Digital Twin Integration With Data Fusion for Enhanced Photovoltaic System Management: A Systematic Literature Review

JIANG YUAN[1], JIEMING MA [1] (Senior Member, IEEE), ZHONGBEI TIAN[2], AND KA LOK MAN [1]

[1]Department of Computing, Xi'an Jiaotong Liverpool University, Suzhou 215123, China
[2]Department of Electrical Engineering and Electronics, University of Liverpool, L69 3BX Liverpool, U.K.

CORRESPONDING AUTHOR: JIEMING MA (e-mail: Jieming.Ma@xjtlu.edu.cn)

ABSTRACT The integration of Digital Twin (DT) technology into the photovoltaic (PV) sector represents a significant advancement in energy management, optimization, servicing, and maintenance. This comprehensive literature review aims to enhance understanding, categorization, and adoption of DT and data fusion technologies within the PV industry to guide future research endeavors. The review categorizes PV models into three types: digital models, digital shadows, and digital twins, based on their data connection and integration attributes. It recognizes data fusion as the critical enabling technology for the development of complex DT models and proposes a framework for integrating data fusion with DT systems. A detailed examination of prevalent PV modeling methodologies is conducted to delineate their advantages and limitations, serving as a valuable resource for industry practitioners. The paper concludes that digital models and digital shadows are effective for initial PV system forecast and monitoring, while fully integrated DT models offer significant advantages, including real-time analysis, predictive capabilities, and active system optimization. However, implementing and maintaining DT models require advanced data analytics, high computational costs, and robust system security, presenting important challenges to be addressed in future research endeavors.

INDEX TERMS Control systems, modeling, photovoltaic power systems, solar energy.

I. INTRODUCTION

Over the past several decades, the challenges posed by global warming and the energy crisis have spurred the advancement and adoption of alternative, sustainable, and eco-friendly energy sources [1]. Solar energy, an inexhaustible resource, is widely regarded as one of the most promising renewable for power generation[2]. Photovoltaic (PV) cells represent the principal technology for the conversion of solar energy into electrical power [3]. The implementation of PV power generation has fostered considerable economic and environmental benefits, including the reduction of CO_2 emissions and the generation of employment opportunities, thereby heightening public awareness and social engagement on these matters [4].

A. THE PV SYSTEM CHALLENGES

PV cells harness energy from sunlight by utilizing photons to displace electrons within silicon semiconductors, thereby generating an electrical current [5]. Consequently, the power output of PV cells is intrinsically linked to solar irradiance. Moreover, various environmental factors, such as temperature, cloud cover, particulate matter, relative humidity, and others, impact the efficiency of energy generation in PV cells [6], [7]. The meteorological variables that influence the performance of PV cells are inherently unpredictable and subject to constant fluctuations [8]. Additionally, certain PV systems exhibit non-stationary characteristics, wherein PV panels undergo continual movement [9]. This dynamic behavior results in fluctuations and unpredictability
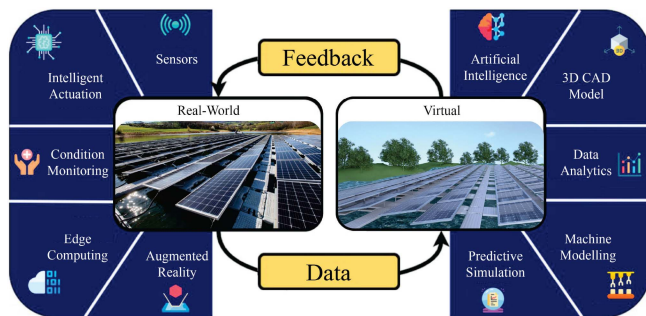
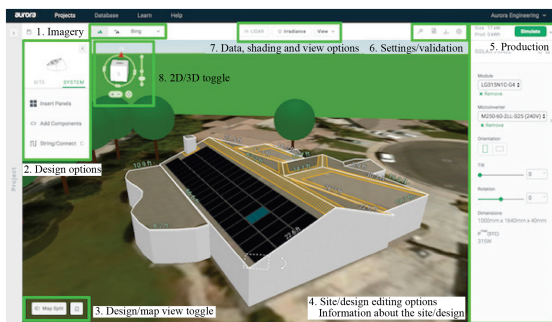**FIGURE 1.** Overview of the concept of DT system.

FIGURE 2. DT system interface [16].

in the irradiance levels captured by the PV panels. The intricate and dynamic nature of weather systems, coupled with the uncertainties surrounding conditions, renders the precise control and maintenance of PV modules a complex endeavor [10].

One potential solution to effectively manage these intricate and volatile PV systems is the implementation of Digital Twin (DT) technology [11]. The concept of DT technology involves the creation of an advanced data-driven virtual model that mirrors a physical entity, and its purpose is to refine, optimize, and sustain real-world operation [12]. Fig. 1 shows an overview of the DT system of a large scale floating PV installation and displays relevant technologies that could be implemented at real-world and virtual sides.

A DT system commonly incorporates an interface that facilitates various functionalities. This interface serves as a means to visually represent the DT model, providing users with the ability to observe its constituent elements and manipulate relevant parameters [13]. Moreover, it encompasses real-time data monitoring capabilities, analytics tools, and simulation functionalities, thereby empowering users to analyze and optimize the behavior of the DT model [14]. In addition, the interface offers alerts and notifications based on predefined conditions, supports collaborative efforts among users, and seamlessly integrates with other software systems. Through the DT interface, users gain the ability to diligently monitor, exercise control, and make well-informed decisions pertaining to the physical object or system that the DT model represents [15]. To illustrate, Fig. 2 showcases a prime example of a DT interface showcasing a PV system, which was

developed using the specialized solar PV design software tool known as Aurora [16].

The application of DT technology can address the challenges posed by PV systems, offering several advantages over conventional simulation methods:

- A DT model is inherently dynamic, capable of adapting to evolving environmental conditions in real-time [17].
- Access to extensive data sets enables DT models to leverage machine learning algorithms for improved performance and results [18].
- By integrating data from multiple sources, a DT model generates a more comprehensive and robust representation [19].

By incorporating these advantages, DT technology presents a promising avenue for overcoming the complexities associated with PV system management.

B. BRIEF HISTORY OF DT TECHNOLOGY
Although the term "Digital Twin" was recently coined in 2011, the underlying concept can be traced back to the 1960s [20]. One notable example is NASA's ground-based engineering team, who successfully resolved Apollo 13's oxygen tank explosion problem in 1970 by testing potential solutions in a virtual simulation [21]. In 2003, Michael Grieves was credited with first publicly introducing the concept and model of the DT during a specialized meeting on product life-cycle management at the University of Michigan's Lurie Engineering Center [22]. In a subsequent publication, Grieves delineated the three primary components of a DT:

1) A virtual twin that emulates the behavior of the corresponding physical counterpart, generating identical outputs in response to the same input values.
2) The physical twin, which represents the real-world entity that the virtual twin seeks to emulate, may encompass a product, system, model, or other physical entity.
3) A data flow cycle that facilitates the exchange of information between the physical and virtual twins, enabling each to inform and influence the other.

NASA employed the term "Digital Twin" in 2011 to describe the digital replication of an aircraft's structural behavior [23]. Initially, DT models were utilized as maintenance tools for continuous structural monitoring. Subsequently, they evolved into comprehensive replicas that could simulate the entire life cycle of an aircraft and predict its performance [24]. Over time, DT technology has played an important part in Industry 4.0 and gained widespread adoption across various sectors, including construction, education, business, transportation, power and electronics, human and healthcare, sports, and networking and communications [25]. The technology has garnered increased popularity among industries striving to make their processes more intelligent, adaptable, and optimally responsive to operational conditions. DT systems are highly sought after for their capacity to identify product defects [26], reduce production costs [27], enable real-time monitoring [28], and extend product lifespans through the prediction of product failure [29]. Fig. 3
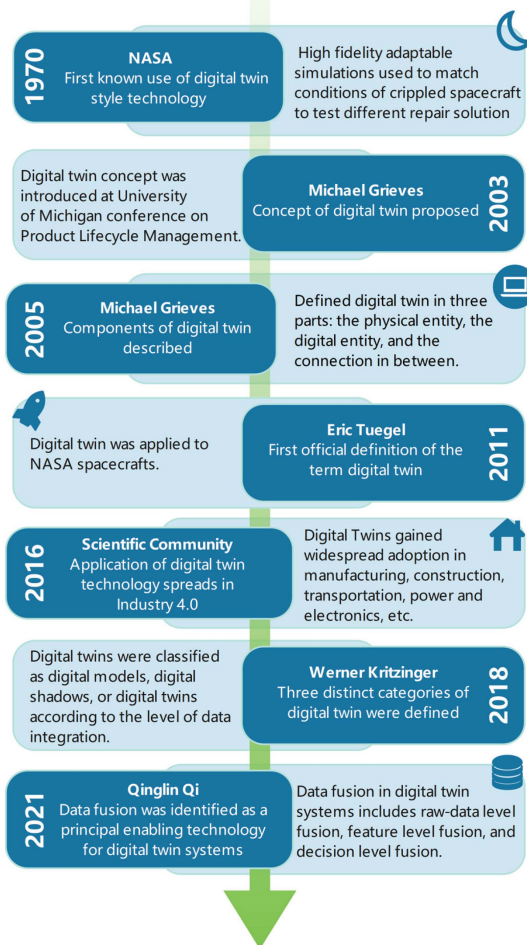
Digital twin history

1970 NASA First known use of digital twin style technology

High fidelity adaptable simulations used to match conditions of crippled spacecraft to test different repair solution

Digital twin concept was introduced at University of Michigan conference on Product Lifecycle Management.

Michael Grieves Concept of digital twin proposed **2003**

2005 Michael Grieves Components of digital twin described

Defined digital twin in three parts: the physical entity, the digital entity, and the connection in between.

Digital twin was applied to NASA spacecrafts.

Eric Tuegel First official definition of the term digital twin **2011**

2016 Scientific Community Application of digital twin technology spreads in Industry 4.0

Digital Twins gained widespread adoption in manufacturing, construction, transportation, power and electronics, etc.

Digital twins were classified as digital models, digital shadows, or digital twins according to the level of data integration.

Werner Kritzinger Three distinct categories of digital twin were defined **2018**

2021 Qinglin Qi Data fusion was identified as a principal enabling technology for digital twin systems

Data fusion in digital twin systems includes raw-data level fusion, feature level fusion, and decision level fusion.

FIGURE 3. A brief history of developments in DT technology.

shows a timeline that catalogs notable developments in DT technology.

DT technology also revolutionized the power electronics field by enabling the creation of highly accurate virtual models of physical systems and components. For example, ABB Ltd. used DT models to monitor and optimize the performance of their electromagnetic flowmeters, leading to enhanced efficiency and reduced energy losses [30]. Engineers at General Electric leveraged digital twins to predict and mitigate potential issues in their power converters before they occurred, optimizing system designs and conducting virtual testing under various operating conditions [31]. This technology also facilitated predictive maintenance, as seen in Siemens' implementation, which reduced downtime and extended the lifespan of critical components in their power management systems [32].

C. RESEARCH AND REVIEW GAP

In recent years, academic literature concerning DT technology has seen a marked increase, though research activities in various fields are not uniformly distributed. The most prevalent areas of focus include manufacturing DT, with over

1,000 papers published since 2010, and architectural DT, with over 400 papers published since 2010. In contrast, there is a relative scarcity of published research on solar energy DT. Consequently, the development of solar energy DT systems represents an emerging domain in the realm of digitization that warrants increased attention, given the pressing need for clean, renewable energy solutions in response to global environmental challenges. Considering the emerging status of solar DT systems, it is essential for researchers to coordinate their efforts to produce meaningful research outputs that directly impact the solar industry and contribute to the urgent demand for clean energy. Therefore, a systematic review of solar DT technology is both timely and necessary to guide future research and development in this critical field.

Numerous review papers have explored the concept of DT modeling. Liu et al. provided a comprehensive analysis of industrial applications, comparing DT models to digital models and digital shadows, and emphasizing the significance of data fusion in DT modeling [33]. However, their focus was on industrial rather than energy applications. The criteria for effective DT modeling in industrial manufacturing differ significantly from those in the energy sector, where predictive analysis is crucial due to the variability in energy availability and consumption [34]. In their review of energy DT models, Amaral et al. categorized these twins by their applications in energy generation, storage, transmission, and consumption, but minimal references to solar energy DT models are distinguished by their unique and complex energy generation methods. Researchers aiming to design solar energy DT systems may find this review less useful due to the lack of specific insights into solar applications. Ghenai et al.'s review discusses some modeling methods for solar energy DT models but offers a limited analysis, primarily highlighting the benefits without addressing the limitations or comparing different methods to understand their underlying mechanisms [35]. Lastly, Kavousi-Fard et al.'s review analyzes DT technology for solar energy, highlighting its applications in optimization, maintenance, security, and resiliency, but fails to compare different DT models within each category [36]. It is crucial to note that DT models utilizing high volumes of diverse data can be structurally different and perform differently from those that do not, highlighting the need for comparative analysis within application categories.

D. REVIEW AIMS AND CONTRIBUTIONS

This review aims to categorize and explore the design and application of DT systems to enhance the efficiency of PV solar energy generation. To accomplish this objective, the paper presents an exhaustive literature review centered around addressing the following research questions:

- How do academic publications define DT technology and classify its various types?
- How do DT systems leverage the data fusion techniques to benefit solar energy generation?
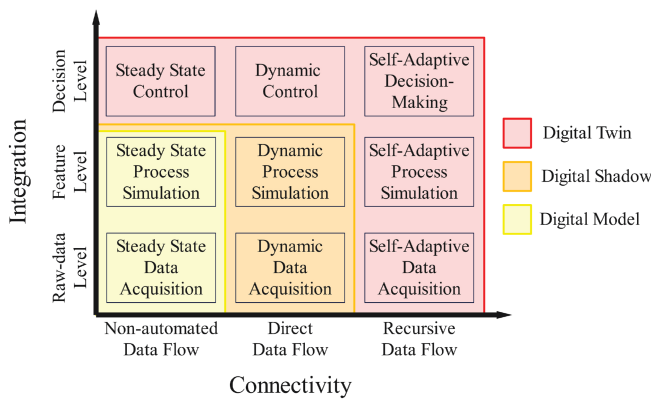
FIGURE 4. Classifying DT models based on data connectivity and data integration.

- What are the advantages and limitations of various PV models?

This review offers the following original contributions to the existing body of literature:
- Identification of the implementation methods for data fusion in solar DT systems.
- Recognition of current limitations in the utilization of DT technology for PV management.

The subsequent sections of this paper follow a structured layout. In Section II, a novel classification framework for DT systems is introduced, utilizing data connectivity and data integration concepts drawn from existing literature. Section III conducts an in-depth analysis of prevalent modeling methods applied in PV systems across various DT categories, carefully evaluating their respective advantages and limitations. Section IV highlights critical challenges encountered by DT systems within the context of the PV industry, while also proposing prospective research directions aimed at surmounting these challenges. Finally, Section V provides a conclusive summary to the paper.

II. DT DEFINITION AND CLASSIFICATION

Numerous academic and industrial publications have sought to define "Digital Twin", resulting in a multitude of ambiguous definitions that provide limited clarity without further elaboration. Examples of these definitions include terms such as "digital replica [37]," "digital counterpart [38]," "virtual counterpart [39]," and "virtual representation [40]." Without specifying the meaning of replica, counterpart, or representation within these definitions, they offer minimal explanatory value. This section defines and classifies DT systems by examining their data connectivity and data integration attributes. Fig. 4 shows the classification framework that categorizes different types of DT systems as Digital Model (DM), Digital Shadow (DS), and Digital Twin [41].

A. DATA CONNECTIVITY ATTRIBUTE

PV models exhibit diverse levels of data connection between their physical and virtual parts. Simple digital simulations are
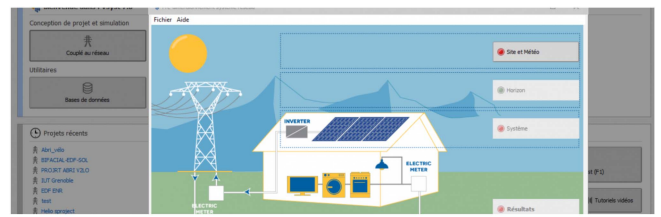
FIGURE 5. Rooftop PV installation model using PVsyst [44].

created without linkage to a physical entity, whereas more complex ones possess total fusion with instantaneous data transfer [42]. The types of data connections present within a DT model can be classified into three distinct categories: non-automated, direct, and recursive data flow.

Non-automated data flow excludes any sort of automated data interchange between physical and virtual components. Such digital representations may encompass comprehensive descriptions of the physical object, including simulation models of proposed manufacturing facilities, mathematical models of innovative products, or other representations of physical objects that do not rely on automatic data integration [41]. For example, software programs are utilized to create 3D models of rooftop PV installation, incorporating the layout of panel, tilt angle, and other physical attributes [43]. An example of rooftop PV installation model is shown in Fig. 5. However, this model does not reflect real-time changes in the environment, such as the rising and falling of the sun or blocking of clouds. Though digital information from real-world systems might be utilized in developing these models, all data transfer is conducted manually. Consequently, changes in the condition of the physical object do not immediately affect its virtual model. Non-automated data flow is used in DM for conducting steady-state analysis of PV systems and is not responsive to real-time changes in environmental data.

Direct data flow is distinguished by the existence of a one-way data flow between an extant physical object and its digital counterpart [45]. Alterations in the state of the physical entity lead to corresponding changes in the digital entity; however, the inverse does not transpire. To illustrate, consider a PV unit within a solar farm. Sensors are strategically placed throughout the unit to collect data such as panel temperature, incident irradiance, or produced power. These data are utilized to create a DS model, which provides a real-time representation of the current status of the PV unit [46]. However, a DS model lacks the ability to respond to changes in the environment. A DS model demonstrates the capability to react in real-time to environmental fluctuations, enabling dynamic analysis of PV systems [47]. While DS models that incorporate direct data flow offer considerable value for dynamic simulation and data acquisition, they do not possess the ability to deliver decision-level control over PV systems.

Recursive data flow is characterized by the complete bidirectional integration of data exchange between an existing physical and digital object [48]. In this arrangement, the digital entity serves as a controlling instance for the physical

entity. A change in the condition of the physical object immediately leads to a transformation in the condition of the virtual model, and vice versa. Suppose an electrical company designs a DT model for a grid connected PV system. This model collects real-time data from sensors embedded in the solar panel, such as voltage, current, and temperature levels [49]. If the DT model predicts a short circuit failure based on these data, it can suggest preventative maintenance. Also, engineers can simulate different scenarios in the DT model to see how the physical PV panel might react, helping them to improve the panel's design or operational protocols. The DT model with recursive data flow is proficient at adapting to environmental conditions by adjusting specific physical attributes of the PV system [11], such as the orientation of sensors, tilt angle of solar panels, and input voltage of solar cells, etc. By recursively optimizing the best solution to maximize power output, a DT model can provide decision-level control for the PV system.

B. DATA INTEGRATION ATTRIBUTE

The pivotal technology driving a DT system is data fusion, enabling the seamless transition of information from raw numeric input to a comprehensive understanding and actionable information [50]. Data fusion in DT systems can be implemented at three different levels: raw-data level, feature level, and decision level fusion [51].

Raw-data level data fusion is the process of combining unsorted data from multiple sources, typically sensors, to create a more comprehensive and accurate representation of a given environment or situation [52]. For example, consider a photovoltaic simulation system that synthesizes historical power data, meteorological information, and statistical analyses into a unified model [53]. This integration effectively addresses the challenges associated with the limited availability and incompleteness of historical photovoltaic output power and meteorological data. This fusion technique proves particularly advantageous in scenarios where visibility is compromised under low light conditions or bad weather conditions [54]. The primary objective of this fusion process is to enhance the overall quality and dependability of the data gathered by reducing ambiguity, redundancy, and interference [55]. Through the amalgamation of information at the raw-data level, the ensuing combined data offers a fuller and more precise portrayal of the physical object. Consequently, the fusion process enables enhanced situational awareness, improved decision-making, and a deeper understanding of the observed phenomena. By employing raw-data level data fusion, DT systems can harness the collective power of multiple sources to generate more reliable and informative representations.

In feature-level data fusion, the data from different sources are first preprocessed and transformed into a set of features, which are characteristic attributes or patterns that can be used for further analysis [56]. This fusion technique plays a crucial role in generating a unified feature set that encompasses the relevant information from diverse sources. For instance, consider a PV simulation that incorporates both light model and thermal model [57]. Each of these models extracts distinct features, such as irradiance from the light model and temperature from the thermal model. By combining the extracted features, a more comprehensive and representative simulation of PV working conditions can be created, exemplifying feature-level fusion [58]. The integration of these features has the potential to enhance the accuracy, reliability, and robustness of the PV simulation. The amalgamation of features enables a more holistic representation of the underlying data, allowing for a deeper understanding and analysis of the observed phenomena. The integration of an extensive set of features derived from various sources not only contributes to creating a simulation that mirrors a broader range of aspects found in reality, but also facilitates more accurate modeling and simulation outcomes.

Decision-level data fusion combines the output from multiple sources, models, or classifiers to arrive at a final decision or result [59]. In decision-level data fusion, each source, model, or classifier first processes the input data independently and produces its own decision or output. Then, these individual decisions are combined or fused to reach a final, more reliable decision [60]. Decision-level fusion is exemplified by an automated system designed for PV fault diagnosis, which utilizes results from multiple distinct classifiers [61]. Each of these classifier undergo separate processing, leading to preliminary diagnoses based on each one. By combining these preliminary diagnoses through decision-level fusion, a final diagnosis can be determined. This fusion technique boosts the precision and dependability of the diagnostic procedure by utilizing the advantages of various information sources [62]. The main goal of decision-level data fusion is to elevate the collective accuracy, reliability, and resilience of the decision-making mechanism, harnessing the synergistic strengths of different models. By integrating outputs from various sources, decision-level reduces the impact of uncertainties and errors associated with individual classifiers [63]. By employing decision-level data fusion, DT systems can harness the collective output of multiple models, leading to improved decision-making processes.

III. COMPARISON OF MODELS, SHADOWS, AND TWINS

This section provides an in-depth analysis of contemporary modeling methodologies relevant to the application and implementation of PV system models. PV models are divided into three categories based on data connectivity: DM, DS, and DT. These categories are arranged hierarchically, with the DS comprising multiple instances of DM, and the DT incorporating multiple instances of DS.

This paper concentrates on the most advanced level of modeling, the DT, exploring its role in enhancing our understanding of PV systems. An examination of DT models in existing literature highlights data fusion as a key element in their development. The integration of DT technology and data fusion techniques significantly enhances the monitoring, analysis, and optimization of the performance of physical assets, systems, or processes. Fig. 6 shows a framework of
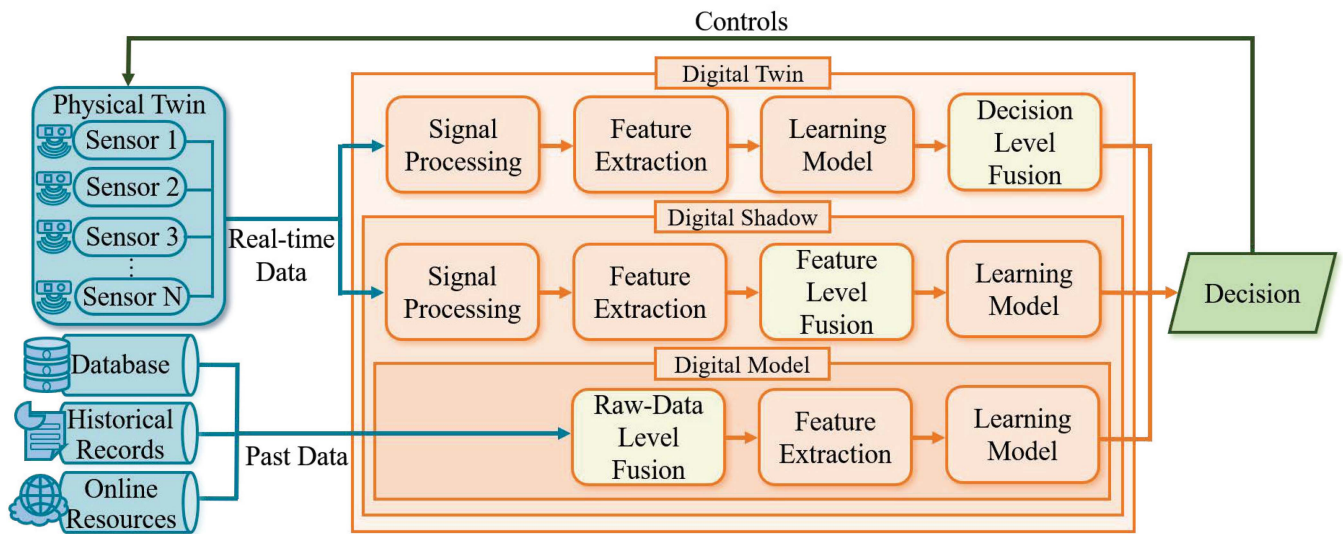
FIGURE 6. Application of data fusion in DT framework.

how different types of data fusion can be implemented in the DT structure.

Differentiation among DT models is based on the type of data fusion employed, categorized into three subdivisions: data acquisition DT, data-driven simulation DT, and data-driven control DT. These correspond to raw-data level data fusion, feature level data fusion, and decision level data fusion, respectively. Furthermore, the full data cycle DT integrates these various levels of data fusion, encapsulating the entire process from raw sensor inputs to system control decisions. This hierarchical and integrated classification approach based on data connectivity and integration attributes enables a more structured and comprehensive analysis and control of solar energy systems. This paper seeks to provide insights into the current applications of DT and data fusion technologies in PV systems, identifying both challenges and opportunities associated with these advancements.

A. DIGITAL MODEL

The DM encapsulates a discrete component of a PV system, isolating it from temporal fluctuations. These components may include the physical architecture, interactions with light, and the thermal and electrical properties of the PV system [64]. The primary attribute of the DM is its static nature, characterized by a set of equations that represent the physical properties of the object it models. The primary aim of DM models is to perform simple PV performance forecasts, which represent 83% of the reviewed DM models. These models focus on how a single environmental change over time affects system performance. Table 1 presents an assortment of DMs used for PV simulation which have been documented extensively in academic research.

DM has proven beneficial for estimating energy gain from FPV (Floating Photovoltaic) systems. The cooling effects of water within these FPV systems were used to project power output prediction contrasting ground-based and FPV

systems [65]. The parameters of the model include irradiance-weighted average temperatures, heat loss coefficients, and wind-dependent heat loss coefficients. This weighted temperature takes into account the higher impact of hours with increased irradiance, due to the significantly higher energy yields that PV panels produce under such conditions. Therefore, temperature data under high irradiance settings are of greater significance. The yearly outputs of the PV systems were inferred from the assessed temperature disparity weighted by irradiance and PVsyst model. The model is used to improve energy yield from the cooling effect of FPV systems compared to the reference PV systems by up to 6%.

DMs help in simulating various configurations and environmental conditions to determine the most efficient setup for solar panels. This can significantly enhance energy output and reduce wastage by predicting how different angles, spacing, and types of solar cells will perform under varying conditions [43]. However, the limitation of the DM lies in their static nature; they aren't typically updated with real-time data. As a result, they might not account for ongoing changes in environmental conditions or degradation of solar panel efficiency over time, which can lead to discrepancies between predicted and actual system performance [66].

B. DIGITAL SHADOW

The DS demonstrates direct data connectivity and provides a real-time representation of the system through the aggregation of multiple DMs. The defining feature of the DS is its ability to reflect the real-time status of a physical object by automatically gathering and analyzing data. DS models employ sensors to establish a direct link between the physical and digital entities, facilitating the transmission of real-time parameters. DS models primarily aim for real-time monitoring, representing 54% of the reviewed DS models. They leverage their direct data connection to facilitate autonomous monitoring and integration with other IoT technologies. Table 1 exhibits a variety

TABLE 1. Overview of PV Digital Models and Digital Shadows

Ref.	Model Type	Data Connection	Aim	Advantages	Limitations
[66]	Conditional probability-based DM	Non-automated	PV performance forecast	Handle the variability and stochastic nature of PV generation and consumption	Complexity of integrating and calibrating the model for different geographic regions and building stocks
[67]	Bayesian parameterization-based DM	Non-automated	PV performance forecast	Addresses uncertainties due to the lack of detailed information about individual PV installations	Sensitive to the accuracy of the initial guess and the estimation of covariance matrices
[68]	Bayesian weighted sum-based DM	Non-automated	PV performance forecast	Mitigate issues such as sunny bias and ensemble underdispersion typically found in raw forecasts	Requires careful selection of training data windows and might involve complex parameter fitting processes
[65]	Thermal-based DM	Non-automated	PV performance forecast	Provides a robust framework for assessing the thermal behavior and energy yield of FPV systems in different climates	Challenge of limited field data, especially across diverse climates, which could impact the generalizability of the findings
[69]	MLP-based DM	Non-automated	PV performance forecast	Addresses the limitations of traditional analytical models by incorporating the adaptability of ANNs to varying data	Lacks interpretability and physical insight into the model's decision-making process
[70]	LSTM-based DM	Non-automated.	PV performance forecast	Captures temporal dependencies and handles the variability in PV power output across different PV plants	High complexity and computational demands
[43]	Lifecycle analysis-based DM	Non-automated	Optimize PV design	Evaluates PV system installations for a wide range of rooftop shapes, orientations, and regional climatic conditions	Relies on specific geographic and climatic data from South Korea, which may not be applicable to other regions or roof types
[71]	Equivalent circuit-based DS	Direct	PV system control	Enables optimal operation through accurate real-time calculations	Reliance on accurate model parameter updates in real time
[72]	Equivalent circuit-based DS	Direct	PV condition monitoring	Allows for the observation of impedance changes and parameter extraction under different conditions	Long procedural time required for data collection and analysis
[73]	Equivalent circuit-based DS	Direct	PV condition monitoring	Enables the real-time monitoring and predictive maintenance of PV systems by utilizing an IoT framework	Necessitates extensive computational iteration cycles to achieve effectiveness, which compromises its real-time capabilities
[74]	Kalman filter-based DS	Direct	PV performance forecast	Enables real-time predictions over various time horizons and is particularly effective in regions with highly variable meteorological conditions	Complex parameter tuning and the assumption of stationary process
[75]	Kalman filter-based DS	Direct	PV performance forecast	Allows for dynamic correction of the prediction model by using measured power to limit prediction error variability	Require accurate future weather prediction data for effective forecasting
[76]	Virtual reality-based DS	Direct	Predict PV performance in new environment	Dynamically assesses the impact of solar irradiance on complex and arbitrarily shaped surfaces	Does not support complex reflections or transparent materials
[77]	Weather forecast-based DS	Direct	PV condition monitoring	Inclusion of real-time data utilization enhances the model's responsiveness to environmental changes	Complexity of the model, which requires detailed weather forecast data
[46]	ANN-based DS	Direct	PV condition monitoring	Integrates with IoT technologies for improved data logging, system updates, and remote management capabilities	Compromised by sensor accuracy issues, which can lead to a error margin of up to 10%

of DS models used for PV simulation and prediction, widely cited in academic literature.

The DS framework facilitates real-time PV monitoring through dynamic alternating current equivalent electric circuit (AC-EEC) modeling of PV modules using impedance spectroscopy [72]. The parameters of the model include series resistance, junction resistance, capacitance, and minority carrier lifetime. This technique allows for detailed characterization of the cells under various conditions such as illumination, shading, and faults, by observing changes in impedance. A key feature is the ability to extract and monitor various parameters such as resistance and capacitance changes that reflect different operational and fault states. This method provides significant insights into internal processes of the photovoltaic cells, which are crucial for improving design, diagnostic processes, and ensuring efficient performance. Moreover, it can be adapted for real-time condition monitoring, offering a non-destructive, insightful tool for ongoing assessment of PV systems.

For PV systems, a DS model offers the advantage of real-time monitoring and performance analysis. By maintaining an ongoing data stream from the physical PV system, a DS model can track energy production, identify panels that are underperforming, and monitor the health of the system [46]. This enables operators to make informed maintenance decisions and quickly address issues like shading or dirt accumulation that can reduce efficiency [73]. The limitation, however, is that a DS model typically does not interact with the system to initiate corrective measures. The DS serves as a passive observer, limiting its utility to monitoring and diagnostic functions without direct intervention capabilities.

C. DIGITAL TWIN

The DT model encompasses multiple instances of DS, each representing a potential future process outcome. The DT model evaluates each DS prediction to identify the most advantageous outcome, thereby optimizing the system towards an optimal state. The fundamental attribute of the

TABLE 2. Overview of PV Digital Twins

Ref.	Model Type	DT Classification	Aim	Advantages	Limitations
[78]	Particle filter-based DT	Data acquisition DT	PV system control	Reduction in dependency on external data simplifies the implementation and lowers the cost and maintenance	Requires optimization to reduce operational energy consumption
[53]	Markov chain-based DT	Data Acquisition DT	PV performance forecast	Merging predictions with past data enhances forecasting accuracy	Require more computational resources, heavily relies on the precision of weather forecasts
[49]	FPGA-based DT	Data Acquisition DT	Fault diagnosis	Integration with distributed power electronic converters eliminates the need for additional hardware, reduces complexity, and supports scalable, cost-effective implementations	High reliance on the precision of simulated attributes to match real field measurements
[79]	ANN-based DT	Data-driven simulation DT	PV performance forecast	Recursive feature elimination with cross-validation enhances model accuracy and interpretability by focusing on the most relevant predictors	Heavily relies on historical data and specific weather conditions. Such dependency may not fully capture unexpected or rare weather events
[80]	Game engine-based DT	Data-driven simulation DT	Predict PV performance in new environment	Models the environmental impact on PV arrays over time, offering a more detailed and realistic planning tool than traditional methods	Limited ability to provide direct access to light collected by a surface
[57]	Runge–Kutta method-based DT	Data-driven simulation DT	Predict PV performance in new environment	Adjusts the coefficients dynamically, providing a more precise simulation of the panel's behavior under different environmental conditions	Ignores multi-reflection effects and forced convection, potentially affecting its practical application in varying environmental conditions
[81]	Electrical behavior-based DT	Data-driven control DT	Fault diagnosis	Efficiently handles different parts of the network by applying the most suitable simulation technique based on the required accuracy	Lower accuracy compared to a full-blown electromagnetic transient program solution
[82]	Bayesian fusion-based DT	Data-driven control DT	PV system control	The Bayesian approach allows the system to intelligently decide between multiple tracking strategies	Computational complexity and dependency on the accuracy of input data for effective performance
[83]	ANN-based DT	Data-driven control DT	Fault diagnosis	Isolates various fault types using two different algorithms tailored to the complexity of the faults	Model's effectiveness depends heavily on the precision of simulated attributes
[84]	Game engine-based DT	Full data cycle DT	Predict PV performance in new environment	Allows for the precise estimation of solar irradiance over various time scales, facilitating strategic decision-making for the installation of photovoltaic systems	Exclusion of complex-shaped buildings and other man-made features
[61]	Ensemble-based DT	Full data cycle DT	Fault diagnosis	Integrates the strengths of individual classifiers, leading to higher classification accuracy under various conditions	Potential overfitting due to the use of stacking ensemble classifiers that might not generalize well to unseen data outside the specific conditions

DT is its capability for continuous optimization, either by manipulating real-world objects or by refining its modeling strategies. Within a DT framework, the virtual entity implements modifications that influence the state of the physical entity. Subsequently, changes in the physical entity are reflected by the virtual counterpart, creating a recursive data connection loop. The DT represents the pinnacle of sophisticated simulation, harnessing data from diverse sources to lead the trend towards more intelligent simulations. DT models are mainly used for fault diagnosis, representing 36% of the reviewed DT models, and PV system control, representing 32% of the reviewed DT models. These models require optimization through various potential actions. While DT models also predict PV performance, representing 29% of the reviewed DT models, they differ significantly from DM models by forecasting performance in entirely new environments rather than just temporal variations in the same environment. These new environments introduce a set of compounding environmental factors, making accurate PV performance predictions more complex. Table 2 introduces an array of DT models in the field of PV control, extensively cited in scholarly literature. To enhance comprehension of DT modeling for PV systems, this paper will categorize DT models based on their approach to data fusion, identified as the data integration attribute of the DT.

1) DATA ACQUISITION DIGITAL TWIN

Data acquisition DT models employ raw-data level fusion, integrating data from multiple sources such as sensors, external databases, historical records, and other online resources. In the context of PV systems, Complex conditions such as fluctuating irradiance, partial shading, and low lighting can render some data inaccurate or unavailable. Additionally, Sensors are susceptible to magnetic interference and degradation associated with aging, factors which may compromise the precision and reliability of data acquisition [85]. The data acquisition DT's key characteristic is its ability to process raw data from various sources, which may be unreliable, incomplete, or contradictory, and to optimize a set of features for dependable modeling.

Pirayawaraporn and colleagues proposed a two axis solar tracking system without sensors that embodies recursive data connectivity at the raw-data level, by merging model predictions with actual measurements [78].The parameters of the model include daily angle, elevation angle, number of particles, mean values of daily and elevation angles, variance values of daily and elevation angles, weights of particles, and normalized weights of particles. The model addresses the challenge posed by the absence of sensor information through the implementation of a robust sampling-based tracking algorithm, known as the particle filter, which is utilized to develop a novel solar tracking strategy. Initially, particles correlating with various orientation angles of the proposed tracking system are generated as inputs. Subsequently, PV power is captured for each particle, and a corresponding weight is calculated to denote each particle's relative importance. Each particle's alignment is successively calculated and revised, incorporating its measurement to determine the likely location of the sun. The tracking approach concludes once every particle aligns to a uniform direction, indicative of the attainment of the optimal PV angle. This method necessitates the physical PV module to recursively explore different potential solutions to arrive at an optimal tracking angle, which improved PV energy generation by 20.1% compared to a fixed flat-plate system.

2) DATA-DRIVEN SIMULATION DIGITAL TWIN

Data-driven simulation DT models utilize feature-level data fusion, where information from various sources is processed and transformed into a set of distinct features displaying unique attributes or patterns. These features are then merged to form a unified set, which is analyzed to understand relationships and correlations among different attributes such as physical structure, lighting, heat, and electrical characteristics of the PV system. The fundamental trait of the data-driven simulation DT is its use of preprocessed data to refine a simulation model or multiple component models, continuously aligning model outputs with actual world data and revising feature connections to better represent the real-world object. An example of data driven simulation DT is shown in Fig. 7.

Qadir et al. applied a recursive feature selection technique to remove the least effective features and choose the best features to estimate the power yield of a combined PV-wind power system [79]. The parameters of the model include solar irradiation, wind speed, ambient temperature, humidity, precipitation, atmospheric pressure, and wind direction. At first, meteorological information is gathered from instruments and refined to eliminate any inaccurate figures that might weaken the framework. Subsequently, feature selection is executed through iterative feature removal employing cross-validation techniques. The dataset undergoes training with artificial neural network predictors, and interconnections among various attributes in the collection are identified. In each iteration, the estimator uses all the features in the data to generate

FIGURE 7. Interface of a data-driven simulation digital twin model of a solar farm [86].

a set of scores at the iteration's conclusion. Each score is associated with a particular feature. For instance, consider a scenario where the goal is to identify the top five features which contribute most significantly to model accuracy from a total of 20 features. In this case, the algorithm begins by recursively eliminating features at each iteration, provided their corresponding scores are below the algorithm's threshold. The primary objective is to identify meaningful patterns among features to enhance performance having MSE (Mean Squared Error) of 0.000000104, MAE (Mean Absolute Error) of 0.00083, R^2 (Coefficient of Determination) of 99.6%, and computation time of 0.02s, significantly outperforming all other models.

3) DATA-DRIVEN CONTROL DIGITAL TWIN

Data-driven Control DT models utilize decision-level data fusion, amalgamating outputs from diverse sources, models, or classifiers to formulate a comprehensive control scheme. Each source or model independently processes input data, producing individual decisions that are then synthesized into a final, robust control strategy.The primary characteristic of a data-driven control digital twin is its ability to leverage outputs and predictions from multiple models to formulate an optimal control strategy that aims to achieve the best possible results in real-world applications. Within the context of PV systems, this DT model considers numerous control variables that affect different aspects of system performance, including adjustments to panel tilt angles, voltage control, and power stability. The Decision-making DT focuses on identifying the optimal control decision to maximize system performance.

Gugulothu et al. adopted bayesian fusion for Maximum Power Point Tracking (MPPT) in PV systems [82]. The parameters of the model include solar irradiance, temperature, voltage, and current of the PV modules. Their approach combined the conventional incremental conductance algorithm with the jaya optimization algorithm, resulting in individual

MPP estimates. These estimates were assigned prior probabilities based on factors such as historical performance and reliability. The likelihoods of observing the actual system output given the MPP estimates were calculated, and bayes' theorem was employed to update the posterior probabilities. The final MPP estimate was obtained by calculating a weighted average of the individual MPP estimates, where the weights were determined by their corresponding posterior probabilities. This methodology enhances the accuracy and robustness of the MPP estimation with tracking speed less than 0.1 s and achieving 99.8% tracking efficiency.

4) FULL DATA CYCLE DIGITAL TWIN

Full data cycle DT models integrates multiple levels of data fusion, managing data from raw sensor signals to the final control decision. Within this comprehensive framework, several simpler models are interconnected to create a data pipeline that streamlines and automates a series of computational processes. These processes typically include data preprocessing, feature extraction, and model inference. As an end-to-end model, the full data cycle DT model encompasses all core functionalities of the DT, such as simulation, monitoring, optimization, and decision-making. This holistic approach renders it highly effective for PV system modeling.

Radhakrishnan et al. proposed a full data cycle DT design for categorizing Power Quality Disturbances (PQDs) in PV enhanced electrical grids [61]. The parameters of the model include value of confidence factor assigned for pruning in the decision tree and the configuration of the ensemble classification model using 10-fold cross-validation. This framework was executed via Matlab-Simulink, and diverse types of PQDs were examined. During the initial processing phase, the discrete wavelet transformation method was applied for feature extraction from assorted PQDs. These isolated characteristics were then employed in the training of foundational classifiers, encompassing logistic regression, naïve bayes, and the J48 decision tree, at the base tier. The outcomes derived from these base classifiers were then utilized in the training of the meta-classifier at a subsequent stage, culminating in a final prediction. In noisy environments, the proposed model achieves up to 27.33% enhanced classification accuracy compared to base classifiers. The utilization of different levels of data fusion in this meta-classifier model is illustrated in Fig. 8.

DT models possess the ability to not only monitor and analyze but also predict and optimize system performance in real time. With a digital twin, operators can simulate the impact of different operational strategies, predict the effects of upcoming weather conditions on energy production, and automatically adjust system parameters to maximize output [79]. The digital twin can also preemptively suggest maintenance or troubleshooting steps before issues become significant [81]. The following points illustrate how DT system and data fusion work together to achieve better results:

1) Data Collection: DT systems depend on data procured from an array of sources, including sensors, databases, and other information systems. Data fusion is instrumental in amalgamating this heterogeneous data to generate a comprehensive and precise representation of the asset or system within the DT model [87].

2) Enhanced Precision: Data fusion methodologies aid in mitigating noise, errors, and discrepancies in the data gathered from disparate sources. This improvement in data quality bolsters the accuracy of DT models, resulting in superior predictions, decision-making, and optimization [88].

3) Real-time Analysis: DT and data fusion technology both facilitate real-time analysis and monitoring of assets or systems. Data fusion techniques guarantee the incorporation of pertinent and current information into the DT model, thereby enabling real-time modifications, predictions, and decision-making [89].

4) Augmented Decision-making: DT systems leverage data fusion to assimilate information from various sources, simplifying the process for decision-makers to access and decipher complex data. This consolidated information empowers them to make well-informed decisions concerning the administration, maintenance, and optimization of the physical asset or system [90].

However, the complexity and cost of implementing and maintaining a digital twin can be substantial, requiring advanced data analytics capabilities and continuous data flow, which might be resource-intensive for smaller operations or less critical applications [82].

IV. DIRECTION FOR FUTURE RESEARCH

A comprehensive review of the literature indicates that DT technology holds significant potential for fostering improved integration and optimization within PV systems. This, in turn, could facilitate increased energy efficiency and greater adoption of renewable energy sources. As a result, it is strongly recommended that solar farms consider integrating DT into their control processes. Nonetheless, the current developmental stage of DT technology presents certain limitations. This section aims to outline the primary constraints identified during the review process and proposes directions for future research to address these challenges. Ultimately, the goal is to further advance the integration of DT technology with PV systems, thereby maximizing its potential benefits.

A. AI-DRIVEN DECISION MAKING

Within the spectrum of DT systems evaluated, an approximate two-thirds can be partitioned into DM at 38% and DS at 24%. While these categories facilitate efficient surveillance and prognostication of PV performance, they inherently lack the capability for autonomous decision-making. Consequently, they cannot serve as the primary control mechanism for their corresponding physical entities.

To foster the evolution of fully autonomous PV systems, the integration of more advanced artificial intelligence methodologies within the DT system is necessary to augment
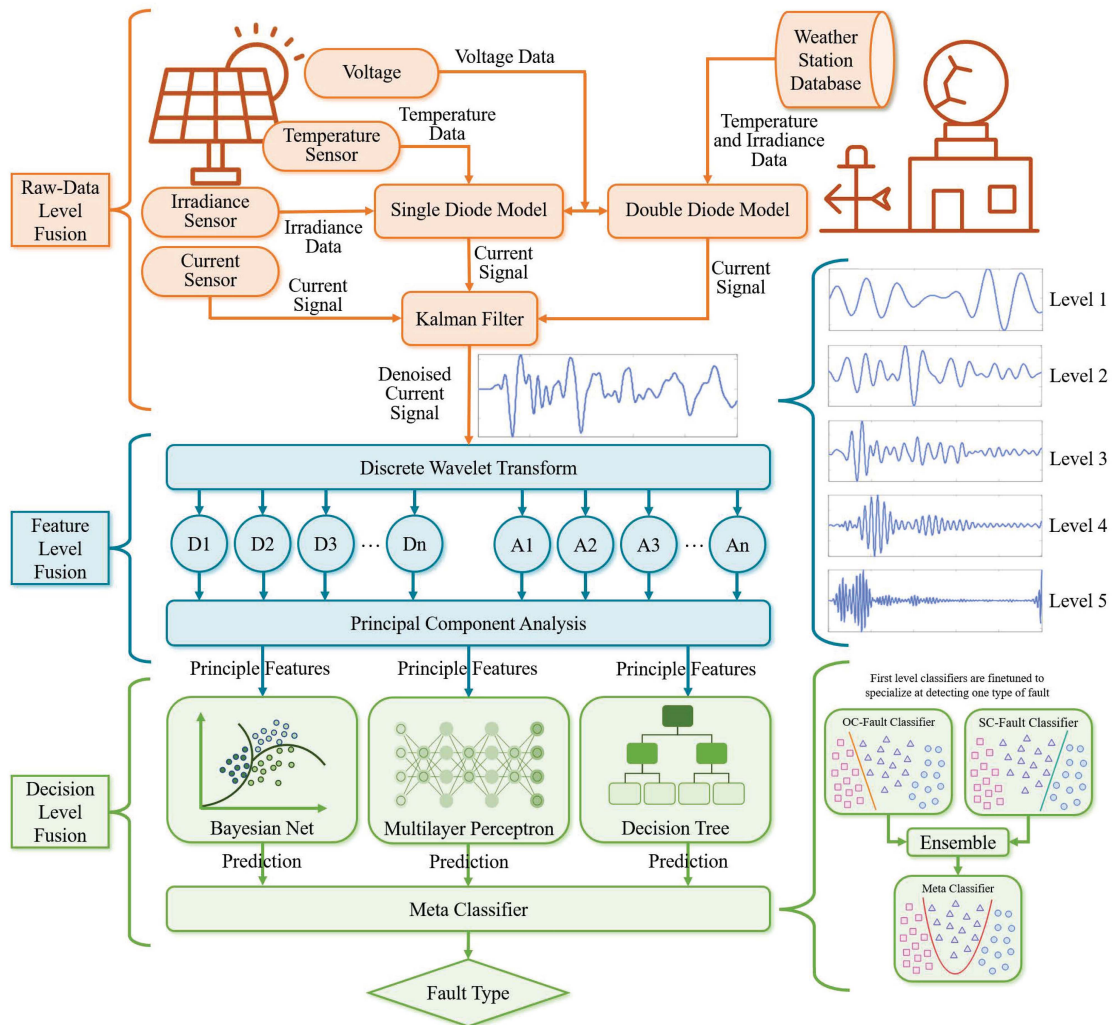
FIGURE 8. A case study of three levels of data fusion used for PV fault identification.

decision-making capabilities. Xia et al. posited that "Reinforcement Learning (RL) has been deployed in the domain of process system engineering to effectively resolve some formidable optimal control challenges [91]." The DT system creates a virtual representation of a tangible physical system, thereby providing an arena where the RL agent can hone its actions. The agent engages with the DT model by initiating actions, garnering feedback in the form of rewards or penalties, and updating its policy to amplify future actions based on this feedback [92]. For example, in a PV powered system, the RL agent could govern operational parameters to minimize energy consumption whilst optimizing output quality [63]. Matulis et al. suggested that "the fusion of digital twins and reinforcement learning provides numerous benefits, such as reduced training time and costs in physical space, or reduced risk of damage to an expensive physical test-bed [93]." This combination of DT technology and RL shapes a powerful instrument for optimizing and understanding intricate physical systems, with the added advantage of allowing real-time adaptation to changes in system dynamics.

B. COMPUTATIONAL REQUIREMENT

The drawbacks inherent to several data fusion methodologies predominantly stem from their intense requirement for computational resources when dealing with substantial data volumes. This necessity for considerable computational power not only results in significant investment but also introduces complications associated with real-time operation within DT systems [40]. Any computational latency may cause the injection of inaccuracies, as sensors would then be recording outdated data. Accumulating over time, these inaccuracies could potentially cause the DT model to lose synchrony with the physical counterpart it is designed to mimic.

This circumstance is especially critical for PV systems, which display a heightened sensitivity to environmental variations; therefore, any latency due to computation is unacceptable. Zhang et al. proposed edge computing, capable of facilitating shared computing resources at the network edge, as a dominant paradigm to address these computational demands [94]. Edge devices in the DT system can be controlled by an on-site microcontroller to manage panel activities and

sensor data collection. However, Chang et al. pointed out that PV "models that can deliver high prediction accuracy typically require extensive computational and storage capacities at run-time, and generally underperform in edge computing systems with limited resources [95]." Edge computing units, capable of executing signal filtering and lightweight algorithms, enable prompt responses to potential hazards while transmitting only relevant data to the cloud for further analysis [96]. This method effectively diminishes the volume of data necessitating transfer. The cloud environment, offering a larger assortment of computational resources, can support decision-level data fusion, thus facilitating the implementation of advanced machine learning algorithms and enhancing the overall efficacy and precision of the system.

C. SYSTEM SECURITY

The DT model of a PV system, being an essential component of the energy system, constitutes a vulnerable target for malicious cyber-attacks [97]. Any disruption or damage inflicted on the power supply could have extensive implications for individuals dependent on its steady operation. However, as Alcaraz et al. indicated, DT's "cybersecurity issues have not been sufficiently explored yet [98]", a statement underscored by the scant number of studies that integrate any security feature into their DT framework. Thus, maintaining security within the PV system's DT model emerges as a crucial challenge that demands immediate resolution. In addressing this security issue effectively, Yaqoob et al. identified that "blockchain possesses the potential to become the most relevant and capable technology to assure transparency, trust, and security in DTs [99]." The integration of a DT system and blockchain technology necessitates a comprehensive understanding of the DT's structure, the identification of critical data points suitable for blockchain integration, and the selection of an appropriate type of blockchain. When considering a PV system, a private blockchain appears to be the most optimal choice due to its inherent advantages in security and reduced computational overhead [100]. The integration of smart contracts can enable the automation of environmental data processing based on predetermined conditions [101]. By adopting these measures, the overall aim is to enhance the security, traceability, and data integrity of the DT models.

V. CONCLUSION

This critical and systematic review of DT technology and research, with a focus on the PV industry, has provided a comprehensive understanding of DT technologies. The review has illustrated the distinct advantages and limitations of different digital twin models by classifying them based on data connectivity and integration. It also summarizes how DT models are applied in various PV systems. Additionally, the review introduces a framework for integrating data fusion into digital twin systems. The critical review of published works reveals that DM models are primarily used to optimize initial designs and predict performance under various conditions but are limited by their static nature and lack of real-time data adaptation.

DS models offer real-time monitoring and performance analysis, aiding in maintenance and diagnostics, yet function only as passive observers without direct interaction with the system. In contrast, DT models deliver comprehensive benefits by monitoring, analyzing, predicting, and actively optimizing system performance in real-time, despite their high implementation and maintenance costs requiring sophisticated data analytics and continuous data integration.

The review has also identified key challenges that must be addressed to promote the widespread adoption of DT technology in the PV industry. Future research should focus on overcoming decision-making, computation, and security challenges. By addressing these obstacles and refining the proposed framework, the PV industry can capitalize on the benefits offered by DT and data fusion technologies, ultimately paving the way for a more sustainable, efficient, and reliable energy landscape.

REFERENCES

[1] D. J. Arent, A. Wise, and R. Gelman, "The status and prospects of renewable energy for combating global warming," *Energy Econ.*, vol. 33, no. 4, pp. 584–593, 2011.

[2] F. M. Guangul and G. T. Chala, "Solar energy as renewable energy source: SWOT analysis," in *Proc. IEEE 4th MEC Int. Conf. Big Data Smart City*, 2019, pp. 1–5.

[3] A. Maleki, A. Haghighi, M. E. H. Assad, I. Mahariq, and M. A. Nazari, "A review on the approaches employed for cooling PV cells," *Sol. Energy*, vol. 209, pp. 170–185, 2020.

[4] V. Benda and L. Černá, "PV cells and modules–State of the art, limits and trends," *Heliyon*, vol. 6, no. 12, 2020, Art. no. e05666.

[5] A. R. Jordehi, "Parameter estimation of solar photovoltaic (PV) cells: A review," *Renewable Sustain. Energy Rev.*, vol. 61, pp. 354–371, 2016.

[6] S. Li, Q. Gu, W. Gong, and B. Ning, "An enhanced adaptive differential evolution algorithm for parameter extraction of photovoltaic models," *Energy Convers. Manage.*, vol. 205, 2020, Art. no. 112443.

[7] M. Hanhart et al., "An integrated 50 V boost controller with digitally-assisted MPPT for submodule PV applications," *IEEE Open J. Power Electron.*, vol. 4, pp. 221–236, 2023.

[8] M. Lave, J. Kleissl, and E. Arias-Castro, "High-frequency irradiance fluctuations and geographic smoothing," *Sol. Energy*, vol. 86, no. 8, pp. 2190–2199, 2012.

[9] F. S. Bagci, K. A. Kim, Y.-C. Liu, and Y.-H. Liu, "Evaluation of power maximization and curtailment control methods for converters in wearable photovoltaic energy harvesting applications," *IEEE Open J. Power Electron.*, vol. 3, pp. 508–520, 2022.

[10] S. Li, W. Gong, and Q. Gu, "A comprehensive survey on meta-heuristic algorithms for parameter extraction of photovoltaic models," *Renewable Sustain. Energy Rev.*, vol. 141, 2021, Art. no. 110828.

[11] Y. Li, Q. Tao, and Y. Gong, "Digital twin simulation for integration of blockchain and internet of things for optimal smart management of PV-based connected microgrids," *Sol. Energy*, vol. 251, pp. 306–314, 2023.

[12] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, "Characterising the digital twin: A systematic literature review," *CIRP J. Manuf. Sci. Technol.*, vol. 29, pp. 36–52, 2020.

[13] A. Löcklin, T. Jung, N. Jazdi, T. Ruppert, and M. Weyrich, "Architecture of a human-digital twin as common interface for operator 4.0 applications," *Procedia CIRP*, vol. 104, pp. 458–463, 2021.

[14] X. Tu, J. Autiosalo, A. Jadid, K. Tammi, and G. Klinker, "A mixed reality interface for a digital twin based crane," *Appl. Sci.*, vol. 11, no. 20, 2021, Art. no. 9480.

[15] S. Yun, J.-H. Park, and W.-T. Kim, "Data-centric middleware based digital twin platform for dependable cyber-physical systems," in *Proc. 9th Int. Conf. Ubiquitous Future Netw.*, 2017, pp. 922–926.

[16] A. R. M. LLP, "Solar software market giants spending is going to boom - folsom," 2021. Accessed: Jul. 06, 2024. [Online]. Available: https://www.openpr.com/news/2316881/solar-software-market-giants-spending-is-going-to-boom-folso

[17] J. Liu, X. Lu, Y. Zhou, J. Cui, S. Wang, and Z. Zhao, "Design of photovoltaic power station intelligent operation and maintenance system based on digital twin," in *Proc. 6th Int. Conf. Robot. Automat. Eng.*, 2021, pp. 206–211.

[18] G. Zhang and X. Wang, "Digital twin modeling for photovoltaic panels based on hybrid neural network," in *Proc. IEEE 1st Int. Conf. Digit. Twins Parallel Intell.*, 2021, pp. 90–93.

[19] A. Livera et al., "Intelligent cloud-based monitoring and control digital twin for photovoltaic power plants," in *Proc. IEEE 49th Photovolt. Specialists Conf.*, 2022, pp. 0267–0274.

[20] L. U. Khan, Z. Han, W. Saad, E. Hossain, M. Guizani, and C. S. Hong, "Digital twin of wireless systems: Overview, taxonomy, challenges, and opportunities," *IEEE Commun. Surveys Tut.*, vol. 24, no. 4, pp. 2230–2254, Fourthquarter 2022.

[21] M. Mashaly, "Connecting the twins: A review on digital twin technology & its networking requirements," *Procedia Comput. Sci.*, vol. 184, pp. 299–305, 2021.

[22] L. Li, B. Lei, and C. Mao, "Digital twin in smart manufacturing," *J. Ind. Inf. Integration*, vol. 26, 2022, Art. no. 100289.

[23] E. J. Tuegel, A. R. Ingraffea, T. G. Eason, and S. M. Spottswood, "Reengineering aircraft structural life prediction using a digital twin," *Int. J. Aerosp. Eng.*, vol. 2011, 2011, Art. no. 154798.

[24] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, "Digital twin in industry: State-of-the-art," *IEEE Trans. Ind. Inform.*, vol. 15, no. 4, pp. 2405–2415, 2018.

[25] C. Semeraro, M. Lezoche, H. Panetto, and M. Dassisti, "Digital twin paradigm: A systematic literature review," *Comput. Ind.*, vol. 130, 2021, Art. no. 103469.

[26] F. Psarommatis, "A generic methodology and a digital twin for zero defect manufacturing (ZDM) performance mapping towards design for ZDM," *J. Manuf. Syst.*, vol. 59, pp. 507–521, 2021.

[27] T. Mukherjee and T. DebRoy, "A digital twin for rapid qualification of 3D printed metallic components," *Appl. Mater. Today*, vol. 14, pp. 59–65, 2019.

[28] J. Yu, P. Liu, and Z. Li, "Hybrid modelling and digital twin development of a steam turbine control stage for online performance monitoring," *Renewable Sustain. Energy Rev.*, vol. 133, 2020, Art. no. 110077.

[29] M. Wang and X. Yin, "Construction and maintenance of urban underground infrastructure with digital technologies," *Automat. Construction*, vol. 141, 2022, Art. no. 104464.

[30] S. Dasgupta and V. Kariwala, "ABB's electromagnetic flowmeter digital twin drives performance," *ABB Rev.*, vol. 2, pp. 58–65, 2018.

[31] A. Vučurović, D. S. Panić, and D. Živković, "Digital twin for a competitive win: The general electric case study," in *Proc. Int. Sci. Conf. Post-Pandemic Econ. Challenges*, 2022, pp. 345–355.

[32] D. A. Guerra-Zubiaga, A. Bondar, G. Escobedo, and A. Schumacher, "Digital twin in a manufacturing integrated system: Siemens TIA and PLM case study," in *Proc. ASME Int. Mech. Eng. Congr. Expo.*, 2019, pp. 1–8.

[33] M. Liu, S. Fang, H. Dong, and C. Xu, "Review of digital twin about concepts, technologies, and industrial applications," *J. Manuf. Syst.*, vol. 58, pp. 346–361, 2021.

[34] J. Do Amaral, C. Dos Santos, J. Montevechi, and A. De Queiroz, "Energy digital twin applications: A review," *Renewable Sustain. Energy Rev.*, vol. 188, 2023, Art. no. 113891.

[35] C. Ghenai, L. A. Husein, M. Al Nahlawi, A. K. Hamid, and M. Bettayeb, "Recent trends of digital twin technologies in the energy sector: A comprehensive review," *Sustain. Energy Technol. Assessments*, vol. 54, 2022, Art. no. 102837.

[36] A. Kavousi-Fard, M. Dabbaghjamanesh, M. Jafari, M. Fotuhi-Firuzabad, Z. Y. Dong, and T. Jin, "Digital twin for mitigating solar energy resources challenges: A perspective review," *Sol. Energy*, vol. 274, 2024, Art. no. 112561.

[37] C. Qian, X. Liu, C. Ripley, M. Qian, F. Liang, and W. Yu, "Digital twin—Cyber replica of physical things: Architecture, applications and future research directions," *Future Internet*, vol. 14, no. 2, 2022, Art. no. 64.

[38] M. Zhang, F. Tao, and A. Nee, "Digital twin enhanced dynamic job-shop scheduling," *J. Manuf. Syst.*, vol. 58, pp. 146–156, 2021.

[39] M. Dietz and G. Pernul, "Digital twin: Empowering enterprises towards a system-of-systems approach," *Bus. Inf. Syst. Eng.*, vol. 62, pp. 179–184, 2020.

[40] A. Rasheed, O. San, and T. Kvamsdal, "Digital twin: Values, challenges and enablers from a modeling perspective," *IEEE Access*, vol. 8, pp. 21980–22012, 2020.

[41] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, "Digital twin in manufacturing: A categorical literature review and classification," *IFAC-PapersOnline*, vol. 51, no. 11, pp. 1016–1022, 2018.

[42] M. P. Almeida, O. Perpinan, and L. Narvarte, "PV power forecast using a nonparametric PV model," *Sol. Energy*, vol. 115, pp. 354–368, 2015.

[43] C. Ban, T. Hong, K. Jeong, C. Koo, and J. Jeong, "A simplified estimation model for determining the optimal rooftop photovoltaic system for gable roofs," *Energy Buildings*, vol. 151, pp. 320–331, 2017.

[44] I. N. del'Énergie Solaire, "Design and simulation of PV systems on PVSYST software," 2024. Accessed: Jul. 06, 2024. [Online]. Available: https://www.ines-solaire.org/en/capacity-building/training/pv13-conception-et-simulation-des-systemes-pv-sur-pvsyst/

[45] T. Bauernhansl, S. Hartleif, and T. Felix, "The digital shadow of production–a concept for the effective and efficient information supply in dynamic industrial environments," *Procedia Cirp*, vol. 72, pp. 69–74, 2018.

[46] S. Samara and E. Natsheh, "Intelligent real-time photovoltaic panel monitoring system using artificial neural networks," *IEEE Access*, vol. 7, pp. 50287–50299, 2019.

[47] A. Ladj, Z. Wang, O. Meski, F. Belkadi, M. Ritou, and C. Da Cunha, "A knowledge-based digital shadow for machining industry in a digital twin perspective," *J. Manuf. Syst.*, vol. 58, pp. 168–179, 2021.

[48] H. Wan, M. David, and W. Derigent, "Modelling digital twins as a recursive multi-agent architecture: Application to energy management of communicating materials," *IFAC-PapersOnLine*, vol. 54, no. 1, pp. 880–885, 2021.

[49] P. Jain, J. Poon, J. P. Singh, C. Spanos, S. R. Sanders, and S. K. Panda, "A digital twin approach for fault diagnosis in distributed photovoltaic systems," *IEEE Trans. Power Electron.*, vol. 35, no. 1, pp. 940–956, Jan. 2020.

[50] B. He, X. Cao, and Y. Hua, "Data fusion-based sustainable digital twin system of intelligent detection robotics," *J. Cleaner Prod.*, vol. 280, 2021, Art. no. 124181.

[51] Q. Qi et al., "Enabling technologies and tools for digital twin," *J. Manuf. Syst.*, vol. 58, pp. 3–21, 2021.

[52] Z. Liu, G. Xiao, H. Liu, and H. Wei, "Multi-sensor measurement and data fusion," *IEEE Instrum. Meas. Mag.*, vol. 25, no. 1, pp. 28–36, Feb. 2022.

[53] X. Yang, M. Xu, S. Xu, and X. Han, "Day-ahead forecasting of photovoltaic output power with similar cloud space fusion based on incomplete historical data mining," *Appl. Energy*, vol. 206, pp. 683–696, 2017.

[54] H. A. M. Maghraby, M. H. Shwehdi, and G. K. Al-Bassam, "Probabilistic assessment of photovoltaic (PV) generation systems," *IEEE Trans. Power Syst.*, vol. 17, no. 1, pp. 205–208, Feb. 2002.

[55] K. Fujii, M. Yamagiwa, and M. Uehara, "A study on operation of photovoltaic grid system using weather forecast," in *Proc. IEEE 3rd Int. Conf. Intell. Netw. Collaborative Syst.*, 2011, pp. 157–163.

[56] A. Downey, F. Ubertini, and S. Laflamme, "Algorithm for damage detection in wind turbine blades using a hybrid dense sensor network with feature level data fusion," *J. Wind Eng. Ind. Aerodynamics*, vol. 168, pp. 288–296, 2017.

[57] P. Hoang, V. Bourdin, Q. Liu, G. Caruso, and V. Archambault, "Coupling optical and thermal models to accurately predict PV panel electricity production," *Sol. Energy Mater. Sol. Cells*, vol. 125, pp. 325–338, 2014.

[58] M. Ma et al., "Multi-features fusion for short-term photovoltaic power prediction," *Intell. Converged Netw.*, vol. 3, no. 4, pp. 311–324, 2022.

[59] Y. Wei, D. Wu, and J. Terpenny, "Decision-level data fusion in quality control and predictive maintenance," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 1, pp. 184–194, Jan. 2021.

[60] F. Castanedo, "A review of data fusion techniques," *Sci. World J.*, vol. 2013, 2013, Art. no. 704504.

[61] P. Radhakrishnan, K. Ramaiyan, A. Vinayagam, and V. Veerasamy, "A stacking ensemble classification model for detection and classification of power quality disturbances in PV integrated power network," *Measurement*, vol. 175, 2021, Art. no. 109025.

[62] X. Guo, Z. Na, D. Ma, Y. Lu, and X. Luo, "Fault diagnosis of photovoltaic system based on machine learning model fusion," in *Proc. IOP Conf. Ser.: Earth Environ. Sci.*, vol. 467, 2020, Art. no. 12073.

[63] H. Liu, X. Xie, D. Wang, X. Ma, and X. Luo, "An intelligent fault diagnosis technology for distributed photovoltaic system based on ensemble voting fusion model," in *Proc. IEEE 4th Conf. Energy Internet Energy System Integration*, 2020, pp. 2146–2150.

[64] A. Sangwongwanich, H. Wang, and F. Blaabjerg, "Reduced-order thermal modeling for photovoltaic inverters considering mission profile dynamics," *IEEE Open J. Power Electron.*, vol. 1, pp. 407–419, 2020.

[65] M. Dörenkömper, A. Wahed, A. Kumar, M. de Jong, J. Kroon, and T. Reindl, "The cooling effect of floating PV in two different climate zones: A comparison of field test data from the Netherlands and Singapore," *Sol. Energy*, vol. 219, pp. 15–23, 2021.

[66] P. A. Leicester, C. I. Goodier, and P. N. Rowley, "Probabilistic analysis of solar photovoltaic self-consumption using Bayesian network models," *IET Renewable Power Gener.*, vol. 10, no. 4, pp. 448–455, 2016.

[67] Y.-M. Saint-Drenan, S. Vogt, S. Killinger, J. M. Bright, R. Fritz, and R. Potthast, "Bayesian parameterisation of a regional photovoltaic model–application to forecasting," *Sol. Energy*, vol. 188, pp. 760–774, 2019.

[68] K. Doubleday, S. Jascourt, W. Kleiber, and B.-M. Hodge, "Probabilistic solar power forecasting using Bayesian model averaging," *IEEE Trans. Sustain. Energy*, vol. 12, no. 1, pp. 325–337, Jan. 2021.

[69] M. Karamirad, M. Omid, R. Alimardani, H. Mousazadeh, and S. N. Heidari, "ANN based simulation and experimental verification of analytical four-and five-parameters models of PV modules," *Simul. Modelling Pract. Theory*, vol. 34, pp. 86–98, 2013.

[70] 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.

[71] T. Ikegami, T. Maezono, F. Nakanishi, Y. Yamagata, and K. Ebihara, "Estimation of equivalent circuit parameters of PV module and its application to optimal operation of PV system," *Sol. Energy Mater. Sol. Cells*, vol. 67, no. 1–4, pp. 389–395, 2001.

[72] O. I. Olayiwola and P. S. Barendse, "Photovoltaic cell/module equivalent electric circuit modeling using impedance spectroscopy," *IEEE Trans. Ind. Appl.*, vol. 56, no. 2, pp. 1690–1701, Mar./Apr. 2020.

[73] J. Yuan, Z. Tian, J. Ma, K. L. Man, and B. Li, "A digital twin approach for modeling electrical characteristics of bifacial solar panels," in *Proc. IEEE Int. Conf. Ind. IoT, Big Data Supply Chain*, 2022, pp. 317–321.

[74] T. Soubdhan, J. Ndong, H. Ould-Baba, and M.-T. Do, "A robust forecasting framework based on the Kalman filtering approach with a twofold parameter tuning procedure: Application to solar and photovoltaic prediction," *Sol. Energy*, vol. 131, pp. 246–259, 2016.

[75] Y. Yang, T. Yu, W. Zhao, and X. Zhu, "Kalman filter photovoltaic power prediction model based on forecasting experience," *Front. Energy Res.*, vol. 9, 2021, Art. no. 682852.

[76] A. J. Veldhuis and A. H. M. E. Reinders, "Real-time irradiance simulation for PV products and building integrated PV in a virtual reality environment," *IEEE J. Photovolt.*, vol. 2, no. 3, pp. 352–358, Jul. 2012.

[77] A. Ameur, A. Berrada, K. Loudiyi, and M. Aggour, "Forecast modeling and performance assessment of solar PV systems," *J. Cleaner Prod.*, vol. 267, 2020, Art. no. 122167.

[78] A. Pirayawaraporn, S. Sappaniran, S. Nooraksa, C. Prommai, N. Chindakham, and C. Jamroen, "Innovative sensorless dual-axis solar tracking system using particle filter," *Appl. Energy*, vol. 338, 2023, Art. no. 120946.

[79] Z. Qadir et al., "Predicting the energy output of hybrid PV–wind renewable energy system using feature selection technique for smart grids," *Energy Rep.*, vol. 7, pp. 8465–8475, 2021.

[80] C. S. B. Clausen, Z. G. Ma, and B. N. Jørgensen, "Can we benefit from game engines to develop digital twins for planning the deployment of photovoltaics?," *Energy Inform.*, vol. 5, no. Suppl 4, pp. 1–26, 2022.

[81] A. Hariri and M. O. Faruque, "A hybrid simulation tool for the study of PV integration impacts on distribution networks," *IEEE Trans. Sustain. Energy*, vol. 8, no. 2, pp. 648–657, Apr. 2017.

[82] R. Gugulothu and B. Nagu, "A Bayesian fusion technique for maximum power point tracking under partial shading condition," *SN Appl. Sci.*, vol. 3, pp. 1–15, 2021.

[83] W. Chine, A. Mellit, V. Lughi, A. Malek, G. Sulligoi, and A. M. Pavan, "A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks," *Renewable Energy*, vol. 90, pp. 501–512, 2016.

[84] G. Buyuksalih, S. Bayburt, A. Baskaraca, H. Karim, and A. Abdul Rahman, "Calculating solar energy potential of buildings and visualization within unity 3D game engine," *Int. Arch. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 42, pp. 39–44, 2017.

[85] M. Metry and R. S. Balog, "An adaptive model predictive controller for current sensorless MPPT in PV systems," *IEEE Open J. Power Electron.*, vol. 1, pp. 445–455, 2020.

[86] B. Tang, "Intelligent station digital platform," 2022. Accessed: Jul. 06, 2024. [Online]. Available: https://guangfu.bjx.com.cn/b2b/product/8528685/product-view-383486.html

[87] M. Xiong and H. Wang, "Digital twin applications in aviation industry: A review," *Int. J. Adv. Manuf. Technol.*, vol. 121, no. 9/10, pp. 5677–5692, 2022.

[88] S. A. Renganathan, K. Harada, and D. N. Mavris, "Aerodynamic data fusion toward the digital twin paradigm," *AIAA J.*, vol. 58, no. 9, pp. 3902–3918, 2020.

[89] A. Procacci, M. Cafiero, S. Sharma, M. M. Kamal, A. Coussement, and A. Parente, "Digital twin for experimental data fusion applied to a semi-industrial furnace fed with H_2-rich fuel mixtures," *Energies*, vol. 16, no. 2, 2023, Art. no. 662.

[90] K. Abutalip, A. Al-Lahham, and A. El Saddik, "Digital twin of atmospheric environment: Sensory data fusion for high-resolution $PM_{2.5}$ estimation and action policies recommendation," *IEEE Access*, vol. 11, pp. 14448–14457, 2023.

[91] K. Xia et al., "A digital twin to train deep reinforcement learning agent for smart manufacturing plants: Environment, interfaces and intelligence," *J. Manuf. Syst.*, vol. 58, pp. 210–230, 2021.

[92] E. Artetxe, J. Uralde, O. Barambones, I. Calvo, and I. Martin, "Maximum power point tracker controller for solar photovoltaic based on reinforcement learning agent with a digital twin," *Mathematics*, vol. 11, no. 9, 2023, Art. no. 2166.

[93] M. Matulis and C. Harvey, "A robot arm digital twin utilising reinforcement learning," *Comput. Graph.*, vol. 95, pp. 106–114, 2021.

[94] K. Zhang, J. Cao, and Y. Zhang, "Adaptive digital twin and multiagent deep reinforcement learning for vehicular edge computing and networks," *IEEE Trans. Ind. Inform.*, vol. 18, no. 2, pp. 1405–1413, Feb. 2022.

[95] X. Chang, W. Li, and A. Y. Zomaya, "A lightweight short-term photovoltaic power prediction for edge computing," *IEEE Trans. Green Commun. Netw.*, vol. 4, no. 4, pp. 946–955, Dec. 2020.

[96] R. Dong, C. She, W. Hardjawana, Y. Li, and B. Vucetic, "Deep learning for hybrid 5G services in mobile edge computing systems: Learn from a digital twin," *IEEE Trans. Wireless Commun.*, vol. 18, no. 10, pp. 4692–4707, Oct. 2019.

[97] J. Zhang et al., "Machine learning-based cyber-attack detection in photovoltaic farms," *IEEE Open J. Power Electron.*, vol. 4, pp. 658–673, 2023.

[98] C. Alcaraz and J. Lopez, "Digital twin: A comprehensive survey of security threats," *IEEE Commun. Surveys Tut.*, vol. 24, no. 3, pp. 1475–1503, Thirdquarter 2022.

[99] I. Yaqoob, K. Salah, M. Uddin, R. Jayaraman, M. Omar, and M. Imran, "Blockchain for digital twins: Recent advances and future research challenges," *IEEE Netw.*, vol. 34, no. 5, pp. 290–298, Sep./Oct. 2020.

[100] S. Pongnumkul, C. Siripanpornchana, and S. Thajchayapong, "Performance analysis of private blockchain platforms in varying workloads," in *Proc. IEEE 26th Int. Conf. Comput. Commun. Netw.*, 2017, pp. 1–6.

[101] J. Lin, M. Pipattanasomporn, and S. Rahman, "Comparative analysis of blockchain-based smart contracts for solar electricity exchanges," in *Proc. IEEE Power & Energy Soc. Innov. Smart Grid Technol. Conf.*, 2019, pp. 1–5.