicate it [43]. Therefore, in order to make the model obtain better predictive ability, it should be ensured that only the most representative and relevant features are retained, while irrelevant and redundant features are excluded.

The most commonly used method to capture the correlation between data is by calculating the correlation coefficient. This coefficient is sensitive to linear relationships between variables, but when there is a nonlinear relationship between the data, it may produce inaccurate results [44]. The maximal information coefficient (MIC) offers a balanced approach to capturing all functional relationships. The underlying idea involves calculating the mutual information between two variables by analyzing their approximate probability density distribution within a grid following any form of meshing of their scatter plots, thereby revealing any correlation between the variables. The normalized mutual information (MI) serves as a measure to quantify the correlation between two variables. It is calculated based on a set of sample pairs associated with the two variables, represented by X and Y, $D = \{(x_i, y_i), i = 1, \ldots, N\}$, where N is the number of samples. The calculation process of normalized mutual information is as follows:

Step 1: Initially, the sample space is partitioned into $m \times n$ grids, denoted as G. Subsequently, the empirical marginal probability densities $p(x)$ and $p(y)$ of X and Y, as well as the joint probability density $p(x, y)$, are estimated [38]. The calculation methodology for MI is as follows:

$$\text{MI}(X, Y | D, G) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log_2 \left(\frac{p(x, y)}{p(x)p(y)} \right) \tag{4}$$

The division of data is diverse. In all possible grids, the maximum MI can be expressed as follows:

$$\mathrm{MI}^*(D, m, n) = \max_G \mathrm{MI}(X, Y | D, G) \tag{5}$$

Step 2: In order to facilitate comparison and analysis, the maximum MI value is normalized, and the normalized value is in the interval of [0, 1].

$$\mathrm{NMI}^*(D, m, n) = \frac{\mathrm{MI}^*(D, m, n)}{\log \min\{m, n\}} \tag{6}$$

Step 3: According to the previous steps, calculate all $\mathrm{NMI}^*(D, m, n)$ that satisfy the condition, this condition is $m \times n < k(N)$, and MIC is the maximum $\mathrm{NMI}^*(D, m, n)$ value in all grids. This process can be expressed by the following formula:

$$\mathrm{MIC}(X, Y) = \max_{m \times n < k(N)}\{\mathrm{NMI}^*(D, m, n)\} \tag{7}$$

Among them, $k(N)$ is defined as a function of the number of samples. When $k(N) = N^{0.6}$, the algorithm works well in practice [38].

In the context of MIC, the value ranges between 0 and 1. High correlation between two variables results in high mutual information values, indicating stronger correlation. Consequently, higher MIC values imply stronger correlation. Conversely, when there is no relationship between the two variables, the value of MIC is 0. The dataset of this paper contains a large amount of meteorological data and radiation data. After calculation and analysis, it is most appropriate to discard the characteristics of MIC less than 0.2. After screening, the characteristics of the input model include global horizontal radiation, diffuse horizontal radiation, wind speed, temperature, radiation global tilted, relative humidity, and radiation diffuse tilted.

2.3. White Shark Optimizer

The WSO is a new meta-heuristic algorithm for biologically inspired global optimization problems. The algorithm is based on the overall situation of white shark predation behavior and the way they track prey. Compared with other existing meta-heuristic methods, the WSO algorithm demonstrates a viable solution in terms of global optimality, evasion of local minima, and overall solution quality. The inspiration of this optimization algorithm comes from three predatory behaviors. The mathematical models of these behaviors are shown below [36].

2.3.1. The Initialization Process of the WSO

In the optimization problem addressed by the WSO algorithm, a set of random initial solutions is generated, with each solution representing the position of a white shark. If there are n white sharks in the population, their positions can be captured by a matrix, effectively modeling the candidate solutions, as shown below:

$$w = \begin{bmatrix} w_1^1 & w_2^1 & \cdots & \cdots & w_d^1 \\ w_1^2 & w_2^2 & \cdots & \cdots & w_d^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_1^n & w_2^n & \cdots & \cdots & w_d^n \end{bmatrix} \tag{8}$$

where w represents the position of the white shark in the given search area, w_d^i specifies the position of the ith white shark in the dth dimension space, and d is the sum of decision variables relevant to the issue at hand. The initial population is established based on the subsequent equation [36]:

$$w_j^i = l_j + r \times (u_j - l_j) \tag{9}$$

Among them, w_j^i is the original vector of the ith white shark in the *j*th, while u_j and l_j denote the maximum and minimum limits of the *j*th dimensional search field. *r* represents

a randomly chosen number within the interval [0, 1]. The fitness function is used to assess the quality of each new alternative candidate solution for white sharks at each new location. If a white shark's current position is superior to its new position, it will remain at its current location. Conversely, if the new position offers an improvement over the current one, the shark's location will be updated accordingly [36].

2.3.2. Speed of Movement to Prey

Because white sharks are a kind of creature with strong desire for survival, they hunt and track their prey most of the time. They usually use auditory, visual, and olfactory senses to track prey. When the white shark detects the presence of its prey through the disturbance created by the movement of the prey, it will move to the prey, and this process can be expressed in Equation (10).

$$v_{k+1}^i = u\left[v_k^i + p_1\left(w_{gbestk} - w_k^i\right) \times c_1 + p_2\left(w_{best}^{v_k^i} - w_k^i\right) \times c_2\right] \tag{10}$$

For the ith shark, v_{k+1}^i represents the updated velocity vector for the ith white shark at step (k+1). w_{gbestk} indicates the globally optimal position vector identified in the kth iteration of any white shark. w_k^i signifies the present position vector of the ith white shark at the kth step, $w_{best}^{v_k^i}$ refers to the optimal discovery position vector for the population marker by the ith shark, and v_i signifies the exponential vector for those white sharks occupying a higher position, as defined in Equation (11). p_1 and p_2 represent the force of white sharks, which regulate the impacts of w_{gbestk} and $w_{best}^{v_k^i}$ on w_k^i, and are calculated as Equations (12) and (13), and u symbolizes the contraction factor instrumental in examining the convergence patterns of great white sharks within the WSO, and its definition is shown in Equation (14) [36].

$$v = [n \times rand(1, n)] + 1 \tag{11}$$

where rand(1, n) denotes an evenly distributed random vector ranging between 0 and 1.

$$p_1 = p_{\max} + (p_{\max} - p_{\min}) \times e^{-(4k/K)^2} \tag{12}$$

$$p_2 = p_{\min} + (p_{\max} - p_{\min}) \times e^{-(4k/K)^2} \tag{13}$$

where k denotes the number of current iterations, and K signifies the maximum iteration limit. $p_{\min}$ and $p_{\max}$ correspond to the beginning and modified velocities of the white shark, which are essential for optimal movement. After analysis, it was found that the values of $p_{\min}$ and $p_{\max}$ were 0.5 and 1.5, respectively [36].

$$\mu = \frac{2}{\left|2 - \tau - \sqrt{\tau^2 - 4\tau}\right|} \tag{14}$$

where τ represents the acceleration coefficient, set at 4.125, a value determined through comprehensive analysis [36].

2.3.3. Advance towards the Ideal Prey

Great white sharks predominantly search for potential prey to secure the best food sources. Therefore, their positions are constantly changing. They typically move towards prey upon detecting wave sounds caused by prey movement or sensing their scent. Occasionally, prey might move from its original spot, either due to the shark's movement or in search of food. Usually, the prey will leave a smell when it leaves, so the white shark can use this clue. In such instances, white sharks might randomly search for prey, like a school

of fish feeding. Under these circumstances, we apply the position update strategy outlined in Equation (15) to model the white shark's movement towards prey [36].

$$w_{k+1}^i = \begin{cases} w_k^i \cdot \neg \oplus w_o + u \cdot a + l \cdot b; & rand < mv \\ w_k^i + v_k^i/f; & rand \geq mv \end{cases} \tag{15}$$

where w_{k+1}^i denotes the updated location vector of the *i*th white shark in the (*k*+1)th iteration, *a* and *b* are binary vectors, with their definitions provided in Equations (16) and (17). *l* and *u* indicate the minimum and maximum limits of the search space. w_o is the logic vector, as shown in Equation (18). *f* represents the wave motion frequency utilized by the white shark, as described in Equation (19). *rand* is a randomly produced number in the interval 0 to 1.

$$a = \text{sgn}(w_k^i - u) > 0 \tag{16}$$

$$b = \text{sgn}(w_k^i - l) < 0 \tag{17}$$

$$w_o = \oplus(a, b) \tag{18}$$

Equations (16) and (17) enable the white shark to comprehensively investigate every possible region within the search area.

$$f = f_{\min} + \frac{f_{\max} - f_{\min}}{f_{\max} + f_{\min}} \tag{19}$$

Here, $f_{\min}$ and $f_{\max}$ represent the minimum and maximum frequencies of wave motion. Through accurate analysis and testing of various problems, it is found that when the values of $f_{\min}$ and $f_{\max}$ are 0.07 and 0.75, respectively, good results can usually be obtained [36].

$$mv = \frac{1}{\left(a_0 + e^{(K/2-k)/a_1}\right)} \tag{20}$$

where *mv* can affect the search capability and a_0 and a_1 represent two positive constants that regulate the exploratory and exploitative behaviors.

2.3.4. Moving toward the Best White Shark

Great white sharks can keep close to the ideal prey location. This behavior can be expressed by Equation (21).

$$w'^i_{k+1} = w_{gbest_k} + r_1 \vec{D}_w \text{sgn}(r_2 - 0.5) \quad r_3 < s_s \tag{21}$$

where w'^i_{k+1} indicates the updated location of the *i*th white shark relative to the prey location, while $\text{sgn}(r_2 - 0.5)$ takes values of 1 or −1 to alter the search direction. The variables r_1, r_2, and r_3 are randomly produced numbers chosen from the interval [0, 1]. $\vec{D}_w$ quantifies the distance between the white shark and the prey, as outlined in Equation (22). s_s is a parameter reflecting the intensity of the white shark's olfactory and visual senses when tracking other sharks near the optimal prey, outlined in Equation (23) [36].

$$\vec{D}_w = |rand \times (w_{gbest_k} - w_k^i)| \tag{22}$$

where *rand* denotes a randomly chosen number with the range [0, 1], and w_k^i indicates the white shark's present location relative to w_{gbest_k}.

$$s_s = |1 - e^{(-a_2 \times k/K)}| \tag{23}$$

where a_2 can influence the exploration and exploitation behavior. For the problem addressed in this study, the value of a_2 is set at 0.0005.

2.3.5. Fish School Behavior

To create a mathematical model of white shark behavior, the best two solutions are retained, and the locations of other white sharks are adjusted according to these solutions. The behavior of white sharks is then described by the subsequent formula:

$$w_{k+1}^i = \frac{w_k^i + w'^i_{k+1}}{2 \times rand} \tag{24}$$

Equation (24) demonstrates that a great white shark can adjust its position based on the location of another shark that has achieved an optimal position near the prey. Consequently, the final position of this shark will likely be very close to the best prey within the search area. The schooling behavior of fish and the tendency of great white sharks to move towards the most successful sharks exemplify their collective behavior. This pattern allows for enhanced exploration and exploitation of the environment.

2.4. Temporal Convolutional Network

A TCN is a convolutional neural network specifically designed for sequence modeling tasks that require causal constraints, such as time series prediction. In the field of deep learning, commonly used models include RNN and its variants, including LSTM and Gated Recurrent Unit (GRU). TCN offers a more straightforward and simpler architecture compared to recurrent frameworks like LSTM and GRU [28].

TCN is optimized on the basis of a traditional CNN. Distinguished from conventional CNNs, TCNs incorporate unique features such as causal convolution, dilation factor, and residual block. The architecture of the TCN is detailed in Figure 3.
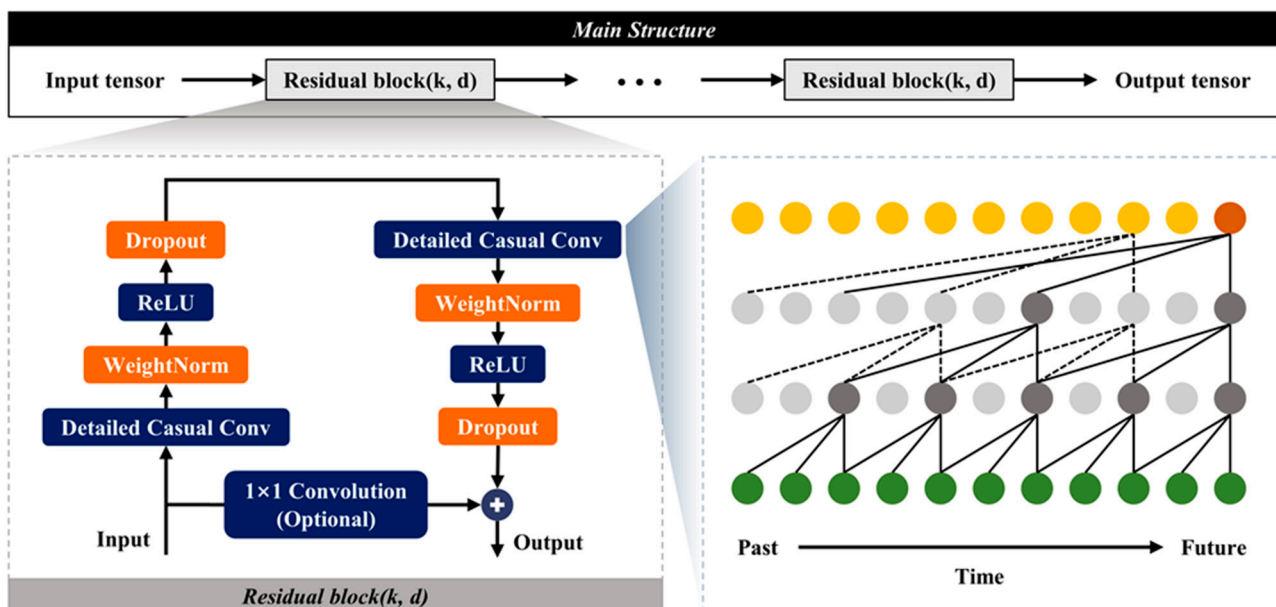

Figure 3. Structure of the TCN.

To maintain an output length identical to the input length, the Temporal Convolutional Network (TCN) employs a 1D fully convolutional network (FCN) framework. This methodology ensures that each hidden layer's length remains consistent with that of the input layer by utilizing zero padding to keep lengths equal through successive layers. To avoid future information from influencing past data, TCN implements causal convolution, where the output at any specific time t is dependent solely on the input at time t and previous times.

The basic causal convolution in a network can only access historical information linearly proportional to its depth, posing a challenge for sequential tasks requiring longer

historical contexts. To address this limitation, dilated convolution is introduced. This technique expands the reach of the convolution operation, allowing it to cover a broader span of input data without increasing the network depth, thereby facilitating the processing of extended historical sequences. For the filter $f : \{0, 1, \ldots, k - 1\}$, the dilated convolution operation F on the element s of the one-dimensional sequence ($x \in R^n$) can be expressed as follows:

$$F(s) = \left(x^*{}_d f\right)(s) = \sum_{i=1}^{k-1} f(i) \cdot x_{s-d \cdot i} \tag{25}$$

where d represents the dilation factor, k is the filter size, $s - d \cdot i$ indicates the direction of the past, and i is the number of filters. When $d = 1$, dilated convolution reverts to traditional convolution. To broaden the network's receptive field, selecting a larger filter size k and increasing the dilation factor d are effective strategies. A larger receptive field allows the network to delve deeper into historical data [45]. Figure 3 illustrates the structure of the dilated causal convolution in the TCN, with a kernel size $k = 3$ and dilation factors $d = [1, 2, 4]$.

Although the dilation factor is employed to enhance model reach, practical applications may still encounter significant model depths, leading to challenges such as gradient vanishing. To address this, a residual block structure similar to that in ResNet can be incorporated in place of simple connections between layers in the TCN. This adjustment helps the model more effectively counter issues like gradient vanishing.

2.5. Model Evaluation

To assess the precision of the prediction model, this study employs multiple statistical metrics. These include the Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2), which are calculated using the following formulas.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{26}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} \left| y_i - \hat{y}_i \right| \tag{27}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{28}$$

$$R^2 = 1 - \frac{\sum\limits_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum\limits_{i=1}^{n} (y_i - \overline{y})^2} \tag{29}$$

where n represents the number of samples, y_i represents the actual value, $\hat{y}_i$ then represents the predicted value, and $\overline{y}$ represents the average of the actual values.

3. Results and Discussion

In this section, RMSE, MAE, MSE, and R^2 are used to evaluate the performance of all the experimental models on three different PV systems; the experimental models include MIC-WSO-TCN, MIC-WSO-LSTM, MIC-WSO-BP, and MIC-TCN, which are represented by TCN2, LSTM, BP, and TCN1 in the figure. In order to evaluate the performance and accuracy of the proposed model, we use the test dataset to calculate the prediction error of each season (spring, summer, autumn, winter). The dataset contains the actual photovoltaic power generation measured in one year, and the power generation data are collected every 5 min. In addition, in order to illustrate the effectiveness of the MIC feature selection method, the results of the prediction model with feature selection and the prediction model without feature selection are compared. Taking the dataset of the dual-axis photovoltaic

system in the summer as an example, the test dataset includes not only sunny days but also various complex weather conditions such as rainy days and cloudy days. Such a dataset is closer to the actual demand and can more comprehensively evaluate the predictive ability of the model. If the CPU is used to train deep learning models with large datasets and multiple layers, it will take a lot of time. Therefore, this experiment chooses to use GPU to speed up the training speed, and the GPU memory used in the experiment is 8 GB. Here are the parameters used to train the model:

(1) The number of epochs selected is 100.
(2) The batch size is equal to 200, which can reduce the training time without affecting the accuracy of the model.
(3) The learning rate is 0.0015.
(4) The historical sequence input of the model is 15.
(5) The loss function is RMSE.

Based on the data in Table 1, when other conditions are consistent, the model using MIC feature selection shows lower RMSE, MAE, and MSE values than the model without feature selection, and the R^2 value is also higher. The RMSE of the model after feature selection decreased from 1.395 to 0.969, and the value of MAE also decreased from 0.528 to 0.386, which decreased by 30.5% and 26.9%, respectively, and the value of MSE decreased from 1.947 to 0.940, which decreased by 51.7%, while the R^2 increased from 0.958 to 0.980, which increased by 23%. It can be seen that the predictive ability of the model has been significantly improved after the addition of the MIC algorithm. Tables 2–5 show the prediction accuracy indexes of summer, autumn, winter, and spring, respectively. In addition, the line charts show the difference between the actual photovoltaic power curve and the prediction curve of three different photovoltaic systems in four seasons. Each figure contains five days of prediction samples.

Table 1. Forecasting results comparison of with and without feature selection.

Index	Without Feature Selection	With Feature Selection
RMSE	1.395	0.969
MAE	0.528	0.386
MSE	1.947	0.940
R^2	0.958	0.980

Table 2. RMSE, MAE, MSE, and R^2 of different models in summer.

Methods	Dual-Axis				Single-Axis				Fixed			
	RMSE (KW)	MAE (KW)	MSE (KW2)	R^2	RMSE (KW)	MAE (KW)	MSE (KW2)	R^2	RMSE (KW)	MAE (KW)	MSE (KW2)	R^2
MIC-TCN	1.383	0.584	1.914	0.959	0.253	0.093	0.064	0.947	0.246	0.092	0.061	0.959
MIC-WSO-BP	1.263	0.507	1.591	0.966	0.226	0.081	0.051	0.958	0.260	0.107	0.068	0.955
MIC-WSO-LSTM	1.267	0.624	1.605	0.966	0.221	0.083	0.049	0.960	0.241	0.101	0.058	0.961
MIC-WSO-TCN	0.969	0.386	0.940	0.980	0.195	0.072	0.038	0.968	0.213	0.081	0.045	0.970

Table 3. RMSE, MAE, MSE, and R^2 of different models in autumn.

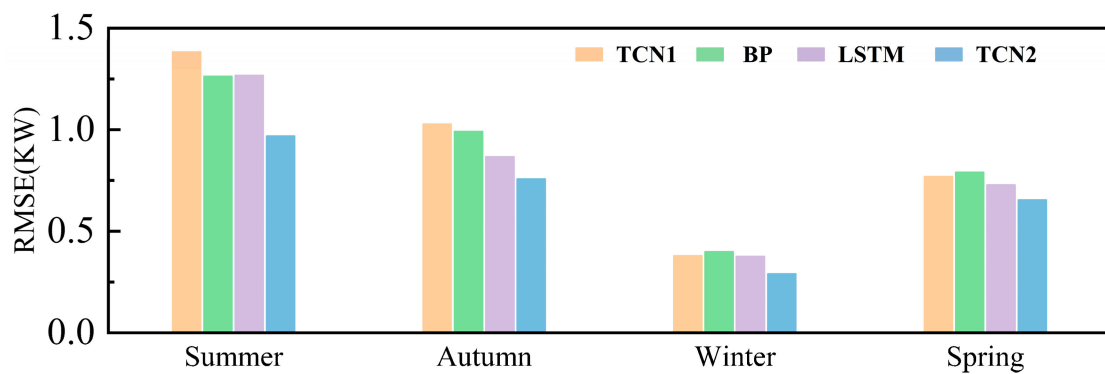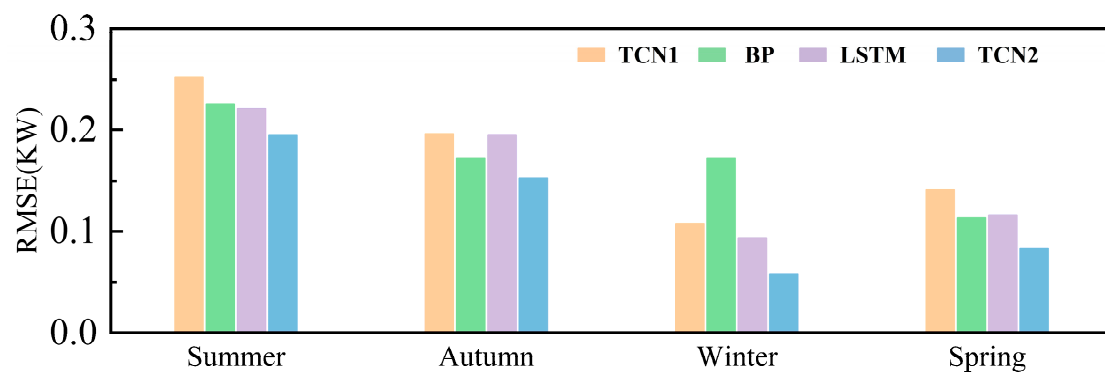| Methods	Dual-Axis				Single-Axis				Fixed			
	RMSE (KW)	MAE (KW)	MSE (KW2)	R^2	RMSE (KW)	MAE (KW)	MSE (KW2)	R^2	RMSE (KW)	MAE (KW)	MSE (KW2)	R^2
MIC-TCN	1.028	0.365	1.057	0.969	0.196	0.066	0.038	0.977	0.168	0.056	0.028	0.980
MIC-WSO-BP	0.991	0.359	0.982	0.972	0.173	0.057	0.030	0.981	0.177	0.107	0.031	0.978
MIC-WSO-LSTM	0.867	0.337	0.752	0.978	0.195	0.064	0.038	0.976	0.161	0.068	0.026	0.982
MIC-WSO-TCN	0.758	0.273	0.575	0.984	0.153	0.053	0.023	0.986	0.148	0.055	0.022	0.985

Table 4. RMSE, MAE, MSE, and R^2 of different models in winter.

Methods	Dual-Axis				Single-Axis				Fixed			
	RMSE (KW)	MAE (KW)	MSE (KW2)	R^2	RMSE (KW)	MAE (KW)	MSE (KW2)	R^2	RMSE (KW)	MAE (KW)	MSE (KW2)	R^2
MIC-TCN	0.380	0.177	0.144	0.997	0.108	0.061	0.012	0.993	0.123	0.066	0.015	0.993
MIC-WSO-BP	0.400	0.194	0.160	0.997	0.173	0.079	0.006	0.996	0.118	0.058	0.014	0.994
MIC-WSO-LSTM	0.376	0.238	0.142	0.998	0.094	0.057	0.009	0.995	0.089	0.045	0.008	0.996
MIC-WSO-TCN	0.291	0.151	0.085	0.999	0.058	0.031	0.003	0.998	0.082	0.044	0.007	0.997

Table 5. RMSE, MAE, MSE, and R^2 of different models in spring.

Methods	Dual-Axis				Single-Axis				Fixed			
	RMSE (KW)	MAE (KW)	MSE (KW2)	R^2	RMSE (KW)	MAE (KW)	MSE (KW2)	R^2	RMSE (KW)	MAE (KW)	MSE (KW2)	R^2
MIC-TCN	0.769	0.331	0.581	0.987	0.141	0.073	0.020	0.981	0.153	0.063	0.023	0.985
MIC-WSO-BP	0.791	0.348	0.625	0.986	0.113	0.050	0.013	0.987	0.136	0.056	0.019	0.988
MIC-WSO-LSTM	0.728	0.357	0.530	0.988	0.116	0.060	0.013	0.987	0.137	0.061	0.019	0.988
MIC-WSO-TCN	0.655	0.304	0.429	0.991	0.083	0.051	0.007	0.993	0.129	0.052	0.016	0.989

Figures 4–6 show the RMSE values of the four hybrid algorithms listed in the tables for three types of photovoltaic systems across the four seasons (summer, autumn, winter, and spring). It can be clearly observed that the proposed hybrid method (MIC-WSO-TCN) achieves lower RMSE values in each season compared to the other three hybrid methods. The RMSE value of the proposed model is generally higher in summer than in the other three seasons, while it is at a relatively low level in winter. This is because the dataset contains more cloudy days in summer (cloudy weather causes significant fluctuations in photovoltaic output) and more sunny days in winter. Even during periods of severe output fluctuations, the proposed model maintains a high level of accuracy and outperforms other comparative models.

Figure 4. Evaluation results of several models (dual-axis).

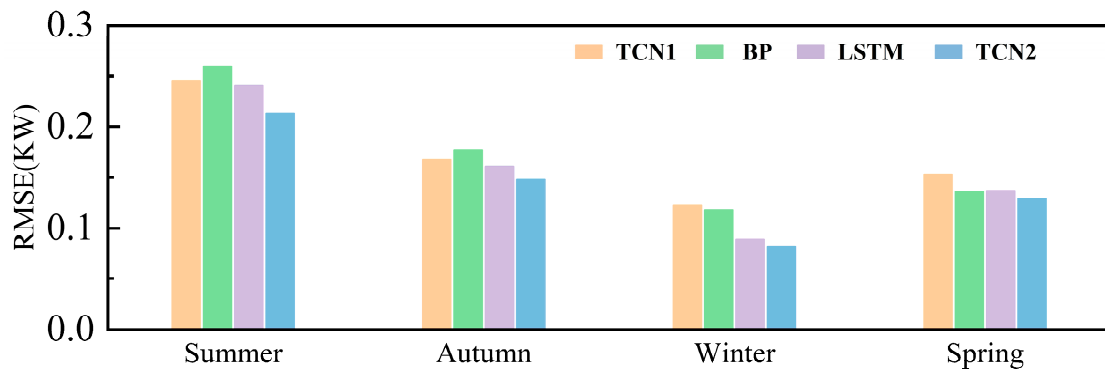**Figure 5.** Evaluation results of several models (single-axis).

Figure 6. Evaluation results of several models (fixed).

Table 2 shows that in summer, compared with the MIC-TCN, MIC-WSO-BP, and MIC-WSO-LSTM methods, MIC-WSO-TCN has the lowest RMSE values on dual, single, and fixed photovoltaic systems, which are 0.969 KW, 0.195 KW, and 0.213 KW, respectively. Compared with MIC-TCN, the RMSE values of MIC-WSO-TCN were reduced by 29.9%, 22.9%, and 13.4% on three different photovoltaic systems, respectively. The MAE values of the proposed MIC-WSO-TCN model are also lower than other models, and the MAE values on the three systems are 0.386 KW, 0.072 KW, and 0.045 KW. This indicates a notable enhancement in the prediction performance of the WSO algorithm upon its integration into the model. Such improvement can be attributed to the simplicity and robustness of WSO, which facilitates the rapid and accurate identification of global solutions for challenging optimization problems, and because WSO does not need to compute the derivatives of the search space of the relevant problem, it can effectively get rid of the local minima problem that exists in the actual problem. Figures 7–9 show the prediction results of the four models in different photovoltaic systems during summer. The figure clearly illustrates the intermittent nature of solar energy during the summer season due to the mostly cloudy weather. Even under such complex fluctuating weather conditions, the prediction results of the method (MIC-WSO-TCN) proposed in this paper can be very close to the actual power curve. Compared with other methods, the proposed hybrid method is also closer to the actual power curve than other methods.
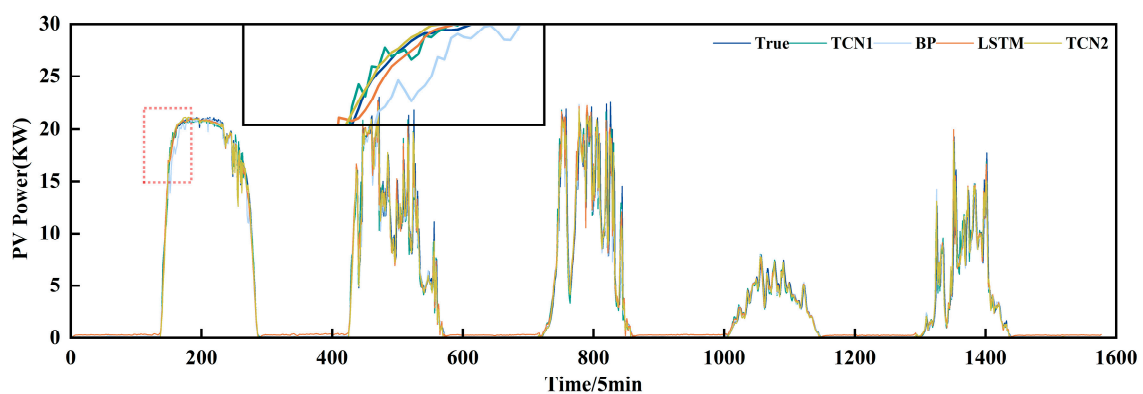

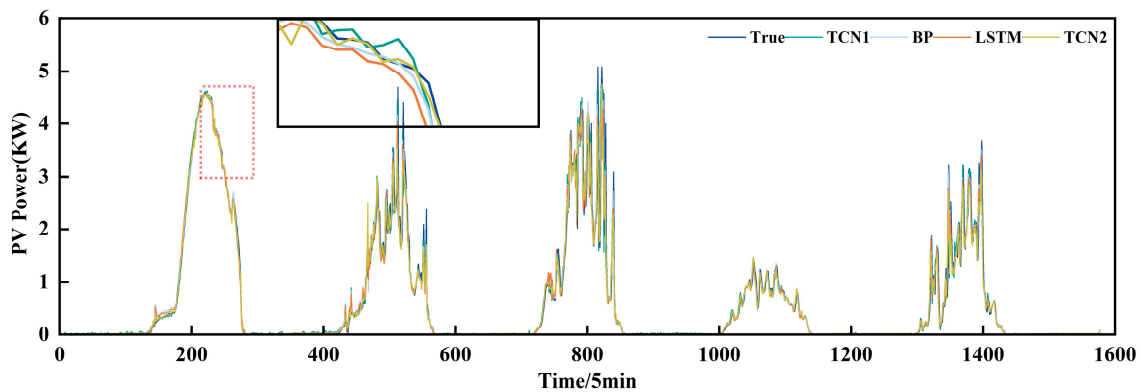**Figure 7.** Summer prediction outcomes for dual-axis PV system.

Figure 8. Summer prediction outcomes for single-axis PV system.
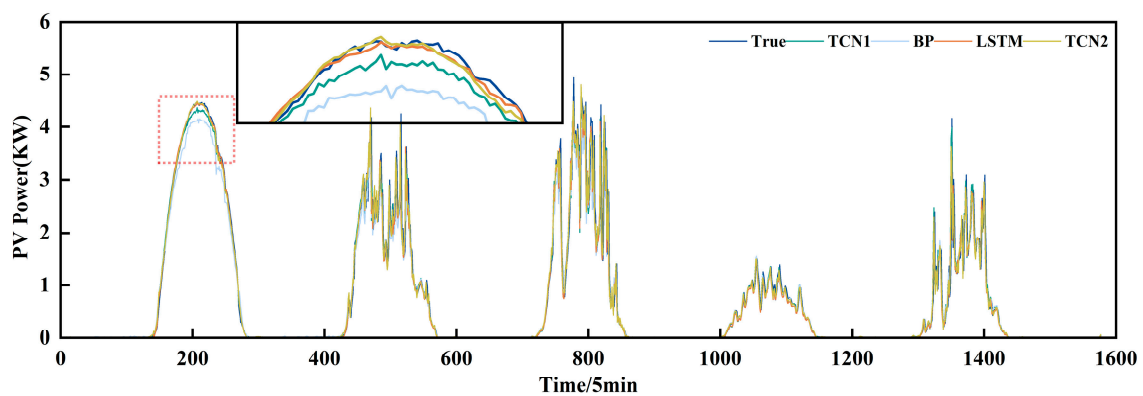

Figure 9. Summer prediction outcomes for fixed PV system.

Figures 10–12 show the prediction results of the four models in different photovoltaic systems during autumn. Table 3 reveals that during autumn, for the dual-axis tracking photovoltaic system, the proposed method achieves an RMSE of 0.758 KW, compared to 0.991 KW for MIC-WSO-BP and 0.867 KW for MIC-WSO-LSTM. This represents a reduction in RMSE of 24.5% and 12.6% relative to these methods, respectively. In fixed photovoltaic systems, the RMSE, MAE, and MSE for the proposed method are 0.148 KW, 0.055 KW, and 0.022 KW^2, respectively. The MAE of the proposed method, MIC-WSO-TCN, is 48.6% and 19.1% lower than that of MIC-WSO-BP and MIC-WSO-LSTM. The proposed method also shows superior performance for single-axis tracking photovoltaic systems. For example, in autumn, the hybrid method achieves RMSE, MAE, and R^2 values of 0.153, 0.053, and 0.986, respectively, while in summer, these values are 0.195, 0.072, and 0.968. This indicates that the prediction accuracy is higher in autumn than in summer. By examining Figures 8 and 11, it can be observed that the weather conditions in autumn are less intermittent than in summer, leading to a more stable power generation curve overall. This suggests that the prediction accuracy of the model improves when there is less fluctuation in photovoltaic power. Consequently, the proposed method delivers better prediction accuracy in autumn compared to summer.
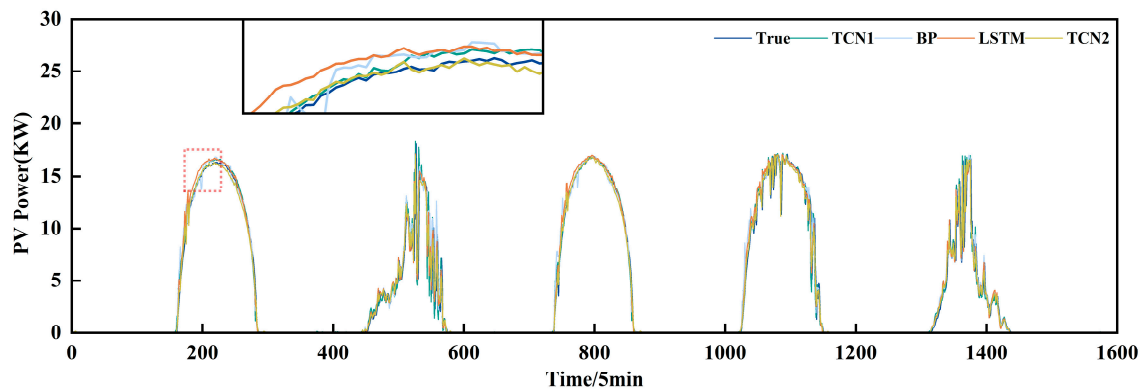
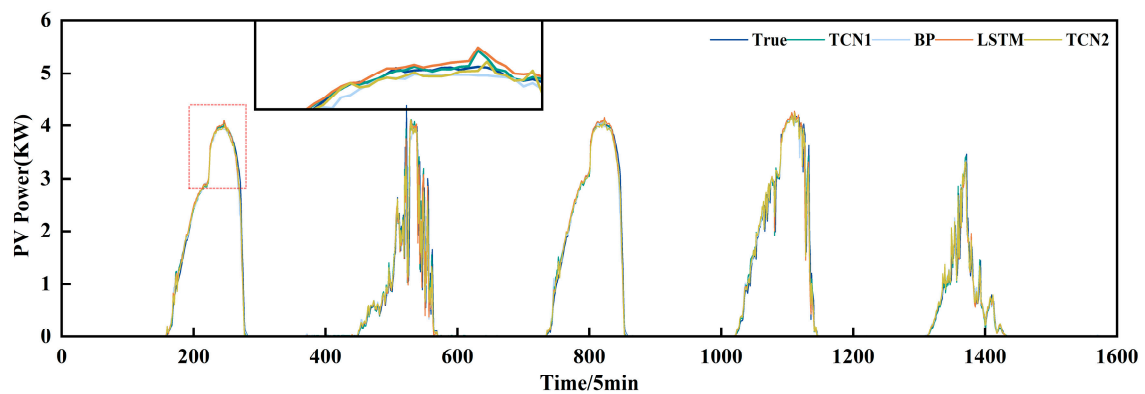**Figure 10.** Autumn prediction outcomes for dual-axis PV system.

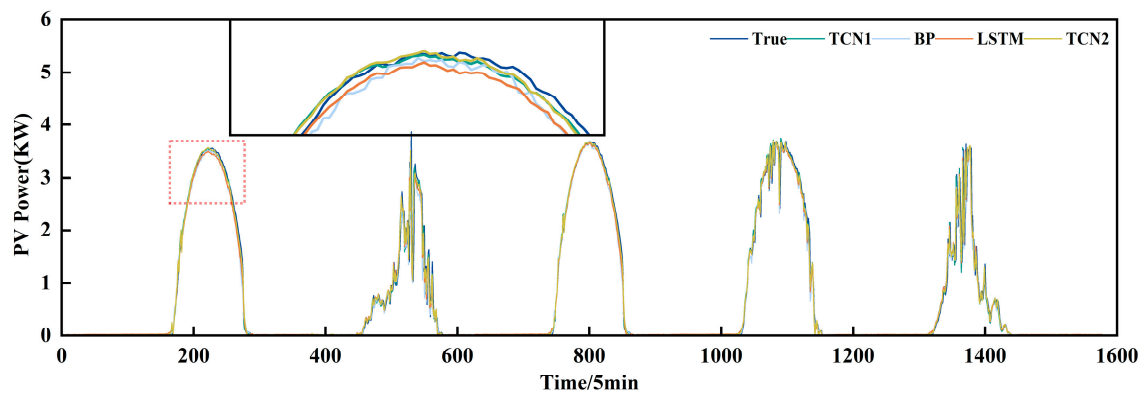**Figure 11.** Autumn prediction outcomes for single-axis PV system.

Figure 12. Autumn prediction outcomes for fixed PV system.

Figures 13–15 show the prediction results of the four models in different photovoltaic systems during winter. Table 4 shows that in winter, the performance of each model is better. For the fixed photovoltaic system, the RMSE values of MIC-TCN, MIC-WSO-BP, and MIC-WSO-LSTM are 0.123, 0.118, and 0.089, respectively, while the RMSE value of MIC-WSO-TCN is 0.082, which is reduced by 33.3%, 30.5%, and 7.9% from these baselines. In addition, the proposed model shows extremely high R2 values in three different photovoltaic systems, which are 0.999, 0.998, and 0.997, respectively. These results show that the model has good fitting ability. Figure 13 shows that there are more sunny days (days where irradiance does not vary much during the day and is essentially cloudless) during the winter months, and that the model can achieve a very high level of prediction accuracy under these weather conditions. Although the prediction accuracy of the comparison methods is relatively high,

the performance of the hybrid method (MIC-WSO-TCN) proposed in this paper is still better than other comparison algorithms. Figures 16–18 show the prediction results of the four models in different photovoltaic systems during spring. As can be seen from Table 5, the model proposed in this paper still performs the best.
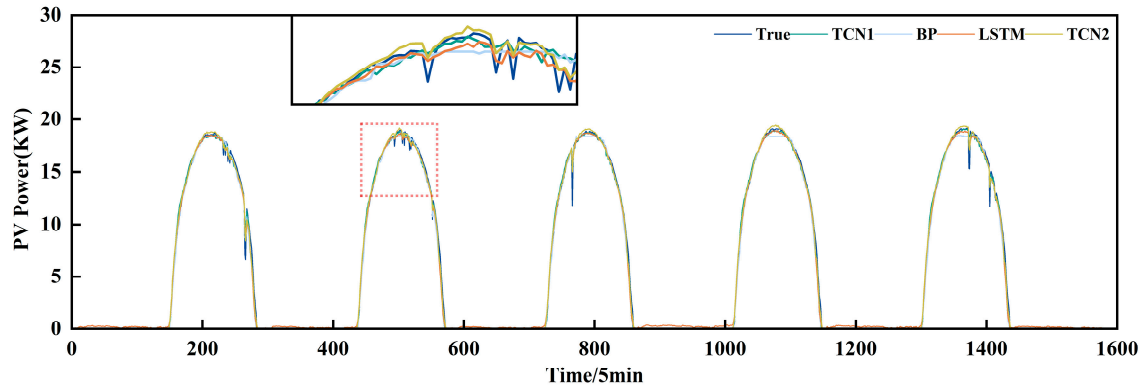

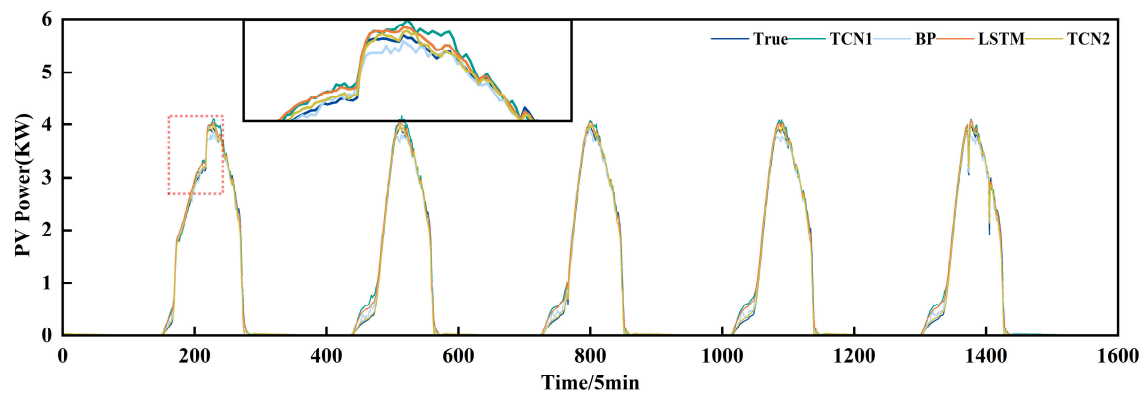**Figure 13.** Winter prediction outcomes for dual-axis PV system.

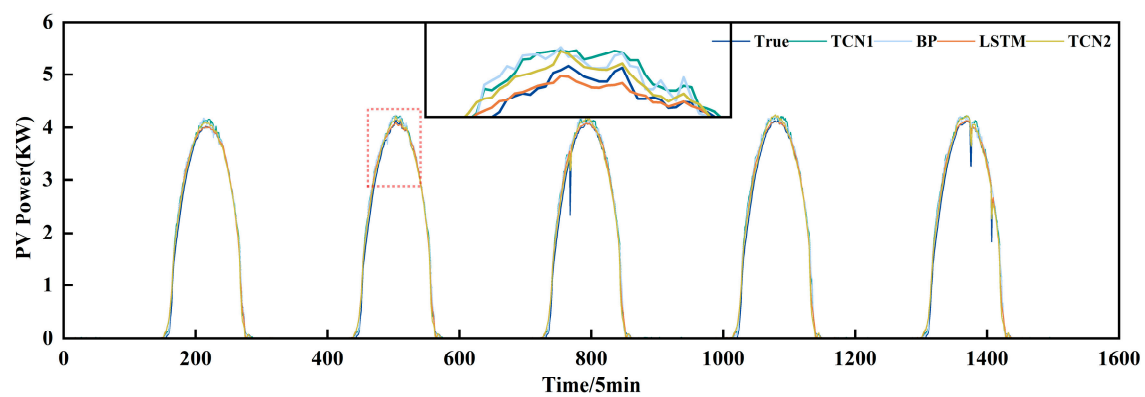**Figure 14.** Winter prediction outcomes for single-axis PV system.

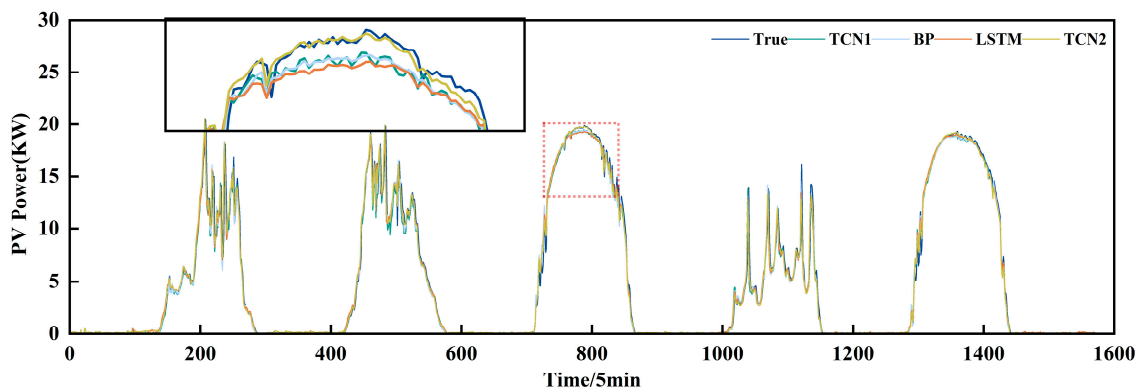**Figure 15.** Winter prediction outcomes for fixed PV system.

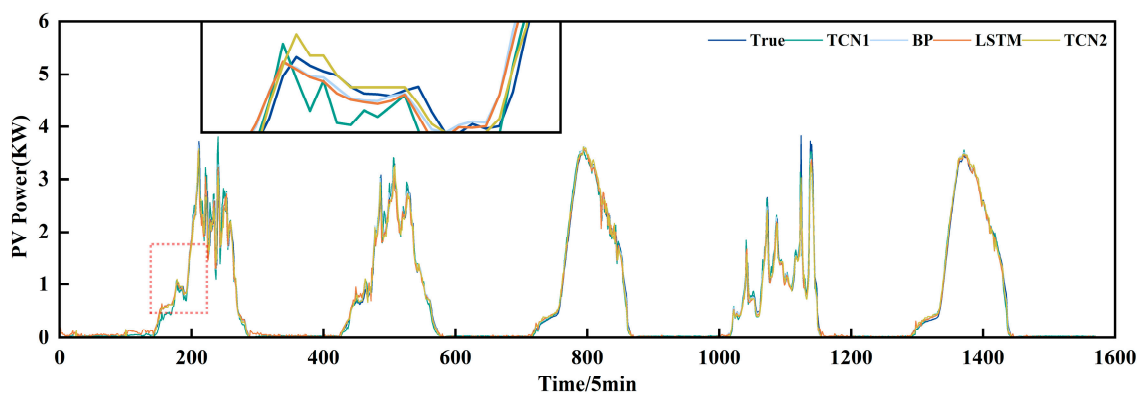**Figure 16.** Spring prediction outcomes for dual-axis PV system.

Figure 17. Spring prediction outcomes for single-axis PV system.
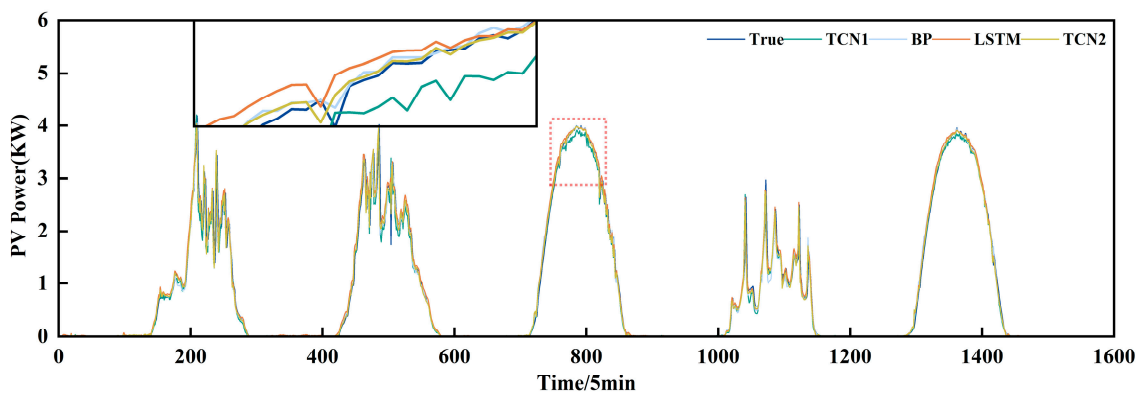

Figure 18. Spring prediction outcomes for fixed PV system.

This study introduces a novel hybrid approach, MIC-WSO-TCN, and this approach is applied to predict the output power for three different PV systems (dual-axis, single-axis, and fixed PV systems) using actual data across all four seasons. The results reveal that the model excels in predicting power generation with high accuracy during periods of minimal power generation fluctuation (such as sunny or cloudless weather) and maintains commendable prediction accuracy even under complex, fluctuating weather conditions.

4. Conclusions

In this study, power output predictions are made for three different PV systems installed in Alice Springs, Australia. These three photovoltaic systems are a dual-axis tracking photovoltaic system, single-axis tracking photovoltaic system, and fixed photovoltaic sys-

tem. This paper proposes a new hybrid deep learning prediction method (MIC-WSO-TCN). To demonstrate its superiority, it is compared with MIC-TCN, WSO-TCN, MIC-WSO-BP, and MIC-WSO-LSTM.

It can be seen from the results that the MIC algorithm and WSO algorithm in the hybrid model are very useful for improving the prediction accuracy. The MIC algorithm is used to explore the linear and nonlinear correlation in the dataset, and the characteristics of high correlation are selected more reasonably. After adding the MIC algorithm, the RMSE decreases from 1.395 to 0.969, and the value of MAE also decreases from 0.528 to 0.386, which decreases by 30.5% and 26.9%, respectively. The novel WSO algorithm is utilized to optimize the performance of the model. Compared with MIC-TCN, the RMSE values of MIC-WSO-TCN are reduced by 29.9%, 22.9%, and 13.4% respectively. Taking the performance of the model in summer as an example, for the dual-axis photovoltaic system, compared with MIC-WSO-BP and MIC-WSO-LSTM, the RMSE values of MIC-WSO-TCN are reduced by 23.3% and 23.5%, and the MAE values are reduced by 23.9% and 38.1%. For the single-axis photovoltaic system, the RMSE values are reduced by 13.7 % and 11.8 %, and the MAE values are reduced by 11.1% and 13.3%. For the fixed PV system, the RMSE values decreased by 18.1% and 11.6%, and the MAE values decreased by 24.3% and 19.8%. Overall, despite the complex operating conditions of real PV systems, the proposed MIC-WSO-TCN model also performs the prediction task well, obtaining low RMSE, MAE, and MSE values and high correlation coefficients, both in the summer season when the power generation fluctuates dramatically, and in the winter season when the power generation is stable. The proposed model proves to be very effective and robust in predicting the output of different types of PV systems.

Author Contributions: Conceptualization, Z.Y.; methodology, X.L.; software, X.L.; validation, X.L.; formal analysis, X.L.; investigation, L.Z.; resources, J.S.; data curation, X.L.; writing—original draft preparation, X.L.; writing—review and editing, P.T.; visualization, Y.N.; supervision, J.S.; project administration, J.S.; funding acquisition, J.S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Research data for this article was obtained from: https://dkasolarcentre.com.au/locations/alice-springs/ (accessed on 20 October 2023).

Conflicts of Interest: The authors have declared no conflicts of interest.

References

1. Iheanetu, K.J. Solar Photovoltaic Power Forecasting: A Review. *Sustainability* **2022**, *14*, 17005. [CrossRef]
2. Mohamad Radzi, P.N.L.; Akhter, M.N.; Mekhilef, S.; Mohamed Shah, N. Review on the Application of Photovoltaic Forecasting Using Machine Learning for Very Short- to Long-Term Forecasting. *Sustainability* **2023**, *15*, 2942. [CrossRef]
3. Rotas, R.; Fotopoulou, M.; Drosatos, P.; Rakopoulos, D.; Nikolopoulos, N. Adaptive Dynamic Building Envelopes with Solar Power Components: Annual Performance Assessment for Two Pilot Sites. *Energies* **2023**, *16*, 2148. [CrossRef]
4. Bhatti, A.R.; Bilal Awan, A.; Alharbi, W.; Salam, Z.; Bin Humayd, A.S.; Praveen, R.P.; Bhattacharya, K. An Improved Approach to Enhance Training Performance of ANN and the Prediction of PV Power for Any Time-Span without the Presence of Real-Time Weather Data. *Sustainability* **2021**, *13*, 11893. [CrossRef]
5. Mei, F.; Pan, Y.; Zhu, K.; Zheng, J. A Hybrid Online Forecasting Model for Ultrashort-Term Photovoltaic Power Generation. *Sustainability* **2018**, *10*, 820. [CrossRef]
6. Wang, K.; Qi, X.; Liu, H. A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network. *Appl. Energy* **2019**, *251*, 113315. [CrossRef]
7. Gu, W.; Ma, T.; Song, A.; Li, M.; Shen, L. Mathematical modelling and performance evaluation of a hybrid photovoltaic-thermoelectric system. *Energy Convers. Manag.* **2019**, *198*, 111800. [CrossRef]

8. Chen, B.; Lin, P.; Lai, Y.; Cheng, S.; Chen, Z.; Wu, L. Very-Short-Term Power Prediction for PV Power Plants Using a Simple and Effective RCC-LSTM Model Based on Short Term Multivariate Historical Datasets. *Electronics* **2020**, *9*, 289. [CrossRef]

9. Wang, F.; Zhen, Z.; Liu, C.; Mi, Z.; Hodge, B.-M.; Shafie-khah, M.; Catalão, J.P.S. Image phase shift invariance based cloud motion displacement vector calculation method for ultra-short-term solar PV power forecasting. *Energy Convers. Manag.* **2018**, *157*, 123–135. [CrossRef]

10. Zaher, A.; Thil, S.; Nou, J.; Traoré, A.; Grieu, S. Comparative study of algorithms for cloud motion estimation using sky-imaging data. *IFAC-PapersOnLine* **2017**, *50*, 5934–5939. [CrossRef]

11. Sharadga, H.; Hajimirza, S.; Balog, R.S. Time series forecasting of solar power generation for large-scale photovoltaic plants. *Renew. Energy* **2020**, *150*, 797–807. [CrossRef]

12. Dolara, A.; Leva, S.; Manzolini, G. Comparison of different physical models for PV power output prediction. *Sol. Energy* **2015**, *119*, 83–99. [CrossRef]

13. Zhou, Y.; Wang, J.; Li, Z.; Lu, H. Short-term photovoltaic power forecasting based on signal decomposition and machine learning optimization. *Energy Convers. Manag.* **2022**, *267*, 115944. [CrossRef]

14. Yagli, G.M.; Yang, D.; Srinivasan, D. Automatic hourly solar forecasting using machine learning models. *Renew. Sustain. Energy Rev.* **2019**, *105*, 487–498. [CrossRef]

15. Ali-Ou-Salah, H.; Oukarfi, B.; Mouhaydine, T. Short-term solar radiation forecasting using a new seasonal clustering technique and artificial neural network. *Int. J. Green Energy* **2022**, *19*, 424–434. [CrossRef]

16. Yousaf, M.Z.; Tahir, M.F.; Raza, A.; Khan, M.A.; Badshah, F. Intelligent Sensors for dc Fault Location Scheme Based on Optimized Intelligent Architecture for HVdc Systems. *Sensors* **2022**, *22*, 9936. [CrossRef] [PubMed]

17. Yousaf, M.Z.; Khalid, S.; Tahir, M.F.; Tzes, A.; Raza, A. A novel dc fault protection scheme based on intelligent network for meshed dc grids. *Int. J. Electr. Power Energy Syst.* **2023**, *154*, 109423. [CrossRef]

18. Yousaf, M.Z.; Mirsaeidi, S.; Khalid, S.; Raza, A.; Zhichu, C.; Rehman, W.U.; Badshah, F. Multisegmented Intelligent Solution for MT-HVDC Grid Protection. *Electronics* **2023**, *12*, 1766. [CrossRef]

19. Hua, L.; Zhang, C.; Peng, T.; Ji, C.; Shahzad Nazir, M. Integrated framework of extreme learning machine (ELM) based on improved atom search optimization for short-term wind speed prediction. *Energy Convers. Manag.* **2022**, *252*, 115102. [CrossRef]

20. Ahmad, A.; Jin, Y.; Zhu, C.; Javed, I.; Waqar Akram, M.; Buttar, N.A. Support vector machine based prediction of photovoltaic module and power station parameters. *Int. J. Green Energy* **2020**, *17*, 219–232. [CrossRef]

21. Yu, W.; Liu, G.; Zhu, L.; Yu, W. Convolutional neural network with feature reconstruction for monitoring mismatched photovoltaic systems. *Sol. Energy* **2020**, *212*, 169–177. [CrossRef]

22. Tahir, M.F.; Yousaf, M.Z.; Tzes, A.; El Moursi, M.S.; El-Fouly, T.H.M. Enhanced solar photovoltaic power prediction using diverse machine learning algorithms with hyperparameter optimization. *Renew. Sustain. Energy Rev.* **2024**, *200*, 114581. [CrossRef]

23. Wang, H.; Yi, H.; Peng, J.; Wang, G.; Liu, Y.; Jiang, H.; Liu, W. Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network. *Energy Convers. Manag.* **2017**, *153*, 409–422. [CrossRef]

24. Tahir, M.F.; Tzes, A.; Yousaf, M.Z. Enhancing PV power forecasting with deep learning and optimizing solar PV project performance with economic viability: A multi-case analysis of 10 MW Masdar project in UAE. *Energy Convers. Manag.* **2024**, *311*, 118549. [CrossRef]

25. Elizabeth Michael, N.; Hasan, S.; Al-Durra, A.; Mishra, M. Short-term solar irradiance forecasting based on a novel Bayesian optimized deep Long Short-Term Memory neural network. *Appl. Energy* **2022**, *324*, 119727. [CrossRef]

26. Ghimire, S.; Bhandari, B.; Casillas-Pérez, D.; Deo, R.C.; Salcedo-Sanz, S. Hybrid deep CNN-SVR algorithm for solar radiation prediction problems in Queensland, Australia. *Eng. Appl. Artif. Intell.* **2022**, *112*, 104860. [CrossRef]

27. Raman, R.; Mewada, B.; Meenakshi, R.; Jayaseelan, G.M.; Sharmila, K.S.; Taqui, S.N.; Al-Ammar, E.A.; Wabaidur, S.M.; Iqbal, A. Forecasting the PV Power Utilizing a Combined Convolutional Neural Network and Long Short-Term Memory Model. *Electr. Power Compon. Syst.* **2024**, *52*, 233–249. [CrossRef]

28. Bai, S.; Kolter, J.Z.; Koltun, V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. *arXiv* **2018**, arXiv:1803.01271.

29. Wang, Y.; Zhang, C.; Fu, Y.; Suo, L.; Song, S.; Peng, T.; Shahzad Nazir, M. Hybrid solar radiation forecasting model with temporal convolutional network using data decomposition and improved artificial ecosystem-based optimization algorithm. *Energy* **2023**, *280*, 128171. [CrossRef]

30. Shao, Z.; Han, J.; Zhao, W.; Zhou, K.; Yang, S. Hybrid model for short-term wind power forecasting based on singular spectrum analysis and a temporal convolutional attention network with an adaptive receptive field. *Energy Convers. Manag.* **2022**, *269*, 116138. [CrossRef]

31. Qiao, X.; Peng, T.; Sun, N.; Zhang, C.; Liu, Q.; Zhang, Y.; Wang, Y.; Shahzad Nazir, M. Metaheuristic evolutionary deep learning model based on temporal convolutional network, improved aquila optimizer and random forest for rainfall-runoff simulation and multi-step runoff prediction. *Expert Syst. Appl.* **2023**, *229*, 120616. [CrossRef]

32. Pousinho, H.M.I.; Mendes, V.M.F.; Catalão, J.P.S. A hybrid PSO–ANFIS approach for short-term wind power prediction in Portugal. *Energy Convers. Manag.* **2011**, *52*, 397–402. [CrossRef]

33. Jaidee, S.; Pora, W. Very Short-Term Solar Power Forecasting Using Genetic Algorithm Based Deep Neural Network. In Proceedings of the 2019 4th International Conference on Information Technology (InCIT), Bangkok, Thailand, 24–25 October 2019; pp. 184–189.

34.	Peng, T.; Zhang, C.; Zhou, J.; Nazir, M.S. An integrated framework of Bi-directional long-short term memory (BiLSTM) based on sine cosine algorithm for hourly solar radiation forecasting. *Energy* **2021**, *221*, 119887. [CrossRef]

35.	Akhter, M.N.; Mekhilef, S.; Mokhlis, H.; Ali, R.; Usama, M.; Muhammad, M.A.; Khairuddin, A.S.M. A hybrid deep learning method for an hour ahead power output forecasting of three different photovoltaic systems. *Appl. Energy* **2022**, *307*, 118185. [CrossRef]

36.	Braik, M.; Hammouri, A.; Atwan, J.; Al-Betar, M.A.; Awadallah, M.A. White Shark Optimizer: A novel bio-inspired meta-heuristic algorithm for global optimization problems. *Knowl. Based Syst.* **2022**, *243*, 108457. [CrossRef]

37.	Fathy, A.; Alanazi, A. An Efficient White Shark Optimizer for Enhancing the Performance of Proton Exchange Membrane Fuel Cells. *Sustainability* **2023**, *15*, 1741. [CrossRef]

38.	Reshef, D.N.; Reshef, Y.A.; Finucane, H.K.; Grossman, S.R.; McVean, G.; Turnbaugh, P.J.; Lander, E.S.; Mitzenmacher, M.; Sabeti, P.C. Detecting Novel Associations in Large Data Sets. *Science* **2011**, *334*, 1518–1524. [CrossRef] [PubMed]

39.	Guo, Z.; Yu, B.; Hao, M.; Wang, W.; Jiang, Y.; Zong, F. A novel hybrid method for flight departure delay prediction using Random Forest Regression and Maximal Information Coefficient. *Aerosp. Sci. Technol.* **2021**, *116*, 106822. [CrossRef]

40.	Ma, Y.; Chen, S.; Khattak, A.J.; Cao, Z.; Zubair, M.; Han, X.; Hu, X. What Affects Emotional Well-Being during Travel? Identifying the Factors by Maximal Information Coefficient. *Int. J. Environ. Res. Public Health* **2022**, *19*, 4326. [CrossRef]

41.	Data Download | DKA Solar Centre. Available online: https://dkasolarcentre.com.au/download?location=alice-springs (accessed on 20 October 2023).

42.	Benson, B.; Pan, W.D.; Prasad, A.; Gary, G.A.; Hu, Q. Forecasting Solar Cycle 25 Using Deep Neural Networks. *Sol. Phys.* **2020**, *295*, 65. [CrossRef]

43.	Butcher, B.; Smith, B.J. Feature Engineering and Selection: A Practical Approach for Predictive Models. *Am. Stat.* **2020**, *74*, 308–309. [CrossRef]

44.	Wang, Y.; Xu, H.; Song, M.; Zhang, F.; Li, Y.; Zhou, S.; Zhang, L. A convolutional Transformer-based truncated Gaussian density network with data denoising for wind speed forecasting. *Appl. Energy* **2023**, *333*, 120601. [CrossRef]

45.	Gan, Z.; Li, C.; Zhou, J.; Tang, G. Temporal convolutional networks interval prediction model for wind speed forecasting. *Electr. Power Syst. Res.* **2021**, *191*, 106865. [CrossRef]