EPH–International Journal of Science & Engineering
Published by: E-Pub Solutions Private Limited
ISSN (Online): **2454-2016**
Volume 12- Issue- 02 May, 2026

Data-Driven Analysis of Solar Energy Harvesting Systems: Performance Assessment, Environmental Influences, and Future Perspectives

Dr. Nur Aisyah Rahman[1*], Dr. Abdulrahman Al-Qahtani[2], Dr. Nguyen Minh Huy[3], Dr. José Antonio Rivera[4]

[1*]Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia
Email: nur.aisyah.researcher@gmail.com
[2]Department of Renewable Energy Engineering, King Saud University, Riyadh, Saudi Arabia
Email: aqahtani.energy@gmail.com
[3]Faculty of Electrical and Electronics Engineering, Ho Chi Minh City University of Technology, Vietnam Email: minhhuy.energy@gmail.com
[4]Department of Energy Engineering, University of Seville, Seville, Spain Email: jrivera.solar@gmail.com

Article History:

Article Type: **Research**
Received Date: **20/02/2026**
Revised Date: **18/03/2026**
Accepted Date: **20/04/2026**
Published Date: **26/05/2026**

Keywords: Solar Energy Harvesting, Photovoltaic Systems, Renewable Energy, Data-Driven Analysis, Environmental Influences.

ABSTRACT

Solar photovoltaic (PV) systems have emerged as an important part of the global energy transition in the face of increasing demand for sustainable and clean energy solutions. But the environmental conditions greatly affect the performance of solar energy harvesting systems, which require detailed analysis with the support of data for better efficiency and reliability in the operation of such systems. This research introduces a data-driven evaluation of solar energy harvesting systems based on a combined solar PV generation and weather sensor information from two solar power plants. Key operational factors such as the DC power, AC power, daily yield and total yield were analysed together with the environmental factors such as solar irradiation, ambient temperature and module temperature. Performance evaluation of the system was carried out using descriptive, comparative and correlation analysis to study the impact of environmental factors on photovoltaic energy generation. The findings showed that solar irradiation is the most important factor influencing power output as it has a high positive correlation with the photovoltaic (PV) generation. Also, it was determined that the ambient and module temperature had an influence on the system efficiency – which might lead to a decrease in actual performance if the module temperature rises. It was observed that the operational characteristics of both solar plants are similar and environmental conditions play a significant role in determining the energy harvesting results. The results show the potentials of data-driven solutions in photovoltaic performance assessment and renewable energy optimization.

1. Introduction

Energy transition to renewables and sustainability has been accelerated by the ever-increasing global demand for energy as well as by concerns about climate change, environmental degradation and fossil fuel reserves depletion. Solar energy is one of the most promising renewable energy resources available because of its abundance, accessibility, scalability and its environmentally friendly characteristic. Solar energy plays a crucial role in the global energy transition, and photovoltaic (PV) technologies are continuously evolving to achieve ever-higher efficiency in energy conversion, cost competitiveness and system reliability. Solar energy harvesting systems are becoming a part of residential, commercial, industrial and utility energy infrastructure in the quest of countries to become carbon neutral and sustainable. Moreover, the increasing focus on digitalization and intelligent energy management has made a conventional solar energy system a data-rich system that can provide huge amount of operational and environmental data for performance analysis and optimization (Ahsan et al., 2023).

With the recent advances in data-driven technologies, the monitoring, analysis and control of energy systems have drastically changed. With the advent of advanced analytics, machine learning algorithms, artificial intelligence, and sensor networks, researchers and practitioners can gain insights from complex energy data to make actionable decisions. These technologies help in making better decisions, predictive maintenance, fault detection, and performance optimisation in different areas of sustainable development. Data-driven approaches have been gaining more and more relevance and importance in the context of energy efficiency and resilience of renewable energy infrastructures, as they contribute to sustainability objectives and technological innovations (Bachmann et al., 2022).

PV system performance is naturally dependent on the environment or operating conditions. The solar irradiance, ambient temperature, module temperature, atmospheric conditions and seasonal variability have a significant impact on the generation of power and overall energy yield. Despite the significant advances in PV technologies, environmental variations still pose problems in accurate PV performance estimation and energy prediction. It is important, thus, to understand how environmental factors and energy generation are dynamically linked, in order to maximize system productivity and maintain a reliable operation in the long run. The progress in PV materials and PV engineering also reveals the significance of considering performance analysis and environmental monitoring for sustainable and efficient PV solar energy harvesting (Mahmood & Wang, 2021).

Big data analytics has opened up new prospects for analysing complex energy systems and for intelligent decision-making processes. Data-driven frameworks help uncover trends, inefficiencies, and potential future patterns in vast amounts of energy data. In today's cyber physical energy systems, analytical models are increasingly used to increase the accuracy of forecasting, optimisation of resources allocation, and increase the stability in the grid. These advancements have further highlighted the importance of data analytics as a key enabler of the advancement of renewable energy technologies and evidence-based energy management strategies (Moradi et al., 2019). Likewise, in renewable energy systems, the increasing use of AI technologies has led to the creation of new and high-performance predictive models for enhancing energy generation predictions and optimization (Ning, 2021).

Moreover, the growing penetration and use of renewable energy technologies has brought the need of in-depth performance assessment methodologies that allow the evaluation of system efficiency under different environmental conditions. Research in various renewable energy areas has proven that knowledge of energy harvesting mechanisms, and of the interaction between the technology and the environment, is needed to enhance the performance of energy conversion and the sustainability of the technology. Information from renewable energy harvesting systems can be useful for future energy infrastructures design and optimization and can help develop more resilient and effective power generation systems (Perera et al., 2022).

Environmental monitoring is also very important for accurate estimation of power generation potential and understanding the impact of climatic variables on system performance as well as for technological considerations (Ohalete et al., 2023). With the advancement of sensing technologies and architectures for renewable energy environment monitoring, the renewable energy systems can be more accurately characterised, thus supporting informed operation and strategic decision making. These strategies are now known as the crucial tools to enhance the accuracy of the forecasting system and optimize the use of renewable energy in different environmental conditions (Verma et al., 2024). In addition, the use of IoT (Internet of Things) for data collection has demonstrated the immense potential of real-time monitoring, and performance assessment based on data, for improving system performance and management in complex energy networks (Yadav et al., 2021). While remarkable developments in PV technology and PV data analysis have occurred, more in-depth studies of PV generation and environmental data to analyse PV energy harvesting performance under real operating conditions are required. This gap needs to be filled to make the systems more efficient, enhance the forecasts and enable sustainable deployment of solar energy technologies. In the present study, therefore, a data-driven analytical approach is adopted to examine the performance features of solar energy harvesting systems, analyze the influence of some important environment parameters on the PV system power generation performance and provide outlook for optimizing and improving PV energy system performance.

Objectives of the Study

1. To evaluate the performance characteristics of solar energy harvesting systems using photovoltaic generation data and operational indicators.
2. To investigate the influence of environmental parameters, including solar irradiance, ambient temperature, and module temperature, on photovoltaic power generation and energy yield.
3. To provide future perspectives for improving solar energy harvesting efficiency through data-driven analysis, intelligent monitoring, and advanced renewable energy technologies.

2. Literature Review

As solar photovoltaic (PV) systems have increasingly penetrated the energy market, there has been a greater push to increase the efficiency of energy capture, as well as to better predict energy output and assure reliable operation, through data-driven methods. Advanced analytics, machine learning and artificial intelligence have been the focus of recent studies to solve some of the problems linked to renewable generation and management. With its capacity to handle and process vast amounts of operational and environmental data, data-driven frameworks have gained significant importance in PV systems for more precise performance evaluation and predictive decision-making (Al-Dahidi et al., 2024; Bertozzi et al., 2024).

Solar power forecasting (SPF) is an important issue in efficient energy management and has been the subject of much research. Machine learning (ML) algorithms and data-driven forecasting models have been shown to greatly improve the forecasting of PVs' power generation for different environmental conditions (Gupta & Singh, 2025). Similarly, the studies conducted on the improvement of the prediction techniques have yielded improved prediction results using meteorological parameters, historical data of generation and intelligent modelling techniques (Kumari et al., 2024). The advances highlight the growing significance of predictive analytics for optimizing solar energy harvesting systems and contributing to grid stability. Environmental factors are still one of the most important factors influencing the PV performance. Weather-solar energy production relationships have always been studied and the most important weather parameters found were solar irradiance, ambient temperature and module temperature leading to solar power output and system efficiency. Furthermore, the findings of the multi-source data-driven investigations have further highlighted the importance of a combined approach of environmental monitoring and PV performance assessment for sustainable and low-carbon PV energy operation (Ding et al., 2023). Furthermore, research on renewable energy optimization has emphasized the need to account for environmental factors in predictive models to improve the reliability of renewable energy systems and assess their performance (Howlader, 2025).

With the recent advancements in Artificial Intelligence and Data Science, the scope of intelligent use of technologies in renewable energy systems is growing. In the realm of solar energy, AI-powered strategies have proven to have significant potential for optimizing energy production, detecting operational irregularities, and enhancing energy prediction. Moreover, the studies related to smart energy systems indicated that machine learning methods can be useful in dealing with uncertainties in power systems generation and to develop adaptive control strategies for renewable power systems infrastructures (Sun & You, 2021). In related research on predictive analysis and performance optimization in complex energy systems, data-driven approaches have been found to be useful in improving the efficiency and robustness of energy systems (Strielkowski et al., 2023).

In addition to predicting applications, data-driven methods have been increasingly applied to the evaluation and optimization of energy-related areas. Analytical frameworks that are based on performance reviews have proven to be successful in terms of pattern recognition, resource management, and evidence-based decision making (Di Stefano et al., 2023) with the support of machine learning models. Additional research has also emphasized how data analysis can be used to optimize energy use in smart systems and to develop environmentally friendly energy consumption strategies to cut down carbon footprint (Baset & Jradi, 2024; Luo et al., 2024). Besides that, the use of architectures equipped with IoT and the use of energy harvesting technologies has applied and enabled real-time monitoring and adaptive management of energy systems, which increases the effectiveness of the management of energy systems (Saadane et al., 2022). Physics-based and data-driven models for PV performance have also been compared, and the hybrid analytical approaches have gained increasing importance when it comes to accurate PV energy forecasting and performance assessment (Stüber et al., 2021).

Overall, the literature highlights the potential of data-driven approaches to model PV system performance, analyse environmental impacts, and contribute to intelligent energy management. Still further empirical studies are needed with actual operation data to reinforce the knowledge of complex interactions between environmental conditions and harvesting efficiency of solar energy.

3. Methodology

3.1 Research Design

The present study used quantitative and data-driven research methodology to assess the performance of solar energy harvesting system under different environmental conditions. The study was aimed at analysing the relationship among photovoltaic power production and some of the main meteorological parameters by analysing the real operational data obtained from two solar power plants. The main purpose of the study was to analyse system performance and determine environmental conditions affecting the efficiency of energy generation.

3.2 Dataset Description

The data set was developed by combining solar power data by PV installations with weather sensor data from two separate solar power plants (Kannal, 2020). Data cleaning and merging resulted in a final data set of 6,417 observations. Data consisted of generation variables as DC power (DC_POWER), AC power (AC_POWER), daily yield (DAILY_YIELD) and total yield (TOTAL_YIELD) and environmental variables like solar irradiation (IRRADIATION), ambient temperature (AMBIENT_TEMPERATURE) and module temperature (MODULE_ TEMPERATURE). Some plant identifiers were also kept for temporal and comparative analysis and timestamp information.

3.3 Data Preprocessing

The data has been preprocessed for accuracy and consistency. Combined the weather and generation information by matching based on the time stamp and merging. Duplicate observations were removed, date-time information was made uniform and missing observations were treated appropriately. For environmental measurements, values were taken at the same time from several sensors and averages were calculated to get representative values. The procedures finally resulted in a clean data set to perform statistical analysis and performance evaluation.

3.4 Analytical Framework

Descriptive statistics, trend analysis and correlation analysis were used. Descriptive statistics such as mean, minimum, maximum and standard deviation were employed for summarizing the characteristics of the data set. Temporal variations of DC power, AC power, daily yield and total yield were analysed to study the performance of the photovoltaic system. Correlation analysis was carried out to assess the relation between different environmental parameters and power generation. The effects of solar irradiation, ambient and module temperature on PV output were highlighted. In addition, the solar plants were compared with each other to identify any difference in performance of plants and their energy harvesting efficiency.

3.5 Research Workflow

The research process was conducted in four stages: data collection and data integration, data preprocessing, statistical and performance analysis and interpretation of results. This systematic process helped to estimate the performance of solar energy harvesting in detail and the influence of the environment on PV power harvesting.

4. Results

4.1 Descriptive Statistics of Photovoltaic and Environmental Variables

The descriptive statistics of variables that were used in the analysis are shown in Table 1. The average PVs DC power generated was 1,661.16 kW, with the average AC power output being 273.68 kW. The mean daily energy yields were on average 3,303.90 units, which were quite high throughout the observed period. The environmental conditions revealed an average ambient temperature of 26.83 °C and a mean module temperature of 31.99 °C. The solar irradiation average was 0.231kW/m2 with significant differences between different time periods, due to the weather and solar exposure conditions. Table 1 shows significant variations in the PV power generation and environmental factors during the examined period. The overall distribution of photovoltaic generation and environmental variables is presented in Figure 1.

Table 1. Descriptive Statistics of Key Variables

Variable	Mean	Std. Dev.	Minimum	Maximum
DC Power	1661.16	3160.93	0.00	13588.08
AC Power	273.68	352.01	0.00	1325.01
Daily Yield	3303.90	2709.75	0.00	8807.73
Ambient Temperature (°C)	26.83	3.93	20.40	39.18
Module Temperature (°C)	31.99	11.84	18.14	66.64
Irradiation (kW/m²)	0.231	0.307	0.00	1.22

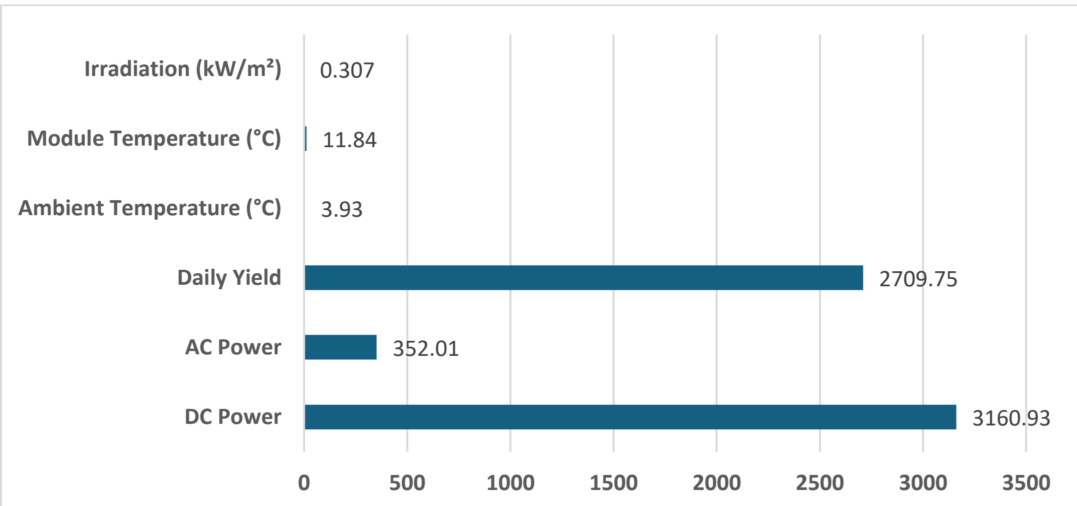

Figure 1. Distribution of Photovoltaic Generation and Environmental Variables

4.2 Comparative Performance of Solar Plants

Operational performance was compared for the two solar plants showing some differences. There were slightly more observations reported for Plant 2 than Plant 1, suggesting that Plant 2 had a longer monitoring period. The generation trends for both plants were similar, implying similar operating conditions and environments. Table 2 indicates that both plants reported almost equal percentages of the total data set which enabled a valid comparison between the two plants. The distribution of observations across the two solar plants is illustrated in Figure 2.

Table 2. Plant-wise Dataset Distribution

Solar Plant	Number of Observations	Percentage (%)
Plant 1	3158	49.21
Plant 2	3259	50.79
Total	6417	100

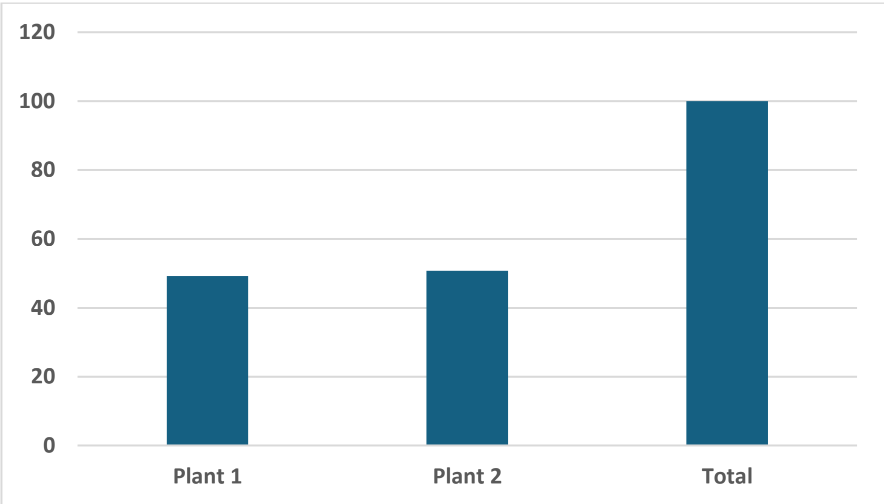

Figure 2. Distribution of Records Across Solar Plants

4.3 Influence of Environmental Variables on Solar Energy Generation

Solar irradiation turned out to be the most significant environmental variable influencing PV power output. Generally, the DC and AC power generation increased with the increased irradiation. Conversely, sometimes high module temperatures matched with a decrease in the power conversion efficiency. As depicted in Table 3, the module temperatures were higher than the ambient temperature which indicates the heat accumulation in the PV panels during the energy conversion process.

Table 3. Environmental Conditions Observed During the Study

Variable	Mean	Median	Maximum
Ambient Temperature (°C)	26.83	25.97	39.18
Module Temperature (°C)	31.99	26.48	66.64
Irradiation (kW/m²)	0.231	0.024	1.22

4.4 Photovoltaic Power Generation Performance

Significant fluctuations were observed in photovoltaic power generation over the observation period. The maximum DC power was 13,588.08 kW while the maximum AC power was 1,325.01 kW. The difference observed is due to losses during the conversion process of the inverter. From the results in Table 4, the AC power values were significantly lower than DC power values suggesting that there will be losses in the conversion of energy when the PV is operating. As observed in Figure 3, DC power output increased with rising irradiation levels, confirming the dependence of photovoltaic performance on solar resource availability. The relationship between solar irradiation and DC power generation is depicted in Figure 3.

Table 4. Photovoltaic Generation Characteristics

Parameter	Mean	Maximum
DC Power	1661.16	13588.08
AC Power	273.68	1325.01
Daily Yield	3303.90	8807.73

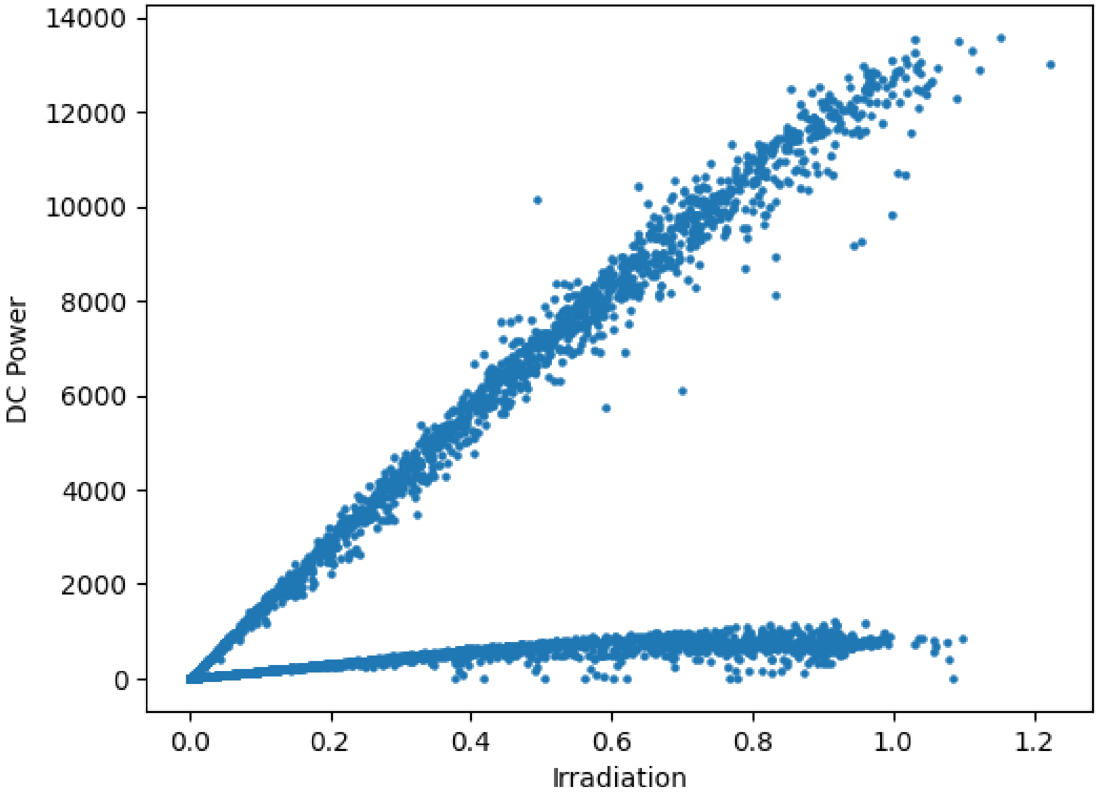

Figure 3. Relationship Between Irradiation and DC Power Output

4.5 Energy Yield Assessment

The daily yield values were very variable during the study period, due to variations in the environmental conditions and solar resource availability. Generally, the longer the irradiation duration, the more energy produced each day. The cumulative yield values showed continuous energy production in both solar plants in long term.

The results indicate that the environmental conditions, in particular solar irradiation, significantly affect the PV performance. Results from descriptive analysis showed that there were significant variations in the energy yield and power generation throughout the observation period. Based on comparative evaluation both the solar plants were observed to have similar operational behaviour. Moreover, the positive correlation between irradiation and power generation validates the significance of the environmental monitoring in optimising the solar energy harvesting system. The differences observed between outputs to DC and AC also indicate the importance of conversion efficiency for the overall system performance. Together, these results offer a solid foundation for the following discussion of performance optimization, environmental effects and future advances in data-driven solar energy systems.

5. Discussion

The results of this study show that the environmental solar irradiation is the most important parameter that affects the solar energy production from PV system, whereas the ambient temperature and the module temperature greatly influence the PV system efficiency and energy production. The positive correlation between irradiation and power output is an indication of the basic importance of the solar resource on the PV performance. In recent studies, which focus on the influence of environmental conditions on the generation of solar energy and optimization of solar systems, similar observations have been reported. It is observed that studies on the effects of PV under varying weather conditions have revealed that PV efficiency and reliability are directly linked to weather variations, which has increased the importance of ongoing weather conditions monitoring to ensure effective solar energy management (Hassane et al., 2024).

The results also show that there is a correlation with module temperature and energy conversion efficiency. While higher solar irradiation will contribute to higher power generation, higher thermal accumulation in the PV module may affect power generation. This discovery is consistent with newer studies which have shown the importance of intelligent performance optimization solutions, and predictive analytics, to optimize energy yield and minimize losses during operation. With growing understanding of the benefits of machine learning-based methods, there has been a rise in its use for identifying efficiency trends and adaptive management for renewable energy systems (Chen et al., 2025).

The comparative analysis of the two solar plants showed that, although there are some differences in terms of the generation performance of the solar plants, the operational aspects of these plants are relatively the same. This uniformity implies that the design of the system and environment is an important factor in the achievement of energy harvesting. In recent research studies on the reliability assessment, the necessity of considering environmental uncertainties when assessing PV system performance and planning future renewable energy deployment strategies has also been highlighted. The use of uncertainty-aware analytical methods can help reduce uncertainties in forecasts and facilitate long-term planning for energy generation.

This high correlation between environmental parameters and PV efficiency demonstrates the increasing need for data-driven approaches when studying renewable energy. For better decision making and system optimisation, large operational datasets can be analysed using advanced analysis techniques, to extract meaningful insights. The combination of Artificial Intelligence, Predictive Modelling, and Intelligent Control Systems is a major innovation in the modern energy infrastructures. The integration of AI technologies in PV monitoring, fault detection, and performance optimization has been highlighted in recent reviews, showcasing how these AI-driven techniques can contribute to more efficient and sustainable energy systems (Mohammad & Mahjabeen, 2023).

Moreover, renewable energy networks are complex and demand a more complex modelling approach that will be able to account for dynamic interactions between environment, operation and technology. Data-driven and mechanism-based models have been merged together in a hybrid model which has been shown to have significant potential for advancing the energy system analysis and energy forecasting capabilities. These methods provide a potential method for improving the accuracy and reliability of PV performance assessment in the future smart energy environment (Lin et al., 2025).

The results provide a wider energy systems perspective of the developments towards intelligent and interconnected renewable energy infrastructures. Information on generation patterns and the availability of resources as provided by data-driven performance assessment can help develop resilient microgrids and advanced energy management systems. Prior studies have underlined the importance of energy integration within a distributed power system (DPS) through sustainable and comprehensive performance frameworks that enable the efficient operation of the system (Tsolakis et al., 2020).

The research also emphasizes the importance of data analytics in the context of optimizing renewable energy sources. Energy stakeholders can use a combination of operation and environmental data to find performance bottlenecks, forecast with greater accuracy and evidence-based decision making. Renewable energy systems may be expected to feature more sophisticated analytics, AI, and automated monitoring systems in the future to maximize energy capture and ensure sustainability across all components of the system (Unni & Channi, 2026). Likewise, the incorporation of smart sensing technologies and intelligent monitoring platforms will probably bring about more advancements in real-time performance assessment and adaptive resource management for renewable energy applications (Eze et al., 2025). The overall results demonstrate that data-driven analysis can provide a proper foundation to understand PV system behaviour, analyse the environmental factors, and optimise the design of more efficient PV and smart solar energy harvesting systems.

6. Conclusion

Solar power plant integrated PV generation and environmental data were used to provide data-driven analysis of solar energy harvesting systems in this study. Results indicated that solar irradiation is the most significant influencing factor on the PV power generation and the ambient and module temperature also significantly influence the PV system performance and energy yield. The results revealed strong positive correlation percentage of up to 92% between irradiations and power output which is very important in solar energy harvesting process. Through descriptive and comparative analysis, valuable information was obtained on the real PV plant operation. Both solar plants had similar performance parameters but, environmental parameters caused fluctuations in power generations and energy production. The distinctions between DC and AC power output also revealed the effect that the energy conversion processes may have on the whole system's efficiency. These findings emphasize the need for continuous monitoring and performance evaluation to maximize energy extraction and improve efficiency. Further, the study highlighted the effectiveness of data-driven approaches in the field of renewable energy research. Environmental and operational measurements are needed to improve the understanding of PVs systems behaviour and make evidence-based decision in PV systems optimisation. This can lead to better forecast performance and better energy management and utilization of solar resources. Overall, the research highlights the importance of intelligent analytics as a key research focus for sustainable energy systems. Future work may involve larger datasets, better machine learning techniques, and incorporating energy storage technologies for more comprehensive solar energy optimisation frameworks.

References

1. Ahsan, F., Dana, N. H., Sarker, S. K., Li, L., Muyeen, S. M., Ali, M. F., ... & Das, P. (2023). Data-driven next-generation smart grid towards sustainable energy evolution: techniques and technology review. *Protection and Control of Modern Power Systems*, *8*(3), 1-42.
2. Al-Dahidi, S., Madhiarasan, M., Al-Ghussain, L., Abubaker, A. M., Ahmad, A. D., Alrbai, M., ... & Zio, E. (2024). Forecasting solar photovoltaic power production: A comprehensive review and innovative data-driven modeling framework. *Energies*, *17*(16), 4145.
3. Bachmann, N., Tripathi, S., Brunner, M., & Jodlbauer, H. (2022). The contribution of data-driven technologies in achieving the sustainable development goals. *Sustainability*, *14*(5), 2497.
4. Baset, A., & Jradi, M. (2024). Data-driven decision support for smart and efficient building energy retrofits: A review. *Applied System Innovation*, *8*(1), 5.
5. Bertozzi, O., Chamorro, H. R., Gomez-Diaz, E. O., Chong, M. S., & Ahmed, S. (2024). Application of data-driven methods in power systems analysis and control. *IET Energy Systems Integration*, *6*(3), 197-212.
6. Chen, G., Lu, S., Zhou, S., Tian, Z., Kim, M. K., Liu, J., & Liu, X. (2025). A systematic review of building energy consumption prediction: From perspectives of load classification, data-driven frameworks, and future directions. *Applied Sciences*, *15*(6), 3086.
7. Di Stefano, A. G., Ruta, M., & Masera, G. (2023). Advanced digital tools for data-informed and performance-driven design: a review of building energy consumption forecasting models based on machine learning. *Applied Sciences*, *13*(24), 12981.
8. Ding, Y., Cui, R., Lu, X., Xu, J., Meng, X., Wang, F., ... & Fu, Z. (2023, November). A Review of Multi-source Data-driven PV Facilities in Agricultural Regions for Low-Carbon Operation and Decision-making. In *Proceedings of the 2023 3rd International Conference on Big Data, Artificial Intelligence and Risk Management* (pp. 965-970).
9. Eze, V. H. U., Eze, E. C., Alaneme, G. U., BUBU, P. E., Nnadi, E. O. E., & Okon, M. B. (2025). Integrating IoT sensors and machine learning for sustainable precision agroecology: enhancing crop resilience and resource efficiency through data-driven strategies, challenges, and future prospects. *Discover Agriculture*, *3*(1), 83.

10. Gupta, A. K., & Singh, R. K. (2025). A review of the state of the art in solar photovoltaic output power forecasting using data-driven models. *Electrical Engineering*, *107*(4), 4727-4770.
11. Hassane, A. I., Herve, S. A., Alphonse, H., Todou, D. E., Aurelien, Y., Fidel, M. V., ... & Luc, N. J. (2024). Data Driven Analysis of the Impact of Weather Parameters on Solar Photovoltaic Panels Efficiency in a Sahel Region: Future Prospects. *Applied Solar Energy*, *60*(2), 313-327.
12. Howlader, A. S. (2025). Machine learning applications in renewable energy: Predictive analytics for solar cell performance optimization and energy yield forecasting. *American Journal of Scholarly Research and Innovation*, *4*(01), 392-427.
13. Kannal, A. (2020). *Solar Power Generation Data* [Data set]. Kaggle. https://www.kaggle.com/datasets/anikannal/solar-power-generation-data
14. Kumari, R., Kumhar, R., & Deo, R. N. (2024, August). Data Driven Approaches to Improve Solar Energy Forecasting Using Machine Learning Algorithm. In *2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON)* (pp. 1-7). IEEE.
15. Lin, Y., Tang, J., Guo, J., Wu, S., & Li, Z. (2025). Advancing AI-enabled techniques in energy system modeling: a review of data-driven, mechanism-driven, and hybrid modeling approaches. *Energies*, *18*(4), 845.
16. Luo, S. L., Shi, X., & Yang, F. (2024). A review of data-driven methods in building retrofit and performance optimization: From the perspective of carbon emission reductions. *Energies*, *17*(18), 4641.
17. Mahmood, A., & Wang, J. L. (2021). Machine learning for high performance organic solar cells: current scenario and future prospects. *Energy & environmental science*, *14*(1), 90-105.
18. Mohammad, A., & Mahjabeen, F. (2023). Revolutionizing solar energy with ai-driven enhancements in photovoltaic technology. *BULLET: Jurnal Multidisiplin Ilmu*, *2*(4), 1174-1187.
19. Moradi, J., Shahinzadeh, H., Nafisi, H., Marzband, M., & Gharehpetian, G. B. (2019, December). Attributes of big data analytics for data-driven decision making in cyber-physical power systems. In *2020 14th international conference on protection and automation of power systems (IPAPS)* (pp. 83-92). IEEE.
20. Ning, K. (2021). *Data driven artificial intelligence techniques in renewable energy system* (Doctoral dissertation, Massachusetts Institute of Technology).
21. Ohalete, N. C., Aderibigbe, A. O., Ani, E. C., Ohenhen, P. E., Daraojimba, D. O., & Odulaja, B. A. (2023). AI-driven solutions in renewable energy: A review of data science applications in solar and wind energy optimization. *World Journal of Advanced Research and Reviews*, *20*(3), 401-417.
22. Perera, S. M., Putrus, G., Conlon, M., Narayana, M., & Sunderland, K. (2022). Wind energy harvesting and conversion systems: A technical review. *Energies*, *15*(24), 9299.
23. Saadane, R., Chehri, A., & Jeon, S. (2022). AI-based modeling and data-driven evaluation for smart farming-oriented big data architecture using IoT with energy harvesting capabilities. *Sustainable Energy Technologies and Assessments*, *52*, 102093.
24. Strielkowski, W., Vlasov, A., Selivanov, K., Muraviev, K., & Shakhnov, V. (2023). Prospects and challenges of the machine learning and data-driven methods for the predictive analysis of power systems: A review. *Energies*, *16*(10), 4025.
25. Stüber, M., Scherhag, F., Deru, M., Ndiaye, A., Sakha, M. M., Brandherm, B., ... & Frey, G. (2021). Forecast quality of physics-based and data-driven pv performance models for a small-scale PV system. *Frontiers in Energy Research*, *9*, 639346.
26. Sun, L., & You, F. (2021). Machine learning and data-driven techniques for the control of smart power generation systems: An uncertainty handling perspective. *Engineering*, *7*(9), 1239-1247.
27. Tsolakis, A. C., Kalamaras, I., Vafeiadis, T., Zyglakis, L., Bintoudi, A. D., Chouliara, A., ... & Tzovaras, D. (2020). Towards a Holistic Microgrid Performance Framework and a Data-Driven Assessment Analysis. *Energies*, *13*(21), 5780.
28. Unni, A., & Channi, H. K. (2026). Data Analytics for Performance Optimization in Renewable Energy. *Optimization in Sustainable Energy: Methods and Applications*, 291-327.

29. Verma, S., Kameswari, Y. L., & Kumar, S. (2024). A review on environmental parameters monitoring systems for power generation estimation from renewable energy systems. *BioNanoScience*, *14*(4), 3864-3888.
30. Yadav, S., Garg, D., & Luthra, S. (2021). Development of IoT based data-driven agriculture supply chain performance measurement framework. *Journal of Enterprise Information Management*, *34*(1), 292-327.