Journal of
Marine Science
and Engineering

MDPI

Article

Short-Term Photovoltaic Power Prediction Based on a Digital Twin Model

Chuan Xiang [1,*], Bohan Li [1], Pengfei Shi [1], Tiankai Yang [1] and Bing Han [2]

[1] College of Marine Electrical Engineering, Dalian Maritime University, Dalian 116026, China; libohan996@gmail.com (B.L.); pfshi@dlmu.edu.cn (P.S.); tiankai900@dlmu.edu.cn (T.Y.)
[2] National Engineering Research Center of Ship & Shipping Control System, Shanghai Ship and Shipping Research Institute Co., Ltd., Shanghai 200135, China; han.bing@coscoshipping.com
* Correspondence: cxiang@dlmu.edu.cn

Abstract: Due to the influence of meteorological conditions, shipboard photovoltaic (PV) systems have problems such as large fluctuation and inaccurate prediction of the output power. In this paper, a short-term PV power prediction method based on a novel digital twin (DT) model and BiLSTM is proposed. Firstly, a PV mechanism model and a data-driven model were established, in which the data-driven model was updated iteratively in real time using the sliding time window update method; then, these two models were converged to construct a PV DT model according to the DS evidence theory. Secondly, a BiLSTM model was built to make short-term predictions of the PV power using the augmented dataset of the DT model as an input. Finally, the method was tested and verified by experiments and further compared with main PV prediction methods. The research results indicate the following: firstly, the absolute error of the DT model was smaller than that of the mechanism model and the data-driven model, being as low as 5.62 W after the data update of the data-driven model; thus, the DT model realized data augmentation and high fidelity. Secondly, compared to several main PV prediction models, the PV DT model combined with BiLSTM had the lowest RMSE, MAE, and MAPE; the best followability; and the smallest absolute error under different weather conditions, which was especially obvious under cloudy weather conditions. In summary, the method can accurately predict the shipboard PV power, which has great theoretical significance and application value for improving the economy and reliability of solar ship operation.

Keywords: PV power prediction; digital twin model; mechanism model; data-driven model; BiLSTM

check for updates

Citation: Xiang, C.; Li, B.; Shi, P.; Yang, T.; Han, B. Short-Term Photovoltaic Power Prediction Based on a Digital Twin Model. *J. Mar. Sci. Eng.* **2024**, *12*, 1219. https://doi.org/10.3390/jmse12071219

Academic Editor: Erkan Oterkus

Received: 21 June 2024
Revised: 11 July 2024
Accepted: 17 July 2024
Published: 19 July 2024

1. Introduction

The United Nations Conference on Trade and Development reported that greenhouse gas (GHG) emissions from shipping have increased by 20% in the last decade, emitting as much as 1.1 billion tons in 2023 and accounting for nearly 3% of global GHG emissions. The International Maritime Organization (IMO) has mandated that GHG emissions from shipping be reduced by at least 50% by 2050. Therefore, renewable energy sources such as solar and wind have been increasingly used in ship power systems [1]. In particular, solar energy has been more widely applied on ships due to its cost-effectiveness and government supportive policies [2]. For example, PV systems are installed on conventional diesel ships to form an integrated energy supply system to reduce the use of fossil fuels, as well as to be used as auxiliary energy sources to provide power for lighting and other auxiliary equipment [3–5]. Studies show that the installation of PV systems on river cruise ships and dry bulk carriers not only reduces fuel consumption but also extends the service life of ship equipment [6]. The installation of PV systems on ships can reduce the dependence on traditional energy sources, lower carbon emissions, and satisfy the EU carbon emission reduction regulations, which is of great significance in promoting the green transformation of shipping [7]. However, shipboard photovoltaic (PV) systems are affected by meteorological conditions and the ship's sailing attitude, resulting in large fluctuations in PV output power, which will impact the stable operation of ship power systems in island mode. Therefore, accurate and fast prediction of the PV output power to provide a reference for real-time scheduling of energy storage systems (ESSs) is crucial to ensure the stable operation of solar ships [8,9].

The main PV models currently available are mechanism and data-driven models. In terms of mechanism models, they have evolved from single-diode models to dual-diode models. Marcelo, G.V. [10] built a single-diode PV model which takes into account the effect of series and parallel resistances and ensures that the maximum output power of the model matches the maximum power of actual PV cells. Yan, J. [11] utilized a single-diode model to simulate PV power generation and proposed a maximum power point tracking technique based on an improved perturb and observe method, which significantly improved the efficiency of the PV energy utilization. Raya-Armenta [12] introduced two new physical equations to represent series and parallel resistances, which significantly improved the model accuracy. Mohammad, H. [13] proposed a parameter extraction method for a five-parameter dual-diode model of PV cells based on information provided by PV manufacturers. Although the above mechanism models provide a theoretical understanding of PV cells, they only focus on the internal characteristics of PV cells in the modeling process and cannot adapt to system component changes, PV panel aging, and dynamic changes in meteorological conditions; thus, their accuracy is usually low in practical applications. In terms of data-driven models, they do not need to be aware of the internal characteristics of PV cells and, in most cases, have a higher modeling accuracy than mechanism models [14]. Patra, J.C. [15] studied a novel Chebyshev neural network to simulate a dual-junction PV cell, which outperformed the commercial software ATLAS in predicting the characteristics of DJ solar cells. Based on a convolutional neural network (CNN) and long short-term memory (LSTM), a PV power prediction model was used for power system generation planning and reserve estimation [16]. Rubasinghe, O. [17] developed a novel sequence-to-sequence hybrid CNN-LSTM PV model, which had high accuracy and could be used as a PV output power prediction model. The accuracy of the above data-driven models is usually high, but they rely heavily on historical data and may not be adequate in situations with limited available data, sudden changes in meteorological conditions, or violent ship swings. In addition, data-driven models are not as interpretable as mechanism models with clear physical descriptions and specific mathematical expressions.

Recently, machine learning methods have been used for PV power predictions. Zhu, H. [18] proposed a PV power prediction method combining wavelet decomposition and an artificial neural network (ANN) which takes into account the effects of solar irradiance and the historical output power of PV cells. However, the ANN only established direct mapping between the input data and the output predicted power without considering the temporal correlation of the data series. Zhou, S.Y. [19] predicted PV power based on a recurrent neural network (RNN), which improved time series prediction by keeping the memory of the previous information and incorporating it into the current computation, resulting in a significant increase in model prediction accuracy compared to ANNs [20]. However, RNNs still suffer from gradient vanishing during long sequences training. Akhter, M.N. and Cantillo-Luna, S. [21,22] proposed a novel PV power prediction method, LSTM-RNN, which takes into account meteorological parameters such as wind speed, temperature, and humidity and achieves good prediction results. Although LSTM captures forward feature information in the time series of the PV output power, it ignores reverse feature information, which leads to incomplete information capture and affects the prediction accuracy. Bi-directional long short-term memory (BiLSTM), upgraded from LSTM, can comprehensively capture information from the forward and reverse directions, which is expected to improve the prediction accuracy of the PV power.

Considering the above, the purpose of this paper is to present a novel PV digital twin (DT) model combined with BiLSTM to improve the prediction accuracy of PV power. It is organized as follows: Section 2 introduces the research idea. Section 3 presents the PV DT model. Section 4 expounds on the theoretical principle of BiLSTM. The experimental procedure is described in detail in Section 5. Section 6 experimentally validates the effectiveness of the proposed method. Conclusions are enclosed in Section 7.

2. Research Idea

The research idea of this paper is illustrated in Figure 1. The essential and innovative work is to establish a PV DT model consisting of a mechanism model and a data-driven model, which provides an augmented dataset for a PV power prediction algorithm. Based on experimentally collected meteorological data related to PV power generation, such as temperature, irradiance, humidity, etc., the mechanism model and the data-driven model are first constructed. The PV DT model is then obtained by converging both of them using the Dempster–Shafer (DS) evidence theory, whereby the data-driven model is continuously updated using the sliding time window update method. The augmented dataset is then fed into a BiLSTM neural network for making a short-term prediction of the PV power. Finally, the PV DT model and BiLSTM prediction algorithm are comparatively analyzed and validated.

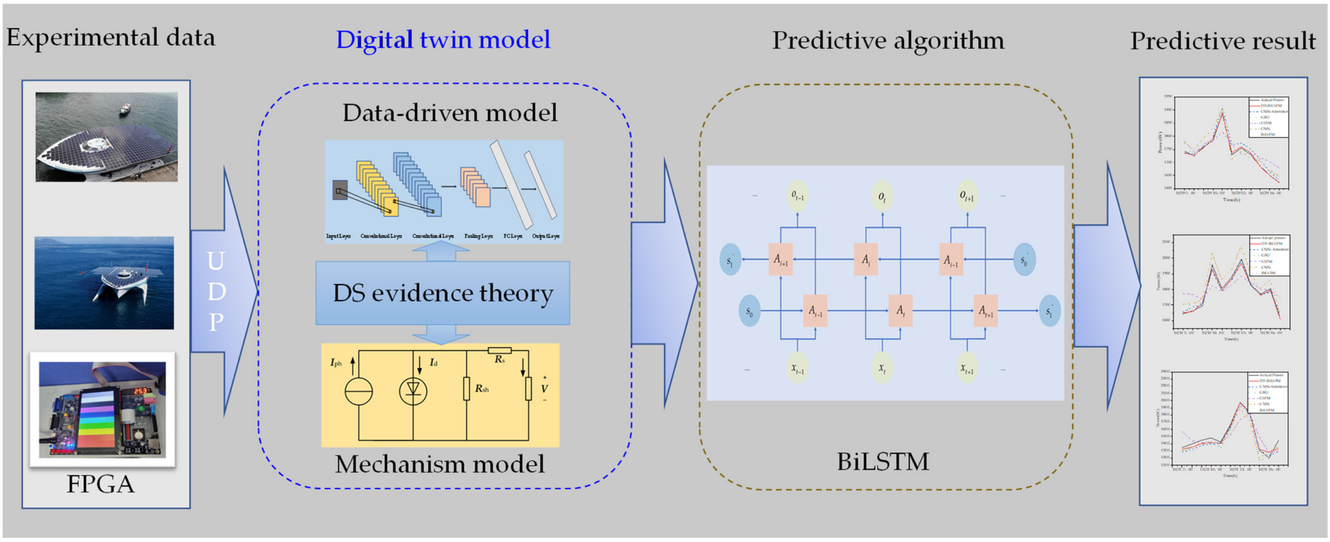


Figure 1. Block diagram of the research idea.

3. PV Digital Twin Model

3.1. Mechanism Model

The output power of an individual PV cell is low, so multiple cells are typically connected in series and parallel to form a PV array for higher output power. The equivalent circuit model of a PV cell is illustrated in Figure 2, where I_{ph} is the photocurrent of the PV cell, R_{sh} is the equivalent parallel resistance, R_s is the equivalent series resistance, I_d is the current through the diode, V is the PV output voltage, and I is the PV output current. Based on Kirchhoff's current law, the output current I of a PV cell can be derived as follows [23]:

$$I = I_{ph} - I_0\left[\exp\left(\frac{V + IR_{sh}}{nkTN_s/q} - 1\right)\right] - \frac{V + IR_{sh}}{R_s} \tag{1}$$

where I_0 is the reverse saturation current of the diode; n is the ideality factor of the diode, $1 \leq n \leq 2$; q is the electric charge, $q = 1.602 \times 10^{-19}$ C; k is the Boltzmann constant, $k = 1.381 \times 10^{-23}$ J/K; N_s is the number of PV cells connected in series; and T is the temperature of the PV cell.

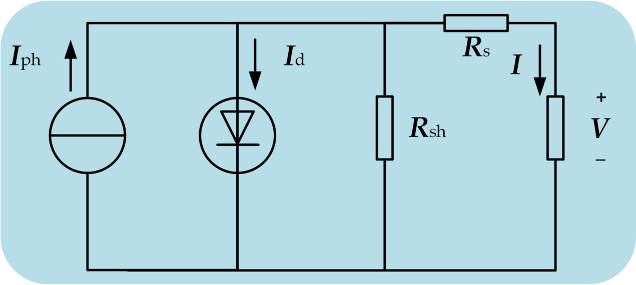


Figure 2. Equivalent circuit of a PV cell.

I_{ph} and I_o are variables affected by meteorological conditions such as irradiance and temperature. Their calculation formulas are as follows [10,13,24–26]:

$$I_{ph} = I_{sco}\left[1 + h_t\left(T - T_{ref}\right)\right]S/S_{ref} \tag{2}$$

$$I_o = b_1 T^3 \exp(-a_1/T) \tag{3}$$

where I_{sco} represents the short-circuit current under standard irradiance and temperature conditions; h_t is the temperature coefficient, $h_t = 6.4 \times 10^{-4}$ 1/°C; T denotes the temperature of the PV cell; T_{ref} represents the reference temperature; a_1 and b_1 are constants, $a_1 = 1.336 \times 10^4$ °C and $b_1 = 235$ A/°C^3; and S and S_{ref} represent the actual irradiance and the reference irradiance, respectively.

Although this model has been widely used for the theoretical analysis of PV output power, the parameters I_{ph}, I_o, R_s, R_{sh}, and n are dependent on temperature and irradiance and are not provided by PV cell suppliers. Therefore, calculation of these parameters is very difficult, which limits the engineering application of the model. In response, a simplified model was proposed by Babu, B. [27], and Formula (1) can be simplified as follows:

$$I = I_{sc}\left\{1 - C_1\left[\exp\left(\frac{V}{C_2 V_{oc}}\right) - 1\right]\right\} \tag{4}$$

The solutions C_1 and C_2 are obtained as follows:

$$C_1 = \left(1 - \frac{I_m}{I_{sc}}\right)\exp\left(-\frac{V_m}{C_2 V_{oc}}\right) \tag{5}$$

$$C_2 = \left(\frac{V_m}{V_{oc}} - 1\right)\left[\ln\left(1 - \frac{I_m}{I_{sc}}\right)\right]^{-1} \tag{6}$$

From the simplified model, the PV output current can be calculated because the values of I_{sc}, V_{oc}, I_m, and V_m are usually available from PV cell suppliers. However, these four parameters are still affected by changes in meteorological conditions and need to be revised. The revision for them with irradiance and temperature is given below. It is used as the PV mechanism model in this study.

$$\Delta T = T - T_{ref} \tag{7}$$

$$\Delta S = \frac{S}{S_{ref}} - 1 \tag{8}$$

$$I'_{sc} = \frac{I_{sc}S}{S_{ref}}(1 + a\Delta T) \tag{9}$$

$$V'_{oc} = V_{oc}(1 - c\Delta T)(1 + b\Delta S) \tag{10}$$

$$I'_m = \frac{I_m S}{S_{ref}}(1 + a\Delta T) \tag{11}$$

$$V'_m = V_m(1 - c\Delta T)(1 + b\Delta S) \tag{12}$$

where I'_{sc}, V'_{oc}, I'_m, and V'_m are the revisions of I_{sc}, V_{oc}, I_m, and V_m, respectively; T is the temperature of the PV cell; S is the irradiance; $T_{ref} = 25\ ^\circ\mathrm{C}$; $S_{ref} = 1000\mathrm{MW/m^2}$; and a, b, and c are constants, $a = 0.0025\ 1/\ ^\circ\mathrm{C}$, $b = 0.5$, and $c = 0.00288\ 1/\ ^\circ\mathrm{C}$.

3.2. Data-Driven Model

Data feature extraction is crucial for PV power prediction. A PV power series has a strong time regularity, while meteorological parameter sequences present a strong randomness, which makes it very difficult to extract the features of PV data series. CNNs, known for their powerful feature extraction capability, have been widely popularized and applied in many fields such as image recognition, pattern classification, object detection, face recognition, and time series data analysis [28,29]. In this paper, a CNN is used to construct the PV data-driven model. A CNN is a type of feedforward neural network that comprises an input layer, convolutional layers, pooling layers, fully connected layers, and an output layer, as shown in Figure 3 [30].

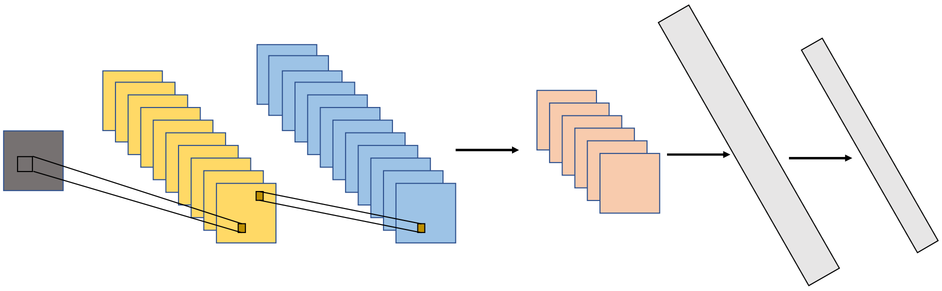


Figure 3. Structural diagram of CNN.

Compared to conventional matrix operations, a significant advantage of CNNs is that they can process data with convolutional operations, which greatly enhances the computational speed. A CNN uses convolutional and pooling layers in parallel to efficiently extract local features from the input data while reducing the dimensionality of these features. This feature extraction mechanism allows the CNN to focus on relevant information and discard some unimportant details, thus improving feature extraction efficiency and performance.

Moreover, a CNN benefits from weight sharing, a key characteristic that reduces the number of parameters and complexity within the model. Weight sharing ensures that the same set of weights is used across different regions of the input data, enabling the network to learn and recognize patterns in a more generalized manner. This not only helps in reducing model complexity but also enhances the network's ability to generalize and make accurate predictions on unknown data.

3.3. Data Update Based on Sliding Time Window Update Method

Traditional PV data-driven models do not update the model data, which cannot effectively extract the features in the latest data and reflect the actual operation of the system, resulting in low fidelity of the model. The PV data-driven model established in this paper uses a sliding time window update method to add the latest PV power generation data to the model, thus improving the fidelity of the model as well as the adaptability of data changes.

In order to ensure the efficiency of model training, the sliding time window update method needs to keep the size of the training data fixed when adding new data samples, i.e., replacing the old data samples with the same number of the latest data samples. To reduce the influence of data further away from the current moment on the description of the model feature, the data samples closest to the current moment are usually selected for model training to retain the latest data information on the timeline. This approach ensures that the data-driven model captures the latest information and adapts to data changes over time, thus improving the accuracy of the data-driven model [31,32].

The process of updating the PV data-driven model M_1 to M_2 using a sliding time window is illustrated in Figure 4. Let us assume that the size of the training data samples is L, and the length of the sliding time window is n. When the number of newly added data samples reaches n, the replacement strategy is executed to form the new training data, and then, the PV data-driven model is retrained to update the model.

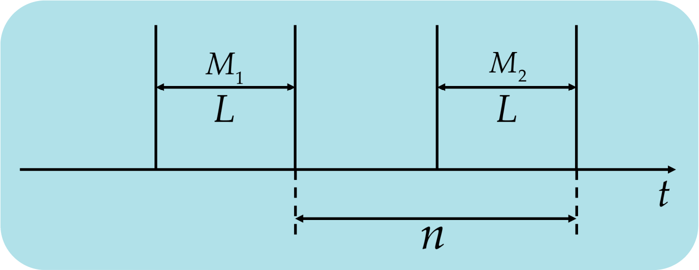


Figure 4. Sliding time window update method.

3.4. Digital Twin Model

The PV mechanism model developed in Section 3.1 has the advantages of stable model error and strong interpretability, but in most cases, its error is large. The PV data-driven model built in Sections 3.2 and 3.3 has a strong feature extraction capability for PV sequences with high modeling accuracy, but its interpretability is poor and the model error increases rapidly when the weather changes. Neither satisfies the requirements of a high-fidelity, high-precision PV model. In this section, the advantages of these two models are combined to construct a PV DT model. The Dempster–Shafer evidence theory is utilized to determine the weights of the mechanism model and the data-driven model and converge them. Meanwhile, the data-driven model is iteratively updated in real time using the sliding time window update method. The principle of the DS evidence theory is described next.

The Dempster–Shafer (DS) evidence theory is a data convergence method used to deal with uncertainty. It is based on a non-empty set of all possible outcomes for a given decision problem, known as the recognition framework [33].

The recognition framework is defined as a Θ finite set of M mutually exclusive and exhaustive proposition sets, denoted as follows [34,35]:

$$\Theta = \{H_1, H_2, \ldots, H_M\} \tag{13}$$

On the basis of the recognition framework, the Power Set 2^Θ is defined as follows:

$$2^\Theta = \{\varnothing, H_1, H_2, \ldots H_M, \{H_1, H_2\}, \{H_1, H_3\}, \ldots \{H_1, H_M\}, \ldots, \{H_1, H_2, \ldots, H_M\}\} \tag{14}$$

where the symbol "$\varnothing$" denotes the empty set, which belongs to any subset of the recognition framework.

The basic probability assignment (BPA) function represents the initial level of belief of an evidence body towards a proposition. It is defined as a mapping from the power set of the system's recognition framework to the interval [0, 1]. The following conditions need to be satisfied:

$$\begin{cases} m(\varnothing) = 0 \\ \sum\limits_{H \subseteq \Theta} m(H) = 1 \end{cases} \tag{15}$$

where $m(H)$ indicates the basic support of the evidence H for the proposition.

The belief function *Bel*() and the plausibility function *Pl*() are the proposition description functions on the recognition framework and the BPA function, respectively. They are denoted as follows:

$$\begin{cases} Bel(H) = \sum\limits_{A \subseteq H} m(A) \\ Pl(H) = \sum\limits_{A \cap H \neq \varnothing} m(A) \end{cases} \tag{16}$$

$$\begin{cases} Bel\{H\} \leq Pl\{H\} \\ Pl\{H\} = 1 - Bel\{\overline{H}\} \end{cases} \tag{17}$$

where $Bel(H)$ and $Pl(H)$ represent the lower and upper bounds of probability, respectively. Therefore, H propositional uncertainty can be described by $[Bel(H), Pl(H)]$.

The convergence condition for data using the DS evidence theory is that the evidence bodies from two different data sources are in the same recognition framework, so the convergence rule is defined as follows:

$$m(H) = \begin{cases} \frac{1}{1-K} \sum\limits_{H_i \cap H_j = H} m_1(H_i) \cdot m_2(H_j) H \neq \varnothing \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad H = \varnothing \end{cases} \tag{18}$$

$$K = \sum_{H_i \cap H_j = \varnothing} m_1(H_i) \cdot m_2(H_j) \tag{19}$$

4. Bi-Directional Long-Short Term Memory

Recently, RNNs and LSTM have been used in PV power prediction [19,21,22]. Classic RNNs have some limitations, such as the difficulty they encounter in remembering long-term dependencies and their inability to effectively relate old information to new inputs. LSTM, a variant of classic RNNs, aims to solve these limitations. It mitigates the gradient vanishing and exploding and improves the ability to preserve long sequences by adding gates within each cell state. These gates, namely the forget gate, input gate, and output gate, play crucial roles in filtering, preserving, and generating information, respectively [36,37]. The structure of LSTM is depicted in Figure 5.

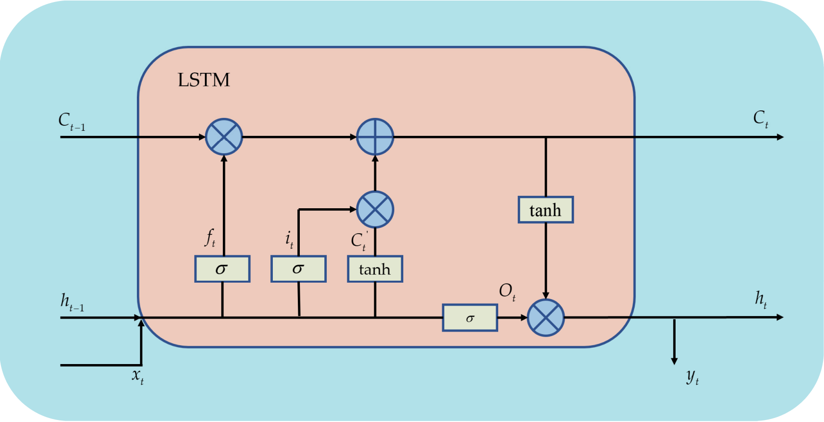


Figure 5. Structural diagram of LSTM.

The forget gate in LSTM networks has the function of selectively filtering and retaining information from the processing of the previous memory cell. The output or activation of the forget gate f_t at the time step t is as follows:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{20}$$

where x_t is the input sequence at the time step t, h_{t-1} is the previous hidden state or memory cell, $\sigma(\cdot)$ is the sigmoid activation function, W_f is the weight matrix for the forget gate, and b_f is the bias term for the forget gate.

The input gate controls the influence of the current input on the memory cell. Its expression is as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{21}$$

$$\widetilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{22}$$

$$C_t = f_t * C_t + i_t * \widetilde{C}_t \tag{23}$$

where $\widetilde{C}_t$ represents the candidate value for the cell state, C_t represents the new cell state, $\tanh(\cdot)$ is the hyperbolic tangent function, W_i represents the weight matrix of the input gate, W_c represents the weight matrix of the cell state, b_i represents the bias of the input gate, b_c represents the bias of the cell state, and i_t represents the state of the input gate.

The output gate controls the output state of the memory cell. Its expression is as follows:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{24}$$

$$h_t = o_t * \tanh(C_t) \tag{25}$$

where W_o represents the weight matrix of the output gate, b_o represents the bias of the output gate, and o_t represents the output state of the output gate.

One limitation of an LSTM neural network is that it relies on the historical information of the forward sequence. To address this limitation, the BiLSTM neural network was introduced. It consists of two LSTM neural networks, one for processing the forward sequence and the other for processing the backward sequence. By integrating information from both directions, the BiLSTM neural network can capture the intrinsic patterns of past and future data, thus enhancing its predictive capability [38–41]. In this paper, it is selected as the PV power prediction algorithm.

In a BiLSTM network, the forward and backward sequences are first processed independently in different hidden layers. The outputs of these hidden layers are then combined and used as inputs to the output layer, thus improving the accuracy of the model. The structure of a typical BiLSTM network is shown in Figure 6.

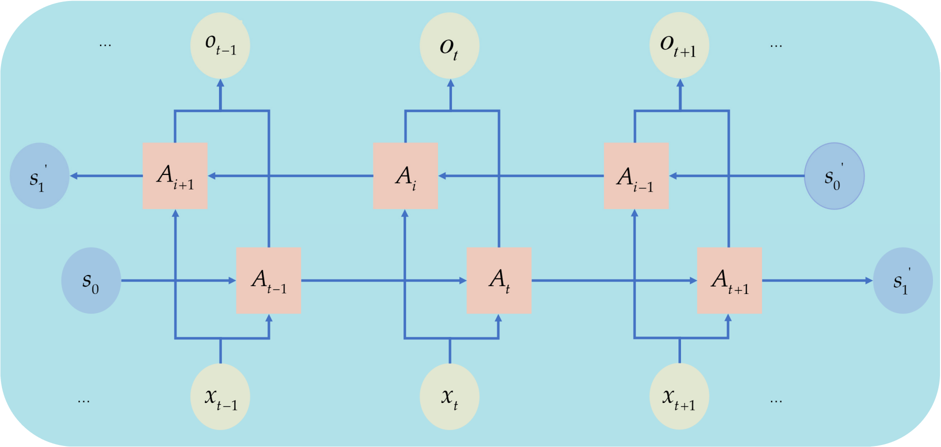


Figure 6. Structural diagram of BiLSTM network.

The hidden state at each layer of the BiLSTM network is composed of three parts: the forward hidden state at the previous time step A_{t-1}, the backward hidden state at the previous time step A_{i-1}, and the input at the current time step x_t. The combination process of the hidden states at each layer can be represented as follows:

$$\begin{cases} A_t = LSTM(x_t, A_{t-1}) \\ A_i = LSTM(x_t, A_{i-1}) \\ A_t = a_t A_t + b_t A_i + c_t \end{cases} \tag{26}$$

where $LSTM()$ represents the operation of the traditional LSTM network, A_t refers to the forward hidden state, A_i refers to the backward hidden state, a_t represents the weight of the output from the hidden layer of the forward propagation unit, b_t represents the weight of the output from the hidden layer of the backward propagation unit, and c_t represents the bias optimization parameter of the hidden layer at current time step.

5. Experimental Procedure

The flowchart of the short-term PV power prediction based on the DT model is shown in Figure 7. As can be seen from Figure 7, the detailed steps are as follows:

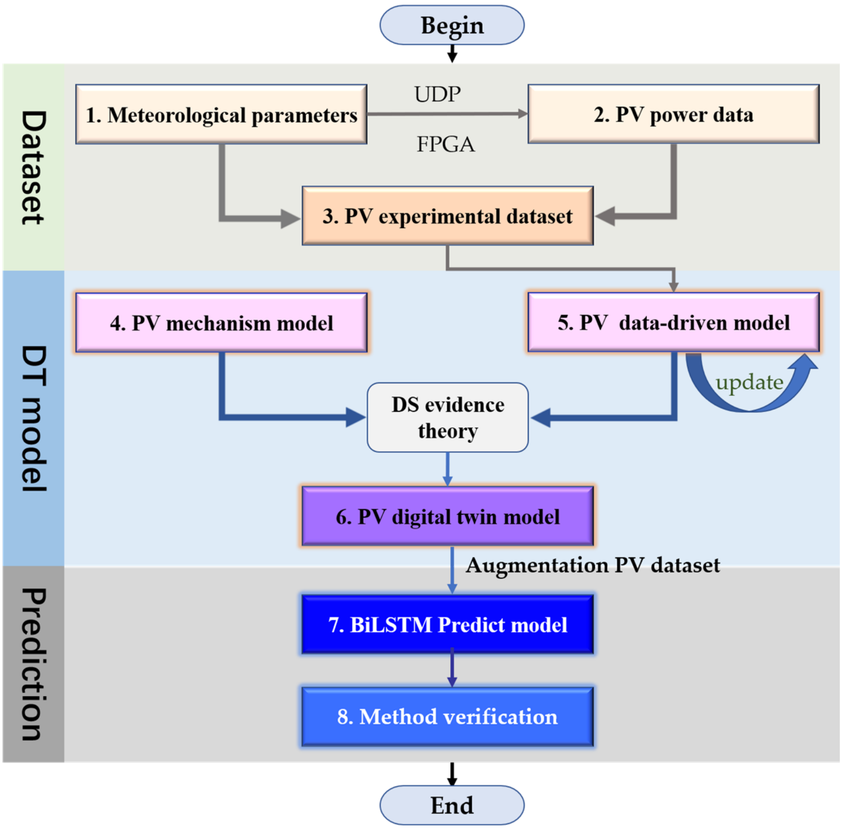


Figure 7. Flowchart of short-term PV power prediction based on the DT model.

Step 1. Collect meteorological parameters such as irradiance, temperature, and humidity via sensors.

Step 2. Transmit meteorological parameters to the FPGA via the UDP communication protocol and then import them into a PV simulator to generate the PV power data.

Step 3. Construct a PV dataset consisting of meteorological parameters and the PV power data.

Step 4. Build the PV mechanism model based on the analysis of the working principle, physical properties, capacity, and series and parallel connections of the PV power generation.

Step 5. Build the PV data-driven model based on CNN, train it with the PV dataset, and update the PV data-driven model using the sliding time window method. This is performed by importing the latest data from the sliding time window into the CNN model for rebuilding the data-driven model.

Step 6. Establish the PV DT model. According to the BPA, the weights of the mechanism model and the data-driven model are first calculated, and then, these two models are converged using DS evidence theory to obtain the DT model.

Step 7. Build the BiLSTM prediction model of the PV power, using the experimental dataset and the augmented dataset of the DT model as inputs to achieve PV power prediction.

Step 8. Evaluate the model performance with metrics such as RMSE, MAE, and MAPE to verify the effectiveness of the method.

6. Experimental Results

6.1. Evaluation Metrics

Evaluation metrics for machine learning methods include mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), symmetric mean absolute percentage error (SMAPE), etc. This study adopts MAE, MAPE, and RMSE to evaluate the model performance. MAE is the average of absolute errors and is a basic indicator for error evaluation. MAPE can further investigate the ratio between errors and actual values. RMSE can accurately identify large or small errors in the predicted data. Their formulas are expressed as follows:

$$MAE = \frac{\sum_{i=1}^{n}\left|X_{real,i} - X_{pred,i}\right|}{n} \tag{27}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{X_{real,i} - X_{pred,i}}{X_{real,i}}\right| \tag{28}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(X_{real,i} - X_{pred,i}\right)^2}{n}} \tag{29}$$

where $X_{real,i}$ is the true value, $X_{pred,i}$ is the predicted value, and n is the data volume.

6.2. Experimental Data

Irradiance, temperature, and humidity are important meteorological parameters that affect PV power generation [36]. In this study, a PV data collection and processing experimental platform was built, as shown in Figure 8, which contained several irradiance, temperature, and humidity sensors; an FPGA; and a PV simulator. Meteorological parameter data were first collected by sensors at intervals of 1 h from 7 a.m. to 5 p.m. every day in October 2022. They were then transferred to the FPGA via the UDP communication protocol, which was formed by programming the FPGA Verilog code. Finally, these data were imported into the PV simulator to calculate the PV power. Therefore, we constructed an experimental dataset consisting of meteorological parameters such as irradiance, temperature, and humidity and PV power. The dataset was divided into four data subsets: subset 1 was from 1 October 7 a.m. to 10 October 7 a.m. and was used to determine the initial BPA in the DS evidence theory; subset 2 and subset 3 were from 10 October 8 a.m. to 19 October 8 a.m. (before the data update) and 19 October 9 a.m. to 28 October 9 a.m. (after the data update) and were used to build and update the data-driven model in the DT model, respectively; subset 4 was from 29 October 7 a.m. to 31 October 5 p.m. and was used to validate the predictive effectiveness of the method.

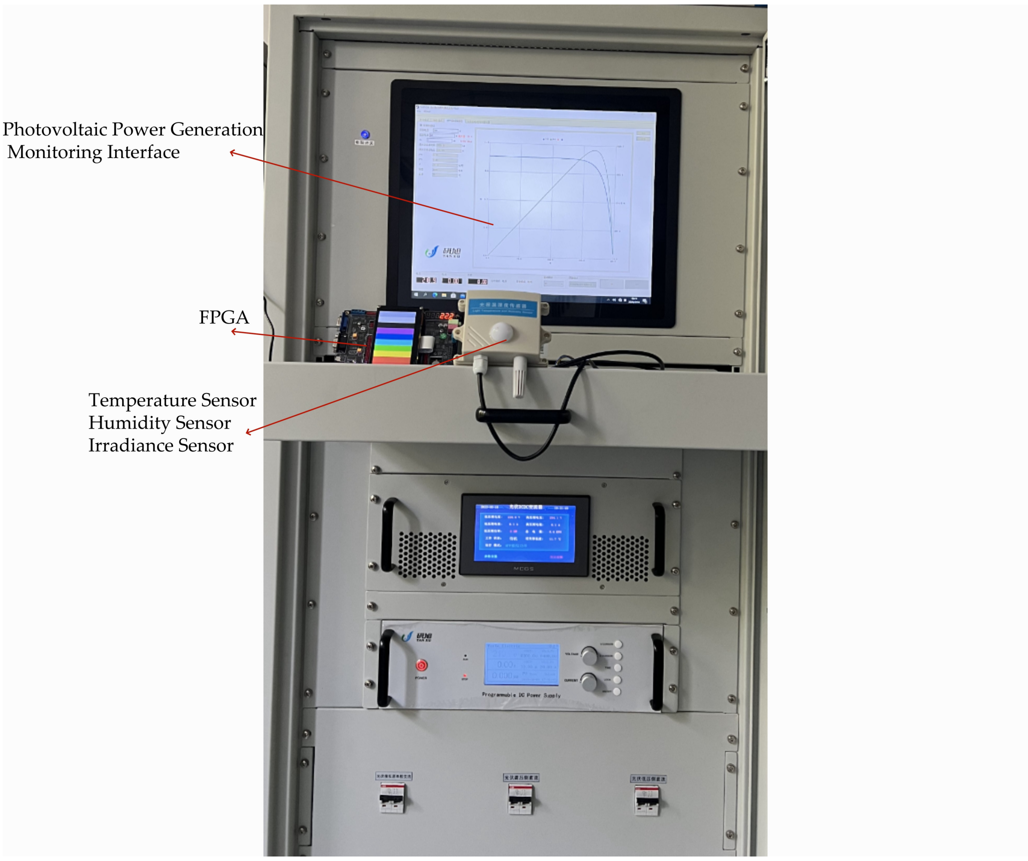


Figure 8. Experimental data collection platform.

The meteorological parameters collected during October 2022 are shown in Figure 9, and the experimental data for 1 October 2022 are shown in Table 1 in detail.

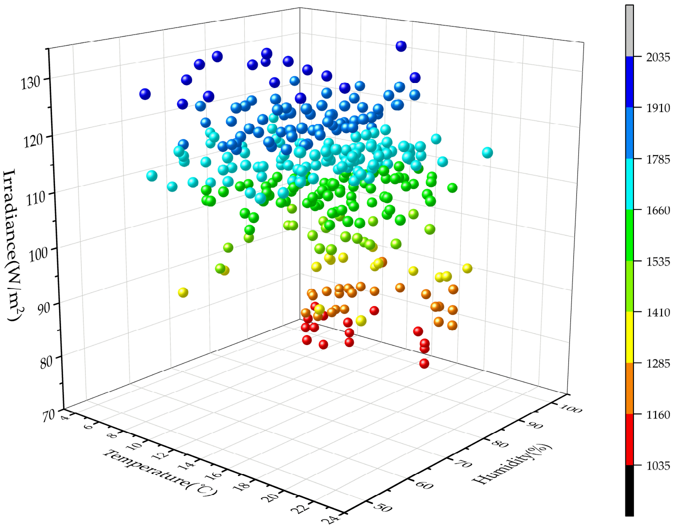


Figure 9. Experimental data during October 2022.

Table 1. Experimental data for 1 October 2022.

Time	Temperature (°C)	Humidity (%)	Irradiance (W/m^2)	PV Power (W)
1 October 2022 7:00	16	73	103	1494.98
1 October 2022 8:00	16	75	108	1586.75
1 October 2022 9:00	17	73	106	1544.16
1 October 2022 10:00	17	85	110	1609.77
1 October 2022 11:00	19	87	115	1665.91
1 October 2022 12:00	21	70	120	1735.27
1 October 2022 13:00	21	69	122	1767.55
1 October 2022 14:00	22	87	118	1696.66
1 October 2022 15:00	20	80	105	1509.54
1 October 2022 16:00	19	75	103	1493.43
1 October 2022 17:00	18	70	90	1300.26

6.3. Data Augmentation Effects of the PV DT Model

According to the DS evidence theory, the convergence of the mechanism model and the data-driven model must first determine the weights of these two models, and the model weights are calculated as the BPA. In this paper, the BPA is determined as the inverse ratio of the error between the output power calculated or derived from the PV mechanism model and data-driven model and the actual output power and the inverse ratio of the RMSE. The BPAs of the mechanism and data-driven models obtained from data subset 1 and data subset 2 are shown in Tables 2 and 3, respectively.

Table 2. BPAs of the mechanism and data-driven models obtained from data subset 1.

	BPA (Inverse Ratio of Error)	BPA (Inverse Ratio of RMSE)
Mechanism model	0.187	0.23
Data-driven model	0.813	0.77

Table 3. BPAs of the mechanism and data-driven models obtained from data subset 2.

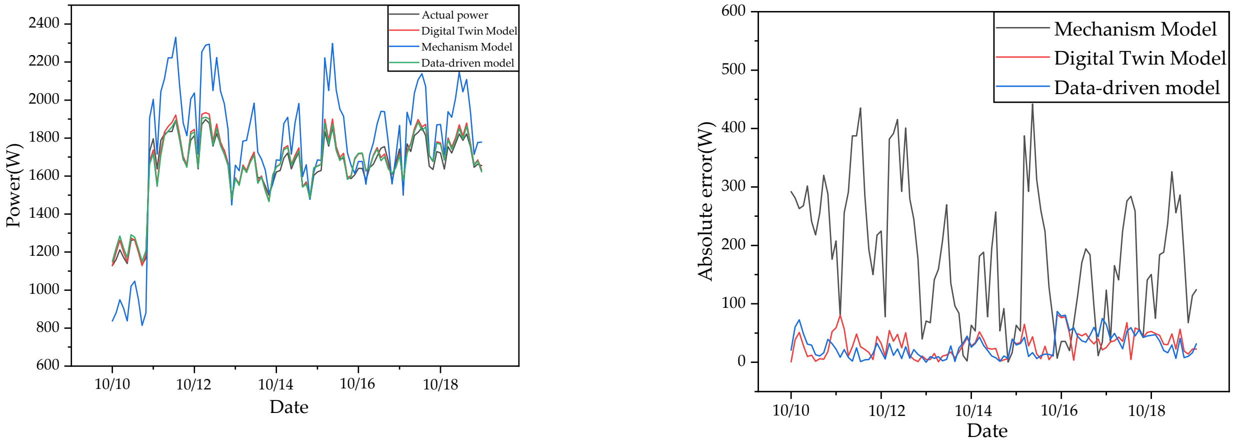
	BPA (Inverse Ratio of Error)	BPA (Inverse Ratio of RMSE)
Mechanism model	0.231	0.20
Data-driven model	0.769	0.80

Based on the BPAs listed in Tables 2 and 3, the weights of the mechanism and data-driven models before and after the data update were calculated, as shown in Table 4.

Table 4. Weights of the mechanism and data-driven models before and after the data update.

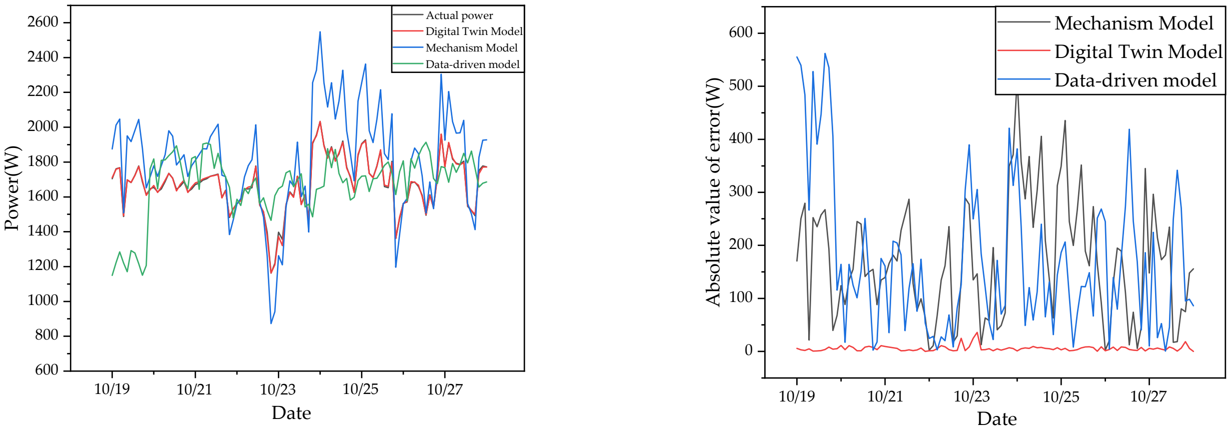
	Weight (before the Data Update)	Weight (after the Data Update)
Mechanism model	6.4%	7.0%
Data-driven model	93.6%	93.0%

According to the model weights listed in Table 4, the PV DT model was obtained by converging the mechanism and data-driven models based on the DS evidence theory. The comparisons of the output powers of the PV mechanism model, data-driven model, and DT model with the actual PV power before the data update and after the data update are illustrated in Figures 10 and 11, respectively.



(**a**) (**b**)

Figure 10. Comparison of the output power of the DT model, mechanism model, and data-driven model with actual PV power before the data update. (**a**) Power. (**b**) Absolute error.



(**a**) (**b**)

Figure 11. Comparison of the output power of the DT model, mechanism model, and data-driven model with actual PV power after the data update. (**a**) Power. (**b**) Absolute error.

From Figures 10 and 11, it can be seen that compared with the actual PV output power, the error of the mechanism model was the largest, and the error of the data-driven model was smaller, but the power curve of the DT model was well fitted, with minimum error. Further analysis of the absolute error revealed that the errors of both the data-driven model and the DT model were further reduced after the data update compared to before the data update; in particular, the error of the DT model had a lowest average value of only 5.62 W, as shown in Figure 11b. Therefore, the DT model built in this study realized data augmentation with high fidelity.

6.4. PV Power Prediction Results

In this section, the established PV DT model combined with BiLSTM is used for PV power prediction, which is compared with the main PV power prediction models such as CNN, LSTM, gate recursive unit (GRU), and CNN-Attention. The effectiveness of the proposed method is evaluated and validated through the metrics RMSE, MAE, and MAPE.

The RMSEs of PV prediction models using the original experimental dataset and the augmented dataset of the DT model are shown in Tables 5 and 6, respectively; Table 6 contains the RMSEs before and after the data update of the DT model. For comparative analysis, the RMSEs of the prediction models in Tables 5 and 6 are summarized as shown in Figure 12, where "experimental" and "digital" represent the original experimental dataset and the augmented dataset of the DT model, respectively, and "before" and "after" represent the status before and after the data update of the DT model, respectively.

Table 5. RMSEs of different models with the experimental dataset (Unit: W).

	LSTM	**CNN**	**GRU**	**CNN-Attention**	**BiLSTM**
Sunny	52.8	81.3	13.3	24.1	18.9
Rainy	74.1	109.2	30.2	22.6	31.7
Cloudy	185.7	116.7	81.9	109.8	107.8

Table 6. RMSEs of different models with the augmented dataset of the DT model (Unit: W).

	LSTM		**CNN**		**GRU**		**CNN-Attention**		**BiLSTM**	
	Before	**After**	**Before**	**After**	**Before**	**After**	**Before**	**After**	**Before**	**After**
Sunny	50.7	18.2	80.6	41.7	10.6	10.8	11.9	9.4	10.8	4.8
Rainy	71.8	52.0	85.5	56.9	25.4	23.7	19.8	19.6	22.4	14.4
Cloudy	167.4	103.7	86.8	89.7	77.2	76.6	87.4	84.7	74.8	50.9

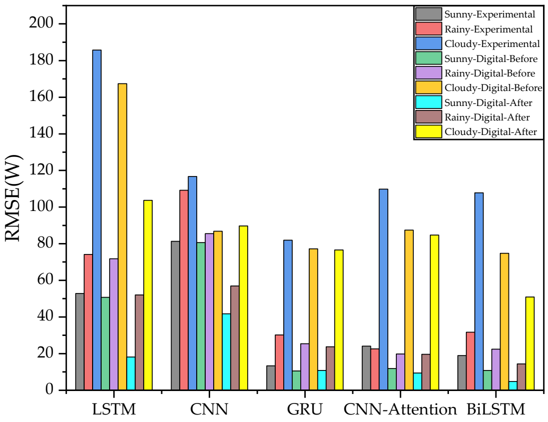


Figure 12. RMSEs of different methods.

As can be seen from Table 5, when the prediction models adopted the experimental dataset for sunny, rainy, and cloudy weather conditions, the RMSEs of the LSTM and CNN models were large, indicating that the predictions of these models deviated greatly from the true values and were poor, whereas the RMSEs of the GRU, CNN-Attention, and BiLSTM models were small and close, indicating that their predictions worked well. In addition, the RMSEs of these models were much larger under cloudy weather conditions than under sunny and rainy conditions, showing that it is difficult to accurately predict PV power in complex and changeable cloudy weather. It can be seen from Table 6 that when the prediction models adopted the augmented dataset, before the data update of the DT model, the RMSEs of the LSTM and CNN models were still large, and the RMSEs of the GRU, CNN-Attention, and BiLSTM models were small and relatively close. However, the RMSEs of all models decreased compared with the experimental dataset, with the RMSE of the BiLSTM model in particular decreasing significantly. After the data update of the DT model, most of the RMSEs of these models were further reduced, and the prediction effects of these models were significantly improved. In particular, compared with the dataset before the DT model update and the experimental dataset, the RMSE of the BiLSTM model decreased by 6 and 14.1 for sunny weather conditions, 6 and 17.3 for rainy weather conditions, and significantly decreased by 23.9 and 56.9 for cloudy weather conditions, respectively. The prediction effect of the BiLSTM model was thus significantly improved. Figure 6 further illustrates these results.

The MAEs and MAPEs of the prediction models are shown in Figures 13 and 14, respectively. Similar to the RMSE, both the MAE and MAPE values of the prediction models decreased significantly with the DT augmented dataset compared to the experimental dataset, and both decreased further after the data update of the DT model. Among these prediction models, the prediction effect of BiLSTM combined with the DT model (after the data update) was still the most prominent; the MAE and MAPE were very low under sunny and rainy weather conditions, and both were also greatly reduced under cloudy conditions.

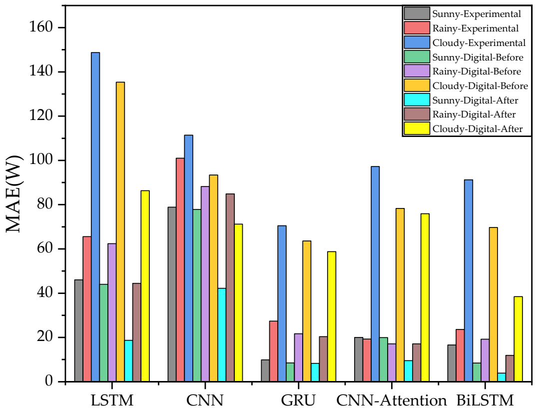


Figure 13. MAEs of different methods.

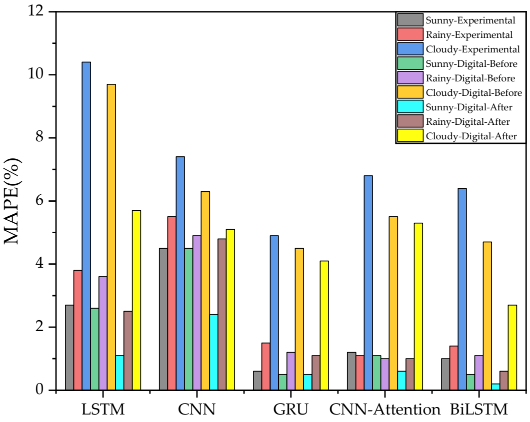
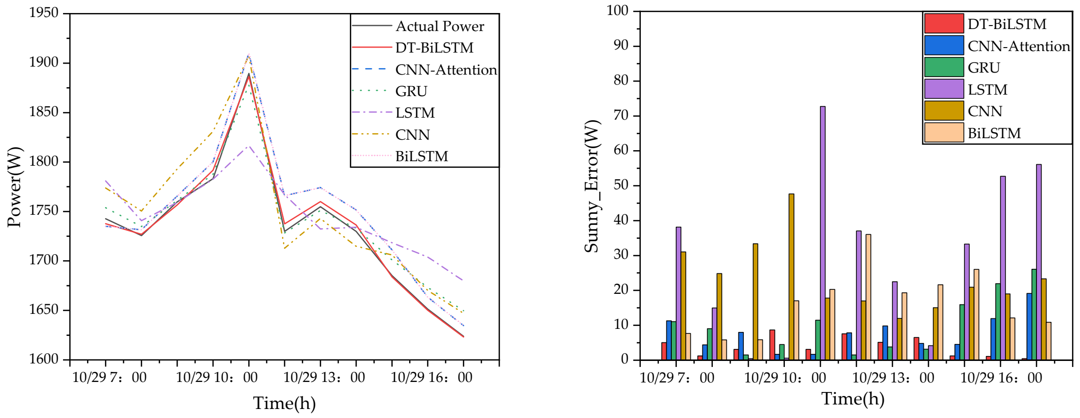


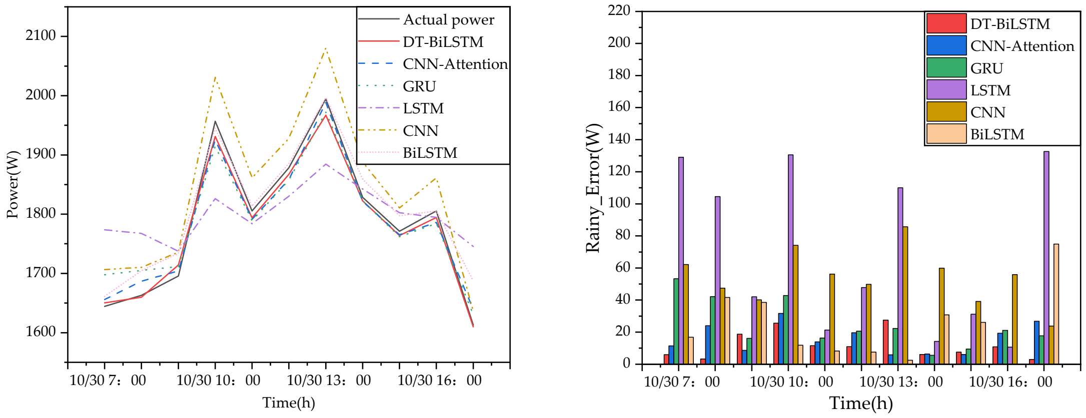
Figure 14. MAPEs of different methods.

The above results and analyses show that the evaluation metrics of RMSE, MAE, and MAPE of the DT model combined with BiLSTM were all reduced to a very low level, and the PV prediction effect was significantly improved. The prediction results and prediction errors under sunny, rainy, and cloudy weather conditions were compared and analyzed to further verify the proposed method, as shown in Figures 15–17, respectively, where the comparison models CNN-Attention, GRU, LSTM, CNN, and BiLSTM used the experimental datasets.



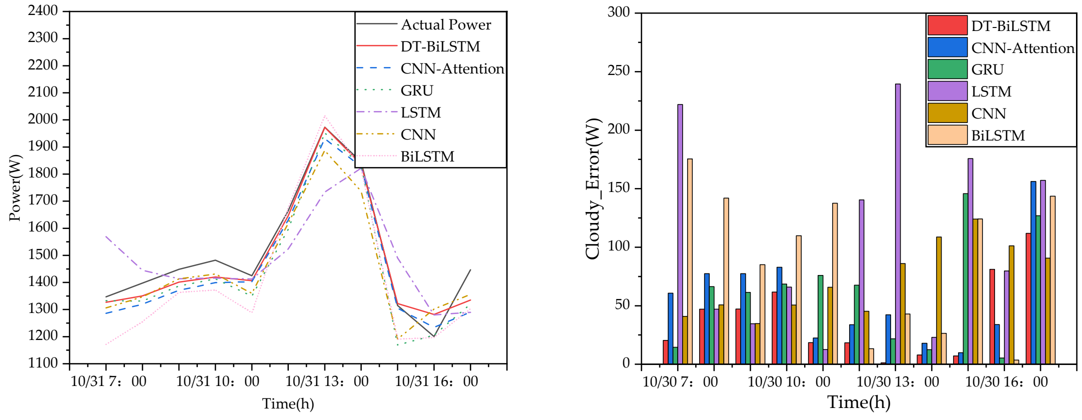
(**a**) (**b**)

Figure 15. Comparison of the predicted and actual PV power under sunny conditions. (**a**) Predicted output power. (**b**) Absolute error.



(**a**) (**b**)

Figure 16. Comparison of the predicted and actual PV power under rainy conditions. (**a**) Predicted output power. (**b**) Absolute error.



(**a**) (**b**)

Figure 17. Comparison of the predicted and actual PV power under cloudy conditions. (**a**) Predicted output power. (**b**) Absolute error.

As we can see from Figures 15–17, the power curve of the DT model combined with BiLSTM fits well with the actual power generation under sunny, rainy, and cloudy weather conditions, and the prediction error is also the smallest, showing excellent prediction performance. Specifically, under sunny weather conditions, during the peak period of PV output (9–14 o'clock), the prediction error of GRU is the smallest, while that of DT-BiLSTM and CNN-Attention is slightly larger; during the trough periods of PV output (7–8 o'clock and 15–17 o'clock), the prediction error of DT-BiLSTM is much smaller than that of GRU and CNN-Attention and reaches a very low level. In summary, DT-BiLSTM has the best prediction effect under sunny weather conditions. Under rainy weather conditions, during the peak period of PV output (9–16 o'clock), DT-BiLSTM makes similar predictions as CNN-Attention with smaller prediction errors; during the trough periods of PV output (7–8 o'clock and 17 o'clock), the prediction error of DT-BiLSTM is significantly smaller than that of CNN-Attention. DT-BiLSTM also shows the best prediction under rainy weather conditions. Under cloudy weather conditions, the power fluctuation is very extreme. During the peak period of PV output (11–15 o'clock), the prediction error of DT-BiLSTM is the smallest, followed by CNN-Attention and GRU. In the trough periods of PV output (7–10 o'clock and 16–17 o'clock), the prediction error of DT-BiLSTM and GRU is the smallest, followed by CNN-Attention. In summary, DT-BiLSTM also achieves the best prediction effect under cloudy weather conditions.

7. Conclusions

This paper proposed a method for the short-term prediction of PV power based on a DT model combined with BiLSTM. The DT model was composed of a mechanism model and a data-driven model converged by DS evidence theory for augmenting the PV power dataset. The BiLSTM network model was used for short-term prediction of the PV power. The following conclusions can be drawn:

1. The PV DT model implements data augmentation, which has the advantages of strong followability, small error deviation, and strong interpretability compared to mechanism and data-driven models, and the fidelity of the data-driven model is further improved after the data update.
2. The PV DT model combined with BiLSTM achieves an accurate prediction of the PV power. Compared with several main PV prediction models, our method had the lowest values for the evaluation indexes of RMSE, MAE, and MAPE under sunny, rainy, and cloudy weather conditions, and the prediction results fit well with the actual values, with a small absolute error value, which was especially obvious under cloudy weather.

In conclusion, the proposed DT model combined with BiLSTM shows excellent prediction accuracy under different weather conditions and can be used for the short-term prediction of ship PV power generation, providing a reference basis for energy management and scheduling of ship power systems and guaranteeing the stable and reliable operation of solar ships. In the future, datasets from different regions and ship routes will be collected to verify the generalizability of the proposed method. PV DT modeling techniques will be further studied in depth. Refined modeling will be carried out on the basis of a single-diode mechanism model with high universality in order to better simulate the actual operation of PV, and the data-driven model will adopt new types of networks with superior performance to improve the fidelity of the PV DT model and the accuracy of PV power prediction.

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

Funding: This work was financially supported by the National Science Foundation of China, grant number 51507025, and the Opening Foundation of Shanghai Ship and Shipping Research Institute Co., Ltd., grant number 85230030.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: Bing Han was employed by the company Shanghai Ship and Shipping Research Institute Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Manolache, A.I.; Andrei, G.; Rusu, L. An Evaluation of the Efficiency of the Floating Solar Panels in the Western Black Sea and the Razim-Sinoe Lagunar System. *J. Mar. Sci. Eng.* **2023**, *11*, 203. [CrossRef]
2. Du, Z.; Chen, Q.; Guan, C.; Chen, H. Improvement and Optimization Configuration of Inland Ship Power and Propulsion System. *J. Mar. Sci. Eng.* **2023**, *11*, 135. [CrossRef]
3. Lee, K.-J.; Shin, D.; Yoo, D.W.; Choi, H.K.; Kim, H.J. Hybrid photovoltaic/diesel green ship operating in standalone and grid-connected mode—Experimental investigation. *Energy* **2013**, *49*, 475–483. [CrossRef]
4. Wang, Z.; Ma, Y.; Sun, Y.; Tang, H.; Cao, M.; Xia, R.; Han, F. Optimizing Energy Management and Case Study of Multi-Energy Coupled Supply for Green Ships. *J. Mar. Sci. Eng.* **2023**, *11*, 1286. [CrossRef]
5. Karatuğ, Ç.; Durmuşoğlu, Y. Design of a solar photovoltaic system for a Ro-Ro ship and estimation of performance analysis: A case study. *Sol. Energy* **2020**, *207*, 1259–1268. [CrossRef]
6. Nyanya, M.N.; Vu, H.B.; Schönborn, A.; Ölçer, A.I. Wind and solar assisted ship propulsion optimisation and its application to a bulk carrier. *Sustain. Energy Technol. Assess.* **2021**, *47*, 101397. [CrossRef]
7. Martínez-López, A.; Ballester-Falcón, P.; Mazorra-Aguiar, L.; Marrero, A. Solar photovoltaic systems for the Short Sea Shipping's compliance with decarbonization regulations in the European Union. *Sustain. Energy Technol. Assess.* **2023**, *60*, 103506. [CrossRef]
8. Ghadimi, N.; Akbarimajd, A.; Shayeghi, H.; Abedinia, O. Two stage forecast engine with feature selection technique and improved meta-heuristic algorithm for electricity load forecasting. *Energy* **2018**, *161*, 130–142. [CrossRef]
9. Rodríguez, F.; Fleetwood, A.; Galarza, A.; Fontán, L. Predicting solar energy generation through artificial neural networks using weather forecasts for microgrid control. *Renew. Energy* **2018**, *126*, 855–864. [CrossRef]
10. Villalva, M.G.; Gazoli, J.R.; Filho, E.R. Comprehensive Approach to Modeling and Simulation of Photovoltaic Arrays. *IEEE Trans. Power Electron.* **2009**, *24*, 1198–1208. [CrossRef]
11. Yan, J.; Qiu, C.; Wang, Y.; Wu, N.; Qu, W.; Zhuang, Y.; Yan, G.; Wang, P.; Zhang, R.; Yan, Y.; et al. Design of Control System for Multistage Distillation Seawater Desalination Device Driven by Photovoltaic-Thermal. *J. Mar. Sci. Eng.* **2023**, *11*, 222. [CrossRef]
12. Raya-Armenta, J.M.; Ortega, P.R.; Bazmohammadi, N.; Spataru, S.V.; Vasquez, J.C.; Guerrero, J.M. An Accurate Physical Model for PV Modules with Improved Approximations of Series-Shunt Resistances. *IEEE J. Photovolt.* **2021**, *11*, 699–707. [CrossRef]
13. Hejri, M.; Mokhtari, H.; Azizian, M.R.; Ghandhari, M.; Söder, L. On the Parameter Extraction of a Five-Parameter Double-Diode Model of Photovoltaic Cells and Modules. *IEEE J. Photovolt.* **2014**, *4*, 915–923. [CrossRef]
14. Ma, X.; Huang, W.H.; Schnabel, E.; Köhl, M.; Brynjarsdóttir, J.; Braid, J.L.; French, R.H. Data-Driven I-V Feature Extraction for Photovoltaic Modules. *IEEE J. Photovolt.* **2019**, *9*, 1405–1412. [CrossRef]
15. Patra, J.C. Chebyshev Neural Network-Based Model for Dual-Junction Solar Cells. *IEEE Trans. Energy Convers.* **2011**, *26*, 132–139. [CrossRef]
16. Ray, B.; Shah, R.; Islam, M.R.; Islam, S. A New Data Driven Long-Term Solar Yield Analysis Model of Photovoltaic Power Plants. *IEEE Access* **2020**, *8*, 136223–136233. [CrossRef]
17. Rubasinghe, O.; Zhang, X.; Chau, T.K.; Chow, Y.H.; Fernando, T. A Novel Sequence to Sequence Data Modelling Based CNN-LSTM Algorithm for Three Years Ahead Monthly Peak Load Forecasting. *IEEE Trans. Power Syst.* **2024**, *39*, 1932–1947. [CrossRef]
18. Zhu, H.; Li, X.; Sun, Q.; Nie, L.; Yao, J.; Zhao, G. A Power Prediction Method for Photovoltaic Power Plant Based on Wavelet Decomposition and Artificial Neural Networks. *Energies* **2016**, *9*, 11. [CrossRef]
19. Zhou, S.Y.; Mao, M.X.; Zhou, L.; Wan, Y.H.; Xi, X.Z. A Shadow Fault Diagnosis Method Based on the Quantitative Analysis of Photovoltaic Output Prediction Error. *IEEE J. Photovolt.* **2020**, *10*, 1158–1165. [CrossRef]
20. Shi, H.; Xu, M.H.; Li, R. Deep Learning for Household Load Forecasting—A Novel Pooling Deep RNN. *IEEE Trans. Smart Grid* **2018**, *9*, 5271–5280. [CrossRef]
21. Akhter, M.N.; Mekhilef, S.; Mokhlis, H.; Almohaimeed, Z.M.; Muhammad, M.A.; Khairuddin, A.S.M.; Akram, R.; Hussain, M.M. An Hour-Ahead PV Power Forecasting Method Based on an RNN-LSTM Model for Three Different PV Plants. *Energies* **2022**, *15*, 2243. [CrossRef]
22. Cantillo-Luna, S.; Moreno-Chuquen, R.; Celeita, D.; Anders, G. Deep and Machine Learning Models to Forecast Photovoltaic Power Generation. *Energies* **2023**, *16*, 4097. [CrossRef]

23. Hejri, M.; Mokhtari, H. On the Comprehensive Parametrization of the Photovoltaic (PV) Cells and Modules. *IEEE J. Photovolt.* **2017**, *7*, 250–258. [CrossRef]
24. Mahmoud, Y.; Xiao, W. Evaluation of Shunt Model for Simulating Photovoltaic Modules. *IEEE J. Photovolt.* **2018**, *8*, 1818–1823. [CrossRef]
25. Babu, B.C.; Gurjar, S. A Novel Simplified Two-Diode Model of Photovoltaic (PV) Module. *IEEE J. Photovolt.* **2014**, *4*, 1156–1161. [CrossRef]
26. Batzelis, E.I. Simple PV Performance Equations Theoretically Well Founded on the Single-Diode Model. *IEEE J. Photovolt.* **2017**, *7*, 1400–1409. [CrossRef]
27. Wang, C.C.; Sun, X.; Qian, H. Study on MPPT of self-seeking optimal perturbation method. *Appl. Electron. Tech.* **2016**, *42*, 142–145.
28. Pan, X.Q.; Zhou, W.S.; Lu, Y.; Sun, N. Prediction of Network Traffic of Smart Cities Based on DE-BP Neural Network. *IEEE Access* **2019**, *7*, 55807–55816. [CrossRef]
29. Lu, K.Y.; Zhang, W.J.; Sun, B. Multidimensional Data-Driven Life Prediction Method for White LEDs Based on BP-NN and Improved-Adaboost Algorithm. *IEEE Access* **2017**, *5*, 21660–21668. [CrossRef]
30. Zhang, W.; Jin, S.; Bian, G.; Cui, Y.; Peng, C.; Xia, H. A Method for Sound Speed Profile Prediction Based on CNN-BiLSTM-Attention Network. *J. Mar. Sci. Eng.* **2024**, *12*, 414. [CrossRef]
31. Xu, J.; Ding, W.; Gong, Q.S.; Hu, X.Y.; Yu, H.Q. A Super Point Detection Algorithm Under Sliding Time Windows Based on Rough and Linear Estimators. *IEEE Access* **2019**, *7*, 43414–43427. [CrossRef]
32. Tang, Z.F.; Ma, J.W.; Liu, W.X.; Zhang, P.F.; Lv, F.Z. Sliding Window Dynamic Time-Series Warping-Based Ultrasonic Guided Wave Temperature Compensation and Defect Monitoring Method for Turnout Rail Foot. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2022**, *69*, 2681–2695. [CrossRef] [PubMed]
33. Cui, Q.; Zhu, J.Z.; Shu, J.; Huang, L.; Ma, Z.T. Comprehensive Evaluation of Electric Power Prediction Models Based on D-S Evidence Theory Combined with Multiple Accuracy Indicators. *J. Mod. Power Syst. Clean Energy* **2022**, *10*, 597–605. [CrossRef]
34. Pan, Y.; Zhang, L.M.; Li, Z.W.; Ding, L.Y. Improved Fuzzy Bayesian Network-Based Risk Analysis with Interval-Valued Fuzzy Sets and D–S Evidence Theory. *IEEE Trans. Fuzzy Syst.* **2020**, *28*, 2063–2077. [CrossRef]
35. Zhao, G.Z.; Chen, A.G.; Lu, G.X.; Liu, W. Data fusion algorithm based on fuzzy sets and D-S theory of evidence. *Tsinghua Sci. Technol.* **2020**, *25*, 12–19. [CrossRef]
36. Ma, J.T.; Liu, H.; Peng, C.; Qiu, T.S. Unauthorized Broadcasting Identification: A Deep LSTM Recurrent Learning Approach. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 5981–5983. [CrossRef]
37. Luo, C.Y.; Li, X.T.; Ye, Y.M. PFST-LSTM: A Spatio Temporal LSTM Model with Pseudo flow Prediction for Precipitation Nowcasting. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 843–857. [CrossRef]
38. Ma, C.X.; Dai, G.W.; Zhou, J.B. Short-Term Traffic Flow Prediction for Urban Road Sections Based on Time Series Analysis and LSTM_BILSTM Method. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 5615–5624. [CrossRef]
39. Wu, H.J.; Yang, M.R.; Yang, S.Q.; Lu, H.; Wang, C.Q.; Rao, Y.J. A Novel DAS Signal Recognition Method Based on Spatiotemporal Information Extraction with 1DCNNs-BiLSTM Network. *IEEE Access* **2020**, *8*, 119448–119457. [CrossRef]
40. Boubaker, S.; Benghanem, M.; Mellit, A.; Lefza, A.; Kahouli, O.; Kolsi, L. Deep Neural Networks for Predicting Solar Radiation at Hail Region, Saudi Arabia. *IEEE Access* **2021**, *9*, 36719–36729. [CrossRef]
41. Alkahtani, H.; Aldhyani, T.H.H.; Alsubari, S.N. Application of Artificial Intelligence Model Solar Radiation Prediction for Renewable Energy Systems. *Sustainability* **2023**, *15*, 6973. [CrossRef]