IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Digital Twin Framework For Performance Monitoring And Predictive Analysis Of Photovoltaic Systems Using Machine Learning

Submitted by, [1]Sumi N, 2Rinku Scaria
[1]Student, MTech, Dept.of.EEE, [2]Assistant Professor, Dept.of.EEE
FISAT College (Affiliated to KTU), Angamaly, Kerala, India

Abstract: The growing adoption of photovoltaic (PV) systems requires intelligent monitoring methods to ensure efficient operation and early detection of performance degradation. Digital Twin technology offers a virtual model that replicates the behavior of a physical energy system using real-time operational data. This study presents a digital twin framework for photovoltaic systems that integrates machine learning techniques to estimate power generation and identify abnormal operating conditions. Environmental variables such as solar irradiance, module temperature, voltage, and current are utilized to train predictive models including Random Forest and Long Short-Term Memory (LSTM) networks. The virtual model continuously evaluates system behavior by comparing predicted output with measured PV performance. Significant deviations between the predicted and actual values indicate potential faults or efficiency losses. The results indicate that combining machine learning algorithms with digital twin modeling improves prediction capability and supports proactive maintenance of solar energy systems. The proposed framework can assist in enhancing reliability and operational efficiency of PV installations and may contribute to smarter renewable energy management in future power systems.

Index Terms - Digital Twin, Photovoltaic Systems, Machine Learning, Random Forest, LSTM, Predictive Maintenance.

I. Introduction

The increasing global demand for sustainable energy has accelerated the deployment of renewable energy technologies, particularly photovoltaic (PV) systems. Solar energy is considered one of the most promising renewable sources due to its availability, environmental benefits, and declining installation costs. However, the performance of PV systems is strongly influenced by environmental conditions such as solar irradiance, temperature variations, dust accumulation, and partial shading. These factors can lead to fluctuations in power generation and gradual system degradation. Therefore, continuous monitoring and intelligent analysis are necessary to maintain optimal system performance and reliability.

Traditional PV monitoring approaches mainly rely on basic measurement systems that record electrical parameters such as voltage, current, and power output. Although these systems provide useful operational data, they often lack advanced analytical capabilities for predicting future performance or identifying hidden system faults. As PV installations grow in scale and complexity, more sophisticated monitoring and predictive techniques are required to ensure efficient operation and reduce maintenance costs.

Digital Twin technology has recently emerged as a powerful concept for modeling and monitoring physical systems. A digital twin is a dynamic virtual representation of a physical asset that continuously updates using real-time data from sensors and monitoring devices. By combining simulation models with real operational data, digital twins enable accurate system analysis, predictive diagnostics, and improved decision-making. In the context of photovoltaic systems, a digital twin can replicate the behavior of a solar installation and analyze its performance under varying environmental and operating conditions.

Recent developments in machine learning have further enhanced the capabilities of digital twin systems. Machine learning algorithms can identify complex nonlinear relationships between environmental variables and PV output power. These techniques allow accurate prediction of energy generation and enable early detection of abnormal system behavior. Algorithms such as Random Forest and Long Short-Term Memory (LSTM) networks are particularly suitable for handling nonlinear patterns and time-series data commonly observed in solar energy generation.

Integrating machine learning models with digital twin frameworks offers a promising solution for advanced PV system monitoring and predictive maintenance. By continuously comparing predicted outputs with real system measurements, the digital twin can identify deviations that may indicate faults, degradation, or efficiency losses. This capability helps operators take preventive actions before major failures occur.

This study proposes a digital twin-based framework for photovoltaic systems that incorporates machine learning models to estimate power generation and monitor system performance. The proposed approach aims to improve prediction accuracy, enable early fault detection, and support intelligent management of solar energy systems. The methodology combines real-time data acquisition, predictive modeling, and simulation-based analysis to create a comprehensive monitoring platform for modern PV installations.

II. LITERATURE REVIEW

The integration of advanced monitoring and predictive techniques has become essential for improving the reliability and efficiency of photovoltaic (PV) systems. Researchers have explored various methods including digital twin modeling, machine learning algorithms, and data-driven prediction approaches to enhance PV system performance.

Digital Twin (DT) technology has emerged as an innovative framework for representing physical systems through dynamic virtual models. A digital twin continuously synchronizes with the real system using operational data, enabling real-time monitoring, simulation, and performance analysis. Recent studies indicate that digital twins can significantly improve PV system management by providing predictive insights and enabling early fault detection. The technology allows the virtual model to update continuously according to changes in environmental conditions and system behavior, making it suitable for dynamic renewable energy systems.

Several researchers have investigated the use of digital twin models specifically for photovoltaic installations. In many implementations, the digital model receives environmental parameters such as irradiance and temperature from sensors and compares the simulated output with the measured PV output. If the difference between these outputs exceeds a predefined threshold, the system identifies a possible fault or performance degradation. Such approaches allow operators to detect abnormal system behavior and schedule maintenance activities before significant power losses occur.

Machine learning techniques have also been widely applied for forecasting PV power generation. Because solar energy production depends strongly on weather conditions and other nonlinear factors, traditional analytical models often struggle to achieve accurate predictions. Machine learning algorithms can learn complex relationships between environmental variables and PV output power using historical datasets. Studies reviewing PV power forecasting methods highlight that machine learning approaches significantly improve prediction accuracy and help manage the variability of solar energy generation.

Among machine learning techniques, deep learning models have gained particular attention for solar power forecasting. Neural networks such as Long Short-Term Memory (LSTM) are capable of learning temporal dependencies in time-series data and have shown promising performance in PV power prediction tasks. Recent research has also explored transformer based deep learning architectures within digital twin frameworks, demonstrating improved prediction accuracy for photovoltaic power output when compared with conventional models.

In addition to prediction tasks, artificial intelligence methods have been applied for fault detection and performance assessment in PV systems. Data-driven algorithms can analyze operational data such as voltage, current, and meteorological parameters to identify deviations from expected system behavior. These techniques enable automatic detection of faults such as module degradation, shading, or inverter malfunction, improving the reliability of solar installations.

Despite these advancements, research on digital twins for photovoltaic systems is still in its early stages. Recent reviews indicate that only a small portion of published studies fully implement the essential characteristics of a digital twin, such as bidirectional data communication and real-time synchronization between physical and virtual systems. This highlights the need for more comprehensive frameworks that integrate accurate simulation models, real-time data acquisition, and machine learning algorithms.

Therefore, integrating machine learning techniques within a digital twin architecture presents a promising research direction for improving PV system monitoring and predictive maintenance. By combining real-time data with intelligent prediction models, digital twin systems can provide deeper insights into PV performance and support more efficient operation of renewable energy systems.

III. PROPOSED DIGITAL TWIN ARCHITECTURE

The proposed digital twin architecture for the photovoltaic (PV) system is designed to create a continuous interaction between the physical solar installation and its virtual counterpart. The architecture enables real-time monitoring, predictive analysis, and performance evaluation by combining sensor data, simulation models, and machine learning algorithms. The framework consists of multiple functional layers that work together to replicate and analyze the behavior of the actual PV system.

3.1 Physical System Layer

The physical layer represents the real photovoltaic power generation system installed in the field. It includes the PV modules, power electronic components, environmental sensors, and measurement units responsible for collecting operational data. The PV array converts solar radiation into electrical energy through the photovoltaic effect. The output of the PV system depends on several environmental and electrical parameters such as solar irradiance, module temperature, voltage, and current. Sensors installed within the system measure these parameters continuously. Typical measurements collected from the physical system include:

• Solar irradiance
• Ambient and module temperature
• PV array voltage
• PV array current
• Generated output power

These parameters provide essential information about the operating condition of the PV system. The collected data is transmitted to the data processing unit through communication interfaces or IoT-based monitoring systems.

3.2 Data Acquisition and Communication Layer

The data acquisition layer acts as the interface between the physical PV installation and the digital twin environment. It gathers real-time measurements from sensors and transfers them to the digital platform for further processing and analysis. Data acquisition systems typically include microcontrollers, data loggers, or IoT-enabled devices that continuously record system parameters. The communication network may use wired or wireless technologies such as Wi-Fi, GSM, or cloud-based platforms to transmit the collected data. Before the data is used for analysis, it undergoes preprocessing procedures such as filtering, normalization, and removal of missing or inconsistent values. Proper preprocessing ensures that the input data is reliable and suitable for machine learning models and simulation algorithms.

3.3 Digital Twin Modeling Layer

The digital twin layer represents the virtual model of the photovoltaic system. It replicates the behavior and performance characteristics of the real PV installation using simulation tools and mathematical models.

In this research, the digital twin model can be developed using simulation environments such as MATLAB/Simulink. The virtual model includes representations of PV modules, power electronic converters, and load components. The model simulates the electrical characteristics of the PV array under varying environmental conditions.

The digital twin continuously receives real-time data from the physical system and updates its internal parameters accordingly. This synchronization enables the virtual model to closely mimic the real operating conditions of the PV system. By running simulations using updated data, the digital twin can estimate expected system performance.

3.4 Machine Learning Prediction Layer

To enhance the predictive capability of the digital twin, machine learning algorithms are integrated into the architecture. These algorithms analyze historical and real-time PV system data to learn the relationship between environmental variables and power generation.Two machine learning techniques are used in this framework:

Random Forest Regression

Random Forest is an ensemble learning method that constructs multiple decision trees during training and combines their outputs to improve prediction accuracy. It is particularly effective in handling nonlinear relationships between input variables and PV output power. The algorithm uses environmental parameters such as irradiance and temperature as inputs to estimate the expected power generation.

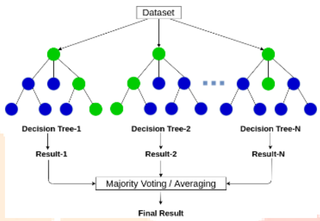

Fig 1: Decision taking in Random Forest

Long Short-Term Memory (LSTM) Neural Network

LSTM is a type of recurrent neural network designed to process sequential data and capture long-term dependencies in time-series datasets. In PV systems, solar irradiance and power output vary over time, making LSTM suitable for predicting future power generation patterns. The model learns temporal relationships within historical data and produces accurate forecasts of PV system output. The outputs from these machine learning models are used as reference values for evaluating the actual system performance.

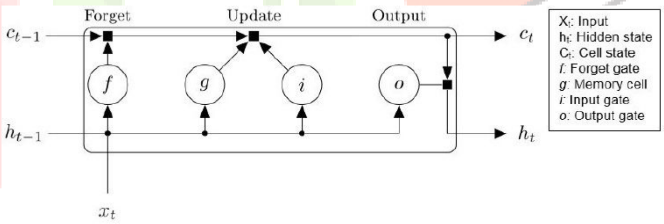

Fig 2: LSTM Algorithm model

3.5 Performance Evaluation and Fault Detection Layer

The performance evaluation layer compares the predicted output from the digital twin with the actual output measured from the PV system. This comparison helps identify discrepancies between expected and real system behavior.
If the difference between predicted power and measured power remains within an acceptable range, the system is considered to be operating normally. However, if the deviation exceeds a predefined threshold, the system may indicate abnormal operating conditions such as:

- Partial shading
- Module degradation
- Dust accumulation on panels
- Sensor malfunction
- Inverter or connection faults

The digital twin system can generate alerts or warnings when such deviations occur, enabling operators to take corrective actions before the problem leads to significant power loss.

3.6 Visualization and Decision Support Layer

The final layer of the architecture provides visualization and decision support functions. Data from the digital twin system can be displayed through dashboards or monitoring interfaces that present real-time system performance indicators.

Operators can monitor parameters such as predicted power, actual power output, efficiency trends, and detected anomalies. Visualization tools help in understanding system behavior and support data-driven maintenance decisions.

Through continuous monitoring and predictive analysis, the digital twin architecture improves the operational reliability of the photovoltaic system and reduces downtime.

3.7 Summary of the Architecture

The proposed digital twin framework integrates physical PV components, data acquisition systems, simulation models, and machine learning algorithms into a unified monitoring platform. The interaction between the physical system and the virtual model enables accurate performance prediction, early fault detection, and improved system management. By combining real-time data with intelligent predictive models, the architecture provides an effective solution for advanced monitoring and optimization of photovoltaic energy systems.

IV. METHODOLOGY

The methodological approach adopted in this work focuses on constructing a virtual representation of a photovoltaic system and enabling it to analyze system behavior using operational data and predictive algorithms. The overall procedure involves several stages, beginning with data acquisition and ending with performance assessment through comparison between predicted and observed outputs.

4.1 Data Collection and Parameter Selection

The initial phase involves gathering operational and environmental information associated with the photovoltaic installation. Parameters that significantly influence solar power generation are considered in this study. These typically include solar irradiance, module temperature, output voltage, current, and generated power. Environmental variables such as irradiance and temperature describe the external conditions affecting the PV modules, while electrical measurements represent the internal operating state of the system. Data may be obtained from monitoring devices, simulation environments, or publicly available solar datasets. The collected information forms the primary input for the subsequent modeling and prediction processes.

4.2 Data Preparation and Processing

Before applying analytical models, the dataset undergoes several preparation steps to improve its reliability and consistency. Raw measurements often contain incomplete entries, noise, or irregular values. Such inconsistencies are addressed through preprocessing procedures such as data cleaning, normalization, and filtering.

Normalization helps ensure that different parameters with varying units and magnitudes are brought to a comparable scale. The processed dataset is then divided into separate subsets for training and testing purposes. The training data is used to develop predictive models, while the testing data evaluates the performance of those models.

4.3 Development of the Digital Twin Model

Following data preparation, a virtual model representing the photovoltaic system is created. The digital twin aims to reproduce the operational characteristics of the physical system under different environmental conditions. Simulation tools can be used to represent PV modules, electrical behavior, and energy conversion processes.

The virtual model receives real-time or historical data inputs corresponding to irradiance and temperature. Based on these inputs, the model estimates the expected electrical output of the PV system. The digital twin therefore acts as a reference system that mirrors the physical installation and predicts its behavior.

4.4 Implementation of Predictive Algorithms

To enhance the analytical capability of the digital twin, machine learning techniques are incorporated into the framework. Predictive algorithms learn relationships between environmental conditions and PV power output using historical datasets.
Regression-based machine learning models are suitable for estimating continuous outputs such as generated power. Algorithms such as Random Forest are effective for capturing nonlinear relationships between input parameters and power generation. Additionally, sequence-based neural networks such as Long Short-Term Memory (LSTM) are capable of identifying temporal patterns in solar generation data. Through a training process, these models adjust their internal parameters to minimize prediction error. Once trained, they can estimate PV power output based on new input conditions.

4.5 Performance Comparison and Deviation Analysis

After generating predictions from the digital twin and machine learning models, the predicted results are compared with the actual system output. This comparison serves as a mechanism for evaluating system performance.
When predicted and measured outputs closely match, the PV system is assumed to be operating normally. However, significant differences between the two values may indicate abnormal operating conditions or potential system faults. By continuously analyzing these deviations, the digital twin framework can assist in identifying performance degradation or operational irregularities.

4.6 Evaluation of Model Performance

The effectiveness of the predictive models is assessed using statistical performance indicators. Metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and prediction accuracy are commonly used to measure how closely the predicted values correspond to the observed data.
Lower error values indicate better model performance and improved predictive capability. The evaluation results help determine the reliability of the digital twin framework in representing the behavior of the photovoltaic system.

4.7 Overall Workflow

The methodological workflow begins with acquiring PV system data, followed by data preprocessing and model development. Machine learning algorithms are then trained using historical datasets to estimate system output. The digital twin generates predicted performance values, which are compared with actual measurements to evaluate system behavior. Through this process, the proposed framework enables predictive monitoring and supports improved operational management of photovoltaic installations.

V. RESULTS AND DISCUSSION

The effectiveness of the proposed digital twin framework was evaluated by comparing the predicted photovoltaic (PV) system output with the actual system performance under varying environmental conditions. The analysis focused on examining how accurately the digital twin model and machine learning algorithms could represent the behavior of the physical PV system.

5.1 Prediction Performance

After training the predictive models using historical PV system data, the models were applied to estimate the output power of the system based on environmental inputs such as solar irradiance and module temperature. The predicted values obtained from the machine learning models were compared with the actual output measurements.

The results indicated that the predicted power profiles closely followed the pattern of the measured PV output. When solar irradiance increased during peak daylight hours, both the predicted and measured power values showed a corresponding rise. Similarly, during periods of reduced sunlight, the generated power decreased in both datasets. This behavior demonstrates that the predictive models were able to capture the general relationship between environmental conditions and PV system performance.

5.2 Comparison of Machine Learning Models

Two different machine learning techniques were used to evaluate their capability in predicting PV power generation. The Random Forest algorithm provided stable predictions across the dataset and showed good performance when handling nonlinear relationships between environmental variables and PV output power.

The Long Short-Term Memory (LSTM) network demonstrated strong performance in capturing temporal variations in solar energy generation. Because PV output changes continuously throughout the day, the ability of LSTM to learn sequential patterns helped improve forecasting accuracy. In many instances, the LSTM model produced slightly closer estimates to the measured values compared with the Random Forest approach, particularly during rapidly changing weather conditions.

5.3 Error Analysis

To evaluate the reliability of the predictive models, several statistical performance indicators were examined. These metrics measure the difference between predicted and actual values and provide an indication of model accuracy. The analysis showed relatively low error values for both models, suggesting that the digital twin framework was able to estimate PV output with reasonable precision. Lower error levels indicate that the virtual model successfully represents the behavior of the physical PV system. The evaluation also demonstrated that incorporating machine learning algorithms improves prediction capability when compared with simple analytical estimation methods.

5.4 System Monitoring and Fault Detection Capability

Another important objective of the digital twin framework is the ability to detect deviations between expected and actual system behavior. By continuously comparing predicted output with measured PV output, the system can identify abnormal operating conditions. During the analysis, instances where the difference between predicted and measured power exceeded the normal operating range were examined. Such deviations may occur due to factors such as partial shading, accumulation of dust on PV panels, or component degradation. When these conditions occur, the digital twin system can generate alerts indicating a potential issue within the PV installation. This monitoring capability provides valuable support for preventive maintenance. Instead of waiting for a major system failure, operators can identify performance problems early and take corrective actions to restore system efficiency.

5.5 Operational Insights from the Digital Twin

The digital twin framework also provides useful insights into the overall performance of the photovoltaic system. By analyzing long-term operational data, the system can reveal patterns related to efficiency variations, environmental influences, and energy production trends. The results demonstrate that the digital twin approach not only enables accurate prediction of PV output but also improves the visibility of system behavior. This enhanced understanding can support better operational planning, optimization of energy generation, and improved maintenance strategies.

5.6 Discussion

The results obtained from this study highlight the potential of integrating digital twin technology with machine learning algorithms for photovoltaic system monitoring. The combination of simulation-based modeling and data-driven prediction techniques allows the system to adapt to changing environmental conditions and maintain accurate performance estimation. Although the proposed framework demonstrates promising results, further improvements can be achieved by incorporating larger datasets, additional environmental parameters, and more advanced machine learning models. Future studies may also explore the integration of Internet of Things (IoT) platforms to enable fully automated real-time monitoring of large-scale solar installations. Overall, the analysis confirms that the proposed digital twin architecture provides an effective method for monitoring PV systems, predicting power generation, and identifying performance deviations in a timely manner.

VI. ADVANTAGES OF THE PROPOSED SYSTEM

The proposed digital twin framework offers several practical benefits for monitoring and managing photovoltaic (PV) systems. By combining real-time data analysis, simulation modeling, and machine learning techniques, the system improves the overall understanding and operation of solar energy installations.

6.1 Improved System Monitoring

One of the main advantages of the proposed framework is its ability to continuously observe the operational condition of the photovoltaic system. The digital twin receives updated information from the physical system and reflects these changes within the virtual model. This continuous synchronization allows operators to observe system behavior in real time and identify irregularities more easily compared to traditional monitoring methods.

6.2 Accurate Power Prediction

The integration of machine learning algorithms enables the system to estimate future PV power generation with improved accuracy. By learning patterns from historical operational data, the predictive models can determine how environmental factors such as irradiance and temperature influence energy production. Accurate forecasting of PV output is useful for planning energy management strategies and maintaining grid stability.

6.3 Early Fault Detection

The framework provides a mechanism for detecting performance deviations by comparing predicted output with actual measured power. If the difference between these values becomes significant, it may indicate abnormal operating conditions within the system. Early identification of such issues allows operators to investigate and resolve problems before they lead to major failures or energy losses.

6.4 Reduced Maintenance Costs

Traditional maintenance approaches often rely on periodic inspections or reactive repair strategies after a failure occurs. In contrast, the proposed digital twin system supports predictive maintenance by continuously analyzing system behavior. By identifying potential issues at an early stage, maintenance activities can be planned more efficiently, reducing unnecessary downtime and repair expenses.

6.5 Better Decision Support

Another benefit of the proposed system is its ability to support informed decision-making. The digital twin provides detailed information about system performance, energy production trends, and potential anomalies. These insights help operators evaluate system efficiency, identify improvement opportunities, and make better operational decisions regarding system management.

6.6 Scalability and Future Integration

The architecture of the digital twin framework is flexible and can be adapted for larger photovoltaic installations or solar farms. The system can also be integrated with emerging technologies such as Internet of Things (IoT) platforms, cloud-based monitoring systems, and smart grid infrastructures. This adaptability makes the approach suitable for future renewable energy management systems.

6.7 Enhanced System Reliability

By combining predictive analysis with real-time monitoring, the proposed framework contributes to improved reliability of PV systems. Continuous evaluation of system behavior helps maintain stable operation and ensures that energy generation remains close to expected levels. This reliability is particularly important for renewable energy systems that operate under variable environmental conditions.

VII. FUTURE SCOPE

Although the proposed digital twin framework demonstrates promising capability in monitoring and predicting the performance of photovoltaic systems, several opportunities exist for further improvement and expansion. Future developments can enhance the accuracy, functionality, and practical applicability of the system. One potential direction is the incorporation of more advanced artificial intelligence techniques. Emerging deep learning models and hybrid algorithms can improve prediction accuracy by capturing more complex relationships between environmental variables and PV system output. By training these models with larger and more diverse datasets, the digital twin can provide more reliable forecasts under varying weather conditions. Another area for future research involves integrating Internet of Things (IoT) technologies with the digital twin framework. IoT-enabled sensors can continuously collect detailed operational data from PV installations and transmit this information to cloud-based platforms. Such integration would enable real-time system monitoring from remote locations and improve the responsiveness of the digital twin model. The framework may also be extended to include additional system parameters that influence photovoltaic performance. Factors such as panel aging, dust accumulation, humidity levels, and shading conditions can significantly affect energy generation. Incorporating these variables into the digital twin model could provide a more comprehensive representation of the physical system. Further research can also explore the application of the digital twin concept to large-scale solar farms and grid-connected renewable energy networks. In such environments, the digital twin could assist in optimizing energy production, coordinating multiple energy sources, and supporting smart grid operations. This would allow better management of distributed renewable energy resources. Another possible improvement is the development of automated decision-support mechanisms. By combining predictive analysis with control strategies, the digital twin system could automatically recommend maintenance actions or operational adjustments when abnormal conditions are detected. Such capabilities would reduce the need for manual intervention and enhance the overall efficiency of PV system management. In addition, the use of cloud computing and edge computing platforms may further strengthen the performance of digital twin systems. Cloud-based environments provide high computational capacity for processing large volumes of data, while edge devices enable faster local analysis and response. Overall, future work can focus on expanding the digital twin framework with advanced data analytics, improved sensing technologies, and intelligent control strategies. These enhancements would further increase the reliability, efficiency, and sustainability of photovoltaic energy systems.

VIII. CONCLUSION

This study presented a digital twin–based framework for monitoring and analyzing the performance of photovoltaic (PV) systems. The approach establishes a virtual representation of the physical PV installation that operates using real-time operational data and predictive algorithms. By integrating simulation models with machine learning techniques, the framework enables continuous observation of system behavior and estimation of expected power generation. Environmental parameters such as solar irradiance and temperature, along with electrical measurements including voltage and current, were utilized to develop predictive models capable of estimating PV output power. Machine learning algorithms were employed to capture the relationship between these variables and the energy generation characteristics of the system. The predictions generated by these models were compared with actual system measurements in order to evaluate operational performance. The analysis demonstrated that the digital twin model can effectively represent the behavior of the physical PV system under varying environmental conditions. The comparison between predicted and measured values allows the identification of deviations that may indicate performance degradation or potential faults. This capability provides a useful mechanism for early fault detection and supports predictive maintenance strategies. Another important outcome of the proposed framework is its ability to enhance system visibility and operational understanding. Continuous monitoring and predictive analysis provide valuable insights into energy production trends and system efficiency. These insights can assist system operators in making informed decisions regarding maintenance planning and operational optimization.

VIII. ACKNOWLEDGMENT

We express our sincere gratitude to the Federal Institute of Science and Technology (FISAT), Angamaly, Kerala, India, for providing the facilities and resources to conduct this research. Special thanks to the Department of Electrical and Electronics Engineering for their guidance and support throughout this study. I extend my heartfelt appreciation to Dr. Rinku Scaria for her mentorship and constructive feedback, which were instrumental in shaping this paper. We also acknowledge the collaboration and insightful discussions with our colleagues and peers, which enriched the scope of this research. Finally, we thank our families and friends for their unwavering encouragement and support during this endeavor.

REFERENCES

[1] M. Grieves and J. Vickers, "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems," in Transdisciplinary Perspectives on Complex Systems, Springer, 2017, pp. 85–113.

[2] S. Boschert and R. Rosen, "Digital twin—The simulation aspect," in Mechatronic Futures, Springer, 2016, pp. 59–74.

[3] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. Martinez-de-Pison, and F. Antonanzas Torres, "Review of photovoltaic power forecasting," Solar Energy, vol. 136, pp. 78–111, 2016.

[4] Y. Liu, S. Qin, and Y. Zhang, "Application of machine learning techniques in photovoltaic power prediction: A review," Renewable and Sustainable Energy Reviews, vol. 136, 2021.

[5] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[6] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[7] H. T. Nguyen and J. B. Song, "Solar power forecasting using deep learning approaches: A review," Energies, vol. 12, no. 5, 2019.

[8] A. Mellit and S. Kalogirou, "Artificial intelligence techniques for photovoltaic applications: A review," Progress in Energy and Combustion Science, vol. 34, no. 5, pp. 574–632, 2008.

[9] T. Tao, Q. Qi, A. Liu, and A. Kusiak, "Data-driven smart manufacturing," Journal of Manufacturing Systems, vol. 48, pp. 157–169, 2018.

[10] R. Isermann, "Fault-diagnosis systems: An introduction from fault detection to fault tolerance," Springer, 2006.

[11] M. A. Hossain, H. R. Pota, and M. A. Mahmud, "Prediction of photovoltaic power generation using machine learning algorithms," IEEE Access, vol. 7, pp. 94650–94663, 2019.

[12] A. Kusiak and H. Zheng, "Prediction of wind farm power ramp rates: A data-mining approach," Journal of Solar Energy Engineering, vol. 131, no. 3, 2009.

[13] P. Pinson, "Very-short-term probabilistic forecasting of wind power with generalized logit–normal distributions," Journal of the Royal Statistical Society, vol. 175, pp. 147–169, 2012.
[14] S. Yang, P. Jirutitijaroen, and W. Walsh, "Hourly solar irradiance time series forecasting using cloud cover index," Solar Energy, vol. 86, no. 12, pp. 3531–3543, 2012.
[15] A. Alzahrani, F. Shamsi, and M. Ferdowsi, "Photovoltaic system monitoring and fault detection using machine learning techniques," Renewable Energy, vol. 145, pp. 1711–1721, 2020.
[16] Q. Qi and F. Tao, "Digital twin and big data towards smart manufacturing and Industry 4.0," Engineering, vol. 5, no. 4, pp. 653–661, 2019.
[17] J. Kennedy and R. Eberhart, "Particle swarm optimization," in Proceedings of the IEEE International Conference on Neural Networks, 1995, pp. 1942–1948.
[18] N. Lu, Y. Wang, and L. Wu, "Solar power prediction using machine learning algorithms," Energy Procedia, vol. 158, pp. 654–659, 2019.
[19] M. Paolanti, M. Pietrini, and A. Mancini, "Machine learning approach for photovoltaic power prediction," Energy Reports, vol. 6, pp. 273–279, 2020.
[20] F. Tao and M. Zhang, "Digital twin shop-floor: A new shop-floor paradigm towards smart manufacturing," IEEE Access, vol. 5, pp. 20418–20427, 2017.