This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Real-Time Power Prediction for Bifacial PV Systems in Varied Shading Conditions: A Circuit-LSTM Approach Within a Digital Twin Framework

Dou Hong, *Student Member, IEEE*, Jieming Ma, *Senior Member, IEEE*, Kangshi Wang, *Member, IEEE*, Ka Lok Man, *Member, IEEE*, Huiqing Wen, *Member, IEEE*, and Prudence Wong, *Member, IEEE*

***Abstract*—The prediction of bifacial photovoltaic (bPV) system performance under variable conditions has persistently challenged researchers and practitioners alike, largely due to the unstable and imprecise irradiance measurements and the extensive training processes required for machine learning-based methods. Addressing these issues, this study introduces an innovative digital twin system that integrates a novel circuit-long short-term memory (LSTM) model with the newly proposed triangle-shading pattern estimation method, eliminating dependencies on direct irradiance measurements and historical data. Our approach uniquely combines the adaptability of LSTM networks with circuit models, facilitating real-time power prediction with unprecedented accuracy and efficiency. Comprehensive evaluations across various shading scenarios demonstrate the proposed model's superior performance, consistently reducing mean absolute error, mean squared error, and root mean squared error by over 50% compared with existing methods. This breakthrough offers a scalable, cost-effective solution for optimizing the deployment and management of bPV systems, marking a significant advancement in the field of photovoltaic research.**

***Index Terms*—Bifacial photovoltaic (bPV) systems, bifacial PV circuit model, digital twin systems (DTS), long short-term memory (LSTM) network, power prediction, shading pattern estimation.**

I. Introduction

THE urgent need to counter global warming and air pollution has accelerated the search for sustainable alternative energy sources. Photovoltaic (PV) systems are a promising source of renewable energy due to their sustainability and abundance [1]. Bifacial PV (bPV) modules utilize transparent glass, permitting light to traverse and reflect off the ground underneath. This reflected irradiance can be captured by the module's rear side. Moreover, this transparent covering results in reduced working temperatures due to lesser infrared absorption compared with the traditional backsheet laminate [2]. The combination of increased irradiance and cooler temperatures enhances power generation in comparison to monofacial PV panels. According to the International Technology Roadmap for PV, bPV installations are expected to increase their market share by 40% by 2033 [3].

Despite the increasing adoption of bPV technology, the evaluation of bPV systems remains a complex issue due to the difficulty of accurately identifying the received irradiance from various light sources. The dynamic nature of irradiance creates significant challenges in controlling and managing PV systems, especially in partial shading conditions (PSCs), resulting in divergent evaluations of PV panels [4]. The challenges intensify with bPV systems, as predicting back-side irradiance is particularly complex due to the intricate influence of albedo. Therefore, a necessity exists for a versatile real-time performance prediction system capable of operating effectively under diverse conditions to support the successful deployment of bPV systems.

The performance of a bPV system can be predicted by circuit-based and data-driven approaches. Equivalent electric circuit models are commonly used to predict electrical characteristics [4], [5], [6]. Based on the similarity of the characteristics between monofacial and bifacial PV panels, several researchers applied an improved single-diode model (SDM) with adjusted parameters [7]. An equivalent irradiance method applied SDM is proposed to estimate the irradiance based on measured front and rear irradiance [8]. However, these equivalent methods predict output-based solely on estimated parameters. To better present the characteristics of bPV panels and increase the accuracy, researchers applied two connected SDMs presenting each side of the panel [9], [10], [11]. In the literature, the generated power of a bPV panel is the sum of the power generated from both sides of the PV modules. However, circuit-based modeling methods require accurate environmental information [12], which is challenging to obtain, and unreliable during dynamic shading conditions [13].

The data-driven prediction approaches rely on weather data, historical output records, or constructors' datasheets. With the

Manuscript received 23 March 2024; accepted 15 April 2024. This work was supported in part by Suzhou Science and Technology Project-Key Industrial Technology Innovation under Grant SYG202122, in part by the Qing Lan Project of Jiangsu Province, Future Network Scientific Research Fund Project under Grant FNSRFP-2021-YB-41, in part by the XJTLU Postgraduate Research Scholarship under Grant PGRS1906004, and in part by the XJTLU AI University Research Centre and Jiangsu (Provincial) Data Science and Cognitive Computational Engineering Research Centre at XJTLU . *(Corresponding author: Jieming Ma.)*

Dou Hong, Jieming Ma, Kangshi Wang, Ka Lok Man, and Huiqing Wen are with the School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou 215123, China (e-mail: jieming.ma@xjtlu.edu.cn).

Prudence Wong is with the School of Computer Science, University of Liverpool, L69 3BX Liverpool, U.K.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JPHOTOV.2024.3393001.

Digital Object Identifier 10.1109/JPHOTOV.2024.3393001

2156-3381 © 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on May 30,2024 at 09:38:08 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

constructors' datasheet, the generated power is calculated with an analyzed method adjusting maximum power point (MPP) or current at MPP under standard testing conditions (STC) [14]. However, the limitations of these models become evident when dealing with limited data. To bolster their precision, the realm of power prediction has recently turned to the application of machine learning methods [15]. It is worth noting that while these methods have been integrated into power prediction, their application within bPV systems remains relatively unexplored. One notable avenue of exploration involves leveraging the potential of recurrent neural networks (RNNs) and long short-term memory (LSTM) models in tandem. The literature showcases instances of such deployment in power prediction, as highlighted in [16], [17], [18], and [19]. An interesting development is the fusion of LSTM and RNN methodologies, resulting in an LSTM-RNN hybrid model, as chronicled in [20]. This innovation aims to evaluate the effectiveness of various forecasting methods in predicting the performance of PV plants. Furthermore, progress has been achieved by combining LSTM with the Pearson coefficient for feature selection [21]. Qi et al. [22] developed a hybrid model that combines a convolutional neural network (CNN) with an LSTM, enhanced by power series and ramp series. However, machine learning-based methods construct complex models with numerous parameters, resulting in a significant computational burden. Moreover, the accuracy of these models primarily depends on the size of the training dataset, highlighting the essential need for historical data during model training. Meanwhile, both the circuit-based models and data-driven models rely on general irradiance information and are unlikely to maintain accurate power predictions during dynamic shading conditions due to their limited knowledge of local irradiance, especially for PSCs.

To address the challenges of real-time power prediction for bPV systems under dynamic conditions, we introduce a digital twin system (DTS) for bPV systems that efficiently integrates the triangle shading pattern estimation (T-SPE) method for environment perception and the circuit-long short-term memory (Circuit-LSTM) power prediction model. The primary contributions of this article are as follows.

1) *Innovative digital twin framework:* A DTS for bPV systems is proposed, which integrates multiple methodologies and efficiently consolidates diverse data streams for real-time scenarios.
2) *T-SPE method for environment perception:* The T-SPE method provides an environment perception mechanism, enabling the estimation of front-side irradiance without direct measurements. Its computational efficiency ensures precise, real-time insights, pivotal for power forecasting.
3) *Circuit-LSTM power prediction model:* This method embodies the principle of multimodel coupling by merging circuit models with LSTM. The hybrid model does not require historical data and effectively analyzes both long-term and short-term data patterns.

The rest of this article is organized as follows. The electrical characteristics of the bPV module are given in Section I. Section II describes the structure of the proposed DTS with the T-SPE method and Circuit-LSTM power prediction model. In Section IV, case studies and numerical results are discussed to show the superiority of the proposed prediction model. Finally, Section V concludes this article.

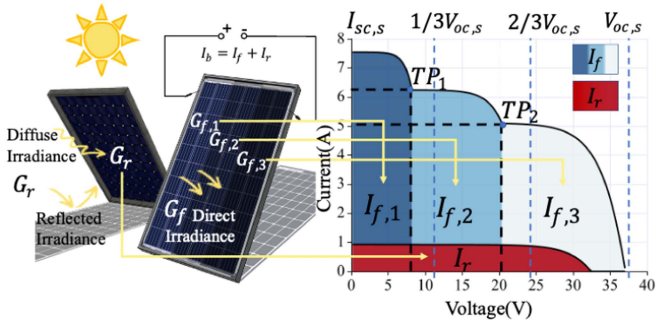


Fig. 1. BPV panel under partial shading: irradiance sources and corresponding I–V curve steps.

II. Electrical Characteristics of bPV Modules Under PSCs

Understanding the electrical characteristics of bPV modules under PSCs is crucial for predicting their performance. The bPV panels can absorb irradiance from both sides for power generation. The main source absorbed by the front side of the PV module is direct irradiance, which mainly depends on the solar altitude angle and atmospheric transparency. The rear sides receive diffused and reflected irradiance, which is normally weaker than direct irradiance as shown in Fig. 1.

A distinctive feature of bPV modules is their dual-sided energy generation. Consequently, the output from a bPV module is the cumulative energy produced from both its front and rear sides as described in [8]

$$I = I_f + I_r \tag{1}$$

where the I is the output current of a bPV module, and I_f and I_r are the current generated from the front side and the rear side. The performance of each side can be modeled as a traditional PV model based on received irradiance G_f, G_r, and temperatures

$$\begin{cases} I_f = (I_{\mathrm{sc},f} + K_i \Delta T) G_f / G_s \\ I_r = (I_{\mathrm{sc},r} + K_i \Delta T) G_r / G_s \end{cases} \tag{2}$$

where $I_{\mathrm{sc},f}$ is the short circuit current of the front sides, $I_{\mathrm{sc},r}$ represents the short circuit current of the rear sides, K_i is the short circuit current coefficient, ΔT is the temperature difference between experimental temperature (T) and standard temperature (25°C), G_f and G_r denote the irradiance received by the front and rear sides of the panel, and G_s is the standard irradiance (1000 $\mathrm{W/m^2}$).

In our PV system modeling, each module within a PV string comprises a series of PV cells alongside one bypass diode, reflecting a strategic simplification to facilitate precise performance prediction at the string level. This configuration allows for a focused analysis of nonactive diodes based on specific shading conditions affecting the modules, thereby refining our understanding of the system's overall behavior under varied environmental influences.

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on May 30,2024 at 09:38:08 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

The characteristics delineated previously portray the performance of the bPV system under ideal uniform irradiance conditions (UIC). However, real-world scenarios introduce complexities, with factors like shading from adjacent elements, such as trees, buildings, and clouds, pushing the PV system into PSCs. In PSC scenarios, PV strings grapple with fluctuating irradiance levels on their front side. This variability can result in imbalanced electrical loads and potential damage, particularly in cells experiencing minimal irradiation that are susceptible to hotspot formation. While the bypass diode mitigates these challenges, it simultaneously introduces complexities reflected in multistair current–voltage (I–V) and multipeak power–voltage (P–V) characteristic curves. The bypass diode possesses a forward voltage V_B, which serves as the base voltage, introducing a voltage drop V_D in the PV system upon activation [23]. In addition, bPV configurations must also consider the shadowing effects on the rear side, further complicating performance evaluation and optimization. Fig. 1 depicts I–V characteristic of a bPV string under three different irradiance conditions. These conditions result in distinct turning points (TPs) that region the curve into three separate regions. Within each region, the active cells have a consistent shading pattern. The location of the TPs and the voltage drop V_D, in the PV system, are affected by the forward voltage of the bypass diode V_B. As a result, (1) is reformulated to (3), which describes the I–V correlation for the ith region [6]

$$\begin{cases} I_{S,i} = I_{\text{ph},i} - I_{o,i}\left[\exp\left(\frac{V_{S,i}+I_{S,i}R_s}{A_i N_{SP,i} V_{t,i}}\right) - 1\right] - \frac{V_{D,i}}{R_p} \\ V_{D,i} = V_S + I_S R_s - V_{\text{tp},i} + V_{BD} + N_{B,i} V_{B,i} \\ I_{\text{ph},i} = I_{f,i} + I_{r,i} \end{cases} \tag{3}$$

where $I_{S,i}$ is the current flowing through the ith string; $I_{\text{ph},i}$ is the photo-generated current, which is the sum of currents generated from the front ($I_{f,i}$) and rear ($I_{r,i}$) sides due to irradiance; $I_{o,i}$ represents the reverse saturation current of the module; $V_{S,i}$ is the voltage across the ith string; $V_{D,i}$ defines the voltage drop across the diode; R_s is the series resistance and R_p stands for the parallel or shunt resistance. A_i is the ideality factor, indicating how closely the diode follows ideal behavior. $N_{SP,i}$ is the number of series-connected cells in the ith string, while $V_{t,i}$ is the thermal voltage; $V_{\text{tp},i}$ signifies the voltage at TPs, V_{BD} is the forward voltage of a blocking voltage, and $N_{B,i}$ and $V_{B,i}$ represent the number of bypass diodes and the forward voltage across each diode, respectively.

III. Proposed Digital Twin for bPV Modules Power Prediction

Digital twins, as virtual replicas of physical systems, are becoming increasingly pivotal in enhancing the design, monitoring, and maintenance processes of various engineering systems. The main part of the proposed framework lies in the seamless integration of the T-SPE method and Circuit-LSTM models. The T-SPE method fuses physical and environmental data, enabling comprehensive environment perception for a nuanced understanding of bPV module behavior. Meanwhile, real-time fused data within the DTS facilitates dynamic data processing, leading to real-time predictions. The DTS also provides a management interface for real-time observation and management of data and

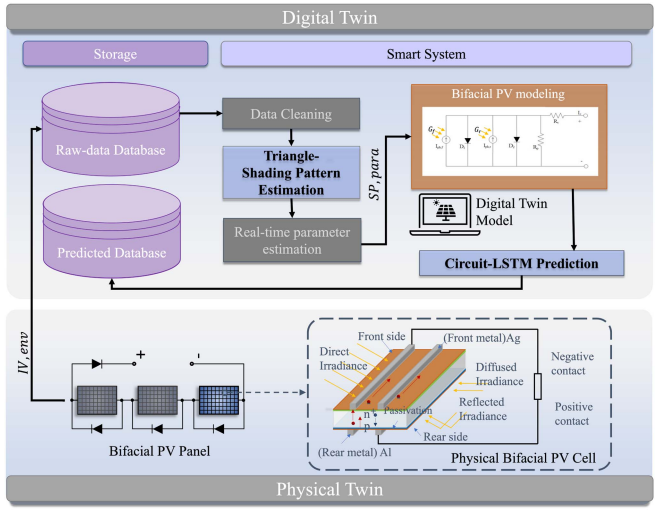


Fig. 2. Integrated workflow of the digital twin framework for bPV modules.

predictions. Therefore, the proposed DTS for bPV modules promises significant advancements by leveraging principles of data fusion, environment perception, and multimodel coupling. This section details the intricate design and mechanism of a digital twin precisely tuned for bPV modules, introducing a proposed shading pattern estimation technique and the integrative prowess of the Circuit-LSTM method for comprehensive power prediction.

A. Overview of the Digital Twin for bPV Modules

The digital twin for bPV modules, depicted in Fig. 2, bridges the bPV modules' digital twin with its physical counterpart, emphasizing the relationship between data storage, fused methods, and the real-world PV system. This integration enables real-time monitoring and predictive insights.

The raw-data database within the storage section captures real-time I–V and environmental outputs with G_r and T, from the physical domain in 1-min intervals, which comprises a bPV system with three panels operating under PSCs. This data originates from a bPV system comprising three panels operating under PSCs. Subsequently, the smart system undertakes rigorous data-cleaning procedures to ensure accuracy and relevance. This involves filtering out missing or erroneous entries and addressing outliers, thereby maintaining the integrity and reliability of both the I–V data and associated parameters.

With the data refined, the T-SPE method harnesses the acquired data to discern the shading patterns that sway the bPV module's performance, an imperative step in environment perception. Following this, the real-time parameter estimation further refines the process, pinpointing the model parameters crucial for the accurate representation of the bPV model. With the established shading patterns and the essential model parameters, a bPV model is crafted as the circuit diagram, capturing the intricacies of the bPV model. Building upon this foundation, the Circuit-LSTM couples the circuit-based model with the LSTM methodology in the digital version. The hybrid model responds

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on May 30,2024 at 09:38:08 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

to real-time conditions and stores the results in the predicted database.

B. Proposed T-SPE Algorithm

In the realm of bPV systems, understanding shading patterns is crucial for accurately predicting system performance and represents a form of environmental perception in DTS. Environmental perception refers to the process by which a system interprets its surrounding conditions, incorporating sensory data, such as irradiance levels and temperature variations, to inform decision-making and optimize performance. However, in the bPV system, the front irradiance sensors are affected by dynamic shading, such as clouds, and that is highly location-related, which brings challenges for front irradiance measurement. To tackle this challenge, we present the T-SPE method for shading pattern estimation with only rear irradiance and temperature based on the fact that the rear irradiance sensor shows less variation throughout the whole plant.

The proposed T-SPE method leverages the Douglas–Peucker (D-P) algorithm to accurately estimate the shading pattern of a bPV array using characteristic points on the I–V curve. The D-P algorithm, originally conceived for polyline simplification, operates on the premise that a subset of a polyline's vertices can approximate it within a specified tolerance, ϵ [24]. This approach allows for identifying characteristic points recursively. Reference lines can be drawn, each potentially harboring a TP. Given that TPs materialize around $V_{\text{tp},i} = i/(n)V_{\text{oc}} - V_D$, we define a set of selected points SPs as candidates of TPs $\text{SPs} = \{\text{SP}_i \,|\, i \le n\}$, with $V_{\text{sp},i} = i/(n)V_{\text{oc}} - V_D$, where n is the number of modules in the bPV string. To distinguish TP from SPs, the reference points are chosen between adjacent SPs, which are located at $V_{\text{rp},i} = (V_{\text{sp},i-1} + V_{\text{sp},i})/2$. Reference lines, constructed with two adjacent reference point RP_i and RP_{i+1}, are denoted by $\{\vec{\text{RL}}_i : \text{RP}_i\vec{\text{RP}}_{i+1} \,|\, i \le n\}$.

Building upon this foundation, the I–V curve of a bPV string can be analyzed in regions delineated by TPs. For a selected point SP_i to be earmarked as a TP, it should fulfill two conditions. First, the magnitude of this distance must exceed the threshold ϵ, set as one in this case. The directed distance from a selected point to its associated reference line is computed using

$$\text{dis}(\text{SP}_i) = \frac{1}{2} \cdot \frac{\det(\vec{v_i}, \vec{w_i})}{\|\vec{\text{RL}_i}\|} \tag{4}$$

where $\vec{v_i} = \vec{\text{RP}_i} - \vec{\text{SP}_i}$, $\vec{w_i} = \vec{\text{RP}_{i+1}} - \vec{\text{SP}_i}$, and $\vec{\text{RL}_i} = \vec{\text{RP}_i} - \vec{\text{RP}_{i+1}}$. As illustrated in Fig. 3, both $\text{SP}_{1,p}$ and $\text{SP}_{2,u}$ are notable due to their significant distances from their respective reference lines $\vec{\text{RL}_i}$. Second, the vectorial distance from SP_i to its corresponding reference line $\vec{\text{RL}_i}$ must be positive, signifying its location on the lower left. Therefore, only $\text{SP}_{1,p}$ is identified as the TP.

After the TPs are identified, the related characteristics are obtained and analyzed based on the locus of TPs and stored in the matrix SI with the characteristics mentioned in Section II. The matrix is instrumental in storing the irradiance values derived from different light sources impacting both sides of the bPV panel. Under PSC, the front side of the PV module experiences notable irradiance fluctuations, with increases of approximately 200 W/m^2 per hour due to shifts in the solar altitude angle, while the rear side daily changes by less than 100 W/m^2 [25]. This irradiance disparity makes accurate measurement of the front-side irradiance complex, especially under PSC. The proposed methodology allows for the determination of front-side irradiance using rear irradiance, temperature, and total current output, as represented in (5)

$$\begin{cases} G_{f,i} = I_{f,i}/(I_{sc,f} + K_i\Delta T)G_s \\ I_{f,i} = I_{\text{ph},i} - I_{r,i} \\ I_{\text{ph},i} = I_{s,i} + I_{o,i}\left[\exp\left(\frac{V_{S,i}+I_{S,i}R_s}{A_i N_{SP,i} V_{t,i}}\right) - 1\right] - \frac{V_{D,i}}{R_p} \\ I_{r,i} = (I_{\text{sc},r} + K_i\Delta T)G_r/G_s. \end{cases} \tag{5}$$

To provide a comprehensive representation of the shading patterns, the number of activated modules N_{SP} and nonactive diodes N_B in each region is recorded in SI. N_{SP}, can be derived from the voltage differences at adjacent TPs. This is captured by the relationship shown in (6)

$$N_{\text{SP},i} = \text{round}((V_{\text{TP},i} - V_{\text{TP},i-1})/V_{\text{oc}}). \tag{6}$$

The number of nonactive diodes N_B can be calculated from N_{SP} as (7)

$$N_{B,i} = \sum_{i=i+1}^{N} N_{\text{SP},i}. \tag{7}$$

After combining the duplicate data, we can form the shading information matrix SI with $[G_f, G_r, T, N_{\text{SP}}, N_B]$. This matrix serves as a concise representation of the shading patterns and offers insights into the bPV system's performance.

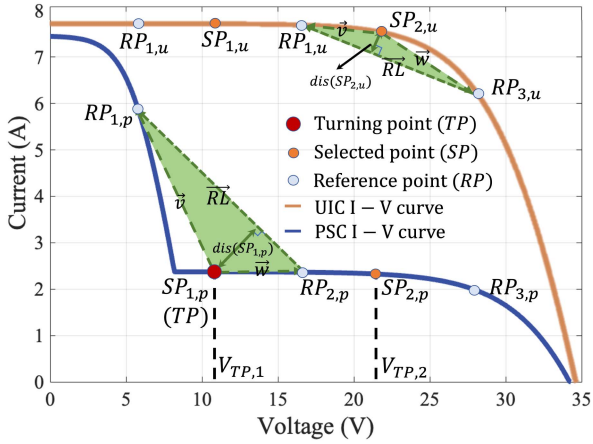


Fig. 3. Illustration of the T-SPE method: identifying TPs on I–V curves under PSC and UIC.

C. Proposed Circuit-LSTM Power Prediction Method

We introduce the Circuit-LSTM model for power prediction within our proposed DTS, exemplifying multimodel coupling to seamlessly integrate physical data and computational intelligence, ensuring accurate bPV string power predictions even without historical data. The Circuit-LSTM model leverages the capabilities of LSTM networks known for their proficiency

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on May 30,2024 at 09:38:08 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

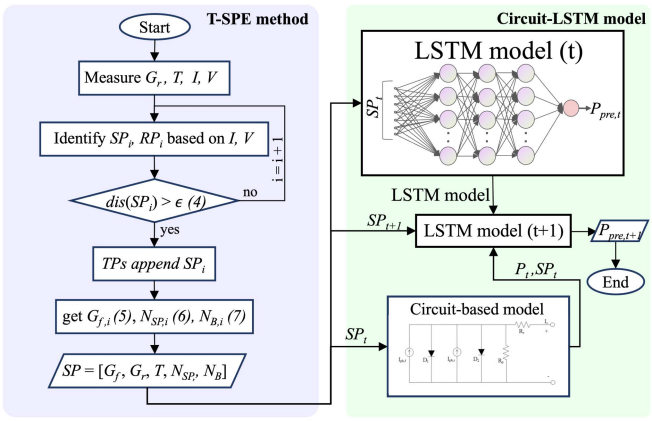


Fig. 4. Flowchart illustrating the operation of the Circuit-LSTM power prediction method.

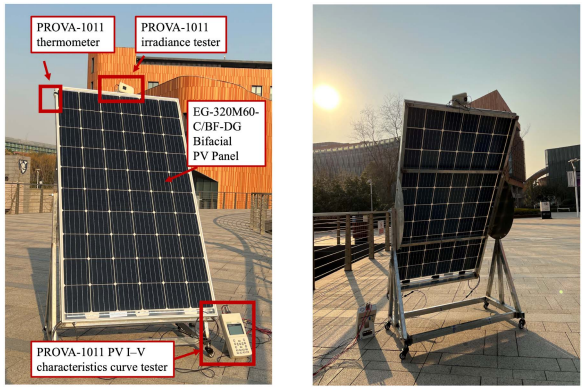


Fig. 5. Experimental setups: (a) front view and (b) back view.

in capturing temporal dependencies in sequential data. Our proposed DTS achieves robust power predictions even without extensive historical data by integrating the LSTM's ability to remember long-term patterns with the circuit-based model's structural understanding of the bPV system. Fig. 4 offers a visual representation of this intricate model, tracing from environment perception from sensor output and shading patterns estimated via T-SPE to the harmonized predictions produced by the collaborative dynamism of the circuit model and LSTM.

1) Accumulation Phase: Initially, the system predominantly operates with the circuit-based bPV model in the accumulation phase. A formidable challenge in this domain is acquiring real-time environmental data in dynamic shading conditions. To circumvent this, the data fusion process integrates the shading information matrix SI derived from the T-SPE method, enabling the circuit model to deliver preliminary power predictions without historical data.

2) Iteration Phase With Feedback Mechanism: In the Circuit-LSTM model's Iteration Phase, the LSTM model primarily uses the SI matrix derived from the T-SPE method as its input. This SI matrix, encapsulating detailed shading patterns affecting the bPV system, enables the LSTM to generate power predictions by analyzing temporal trends in conjunction with current shading dynamics. This process allows the LSTM to forecast power output with a nuanced understanding of environmental factors that directly impact the bPV system's efficiency.

Simultaneously, the output of the circuit model, which provides real-time power predictions based on current environmental conditions, is pivotal in updating the LSTM network for future rounds of prediction. After each LSTM prediction cycle, the circuit model's latest outputs adjust the LSTM's internal parameters, ensuring that subsequent forecasts are informed by the most recent physical data. This iterative updating mechanism enriches the LSTM's learning process, enabling it to adapt its predictions in alignment with ongoing changes in the bPV system's performance and environmental conditions.

In essence, the Circuit-LSTM model exemplifies a groundbreaking approach in the realm of bPV power prediction, leveraging the strength of multimodel coupling to enhance predictive accuracy. Through its innovative iteration phase with feedback mechanism, it dynamically integrates the immediate insights from the circuit model with the predictive prowess of the LSTM, ensuring that each prediction is as informed and current as possible.

IV. Experimental Results

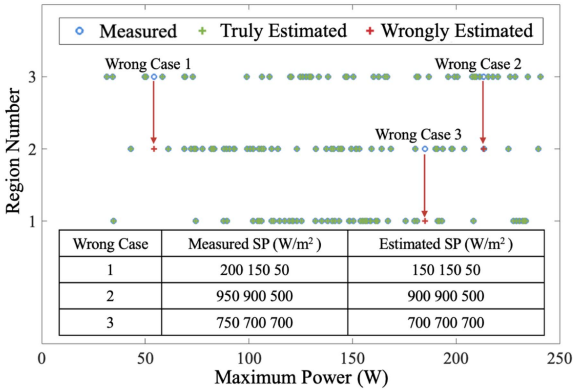
To validate the proposed DTS, we performed a series of computational experiments. The simulations were carried out on a consistent computing environment, specifically a computer equipped with an Apple M1 Pro CPU, 32 GB RAM, running macOS Ventura 13.4.1, and MATLAB R2023a. The experimental data were sourced in Jiangsu, China, from an EG-320M60-C/BF-DG PV panel, which can be regarded with 3 PV modules in 1 PV string. Under the STC of 1000 W/m^2 and 25°C, the panel exhibits a V_{oc} of 31.53 V, an I_{sc} of 8.10 A, a K_i value of 0.039%/°C, and a K_v value of $-0.295\%/^{\circ}$C. The experimental configurations are depicted in Fig. 5.

The proposed T-SPE method is pivotal in the DTS, demonstrating the strength of environment perception. To understand the effectiveness and accuracy of our DTS, we evaluated the performance of the T-SPE method in predicting shading information matrices SI under various environmental conditions as Fig. 6. This figure provides the simulation results of the T-SPE method for 150 randomly generated shading conditions grouped in three different shading patterns, 50 cases for each shading pattern. These conditions were spread across an irradiance range of 100–1300 W/m^2 to reflect a broad spectrum of possible environmental situations. For the generation of these conditions, we utilized a randomization process that introduced a minimum irradiance variation of 150 W/m^2 between individual test cases, particularly for the PSC subsets with 2 or 3 regions. This ensured that each generated shading pattern was distinct and recognizable, thereby providing a stringent test of the T-SPE method's capabilities.

During the simulation, there were three cases that were wrongly estimated, resulting in an accuracy of 98%. This issue arises when neighboring regions exhibit similar irradiance levels. For example in Wrong Case 2, the front irradiance matrix

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on May 30,2024 at 09:38:08 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



Wrong Case	Measured SP (W/m^2)	Estimated SP (W/m^2)
1	200 150 50	150 150 50
2	950 900 500	900 900 500
3	750 700 700	700 700 700

Fig. 6. Performance of shading pattern prediction under various environmental conditions.

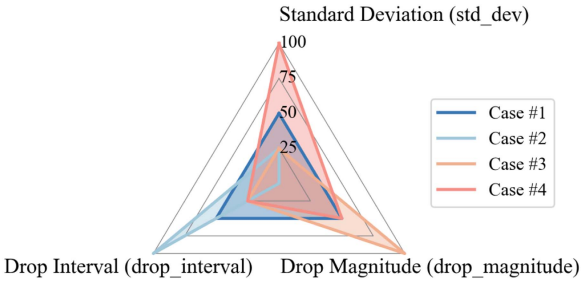


Fig. 7. Visualization of criteria std _ dev, drop _ magnitude, and drop _ interval for four simulated shading patterns: Case #1, Case #2, Case #3, and Case #4.

of the bPV string with three regions (950W/m^2, 900 W/m^2, and 500 W/m^2), and the T-SPE method wrongly predicted the matrix with two irradiance levels (950W/m^2, 900 W/m^2, and 500 W/m^2), but still resulting in the same predicted power of 213 W. The results validate the effectiveness of the T-SPE method in estimating front-side irradiance without the need for its measurement, making it suitable for real-time applications in bPV systems.

In order to validate the efficacy of the circuit-LSTM method integrated, we conducted simulation cases that encompass four distinct shading scenarios based on the following criteria.

1) *Standard deviation (std _ dev):* Indicates irradiance variability. Higher values hint at turbulent atmospheric conditions, whereas lower values suggest a stable irradiance profile.
2) *Drop magnitude (drop _ magnitude):* Measures the extent of sudden irradiance decreases. Zero implies clear skies, while higher values point to significant shading events, such as from cloud cover.
3) *Drop interval (drop _ interval):* Dictates shading event frequency. Shorter intervals indicate frequent shading, whereas longer ones suggest occasional obstructions.

These criteria replicate real-world scenarios bPV systems face, accounting for environmental variability. The simulated configurations for each shading pattern are presented in Fig. 7.

The simulated shading patterns are delineated as follows: Case #1 emulates an environment with moderate irradiance

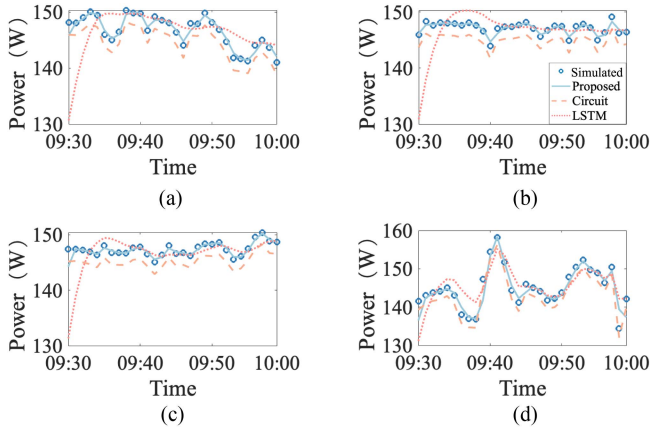


Fig. 8. Comparative power prediction performance across diverse shading patterns among proposed Circuit-LSTM model, circuit model, and LSTM model under (a) Case #1, (b) Case #2, (c) Case #3, and (d) Case #4.

variability, average-magnitude shading events, and medium frequency occurrences; Case #2 portrays a stable irradiance scenario devoid of shading events, epitomizing the calmest condition; Case #3 is characterized by average variability, pronounced shading occurrences, and high-frequency events; and Case #4 represents a highly turbulent environment marked by significant variability, average-magnitude shading events, and frequent occurrences.

These diverse scenarios ensure our system is well-calibrated to handle a range of real-world conditions. As illustrated in Fig. 8, the strength of the circuit model has been proved due to its consistency, meanwhile, the hybrid model still demonstrated its adaptability to moderate fluctuations and shading events, also in stable conditions. While in the relatively stable scenario, the LSTM model struggles in the beginning, then its adaptability becomes evident as it closely follows the simulated data trends. The possible reason is that LSTM models have an initial state that may not align perfectly with the data at the beginning of the sequence. Case #4 with a turbulent environment and high variability, tested all models' limits. The hybrid model's advantage is evident, providing accurate predictions amidst this variability.

Fig. 9 presents a meticulous comparison of the proposed Circuit-LSTM model with other prediction models. For a comprehensive evaluation, we have incorporated the modern time series-based LSTM, CNN-LSTM model [22], RNN-LSTM model [26], and the traditional machine learning approach, artificial neural network (ANN) model [27]. The benchmark for this comparison hinges on the power output prediction of a bPV system located in Jiangsu, China, on 31 October, 2022, with Case #5, a sunrise condition with fewer clouds, and Case #6, a more cloudy moving scenario. To ensure a fair and robust comparison, specific hyperparameters have been meticulously selected and applied across all models, with their details provided in Table I. The chosen hyperparameters are pivotal in optimizing each model's ability to accurately predict power output, reflecting a balance between computational efficiency and prediction precision. The same hyperparameters underpin the models in

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on May 30,2024 at 09:38:08 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

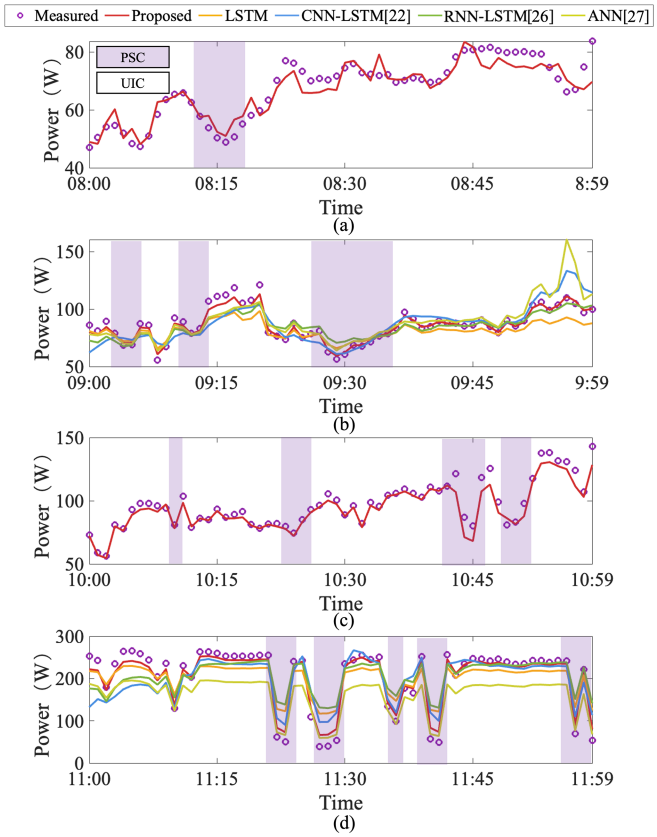


Fig. 9. Power prediction performance comparison among the proposed Circuit-LSTM model, LSTM, CNN-LSTM [22], RNN-LSTM [26], and ANN model [27]: (a) accumulation phase, (b) iteration phase in Case #5, (c) accumulation phase, and (d) iteration phase in Case #6.

TABLE I
HYPERPARAMETERS FOR MODEL COMPARISONS IN BPV SYSTEM POWER PREDICTION

Parameter	Proposed	LSTM	CNN-LSTM	RNN-LSTM	ANN
LSTM Units	32	32	64, 128	64, 128	N/A
Convolution Layers	N/A	N/A	2	N/A	N/A
Fully Connected Layers	1	1	3	1	3
Activation Function	N/A	N/A	ReLU	N/A	ReLU
Dropout	N/A	N/A	0.1	N/A	N/A
Max Epochs	150	150	150	150	150
Initial Learn Rate	N/A	N/A	N/A	0.001	0.001
Gradient Threshold	2	2	2	2	N/A

additional experiments, including two-day predictive scenarios, ensuring consistency and comparability of our results.

A significant observation is made in the UIC environment of Case #5. Notably, the Circuit-LSTM model has the unique ability to predict without relying on historical data. This feature sets it apart from other prediction models, as evident in Table II for the Accumulation Phase, other models are denoted with "N/A," elucidating their inherent limitation of sole training without simultaneous prediction during the accumulation phase.

TABLE II
COMPARISON OF RMSES (W) FOR POWER PREDICTION RESULTS

Methods	Phases in Case #5		Phases in Case #6	
	Accumulation	Iteration	Accumulation	Iteration
Proposed	4.042	5.015	7.540	21.234
LSTM	N/A	11.268	N/A	38.965
CNN-LSTM [22]	N/A	10.763	N/A	40.739
RNN-LSTM [26]	N/A	21.771	N/A	79.658
ANN [27]	N/A	12.854	N/A	53.151

Furthermore, under the UIC environment, other models except the proposed model consistently tend to underestimate predictions, especially when power levels are above the mean value. Such predictions gravitate toward the average rather than the true value, leading to escalating errors over time. This phenomenon might be attributed to these models' predilection for conservative predictions, potentially due to their inability to handle nonlinearities and intricate patterns in the data. This also hints at the challenges faced when not effectively fusing data sources and knowledge from various domains.

Transitioning to the hybrid UIC and PSC environment of Case #6, while the LSTM and CNN-LSTM models show good performance, the proposed Circuit-LSTM model still outperforms all models. This could be credited to the Circuit-LSTM's superior data processing capabilities, allowing it to capture intrinsic patterns and variations in the data across both environments more proficiently. The success of the Circuit-LSTM model in this environment underscores its efficacy in multimodel coupling, offering accurate predictions.

In a comprehensive two-day experiment spanning 5–6 February, 2023, the study encompassed data between 6:00–18:00 on each day as presented in Fig. 10. Notably, 5th February, is a sunny day, while, the next day is cloudy with partial shading, which causes a salient observation. Following an initial midday zenith, the power output displayed a sharp decrement, only to recover subsequently. At this juncture, the traditional models, namely, LSTM, CNN-LSTM, RNN-LSTM, and ANN, manifested conspicuous errors, although their predictions remained congruent. This observed deviation can be ascribed to potential anomalies in light measurement values, further exacerbated by the complexities of PSCs. The proposed Circuit-LSTM method, embedded in the philosophy of data fusion, deviates from traditional approaches. It merges the electrical domain with environmental factors, minimizing the sole reliance on direct sunlight measurements.

For a more granular understanding of model performance, error metrics were collated. A heatmap representation in the top-left corner of subfigures in Fig. 10 elucidates these metrics, offering a visual comparison of the models. In both cases, the proposed Circuit-LSTM method consistently outperforms the LSTM, CNN-LSTM, RNN-LSTM, and ANN methods, with significant improvements of over 50% in MAE, MSE, and RMSE, indicating its superior predictive accuracy as given in Table III.

For managing a real-time bPV power prediction system, the web interface of the proposed DTS, is displayed in Fig. 11. In the left and right top corners, there presents a live camera feed and

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on May 30,2024 at 09:38:08 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

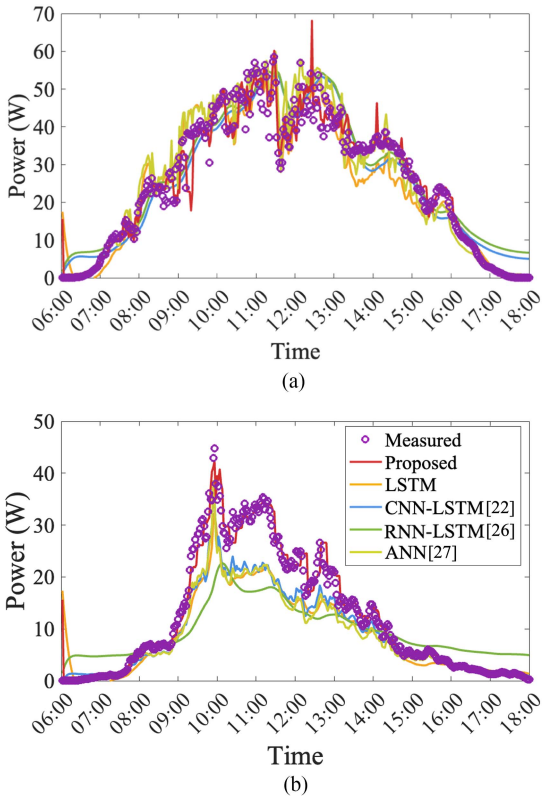


Fig. 10. Two-day power prediction analysis: (a) sunny day and (b) cloudy day.

TABLE III
PERFORMANCE METRICS ON SUNNY AND CLOUDY DAYS FOR DIFFERENT MODELS

		Proposed	LSTM	CNN -LSTM [22]	LSTM -RNN [26]	ANN [27]
Sunny Day	MAE (W)	2.147	3.996	4.544	4.545	3.365
	RMSE (W)	3.692	5.074	5.636	5.729	4.666
	MSE (W)	13.628	25.746	31.770	32.819	21.772
	R^2 (W)	0.953	0.911	0.890	0.887	0.925
Cloudy Day	MAE (W)	1.059	3.816	3.135	5.580	3.526
	RMSE (W)	1.965	5.551	4.709	7.426	5.185
	MSE (W)	3.863	30.811	22.175	55.143	26.886
	R^2 (W)	0.971	0.768	0.833	0.585	0.798

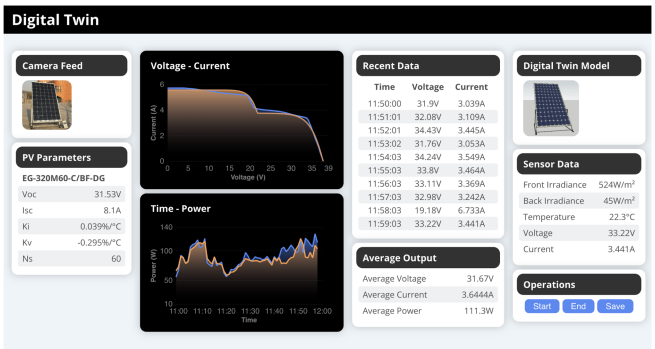


Fig. 11. User interface of the DTS.

the virtual bPV model. On the left-hand segment of the interface, users can access a comprehensive suite of information about the PV system's technical parameters, presented in an intuitive and accessible manner. The main highlighted sections include real-time representations of the I–V relationship and a time-based power graph that encapsulates the system's performance. Moreover, the real-time sensor data and the recent average electrical output generated by the PV system are systematically displayed, allowing users to promptly review and analyze the system's efficiency.

Once initiated, the interface starts computations for the time-based power prediction curve and updates the current data in recent data matrixes, reflecting the efficient data fusion at work. After the program concludes, users can extract the simulated PV data, downloadable as a CSV datasheet, facilitating further analysis using myriad data tools.

V. CONCLUSION

This study introduces a digital twin framework for real-time bPV systems performance prediction under various shading conditions. Through efficient data fusion, it offers real-time environment perception with the proposed T-SPE method, which negates direct irradiance measurements and enhances prediction accuracy. The proposed Circuit-LSTM model is unique with the ability to predict without historical data during the accumulation phase sets it apart from counterparts, such as LSTM, CNN-LSTM, RNN-LSTM, and ANN, which lack this capability. In the iteration phase, our proposed Circuit-LSTM method consistently outperforms the CNN-LSTM and RNN-LSTM methods, with notable improvements of over 50% in MAE, MSE, and RMSE, accompanied by an impressive R^2 value exceeding 0.97 and 0.95 of the whole day prediction results on February 5–6. The integrated digital twin framework, combined with the fusion of the T-SPE method and Circuit-LSTM, offers an enhancement in solar power prediction, setting a potential new benchmark for future research.

REFERENCES

[1] R. B. Bollipo, S. Mikkili, and P. K. Bonthagorla, "Hybrid, optimal, intelligent and classical PV MPPT techniques: A review," *CSEE J. Power Energy Syst.*, vol. 7, no. 1, pp. 9–33, 2021.
[2] R. Kopecek and J. Libal, "Bifacial photovoltaics 2021: Status, opportunities and challenges," *Energies*, vol. 14, no. 8, 2021, Art. no. 2076.
[3] M. Fischer, *International Technology Roadmap for Photovoltaics (ITRPV)* 14th ed.. Germany: PV Production Dept., VDMA German Eng. Feder., Mar. 2023.
[4] E. M. Ahmed et al., "An accurate model for bifacial photovoltaic panels," *Sustainability*, vol. 15, no. 1, Dec. 2022, Art. no. 509.
[5] A. Kumar and R. Agarwal, "Mathematical modeling and analysis of single-diode, double-diode and triple diode based pv module," in *Proc. Int. Conf. Advance. Technol.*, 2023, pp. 1–6.
[6] J. Ma et al., "Analytical modeling and parameter estimation of photovoltaic strings under partial shading conditions," *Sol. Energy Mater. Sol. Cells*, vol. 235, 2022, Art. no. 111494.
[7] C. Rodríguez Gallegos et al., "Monofacial vs bifacial SI-based PV modules: Which one is more cost-effective?," *Sol. Energy*, vol. 176, pp. 412–438, 2018.

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on May 30,2024 at 09:38:08 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[8] Y. Zhang, Q. Gao, Y. Yu, and Z. Liu, "Comparison of double-side and equivalent single-side illumination methods for measuring the I-V characteristics of bifacial photovoltaic devices," *IEEE J. Photovolt.*, vol. 8, no. 2, pp. 397–403, Mar. 2018.

[9] B. G. Bhang et al., "Power performance of bifacial c-Si PV modules with different shading ratios," *IEEE J. Photovolt.*, vol. 9, no. 5, pp. 1413–1420, Sep. 2019.

[10] S. Bouchakour et al., "Modelling and simulation of bifacial PV production using monofacial electrical models," *Energies*, vol. 14, no. 14, 2021, Art. no. 4224.

[11] M. S. Mahmud et al., "Solar highway in Bangladesh using bifacial PV," in *Proc. IEEE Int. Conf. System Computation Autom. Netw.*, 2018, pp. 1–7.

[12] A. Kadyri, K. Kandoussi, and O. Souhar, "An approach on mathematical modeling of PV module with sensitivity analysis: A case study," *J. Comput. Electron.*, vol. 21, no. 6, pp. 1365–1372, Dec. 2022.

[13] Y. Zhang and L. Kong, "Photovoltaic power prediction based on hybrid modeling of neural network and stochastic differential equation," *ISA Trans.*, vol. 128, pp. 181–206, 2022.

[14] D. Riley et al., "A performance model for bifacial PV modules," in *Proc. IEEE 44th Photovoltaic Specialist Conf.*, 2017, pp. 3348–3353.

[15] C. Hajjaj et al., "Comparing photovoltaic power prediction: Ground-based measurements vs. satellite data using an ANN model," *IEEE J. Photovolt.*, vol. 13, no. 6, pp. 998–1006, Nov. 2023.

[16] S. Souabi, A. Chakir, and M. Tabaa, "Data-driven prediction models of photovoltaic energy for smart grid applications," *Energy Rep.*, vol. 9, pp. 90–105, 2023.

[17] M. Madhiarasan, D. T. Cotfas, and P. A. Cotfas, "Medium-term forecasting of global horizontal solar radiation in Brasov using multivariate transformer," in *Proc. Int. Symp. Electron. Telecommun.*, 2022, pp. 1–4.

[18] M. Gao, J. Li, F. Hong, and D. Long, "Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using lstm," *Energy*, vol. 187, 2019, Art. no. 115838.

[19] K. Wang, X. Qi, and H. Liu, "A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network," *Appl. Energy*, vol. 251, 2019, Art. no. 113315.

[20] F. Wang et al., "A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework," *Energy Convers. Manage.*, vol. 212, 2020, Art. no. 112766.

[21] H. Chen and X. Chang, "Photovoltaic power prediction of LSTM model based on Pearson feature selection," *Energy Reports*, vol. 7, pp. 1047–1054, 2021.

[22] X. Qi, Q. Chen, and J. Zhang, "Short-term prediction of PV power based on fusions of power series and ramp series," *Electric Power Syst. Res.*, vol. 222, 2023, Art. no. 109499.

[23] J. Teo, R. H. Tan, V. Mok, V. K. Ramachandaramurthy, and C. Tan, "Impact of bypass diode forward voltage on maximum power of a photovoltaic system under partial shading conditions," *Energy*, vol. 191, 2020, Art. no. 116491.

[24] M. Visvalingam and J. D. Whyatt, "The douglapeucker algorithm for line simplification: Revaluation through visualization," *Comput. Graph. Forum*, vol. 9, no. 3, pp. 213–225, 1990.

[25] B. Marion et al., "A practical irradiance model for bifacial PV modules," in *Proc. IEEE 44th Photovoltaic Specialist Conf.*, 2017, pp. 1537–1542.

[26] M. N. Akhter et al., "An hour-ahead PV power forecasting method based on an RNN-LSTM model for three different PV plants," *Energies*, vol. 15, no. 6, 2022, Art. no. 2243.

[27] A. Geetha et al., "Prediction of hourly solar radiation in Tamil Nadu using ANN model with different learning algorithms," *Energy Rep.*, vol. 8, pp. 664–671, 2022.

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on May 30,2024 at 09:38:08 UTC from IEEE Xplore. Restrictions apply.