*energies*

MDPI

*Article*

# High-Accuracy Photovoltaic Power Prediction under Varying Meteorological Conditions: Enhanced and Improved Beluga Whale Optimization Extreme Learning Machine

**Wei Du [1], Shi-Tao Peng [1], Pei-Sen Wu [1] and Ming-Lang Tseng [2,3,4,5,*]**

[1] Key Laboratory of Environmental Protection in Water Transport Engineering Ministry of Transport, Tianjin Research Institute for Water Transport Engineering, No. 2618 Xingang Erhao Road, Binhai New District, Tianjin 300456, China; wei_du@126.com (W.D.); pengshitaotj@126.com (S.-T.P.); wups@tiwte.ac.cn (P.-S.W.)
[2] Institute of Innovation and Circular Economy, Asia University, Taichung 413, Taiwan
[3] Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 404327, Taiwan
[4] UKM Graduate School of Business, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia
[5] Department of Industrial Engineering, Khon Kaen University, Khon Kaen 40002, Thailand
* Correspondence: tsengminglang@gmail.com or tsengminglang@asia.edu.tw

**Abstract:** Accurate photovoltaic (PV) power prediction plays a crucial role in promoting energy structure transformation and reducing greenhouse gas emissions. This study aims to improve the accuracy of PV power generation prediction. Extreme learning machine (ELM) was used as the core model, and enhanced and improved beluga whale optimization (EIBWO) was proposed to optimize the internal parameters of ELM, thereby improving its prediction accuracy for PV power generation. Firstly, this study introduced the chaotic mapping strategy, sine dynamic adaptive factor, and disturbance strategy to beluga whale optimization, and EIBWO was proposed with high convergence accuracy and strong optimization ability. It was verified through standard testing functions that EIBWO performed better than comparative algorithms. Secondly, EIBWO was used to optimize the internal parameters of ELM and establish a PV power prediction model based on enhanced and improved beluga whale optimization algorithm–optimization extreme learning machine (EIBWO-ELM). Finally, the measured data of the PV output were used for verification, and the results show that the PV power prediction results of EIBWO-ELM were more accurate regardless of whether it was cloudy or sunny. The R2 of EIBWO-ELM exceeded 0.99, highlighting its efficient ability to adapt to PV power generation. The prediction accuracy of EIBWO-ELM is better than that of comparative models. Compared with existing models, EIBWO-ELM significantly improves the predictive reliability and economic benefits of PV power generation. This study not only provides a technological foundation for the optimization of intelligent energy systems but also contributes to the sustainable development of clean energy.

**Keywords:** photovoltaic power prediction; enhanced and improved beluga whale optimization; varying meteorological conditions; extreme learning machine



**Citation:** Du, W.; Peng, S.-T.; Wu, P.-S.; Tseng, M.-L. High-Accuracy Photovoltaic Power Prediction under Varying Meteorological Conditions: Enhanced and Improved Beluga Whale Optimization Extreme Learning Machine. *Energies* **2024**, *17*, 2309. https://doi.org/10.3390/en17102309

Academic Editor: Tapas Mallick

Received: 15 April 2024
Revised: 3 May 2024
Accepted: 8 May 2024
Published: 10 May 2024



## 1. Introduction

Photovoltaic (PV) power generation stands out for its environmental friendliness, reliability, and widespread availability [1]. However, the intermittent nature of solar radiation and the variability of weather conditions present challenges to power systems' stability and reliability [2]. The inherent intermittency, randomness, and volatility of PV power generation highlight the need for innovative solutions to ensure grid stability and energy reliability in the face of these fluctuations. The output of PV power generation is highly dependent on meteorological conditions such as solar radiation intensity and temperature, which makes the volatility and uncertainty of power generation a major

challenge for grid scheduling and energy management [3]. In this context, high-precision PV power prediction under varying meteorological conditions is particularly important. It can not only ensure the stable operation of the power grid and optimize energy distribution but also improve the economic and cost-effectiveness of PV power generation, promote the effective integration of renewable energy, and address the impact of climate change [4]. In addition, with the advancement of technologies such as numerical weather forecasting, artificial intelligence, and big data analysis, the implementation of high-precision prediction has become more feasible, providing strong technical support for the future development of PV power generation.

Predicting PV power under varying meteorological conditions plays an important role in ensuring the reliability of energy supply, improving energy utilization efficiency, promoting the widespread application of renewable energy, and supporting sustainable development [5]. In existing research, under varying meteorological conditions, PV power prediction can be divided into physical modeling methods, statistical prediction methods, probability prediction methods, and intelligent prediction methods. These methods each have their own advantages and disadvantages [6]. The physical modeling method is suitable for building new PV power stations, but the model is relatively complex. Statistical prediction methods do not require a large amount of historical data but are sensitive to changes in meteorological environmental factors [7]. Intelligent prediction methods have high accuracy but may require a large amount of historical data and computational resources. With the advancement of technology, these methods are constantly being integrated and optimized to adapt to changing meteorological conditions and improve the accuracy of predictions. The main work and contributions of this study are as follows.

(1) Quantitative analysis of meteorological impacts: This study introduces the Pearson coefficient quantitative analysis framework, which carefully examines the complex relationship between meteorological indicators and photovoltaic power generation under different conditions (including sunny and cloudy). The analysis provides a deeper insight into the differences in photovoltaic performance under different weather patterns.
(2) Advanced optimization algorithm: This study proposes an innovatively enhanced and improved beluga whale optimization (EIBWO) algorithm, which demonstrates superior optimization and convergence capabilities compared to existing optimization algorithms. EIBWO has better optimization ability and convergence accuracy, which can efficiently solve optimization problems and help achieve more precise and efficient parameter adjustment processes for ELM.
(3) High-precision prediction model: By integrating EIBWO with ELM, this study develops a new output power prediction model. Whether under sunny or cloudy conditions, the coefficient of determination ($R^2$) of EIBWO-ELM is above 0.99, indicating its accuracy in predicting photovoltaic power generation. This level of accuracy represents further development in the predictability and reliability of photovoltaic energy systems.

This study not only has a profound impact on the integration and management of renewable energy in the wider power grid but also contributes to the global shift towards sustainable energy solutions. The remaining structural arrangements are as follows: Section 2 provides a literature review on existing research methods for predicting PV power. Section 3 elaborates on the method proposed by this study. Section 4 establishes a PV power generation prediction model. Section 5 validates and analyzes the PV prediction results. The discussion of this study is presented in Section 6. Finally, the concluding remarks are provided in Section 7.

## 2. Related Works

### *2.1. Photovoltaic Power Prediction*

PV output power exhibits obvious non-stationary random characteristics due to the close correlation between PV output and many influencing factors [8]. The fluctuation of PV

power significantly reduces the safety and reliability of energy system operation for energy systems containing PV power generation [9,10]. To cope with the uncertainty of PV output, the system needs to reserve a certain amount of spare capacity to maintain power balance and ensure safe and reliable operation of the system. Therefore, accurate prediction of the power output of PV power generation is a key means to effectively dealing with the inherent randomness and instability of PV grid connection [11]. Due to the nonlinearity, indirect volatility, and uncertainty of power output in PV power generation systems, improving the characterization ability of existing models and establishing high-precision PV power prediction models is a major challenge [12,13]. At present, domestic and foreign scholars have conducted extensive and in-depth research on PV power prediction methods. The classification of PV power prediction methods based on different classification criteria is shown in Figure 1 [14,15].

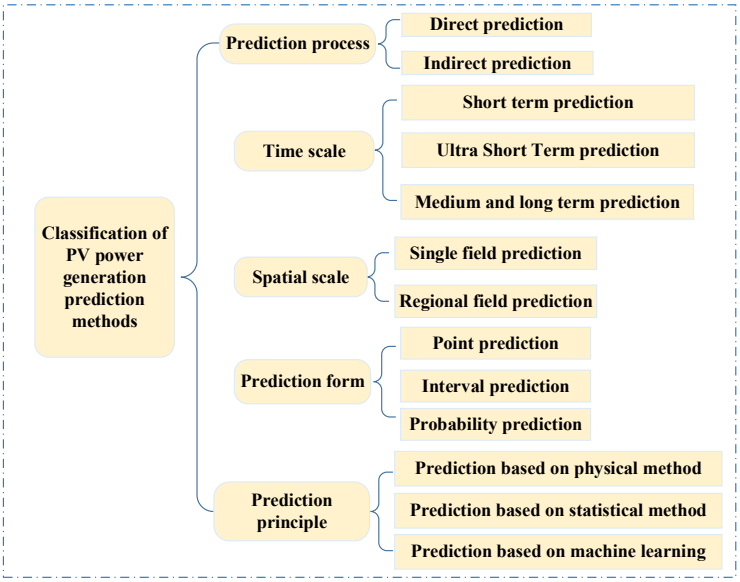


**Figure 1.** Classification of PV power prediction methods.

As shown in Figure 1, according to the prediction principle, PV power prediction is mainly divided into physical methods, statistical methods, and machine learning methods. Ref. [16] proposed a PV power plant prediction method using physical hybrid neural networks, aiming to solve the problem of integrating renewable energy in smart grids. Based on statistical methods, Ref. [17] combined the autoregressive moving average with nonlinear autoregressive models to establish a hybrid model for effective prediction of solar radiation. With the continuous deepening of scientific research, domestic and foreign researchers have proposed many machine learning-based prediction models and applied them to the field of PV power generation prediction, achieving good research results [18]. Machine learning models have strong nonlinear mapping and generalization abilities [19].

These methods train the model based on historical data and establish complex mapping relationships between input and output features. Finally, the trained model is used to predict photovoltaic power. Although machine learning models have many advantages in the field of PV power prediction, model training and testing are sensitive to the setting of random parameters [20]. Ref. [21] used seven machine learning models to predict the power generation of PV plants. The above seven machine learning models can achieve the fitting of PV power generation. Ref. [22] used an inheritance machine learning model for PV power prediction. The results were verified by actual measurement data of PV power plants, indicating that the proposed model has good predictive performance.

Existing studies usually use artificial intelligence algorithms to address the problem of random parameter perturbations in machine learning models and improve the predictive performance of the model through parameter optimization.

### *2.2. Enhanced and Improved Beluga Whale Optimization Extreme Learning Machine*

Common machine learning models include support vector machine (SVM) [23], long short-term memory [24], back propagation neural network (BPNN) [25], and so on. SVM is a small-sample learning method suitable for processing samples with small amounts of data, so it is widely used in the field of ultra-short-term and short-term PV power generation power prediction. BPNN has strong nonlinear mapping ability, high self-learning ability, adaptive ability, and fault tolerance. However, traditional BPNNs are slightly inferior in prediction accuracy and need to be improved through optimization.

ELM belongs to an improved neural network model [26]. ELM has been widely applied in multiple prediction research fields due to its unique advantages of fast learning speed and good generalization ability. In financial market analysis, ELM can be used to predict stock prices and market trends [27]. A predictive model for financial market system risk based on ELM was constructed and the stability of the model was verified through examples. In terms of healthcare, ELM can help predict the development of diseases and the recovery process of patients [28]. ELM has also achieved results in the field of traffic flow prediction. Ref. [29] used ELM to predict traffic flow data and conducted comparative analysis using different time flows. In the above studies, ELM has demonstrated its potential and effectiveness in processing complex data and solving practical problems.

In this study, ELM was applied to the method of power generation prediction of photovoltaic energy. Ref. [30] applied ELM to practical case studies of PV power plants. After training the ELM model, it was used to predict power generation. The experimental results indicate that the prediction results of ELM were superior to those of BP. However, the connection weights and thresholds of ELM were randomly generated, which affected the predictive ability of ELM [31]. Therefore, Ref. [32] constructed a new prediction model for predicting power generation output under various meteorological conditions. The results show that the model had better prediction performance than traditional ELM.

PV output power prediction is relatively complex, mainly due to the difficulty in controlling the uncertainty of PV power [33]. If a model is directly used for prediction, it is difficult to obtain ideal prediction results. Therefore, scholars have integrated different algorithms to improve the model's ability to characterize the uncertainty of PV output power [34]. Beluga whale optimization (BWO) is a new type of artificial intelligence optimization algorithm with good convergence ability [35]. Ref. [36] proposed a high-performance improved beluga whale optimization algorithm to optimize the DG multi-objective hierarchical optimization planning model. The proposed method effectively solved the nonlinear problem of the model. Ref. [37] proposed a multi-strategy to improve the beluga whale optimization method. In addition, the IEEE CEC benchmark test function verifies that the proposed method can be used to solve engineering optimization problems. Therefore, the chaotic mapping strategy, sine dynamic adaptive factor, and disturbance strategy are introduced to BWO in this study, and the EIBWO algorithm is proposed to optimize machine learning models.

## 3. Method

### *3.1. Beluga Whale Optimization Algorithm*

BWO is a population-based mechanism; therefore, each beluga whale is a candidate solution that is updated during the optimization process. The matrix of the search agent location is modeled as shown in Equation (1).

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d} \end{bmatrix} \quad (1)$$

where $n$ is the population size, and $d$ is the dimension. In addition, the corresponding fitness values are as follows in Equation (2).

$$F_X = \begin{bmatrix} f(x_{1,1}, x_{1,2}, \cdots x_{1,d}) \\ f(x_{2,1}, x_{2,2}, \cdots x_{2,d}) \\ \vdots \\ f(x_{n,1}, x_{n,2}, \cdots x_{n,d}) \end{bmatrix} \tag{2}$$

BWO can transition from exploration to development, depending on the balance factor $B_f$. The mathematical model is shown in Equation (3).

$$B_f = B_0(1 - \frac{t}{2T}) \tag{3}$$

where $B_0$ randomly changes between (0, 1) during each iteration. $T$ and $t$ are the maximum and current iteration, respectively. If $B_f > 0.5$, beluga individuals are in the exploration stage. If $B_f \leq 0.5$, beluga individuals are in the development stage.

During the exploration phase, the individual position of the beluga whale is determined by paired swimming, and the updated position is as follows in Equation (4).

$$\begin{cases} X_{i,j}^{t+1} = X_{i,p_j}^t + (X_{r,p_1}^t - X_{i,p_j}^t)(1 + r_1)\sin(2\pi r_2), j = even \\ X_{i,j}^{t+1} = X_{i,p_j}^t + (X_{r,p_1}^t - X_{i,p_j}^t)(1 + r_1)\cos(2\pi r_2), j = odd \end{cases} \tag{4}$$

where $P_j(j = 1, 2, \cdots, d)$ is randomly selected in the dimension of $d$. $X_{i,P_j}^t$ is the $i_{\text{th}}$ beluga whale's new position in the $P_j$ dimension. $X_{i,P_j}^t$ and $X_{r,P_1}^t$ are the $i_{\text{th}}$ beluga whale's and the $r_{\text{th}}$ beluga whale's current positions, respectively. The updated position reflects the synchronized or mirrored behavior of the beluga whale during swimming or diving based on the odd and even selected dimensions. $\sin(2\pi r_2)$ and $\cos(2\pi r_2)$ are used to average the random numbers between fish fins.

Beluga whales hunt their prey by sharing information about their location with each other. Levy flight strategy was introduced in the development phase of BWO to enhance convergence, as shown by Equation (5).

$$X_i^{t+1} = r_3 X_{best}^t - r_4 X_i^t + C_1 \cdot L_F \cdot (X_r^t - X_i^t) \tag{5}$$

where $C_1 = 2r_4(1 - t/T_{\max})$. $C_1$ is used to measure the random jump intensity of Levy flight intensity.

$L_F$ is the Levy flight function, and is expressed in Equations (6) and (7).

$$L_F = 0.05 \times \frac{u \times \sigma}{|v|^{1/\beta}} \tag{6}$$

$$\sigma = \left( \frac{\Gamma(1 + \beta) \times \sin(\pi\beta/2)}{\Gamma((1 + \beta)/2) \times \beta \times 2^{(\beta-1)/2}} \right)^{1/\beta} \tag{7}$$

where $u$ and $v$ are normally distributed random numbers. $\beta$ is a constant with a value of 1.5.

To simulate the behavior of whales falling in each iteration and ensure that the population remains unchanged, the updated position is established using the position of the beluga whale and the step size of whale descent, which is represented in Equation (8).

$$X_i^{t+1} = r_5 X_i^t - r_6 X_r^t + r_7 X_{step} \tag{8}$$

where $X_{step}$ is the step size of a whale falling, which is expressed in Equation (9).

$$X_{step} = (u_b - l_b)\exp(-C_2 \frac{t}{T}) \tag{9}$$

where $C_2 = 2W_f \times n$, $C_2$ is a step factor related to the probability of whale decline and population size. $u_b$ and $l_b$ represent the upper and lower bounds of the variable, respectively.

The probability of a whale falling $W_f$ is calculated as a linear function, and is shown in Equation (10).

$$W_f = 0.1 - 0.05t/T \tag{10}$$

*3.2. EIBWO Proposal and Validation*

During the exploration and development stages of BWO, there was a lack of search capability, resulting in low population diversity and decreased solution accuracy. Based on the shortcomings of the original BWO, this study proposes an enhanced and improved beluga whale optimization (EIBWO). The improvement measures for EIBWO are as follows:

(1) Initialization of beluga whale population based on chaotic mapping strategy

To obtain high-quality positions of the first-generation population, accelerate the convergence speed of EIBWO, and reduce the computational cost, this study uses the randomness and traversal of the chaotic mapping strategy to obtain the positions of the first-generation population. The details are shown in Equations (11) and (12).

$$P_S = P_S^{max} + r_M \times (P_S^{max} - P_S^{min}) \tag{11}$$

$$r_M = \alpha \times r_M \times (1 - r_M) \tag{12}$$

where $r_M$ represents the scale of chaotic mapping. When $\alpha$ = 4, the model is in a chaotic state. $P_S^{min}$ and $P_S^{max}$ represent the minimum and maximum boundaries of the search space, respectively.

(2) Sine dynamic adaptive factor

To improve the local search capability of EIBWO, a sine adaptive factor is added to the position update formula, and its expression is defined in Equation (13).

$$S = 1 + \sin\frac{\pi(2T + t)}{2T} \tag{13}$$

By introducing the sine dynamic adaptive factor $S$ into Equation (4), the improved position update equation can be obtained as follows in Equation (14).

$$\begin{cases} X_{i,j}^{t+1} = S * X_{i,p_j}^t + (X_{r,p_1}^t - X_{i,p_j}^t)(1 + r_1)\sin(2\pi r_2), j = even \\ X_{i,j}^{t+1} = S * X_{i,p_j}^t + (X_{r,p_1}^t - X_{i,p_j}^t)(1 + r_1)\cos(2\pi r_2), j = odd \end{cases} \tag{14}$$

(3) Individual position disturbance strategy in the population of beluga whales

To fully ensure that the population maintains high diversity during the optimization process, a random position perturbation strategy is proposed. The detailed improvement strategy is as follows: During the iteration process of EIBWO, the disturbance frequency $f_r$ is set to a random number $r_p$ within the [0, 1] interval, and the values of $f_r$ and $r_p$ are compared to determine whether the population position has been disturbed.

Specifically, when $f_r \leq r_p$, individual positions are updated according to Equation (14); when $f_r > r_p$, individual positions are updated according to Equations (15) and (16).

$$X_S(x + 1) = a_w \times X_S(x) + randn(0, \sigma^2) \times (X_{best} - X_S(x)) \times e^{f_{best} - f_P} \tag{15}$$

$$a_w = a_{wmax} \times e^{-35 \times (\frac{t}{M_{iter}})^5} + a_{wmin} \tag{16}$$

where $a_w$ is convergence coefficient and is used to guide individuals towards convergence direction, with a value range of [0.3, 0.7]. *randn* (0, $\sigma^2$) follows the Gaussian distribution of a mean of 0 and a variance of $\sigma^2$.

To verify the convergence and optimization effect of EIBWO, the traditional BWO algorithm, multi-verse optimization (MVO), the seagull optimization algorithm (SOA), and particle swarm optimization (PSO) were selected as comparative algorithms. In the same testing environment, this study selected 6 benchmark testing functions to test all algorithms. The detailed benchmark testing functions are shown in Table 1. Some algorithms have unique parameters, and their set values are shown in Table 2.

**Table 1.** Benchmark functions.

| Function | Range | $F_{min}$ |
|---|---|---|
| $F_1(x) = \sum_{i=1}^{n} {x_i}^2$ | $[-100, 100]$ | 0 |
| $F_2(x) = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=1}^{n} \lvert x_i \rvert$ | $[-10, 10]$ | 0 |
| $F_3(x) = \sum_{i=1}^{n} \left( \sum_{j-1}^{i} x_j \right)^2$ | $[-100, 100]$ | 0 |
| $F_4(x) = \max_i \{ \lvert x_i \rvert, 1 \le i \le n \}$ | $[-100, 100]$ | 0 |
| $F_6(x) = \sum_{i=1}^{n} \left[ {x_i}^2 - 10\cos(2\pi x_i) + 10 \right]$ | $[-5.12, 5.12]$ | 0 |
| $F_7(x) = -20\exp(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}) - \exp(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)) + 20 + e$ | $[-32, 32]$ | 0 |

**Table 2.** Parameter setting in some algorithms.

| Algorithm | Value | Iteration | Population Size | Dimension |
|---|---|---|---|---|
| MVO | $V_{max} = 1$, $V_{min} = 0.2$ | 300 | 50 | 30 |
| PSO | $C_{max} = 0.9$, $C_{min} = 0.4$ | 300 | 50 | 30 |
| | $C_1 = C_2 = 2$ | 300 | 50 | 30 |
| SOA | $d = 2$ | 300 | 50 | 30 |
| | $u = 1$, $v = 1$ | 300 | 50 | 30 |

As shown in Table 2, the $V_{max}$ and $V_{min}$ are the wormhole existence rates in MVO. $C_{max}$ and $C_{min}$ are the inertia factors in PSO. $C_1$ and $C_2$ are the acceleration constants in PSO. $d$ is the control factor in SOA. The population size was set to 50. The number of iterations for all algorithms was set to 300. The dimension was set to 30.

To ensure the objectivity of testing, each test function was tested 50 times for each algorithm, calculating the average value and standard deviation of the convergence values after each run. The algorithm test results under different test functions are shown in Table 3.

**Table 3.** Convergence values of algorithms under different test functions.

| Algorithm | $F_1$ | | $F_2$ | | $F_3$ | |
|---|---|---|---|---|---|---|
| | *Ave.* | *Std.* | *Ave.* | *Std.* | *Ave.* | *Std.* |
| EIBWO | $1.45 \times 10^{-228}$ | 0 | $2.44 \times 10^{-196}$ | 0 | 0 | 0 |
| BWO | $3.96 \times 10^{-155}$ | $2.08 \times 10^{-154}$ | $9.96 \times 10^{-80}$ | $5.96 \times 10^{-79}$ | $4.89 \times 10^{-166}$ | 0 |
| SOA | $2.76 \times 10^{-6}$ | $2.94 \times 10^{-6}$ | $4.41 \times 10^{-5}$ | $2.51 \times 10^{-5}$ | $6.23 \times 10^{-2}$ | 0.18 |
| PSO | $1.33 \times 10^{-14}$ | $2.59 \times 10^{-14}$ | $3.22 \times 10^{-8}$ | $3.14 \times 10^{-8}$ | $9.95 \times 10^{-5}$ | $1.20 \times 10^{-4}$ |
| MVO | 2.04 | 0.59 | 12.13 | 31.49 | 314.63 | 143.66 |
| Algorithm | $F_4$ | | $F_5$ | | $F_6$ | |
| | *Ave.* | *Std.* | *Ave.* | *Std.* | *Ave.* | *Std.* |
| EIBWO | $1.19 \times 10^{-190}$ | 0 | 0 | 0 | $8.88 \times 10^{-16}$ | 0 |
| BWO | $2.80 \times 10^{-146}$ | $1.86 \times 10^{-145}$ | 0 | 0 | $8.88 \times 10^{-16}$ | 0 |
| SOA | $7.66 \times 10^{-2}$ | 0.12 | 7.34 | 9.01 | 19.96 | $1.48 \times 10^{-3}$ |
| PSO | $1.45 \times 10^{-4}$ | $1.38 \times 10^{-4}$ | 5.69 | 2.66 | $8.21 \times 10^{-8}$ | $8.34 \times 10^{-8}$ |
| MVO | 0.10 | 0.03 | 125.71 | 36.56 | 2.06 | 0.53 |

From the statistical results in Table 3, it can be seen that EIBWO achieved the optimal convergence value in the case of unimodal test function $F_3$ and multimodal test function $F_5$. Although EIBWO did not obtain the optimal value in other test functions, it can be seen from the standard deviation of the test results being 0 that the EIBWO algorithm had good optimization robustness and obtained stable convergence results.

### *3.3. Extreme Learning Machine*

Due to its excellent learning ability, single-hidden-layer feed-forward neural network (SLFN) is widely used in fields such as lifespan prediction and pattern recognition. However, there are some inherent problems with traditional SLFN, such as the need to use the gradient descent method for multiple iterations during the training process to complete the correction of network thresholds and weights, the sensitivity of SLFN to the selection of learning rate, and the long training time of the network.

To address the issues in traditional SLFN, scholars have developed a new network model—extreme learning machine (ELM). Compared to traditional SLFN, ELM has a faster learning speed and stronger generalization ability. The network structure of ELM is shown in Figure 2 [38,39].

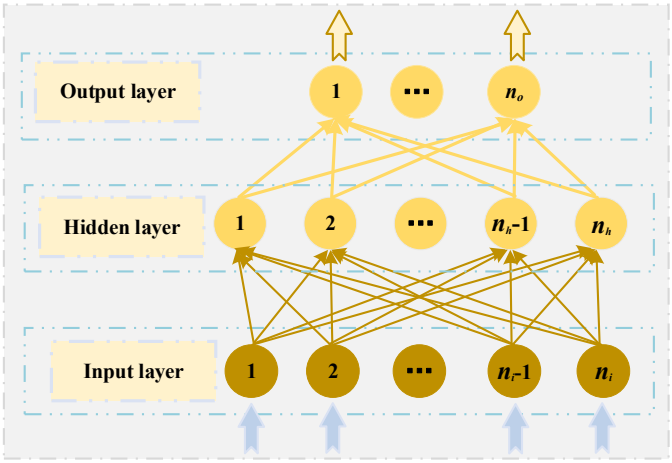


**Figure 2.** The network structure of ELM.

As shown in Figure 2, ELM consists of an input layer, a hidden layer, and an output layer, with neurons in each layer connected in sequence.

It is assumed that the number of neurons in the input layer, hidden layer, and output layer is $n_i$, $n_h$, and $n_o$, respectively. The corresponding number of input and output variables is $n_i$ and $n_o$, respectively. The connection weight matrix between the input layer and the hidden layer is $\boldsymbol{w^i}$, as shown in Equation (17).

$$\boldsymbol{w^i} = \begin{bmatrix} w^i_{11} & w^i_{12} & \cdots & w^i_{1n_i} \\ w^i_{21} & w^i_{22} & \cdots & w^i_{2n_i} \\ \vdots & \vdots & \vdots & \vdots \\ w^i_{n_h1} & w^i_{n_h2} & \cdots & w^i_{n_hn_i} \end{bmatrix}_{n_h \times n_i} \tag{17}$$

where $w^i{}_{jk}$ is the connection weight between the $j_{\text{th}}$ neuron in the hidden layer and the $k_{\text{th}}$ neuron in the input layer.

It is assumed that the connection weight matrix between the output layer and the hidden layer is $\boldsymbol{w^o}$ according to Equation (18).

$$\boldsymbol{w^o} = \begin{bmatrix} w^o_{11} & w^o_{12} & \cdots & w^o_{1n_o} \\ w^o_{21} & w^o_{22} & \cdots & w^o_{2n_o} \\ \vdots & \vdots & \vdots & \vdots \\ w^o_{n_h1} & w^o_{n_h2} & \cdots & w^o_{n_hn_o} \end{bmatrix}_{n_h \times n_o} \tag{18}$$

where $w^o{}_{jk}$ *is* the connection weight between the $j_{\text{th}}$ neuron in the hidden layer and the $k_{\text{th}}$ neuron in the output layer.

It is assumed that the hidden layer neuron threshold matrix is $\boldsymbol{q}$, as shown in Equation (19).

$$\boldsymbol{q} = \left[q_1, q_2, \ldots, q_{n_h}\right]^T \tag{19}$$

For a training set containing $n_m$ samples, the input matrix $\boldsymbol{X}^*$ and output matrix $\boldsymbol{Y}^*$ are as shown in Equations (20) and (21).

$$\boldsymbol{X}^* = \begin{bmatrix} x_{11}^* & x_{12}^* & \cdots & x_{1n_m}^* \\ x_{21}^* & x_{22}^* & \cdots & x_{2n_m}^* \\ \vdots & \vdots & \vdots & \vdots \\ x_{n_i1}^* & x_{n_i2}^* & \cdots & x_{n_in_m}^* \end{bmatrix}_{n_i \times n_m} \tag{20}$$

$$\boldsymbol{Y}^* = \begin{bmatrix} y_{11}^* & y_{12}^* & \cdots & y_{1n_m}^* \\ y_{21}^* & y_{22}^* & \cdots & y_{2n_m}^* \\ \vdots & \vdots & \vdots & \vdots \\ y_{n_o1}^* & y_{n_o2}^* & \cdots & y_{n_on_m}^* \end{bmatrix}_{n_o \times n_m} \tag{21}$$

$G(\cdot)$ is the activation function of the hidden layer; then, the output matrix $O$ is as follows in Equations (22) and (23).

$$\boldsymbol{O} = \left[o_1,\ o_2, \ldots,\ o_{n_o}\right]_{n_o \times n_m} \tag{22}$$

$$\boldsymbol{o}_j = \begin{bmatrix} o_{1j} \\ o_{2j} \\ \vdots \\ o_{n_mj} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{n_h} w_{k1}^o G(w_k^i x_j^* + q_k) \\ \sum_{k=1}^{n_h} w_{k2}^o G(w_k^i x_j^* + q_k) \\ \vdots \\ \sum_{k=1}^{n_h} w_{kn_o}^o G(w_k^i x_j^* + q_k) \end{bmatrix} \tag{23}$$

where $\boldsymbol{w}_k^{\boldsymbol{i}} = \left[w_{k1}^i, \quad w_{k2}^i, \quad \cdots, \quad w_{kn_i}^i\right]^{\mathrm{T}}, \boldsymbol{x}_j^* = \left[x_{1j}^*, \quad x_{2j}^*, \quad \cdots, \quad x_{n_ij}^*\right]^{\mathrm{T}}$.

The above equation can be expressed in Equation (24).

$$\boldsymbol{HY^*} = \boldsymbol{O}^{\mathrm{T}} \tag{24}$$

where $\boldsymbol{H}$ is the output matrix of the hidden layer, as shown in Equation (25).

$$\boldsymbol{H} = \begin{bmatrix} G(\boldsymbol{w}_1^{\boldsymbol{i}} \cdot \boldsymbol{x}_1^* + q_1) & G(\boldsymbol{w}_2^{\boldsymbol{i}} \cdot \boldsymbol{x}_1^* + q_2) & \cdots & G(\boldsymbol{w}_{n_h}^{\boldsymbol{i}} \cdot \boldsymbol{x}_1^* + q_{n_h}) \\ G(\boldsymbol{w}_1^{\boldsymbol{i}} \cdot \boldsymbol{x}_2^* + q_1) & G(\boldsymbol{w}_2^{\boldsymbol{i}} \cdot \boldsymbol{x}_2^* + q_2) & \cdots & G(\boldsymbol{w}_{n_h}^{\boldsymbol{i}} \cdot \boldsymbol{x}_2^* + q_{n_h}) \\ \vdots & \vdots & \vdots & \vdots \\ G(\boldsymbol{w}_1^{\boldsymbol{i}} \cdot \boldsymbol{x}_{n_m}^* + q_1) & G(\boldsymbol{w}_2^{\boldsymbol{i}} \cdot \boldsymbol{x}_{n_m}^* + q_2) & \cdots & G(\boldsymbol{w}_{n_h}^{\boldsymbol{i}} \cdot \boldsymbol{x}_{n_m}^* + q_{n_h}) \end{bmatrix}_{n_m \times n_h} \tag{25}$$

In ELM, if the number of hidden-layer neurons is equal to the number of samples in the training set, it is possible to achieve zero error approximation of the network output to the training set. The details are expressed in Equation (26).

$$\sum_{j=1}^{n_m} \left\| \boldsymbol{o}_j - \boldsymbol{y}_j^* \right\| = 0 \tag{26}$$

where $\boldsymbol{y}_j^* = \left[y_{1j}^*, \quad y_{2j}^*, \quad \cdots, \quad y_{n_oj}^*\right]^{\mathrm{T}} (j = 1, 2, \ldots, n_m)$.

When there is a large number of samples in the training set, in order to reduce computational costs, the number of hidden-layer neurons is usually smaller than the

number of samples in the training set. At this point, the network training error can approach an arbitrary value $\varepsilon$, as shown in Equation (27).

$$\sum_{j=1}^{n_m} \left\| \boldsymbol{o}_j - \boldsymbol{y}_j^* \right\| < \varepsilon \tag{27}$$

The least-squares solution in the following Equation (28) can be solved to obtain $q$.

$$\min_{q} \left\| \boldsymbol{H}\boldsymbol{q} - \boldsymbol{O}^{\mathrm{T}} \right\| \tag{28}$$

The solution is as shown in Equation (29).

$$\hat{q} = \boldsymbol{H}^{+}\boldsymbol{O}^{\mathrm{T}} \tag{29}$$

where $\boldsymbol{H}^{+}$ is the Moore Penrose generalized inverse matrix of $\boldsymbol{H}$.

ELM is the basic model for PV prediction selected in this study. Similar to other machine learning models, the performance of ELM is also affected by random parameters. Considering the strong randomness and volatility of PV power, there is a high requirement for the regression ability of the model. In the process of predicting PV power, improper selection of random parameter values will directly affect the regression effect of ELM. To accurately characterize the uncertainty of PV power in ELM, EIBWO is proposed to optimize the connection weight and threshold in ELM, thereby improving the predictive performance of ELM and enhancing the accuracy of the model in predicting PV power.

## 4. Establishment of PV Power Generation Prediction Model

### *4.1. Selection of Input Variables for Prediction Model*

Due to strong temporal variability, meteorological factors have the most significant interference on photovoltaic power generation. Therefore, when constructing a PV power generation prediction model, meteorological factors are selected as the input, and the output is PV power. However, different meteorological factors have varying degrees of impact on photovoltaic power. It is necessary to measure the degree of correlation between various factors and power generation, select the meteorological factors that have the greatest impact on power generation, and use them as input features for the model.

To quantitatively describe the impact of various meteorological factors on power generation, Pearson coefficients and correlation coefficients are used to comprehensively measure the correlation between various meteorological factors and PV power generation. The Pearson coefficient judgment criteria are shown in Table 4.

**Table 4.** The judgment criteria of the Pearson coefficient.

| Range of $\rho$ | Degree of Association |
|---|---|
| 0.0–0.2 | Extremely weakly correlated or uncorrelated |
| 0.2–0.4 | Weak correlation |
| 0.4–0.6 | Moderate correlation |
| 0.6–0.8 | Strong correlation |
| 0.8–1.0 | Strongly correlated |

The Pearson correlation coefficient $\rho$ is expressed in Equation (30).

$$\rho = \frac{N\sum XY - \sum X \sum Y}{\sqrt{N\sum X^2 - (\sum X)^2}\sqrt{N\sum Y^2 - (\sum Y)^2}} \tag{30}$$

where $N$ represents the number of calculated samples, and $X$ and $Y$ represent the two variables that need to be validated for correlation, namely, the input and output variables of the predictive model, respectively.

Due to the impact of weather conditions on PV power generation, it is recommended to analyze the correlation between different meteorological indicators and PV power generation under sunny and cloudy weather conditions.

The meteorological indicators studied include radiation intensity, environmental temperature, and relative humidity. The relationship curves between output power and different meteorological indicators on sunny and cloudy days are shown in Figure 3.

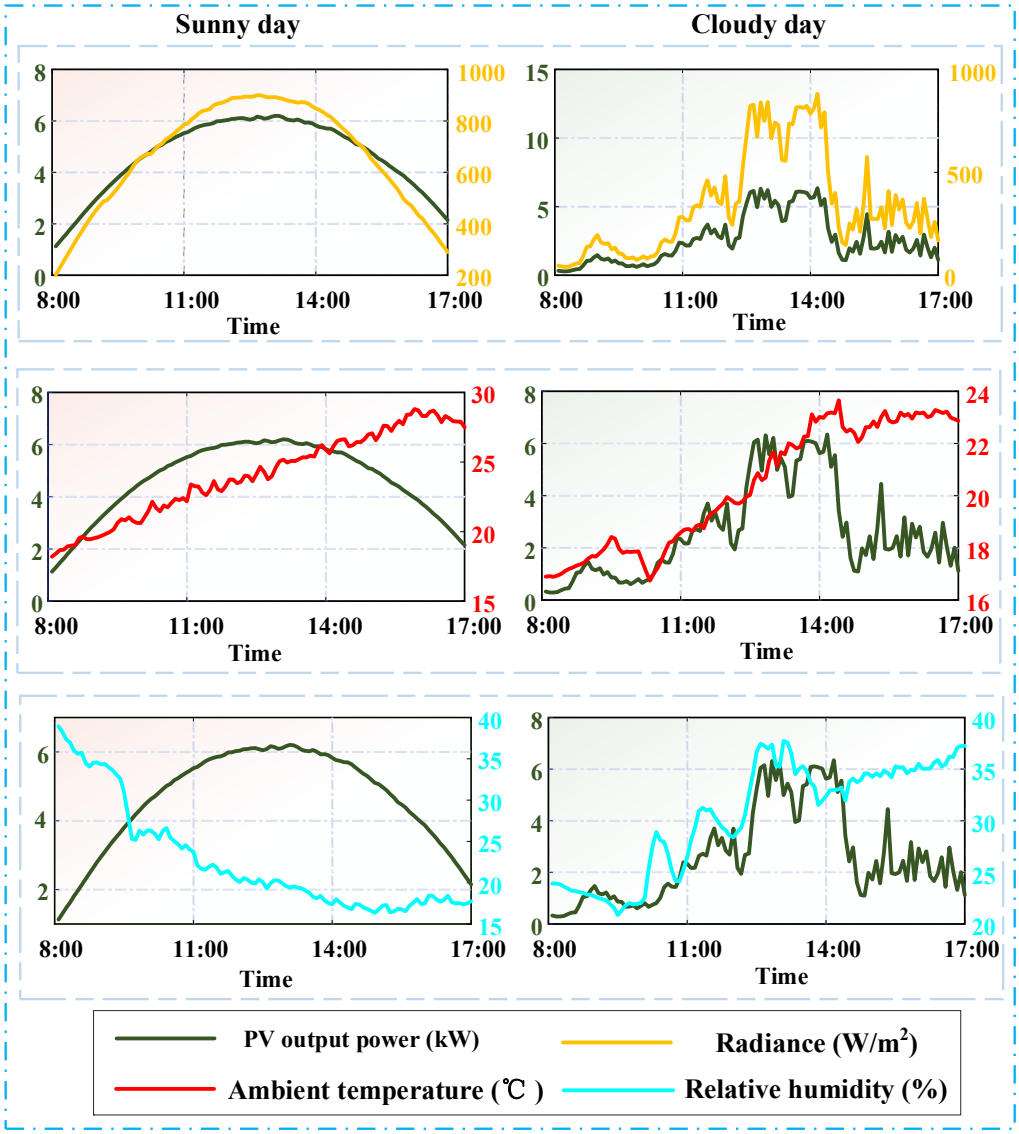


**Figure 3.** The relationship curves between output power and different meteorological indicators.

From Figure 3, it can be seen that under both sunny and cloudy conditions, only the radiation intensity shows a significant correlation with the output power of photovoltaic power generation. The output power changes in real time with the radiation intensity curve, while the relationship between environmental temperature, relative humidity, and output power cannot be clearly obtained from Figure 3. Therefore, the Pearson correlation coefficient was used to calculate the correlation between the three meteorological indicators and the output power of photovoltaic power generation. The values between relative humidity and output power, radiation intensity and output power, and ambient temperature and output power were calculated, as shown in Table 5.

**Table 5.** Correlation coefficient values between different meteorological indexes and output power on sunny and cloudy days.

| Meteorological Indexes | Sunny | Cloudy |
|---|---|---|
| Radiance | 0.99 | 0.99 |
| Ambient temperature | 0.27 | 0.53 |
| Relative humidity | −0.60 | 0.62 |

According to Table 5, regardless of the weather conditions, the correlation coefficient between radiation intensity and output power is very close to 1, indicating a strong correlation between the two. The correlation coefficient between ambient temperature and output power is 0.27 on sunny days, showing a weak correlation, and 0.53 on cloudy days, showing a moderate correlation. The correlation coefficient between relative humidity and output power is −0.60 on sunny days, showing a moderate degree of correlation, and 0.62 on cloudy days, showing a strong correlation.

In summary, regardless of the weather conditions, the three meteorological factors of radiance, ambient temperature, and relative humidity all show a certain degree of correlation with output power. Therefore, all can be used as input variables for the model in predicting photovoltaic power generation.

### *4.2. Evaluation Indicators*

It is difficult for a single evaluation indicator to effectively evaluate the predictive performance of a model, so three indicators were selected to comprehensively measure the predictive performance of the model. The evaluation indicators used are root mean square error (RMSE), mean absolute percentage error (MAPE), and goodness of fit ($R^2$), and are shown in Equations (31) and (33).

$$RMSE = \sqrt{\frac{1}{N_{sample}} \sum_{i=1}^{N_{sample}} (pre_i - act_i)^2} \tag{31}$$

$$MAPE = \frac{100}{N_{sample}} \times \sum_{i=1}^{N_{sample}} \left| \frac{pre_i - act_i}{act_i} \right| \tag{32}$$

$$\begin{cases} R^2 = \dfrac{\sum_{i=1}^{N_{sample}} (pre_i - act_{mean})^2}{\sum_{i=1}^{N_{sample}} (act_i - act_{mean})^2} \\ act_{mean} = \dfrac{1}{N_{sample}} \sum_{i=1}^{N_{sample}} act_i \end{cases} \tag{33}$$

where, $N_{sample}$ is the number of samples, $pre_i$ is the predicted value, and $act_i$ is the real value.

### *4.3. PV Power Generation Prediction Model Based on EIBWO-ELM*

In this study, the enhanced and improved beluga whale optimization algorithm is proposed to optimize ELM. A PV power generation prediction model based on enhanced and improved beluga whale optimization algorithm–optimization extreme learning machine (EIBWO-ELM) is established. The flowchart of EIBWO-ELM is as shown in Figure 4.

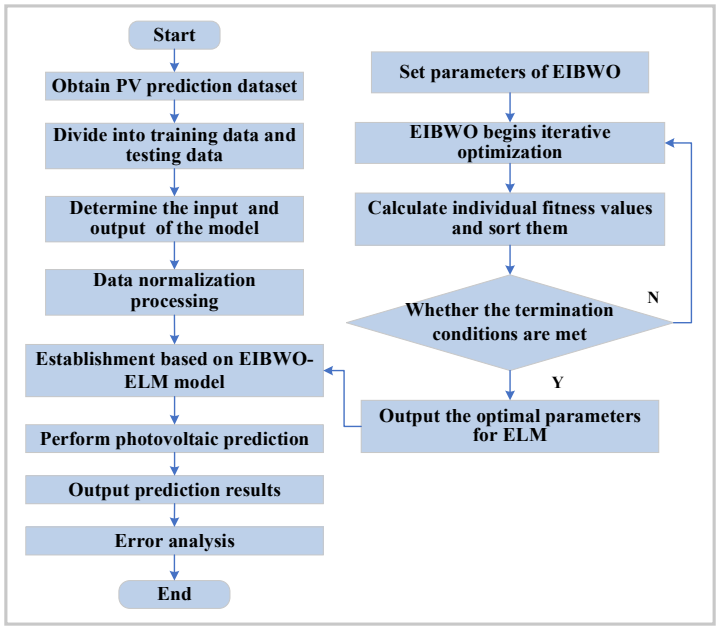


**Figure 4.** The flowchart of the prediction model based on EIBWO-ELM.

Step 1: Classify the photovoltaic data into a training dataset and a testing dataset.
Step 2: Determine the input and output of the prediction model.
Step 3: Process data normalization according to Equation (34).

$$A_N = \frac{A - A_{\min}}{A_{\max} - A_{\min}} \tag{34}$$

where $A$ represents the variable values to be normalized, $A_{\max}$ is the maximum value, and $A_{\min}$ is the minimum value.

Step 4: Initialize the EIBWO algorithm population and calculate its fitness value. In EIBWO-ELM, the variance between the predicted photovoltaic value and the actual value is defined as the fitness value of EIBWO, according to Equation (35).

$$Fit = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} (pre_i - act_i)^2 \tag{35}$$

where $N_{test}$ is the number of samples in the test set.

Step 5: Train the model using training data, and at this point, the EIBWO algorithm begins iterative optimization.

Step 6: Determine whether the termination condition is met. If it is, output the optimal parameters of ELM. Otherwise, return to the previous step.

Step 7: Use test data for testing to obtain the predicted results of photovoltaic power.

Step 8: Evaluate and analyze the predicted results.

## 5. Analysis of Photovoltaic Power Generation Prediction Results

### *5.1. Source and Introduction of Experimental Data*

The photovoltaic power generation data were sourced from publicly available data from the Desert Knowledge Australia Solar Center (DKASC), which records the daily photovoltaic power output, relative humidity, ambient temperature, radiance, and rainfall of the Alice Spring photovoltaic power station, recorded every 5 min.

Due to the susceptibility of photovoltaic power generation to weather conditions, multiple weather conditions such as sunny and cloudy should be considered when making predictions. The five days between 14 August 2021 and 18 August 2022 were sunny. The data during this period were used as research data for sunny days. The research data from

14 August 2022 to 17 August 2022 were used as training data, and the research data from 18 August 2022 were used as testing data. The five days from 7 August 2021 to 11 August 2022 were cloudy. The data during this period were used as research data for cloudy weather. The research data from 7 August 2022 to 10 August 2022 were used as training data, and the research data from 11 August 2022 were used as testing data. The above daily research data are from 8:00 to 17:00, and there are 108 data sample points per day.

### *5.2. Case 1—Sunny Days*

Case 1 is the prediction of PV power generation on sunny days. The EIBWO-ELM, BWO-ELM, PSO-ELM, and ELM prediction models established in this study were used to predict the output power of PV power generation on sunny days. The four prediction models constructed in this study, EIBWO-ELM, BWO-ELM, PSO-ELM, and ELM were used to predict the output power of PV power generation. The actual output power values and the predicted values obtained after model prediction are shown in Figure 5.

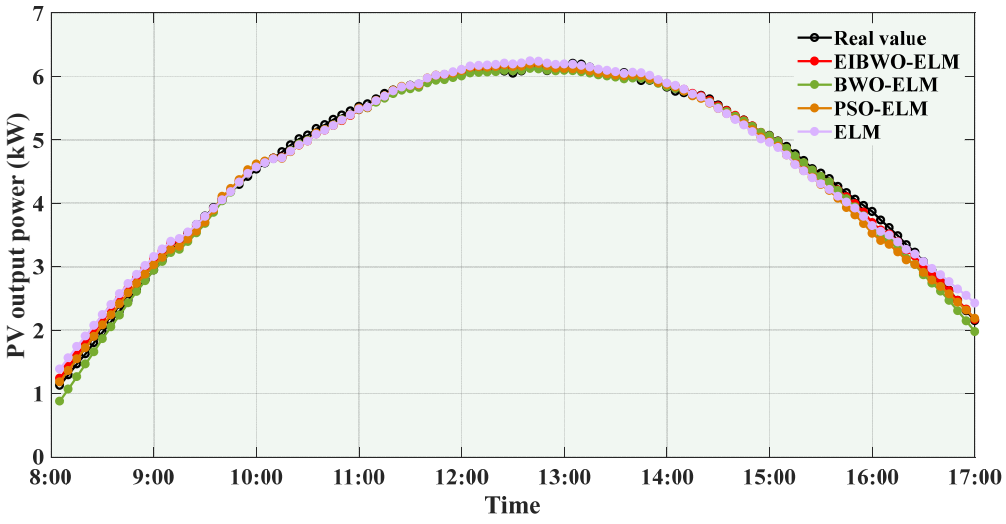


**Figure 5.** The prediction results of four models on sunny days.

From Figure 5, it can be seen that EIBWO-ELM, BWO-ELM, PSO-ELM, and ELM all achieved the characterization of sunny photovoltaic output power. To better represent the error between the predicted and true values of the proposed model, the absolute error of the two at each sample point is presented in the form of a curve. The absolute errors of the four models' prediction results on sunny days are shown in Figure 6.

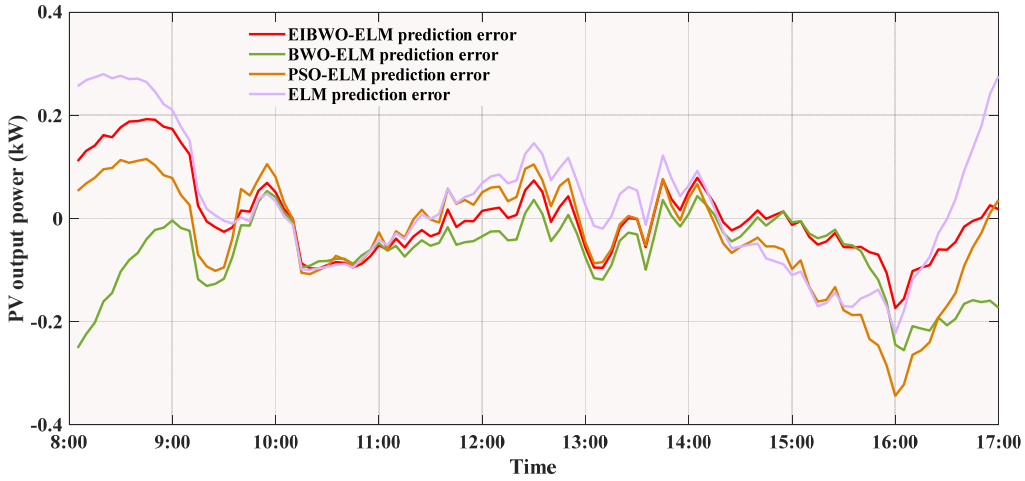


**Figure 6.** The absolute errors of the four models' prediction results on sunny days.

From the absolute error curve in Figure 6, it can be seen that the PV power prediction error of the EIBWO-ELM model was closer to 0 on sunny days. Compared to BWO-ELM, PSO-ELM, and ELM, the prediction error of the EIBWO-ELM model fluctuated more slowly. The relative error between the two at each sample point is represented by a data statistical histogram. The calculation equation for relative error $\delta$ is shown in Equation (36).

$$\delta = \frac{\hat{y}_i - y_i}{y_i} \times 100\% \tag{36}$$

In addition, the statistical results of the number of sample points in each error range of prediction models on sunny days are shown in Figure 7. The detailed statistical results of the data are shown in Table 6.

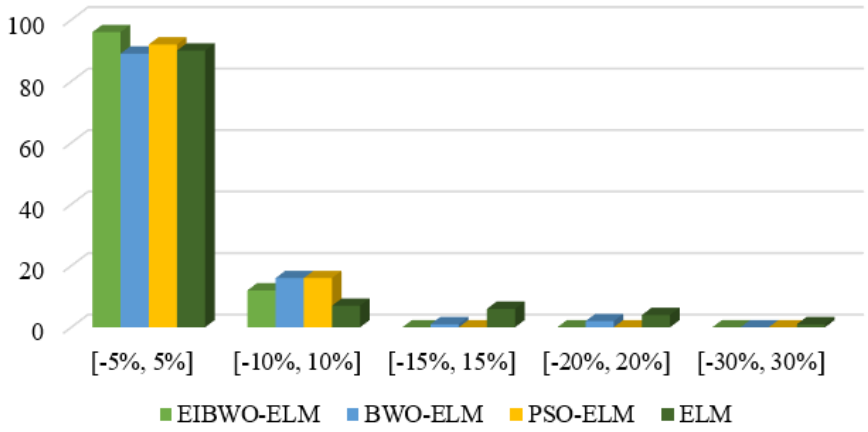


**Figure 7.** Statistical results of the number of sample points in each error range of the prediction models on sunny days.

**Table 6.** Statistical results of the number of sample points in each error range on sunny days.

| Error Range | EIBWO-ELM | | BWO-ELM | |
|---|---|---|---|---|
| | Number | Proportion (%) | Number | Proportion (%) |
| [−5%, 5%] | 96 | 88.89 | 89 | 82.41 |
| [−10%, 10%] | 108 | 100 | 105 | 97.22 |
| [−15%, 15%] | 108 | 100 | 106 | 98.15 |
| [−20%, 20%] | 108 | 100 | 108 | 100 |
| [−30%, 30%] | 108 | 100 | 108 | 100 |
| **Error Range** | **PSO-ELM** | | **ELM** | |
| | **Number** | **Proportion (%)** | **Number** | **Proportion (%)** |
| [−5%, 5%] | 92 | 85.18 | 90 | 83.33 |
| [−10%, 10%] | 108 | 100 | 97 | 89.81 |
| [−15%, 15%] | 108 | 100 | 103 | 95.37 |
| [−20%, 20%] | 108 | 100 | 107 | 99.07 |
| [−30%, 30%] | 108 | 100 | 108 | 100 |

From Table 6, it can be seen that on sunny days, the relative error range of all models was between [−30%, 30%]. There were 96 sample points with a relative error range between [−5%, 5%] with EIBWO-ELM, accounting for 88.89% of the total number of samples. However, the relative error range of ELM prediction results was more dispersed, and there was one sample point within the range of [−30%, 30%]. Therefore, on sunny days, the EIBWO-ELM model had the highest number of sample points in the low error range and had good prediction accuracy. The effectiveness of the proposed EIBWO optimization ELM is proven in this study.

To objectively analyze the prediction results of each model, the evaluation indicators RMSE, MAPE, and $R^2$ were used to comprehensively and objectively evaluate the predictive performance of the proposed EIBWO-ELM. The evaluation index values of the prediction results of each prediction model on sunny days were calculated separately, and the calculation results are shown in Table 7.

**Table 7.** Statistics of prediction result indicators of the models on sunny days.

| Model | RMSE (kW) | MAPE (%) | $R^2$ |
|---|---|---|---|
| EIBWO-ELM | 0.0787 | 1.83 | 0.9968 |
| BWO-ELM | 0.0982 | 2.34 | 0.9951 |
| PSO-ELM | 0.1089 | 2.23 | 0.9940 |
| ELM | 0.1269 | 3.21 | 0.9918 |

From Table 7, it can be seen that the RMSE values of EIBWO-ELM and BWO-ELM were both less than 0.1 kW. However, the RMSE value of ELM was the highest, at 0.1269 kW. In terms of $R^2$, EIBWO-ELM had the highest $R^2$, reaching 0.9968. This indicates a high degree of fit between the predicted values of EIBWO-ELM and the real values. In terms of MAPE, EIBWO-ELM, BWO-ELM, and PSO-ELM decreased by 1.38%, 0.87%, and 0.98%, respectively, compared to ELM. It has been proven that using algorithms to optimize the internal parameters of ELM is the main solution to improving model prediction accuracy. The MAPE of EIBWO-ELM was reduced by 0.51% and 0.4% compared to BWO-ELM and PSO-ELM, respectively. This indicates that the EIBWO proposed in this study has efficient optimization capabilities.

### *5.3. Case 2—Cloudy Days*

Case 2 is the prediction of PV power generation on cloudy days. EIBWO-ELM, BWO-ELM, PSO-ELM, and ELM were used to predict the output power of PV power generation. The actual output power values and the predicted values obtained after model prediction on cloudy days are shown in Figure 8.

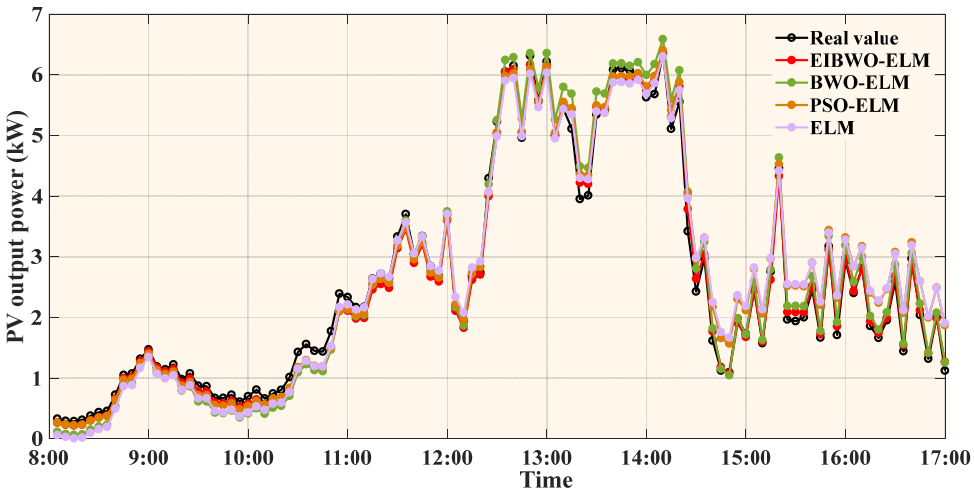


**Figure 8.** The prediction results of the four models on cloudy days.

The prediction curve of EIBWO-ELM was very close to the true value curve, which means that EIBWO-ELM had high prediction accuracy. On the contrary, the difference between the predicted curve of ELM and the true value curve was significant, indicating that the prediction accuracy of ELM was low. The relative error between the predicted values and the true values of the prediction models at each sample point for cloudy days was calculated, and the relative error curve is presented in Figure 9.

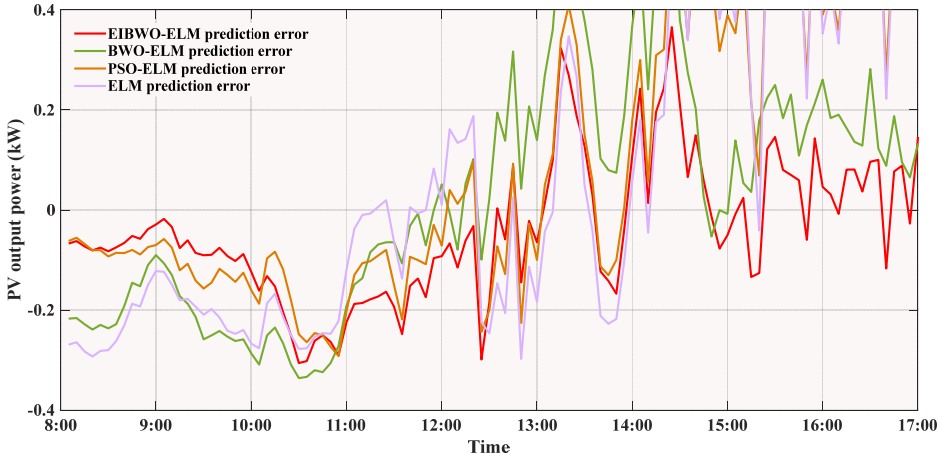


**Figure 9.** The absolute errors of the four models' prediction results on cloudy days.

From Figure 9, it can be intuitively seen that the prediction models performed differently at different sample points. During 8:00–10:00, the predicted relative error curve of EIBWO-ELM was closer to 0, which indicates that during this time period, the predicted values of EIBWO-ELM were very close to the actual observed values and the prediction error was smaller. In contrast, the error curves of other models fluctuated significantly, deviating from the straight line with an error of 0. this means that BWO-ELM, PSO-ELM, and ELM had lower accuracy.

In addition, Figure 10 shows the statistical results of the number of sample points for the four prediction models within each error range, which helps analyze the error distribution of the models and visualize the performance of different models. Table 8 provides the specific number of sample points within each error range, which helped to quantitatively analyze the predictive stability of the prediction model.

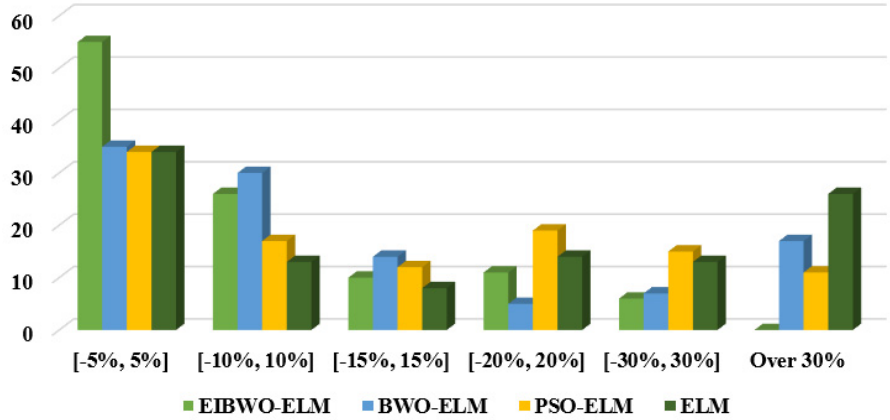


**Figure 10.** Statistical results of the number of sample points in each error range of the prediction models on cloudy days.

In Table 8, within the relative error range of [−10%, 10%], EIBWO-ELM had 81 samples, which accounted for 75% of the total samples. This indicates that EIBWO-ELM provided predictions close to the true values with a high probability on cloudy days, demonstrating excellent prediction accuracy. Furthermore, all sample points of EIBWO-ELM fell within the error range of [−30%, 30%]. In contrast, ELM had a sample point proportion of 24.07% within an error interval of over 30%, which is significantly higher than that of EIBWO-ELM. This indicates that the prediction results of ELM were more prone to significant deviations on cloudy days, and that the prediction stability and accuracy were relatively low. In summary, the prediction performance of EIBWO-ELM was significantly better than that of ELM, especially in the low error range, where the number of sample points had a significant

advantage. The statistics of prediction results indicators of the models on cloudy days are shown in Table 9.

**Table 8.** Statistical results of the number of sample points in each error range on cloudy days.

| Error Range | EIBWO-ELM | | BWO-ELM | |
|---|---|---|---|---|
| | Number | Proportion (%) | Number | Proportion (%) |
| [−5%, 5%] | 55 | 50.93 | 35 | 32.41 |
| [−10%, 10%] | 81 | 75 | 65 | 60.19 |
| [−15%, 15%] | 91 | 84.26 | 79 | 73.15 |
| [−20%, 20%] | 102 | 94.44 | 84 | 77.78 |
| [−30%, 30%] | 108 | 100 | 91 | 84.26 |
| Over 30% | 108 | 100 | 108 | 100 |
| **Error Range** | **PSO-ELM** | | **ELM** | |
| | **Number** | **Proportion (%)** | **Number** | **Proportion (%)** |
| [−5%, 5%] | 34 | 31.48 | 34 | 31.48 |
| [−10%, 10%] | 51 | 47.22 | 47 | 43.52 |
| [−15%, 15%] | 63 | 58.33 | 55 | 50.93 |
| [−20%, 20%] | 82 | 75.93 | 69 | 63.89 |
| [−30%, 30%] | 97 | 89.81 | 82 | 75.93 |
| Over 30% | 108 | 100 | 108 | 100 |

**Table 9.** Statistics of prediction result indicators of the models on cloudy days.

| Model | RMSE (kW) | MAPE (%) | $R^2$ |
|---|---|---|---|
| EIBWO-ELM | 0.1440 | 7.40 | 0.9935 |
| BWO-ELM | 0.2373 | 14.53 | 0.9823 |
| PSO-ELM | 0.3006 | 13.99 | 0.9715 |
| ELM | 0.3237 | 20.11 | 0.9670 |

Table 9 shows that, among the four prediction models, EIBWO-ELM had the smallest RMSE of 0.1440 kW, indicating the smallest prediction error. In contrast, ELM had the highest value of RMSE of 0.3237 kW. The MAPE of EIBWO-ELM was the lowest, at 7.40%, indicating a small difference between its predicted and actual values. The MAPE of EIBWO-ELM decreased by 7.13% and 6.59% compared to BWO-ELM and PSO-ELM, respectively. The $R^2$value of EIBWO-ELM was the highest, at 0.9935. Taking these three indicators into consideration, EIBWO-ELM outperformed the other three models in terms of predictive performance.

In addition, by comparing the evaluation indicators of sunny and cloudy days, it was found that almost all models had better predictive performance on sunny days than on cloudy days. This is because the output power fluctuates greatly on cloudy days, which increases the difficulty of model prediction. However, regardless of whether it was sunny or cloudy, the RMSE and MAPE values of EIBWO-ELM were the smallest, indicating the smallest error. Moreover, the $R^2$ of EIBWO-ELM was higher than 0.99, indicating a high degree of fit. In summary, the proposed EIBWO-ELM had better prediction accuracy than other comparative models and is suitable for predicting photovoltaic power generation in various scenarios.

## 6. Discussion

### *6.1. Photovoltaic Power Prediction under Varying Meteorological Conditions*

High-precision PV prediction under varying meteorological conditions is of great significance for promoting the utilization of renewable energy and the development of clean energy technologies. By adopting advanced prediction models and algorithms, it is possible to more accurately predict PV power generation under varying meteorological

conditions, which is of great significance for power grid management, economic efficiency improvement, and addressing climate change.

This study aims to improve the accuracy of PV power prediction by delving into PV power generation performance under varying meteorological conditions. The aim is to reveal the shortcomings of existing technologies and provide direction for the research into and improvement of new technologies, thereby promoting significant improvements in the performance and efficiency of PV power generation systems. In this study, the Pearson coefficient method is used to quantitatively evaluate the correlation between meteorological parameters such as irradiance, temperature, and relative humidity and PV power generation. This step is of decisive significance for accurately determining the key input characteristics of the prediction model. Through this method, it is possible to more accurately capture the key meteorological factors that affect the efficiency of PV power generation, thereby constructing a more accurate and reliable prediction model.

### *6.2. Enhanced and Improved Beluga Whale Optimization Extreme Learning Machine*

The enhanced and improved beluga whale optimization extreme learning machine (EIBWO-ELM) model proposed in this study has achieved significant results in the field of photovoltaic power generation prediction, thanks to a series of innovations in algorithm design and optimization.

Firstly, EIBWO, which has strong global search capability and fast convergence speed, has been proposed to effectively handle complex problems in PV power generation prediction. In Section 3.2, it is verified that compared to other existing algorithms, EIBWO can quickly find the optimal solution, which not only improves the efficiency but also enhances its stability and reliability in practical applications.

Secondly, by applying EIBWO to optimize ELM, the proposed EIBWO-ELM significantly improves the accuracy and generalization ability of PV power generation prediction. Based on measured data of PV power generation under sunny and cloudy weather conditions, compared with other models, EIBWO-ELM has good prediction accuracy. In addition, the adaptive adjustment mechanism and robustness of EIBWO-ELM enable it to maintain stable predictive performance even in the face of data noise and uncertainty. This optimization not only improves the accuracy of predictions and helps reduce energy waste and economic losses but also enhances the adaptability of the model in the face of uncertainty and variability.

Finally, the technological innovation of this study has brought new knowledge contributions to the field of intelligent algorithms. The successful application of the EIBWO-ELM model not only demonstrates the potential of intelligent algorithms in the field of renewable energy but also provides a new idea and generalized framework for the development of intelligent algorithms and the utilization of renewable energy. With further research in the future, this model is expected to play to its advantages in more fields and make greater contributions to sustainable development and green energy transformation.

## 7. Concluding Remarks

This study proposed a PV power generation prediction model based on EIBWO-ELM, and the main conclusions are as follows:

(1) An EIBWO with high convergence accuracy and optimization robustness was proposed. EIBWO was proposed in this study by introducing the chaotic mapping strategy, sine dynamic adaptive factor, and disturbance strategy into the traditional BWO. The convergence value of the proposed method almost reached the optimal value of 0 through standard testing functions. In addition, the standard deviation of the results of different testing functions through EIBWO was 0, indicating that the proposed method had the best optimization robustness.

(2) A PV power prediction model based on EIBWO-ELM was established. The internal parameters of ELM were optimized by using EIBWO with high convergence accuracy. The proposed EIBWO-ELM was validated through two weather conditions, sunny and

cloudy days, and the results show that it achieved high-precision photovoltaic power prediction. According to the predicted results, the MAPE values of EIBWO-ELM reached 1.83% and 7.40% on sunny and cloudy days, respectively.

(3) EIBWO-ELM provides a new PV prediction method to improve the integration and efficiency of renewable energy in energy portfolios. The good performance of EIBWO-ELM in instance data validation not only demonstrates the potential of intelligent optimization technology in renewable energy prediction but also provides a universal framework for developing similar intelligent algorithms.

However, there are still some limitations: (1) the quality of input data is not fully considered, and (2) the adverse effects of external factors on photovoltaic panels are not fully considered. In the future, there is an urgent need to address existing limitations and propose more universal photovoltaic prediction models.

**Author Contributions:** Writing—original draft, W.D., S.-T.P., P.-S.W. and M.-L.T.; Writing—review & editing, W.D., S.-T.P., P.-S.W. and M.-L.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the following foundation programs: the China National Key R&D Program (grant No. 2022YFB2603100), the Guangxi Science and Technology Major Program (grant No. AA23062054), and the Basic Research Fund of Central Public Interest Scientific Research Institutes in China (grant No. TKS20230301).

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

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

## References

1. Aniello, G.; Shamon, H.; Kuckshinrichs, W. Micro-economic assessment of residential PV and battery systems: The underrated role of financial and fiscal aspects. *Appl. Energy* **2021**, *281*, 115667. [CrossRef]
2. Calil, W.V.; Morozovska, K.; Laneryd, T.; da Costa, E.C.M.; Salles, M.B.D. Determining total cost of ownership and peak efficiency index of dynamically rated transformer at the PV-power plant. *Electr. Power Syst. Res.* **2024**, *229*, 110061. [CrossRef]
3. Anusuya, K.; Vijayakumar, K. Mission profile-based assessment of photovoltaic system reliability for Indian climatic zones. *Energy Sources Part A-Recovery Util. Environ. Eff.* **2024**, *46*, 1779–1799. [CrossRef]
4. Annuk, A.; Jogi, E.; Kalder, J.; Allik, A. Evaluating Electricity Self-Consumption in Different Renewable Energy Supply Conditions. In Proceedings of the 17th International Scientific Conference: Engineering for Rural Development, Jelgava, Latvia, 23–25 May 2018; pp. 1704–1709. [CrossRef]
5. Liu, Z.F.; Liu, Y.Y.; Chen, X.R.; Zhang, S.R.; Luo, X.F.; Li, L.L.; Yang, Y.Z.; You, G.D. A novel deep learning-based evolutionary model with potential attention and memory decay-enhancement strategy for short-term wind power point-interval forecasting. *Appl. Energy* **2024**, *360*, 122785. [CrossRef]
6. Eseye, A.T.; Zhang, J.H.; Zheng, D.H. Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and Meteorological information. *Renew. Energy* **2018**, *118*, 357–367. [CrossRef]
7. Li, L.L.; Wen, S.Y.; Tseng, M.L.; Wang, C.S. Renewable energy prediction: A novel short-term prediction model of photovoltaic output power. *J. Clean. Prod.* **2019**, *228*, 359–375. [CrossRef]
8. Ullah, Z.; Wang, S.R.; Wu, G.; Hasanien, H.M.; Rehman, A.U.; Turky, R.A.; Elkadeem, M.R. Optimal scheduling and techno-economic analysis of electric vehicles by implementing solar-based grid-tied charging station. *Energy* **2023**, *267*, 126560. [CrossRef]
9. Ma, R.F.; Cai, H.; Ji, Q.; Zhai, P.X. The impact of feed-in tariff degression on R&D investment in renewable energy: The case of the solar PV industry. *Energy Policy* **2021**, *151*, 112209. [CrossRef]
10. Kurm, S.; Agarwal, V. Dual Active Bridge Based Reduced Stage Multiport DC/AC Converter for PV-Battery Systems. *IEEE Trans. Ind. Appl.* **2022**, *58*, 2341–2351. [CrossRef]
11. Ahmed, R.; Sreeram, V.; Togneri, R.; Datta, A.; Arif, M.D. Computationally expedient Photovoltaic power Forecasting: A LSTM ensemble method augmented with adaptive weighting and data segmentation technique. *Energy Convers. Manag.* **2022**, *258*, 115563. [CrossRef]
12. Al-Dahidi, S.; Ayadi, O.; Adeeb, J.; Louzazni, M. Assessment of artificial neural networks learning algorithms and training datasets for solar photovoltaic power production prediction. *Front. Energy Res.* **2019**, *7*, 130. [CrossRef]
13. Hajjaj, C.; El Ydrissi, M.; Azouzoute, A.; Oufadel, A.; El Alani, O.; Boujoudar, M.; Abraim, M.; Ghennioui, A. Comparing Photovoltaic Power Prediction: Ground-Based Measurements vs. Satellite Data Using an ANN Model. *IEEE J. Photovolt.* **2023**, *13*, 998–1006. [CrossRef]

14. Tukymbekov, D.; Saymbetov, A.; Nurgaliyev, M.; Kuttybay, N.; Dosymbetova, G.; Svanbayev, Y. Intelligent autonomous street lighting system based on weather forecast using LSTM. *Energy* **2021**, *231*, 120902. [CrossRef]
15. Huang, Y.S.; Wu, Y.H. Short-Term Photovoltaic Power Forecasting Based on a Novel Autoformer Model. *Symmetry* **2023**, *15*, 238. [CrossRef]
16. Dolara, A.; Grimaccia, F.; Leva, S.; Mussetta, M.; Ogliari, E. A Physical Hybrid Artificial Neural Network for Short Term Forecasting of PV Plant Power Output. *Energies* **2015**, *8*, 1138–1153. [CrossRef]
17. Sansa, I.; Boussaada, Z.; Bellaaj, N.M. Solar Radiation Prediction Using a Novel Hybrid Model of ARMA and NARX. *Energies* **2021**, *14*, 6920. [CrossRef]
18. Kim, B.; Suh, D.; Otto, M.O.; Huh, J.S. A Novel Hybrid Spatio-Temporal Forecasting of Multisite Solar Photovoltaic Generation. *Remote Sens.* **2021**, *13*, 2605. [CrossRef]
19. Wan, C.; Lin, J.; Song, Y.H.; Xu, Z.; Yang, G.Y. Probabilistic Forecasting of Photovoltaic Generation: An Efficient Statistical Approach. *IEEE Trans. Power Syst.* **2017**, *32*, 2471–2472. [CrossRef]
20. Pandzic, F.; Capuder, T. Advances in Short-Term Solar Forecasting: A Review and Benchmark of Machine Learning Methods and Relevant Data Sources. *Energies* **2024**, *17*, 97. [CrossRef]
21. Krechowicz, M.; Krechowicz, A.; Licholai, L.; Pawelec, A.; Piotrowski, J.Z.; Stepien, A. Reduc-tion of the Risk of Inaccurate Prediction of Electricity Generation from PV Farms Using Machine Learning. *Energies* **2022**, *15*, 4006. [CrossRef]
22. Raj, V.; Dotse, S.Q.; Sathyajith, M.; Petra, M.I.; Yassin, H. Ensemble Machine Learning for Pre-dicting the Power Output from Different Solar Photovoltaic Systems. *Energies* **2023**, *16*, 671. [CrossRef]
23. Liu, Y.W.; Li, L.L.; Liu, J.Q. Enhancing short-term wind power forecasting accuracy for reliable and safe integration into power systems: A gray relational analysis and optimized support vector regression machine approach. *J. Renew. Sustain. Energy* **2024**, *16*, 013311. [CrossRef]
24. Muhammad, A.; Lee, J.M.; Hong, S.W.; Lee, S.J.; Lee, E.H. Deep Learning Application in Power System with a Case Study on Solar Irradiation Forecasting. In Proceedings of the 2019 1st International Conference on Artificial Intelligence in Information and Communication, Okinawa, Japan, 11–13 February 2019; pp. 275–279. [CrossRef]
25. Li, G.Q.; Wang, H.Z.; Zhang, S.L.; Xin, J.T.; Liu, H.C. Recurrent Neural Networks Based Photovoltaic Power Forecasting Approach. *Energies* **2019**, *12*, 2538. [CrossRef]
26. Liu, Z.H.; Li, Q.Z.; Wang, D.L.; Zhang, G.F.; Wang, W.; Zhao, Y.; Guo, R. Research on the Harmonic Prediction Method of a PV Plant Based on an Improved Kernel Extreme Learning Machine Model. *Electronics* **2024**, *13*, 32. [CrossRef]
27. Das, S.; Sahu, T.P.; Janghel, R.R.; Sahu, B.K. Effective forecasting of stock market price by using extreme learning machine optimized by PSO-based group oriented crow search algorithm. *Neural Comput. Appl.* **2022**, *34*, 555–591. [CrossRef] [PubMed]
28. Matisone, I.; Kaupe, D.; Matisons, R.; Klavina, D.; Jansons, A. Understory changes in mixed elm stands in response to canopy dieback in Latvia. *Balt. For.* **2023**, *29*, 113–122. [CrossRef]
29. Zhang, Z.Y.; Zhang, A.; Sun, C.; Xiang, S.D.; Guan, J.C.; Huang, X.D. Research on Air Traffic Flow Forecast Based on ELM Non-Iterative Algorithm. *Mob. Netw. Appl.* **2021**, *26*, 425–439. [CrossRef]
30. Al-Dahidi, S.; Ayadi, O.; Adeeb, J.; Alrbai, M.; Qawasmeh, B.R. Extreme learning machines for solar photovoltaic power predictions. *Energies* **2018**, *11*, 2725. [CrossRef]
31. Zhou, Y.; Zhou, N.R.; Gong, L.H.; Jiang, M.L. Prediction of photovoltaic power output based on similar day analysis, genetic algorithm and extreme learning machine. *Energy* **2020**, *204*, 117894. [CrossRef]
32. Liu, Z.F.; Li, L.L.; Tseng, M.L.; Lim, M.K. Prediction short-term photovoltaic power using improved chicken swarm optimizer—Extreme learning machine model. *J. Clean. Prod.* **2020**, *248*, 119272. [CrossRef]
33. Zhang, W.Y.; Li, Q.W.; He, Q.F. Application of machine learning methods in photovoltaic output power prediction: A review. *J. Renew. Sustain. Energy* **2022**, *14*, 022701. [CrossRef]
34. Zhang, J.L.; Tan, Z.F.; Wei, Y.M. An adaptive hybrid model for day-ahead photovoltaic output power prediction. *J. Clean. Prod.* **2020**, *244*, 118858. [CrossRef]
35. Zhong, C.T.; Li, G.; Meng, Z. Beluga whale optimization: A novel nature-inspired metaheuristic algorithm. *Knowl. Based Syst.* **2022**, *251*, 109215. [CrossRef]
36. Li, L.L.; Fan, X.D.; Wu, K.J.; Sethanan, K.; Tseng, M.L. Multi-objective distributed generation hierarchical optimal planning in distribution network: Improved beluga whale optimization algorithm. *Expert Syst. Appl.* **2024**, *237*, 121406. [CrossRef]
37. Jia, H.M.; Wen, Q.X.; Wu, D.; Wang, Z.; Wang, Y.H.; Wen, C.S.; Abualigah, L. Modified beluga whale optimization with multi-strategies for solving engineering problems. *J. Comput.* **2023**, *10*, 2065–2093. [CrossRef]
38. Novykov, V.; Bilson, C.; Gepp, A.; Harris, G.; Vanstone, B.J. Empirical validation of ELM trained neural networks for financial modelling. *Neural Comput. Appl.* **2023**, *35*, 1581–1605. [CrossRef]
39. Yan, P.C.; Li, G.D.; Wang, W.C.; Zhao, Y.T.; Wang, J.B.; Wen, Z.M. A Mine Water SAource Prediction Model Based on LIF Technology and BWO-ELM. *J. Fluoresc.* **2024**. [CrossRef]