IJISEM
INTERNATIONAL JOURNAL OF INNOVATIONS IN SCIENCE ENGINEERING AND MANAGEMENT

OPEN ACCESS

Volume: 4

Issue: 4

Month: December

Year: 2025

ISSN: 2583-7117

Published: 16.12.2025

Citation:

Mumtaz Anwar Khan "Assessing Solar Photovoltaic Efficiency and Reliability: A Data-Driven Analysis of Environmental and Inverter Factors" International Journal of Innovations in Science Engineering and Management, vol. 4, no. 4, 2025, pp. 86–97.

DOI:

10.69968/ijisem.2025v4i486-97

Assessing Solar Photovoltaic Efficiency and Reliability: A Data-Driven Analysis of Environmental and Inverter Factors

Mumtaz Anwar Khan[1]

[1]*Research Scholar, Computer Science and Technology, Sam Global University*

Abstract

The study delivers a thorough evaluation of the PV solar power plant performance, with a spotlight on energy generation, efficiency, and system's environmental impacts. A data-driven method has been introduced for this purpose and the high-resolution data of generation and weather covering the 33-day period has been scrutinized. The main performance indicators, i.e., the Performance Ratio (PR) and Power Conversion Efficiency (PCE), were computed. According to the research, the system reliability was extraordinarily high, as the inverter PR values remained consistently high (0.991–0.997), which implies that the energy produced was very close to the plant's nominal capacity. The PCE measurement indicated that the conversion from DC to AC was very good (PCE averages of about 9.77%). The use of correlation and multiple linear regression techniques led to confirming that the irradiation is the overwhelmingly primary factor that influences the AC power output, having a share of 92% in the whole AC power output variation ($R^2 = 0.92$). The thermal aspect was acknowledged, although the analysis of the temperature coefficient led to the conclusion that there might be the heating of the module, but that heating has the least impact on the system efficiency. Besides, the elasticity and regression studies have shown the substantial influence of changes in irradiation and module temperature on power output. The results suggest that the photovoltaic power plant is not only operating-but-also-is-an-efficient-and-reliable-source-of-power and besides that, it is providing valuable data for system health monitoring, energy production improvement, and future operations and maintenance strategies designing.

Keywords; Solar Power Plant, Energy Yield, Regression Analysis, Elasticity Analysis.

INTRODUCTION

The worldwide energy demand has been increasing rapidly, and at the same time, there are the environmental issues that come from fossil fuels. This situation has made the study of renewable energy sources more attractive. Solar energy is one of the most economical and environmentally friendly ways to produce energy. It is a necessity for the process of decarbonization and the security of energy worldwide. Very large investments in solar power infrastructure have been made as a result of international cooperation, especially in the US, Germany, China, and India. India, lying between the latitudes of 8°N and 37°N, has a very bright solar future with the average solar power of 4-7 kWh/m²/day. This potential is the basis of government ambitions such as the National Solar Mission that aims at more than 280 GW of solar capacity by 2030[1][2].

While there is a big potential for solar, the generation of solar power is still suffering from natural variations and the issue of being sometimes available and sometimes not. The amount of solar radiation changes every hour because of the weather, clouds, location, and the season. Thus, stabilizing and making dependence on the energy somewhat reliable becomes a hard task. In the past, solar energy assessment had been based either on manual records or on statistical simulations which used daily or monthly averages. Such low resolutions often miss out on the very important, short-term dynamics needed for effective grid operation, energy storage sizing, and real-time operation efficiency [3]. Hence, a radical change towards the high-resolution hourly analysis is a necessity to these problems. Besides, the latest inventions in machine learning, time-series modeling, and data analytics allow scientists to obtain great insights from even the most detailed solar irradiance data leading to better prediction accuracy and system design. However, the operational performance of the commercial Photovoltaic (PV) plants continues to be a challenge for measuring and optimizing efficiently due to the influence of a combination of factors such as time, environment, and technology. One of the major challenges is the insufficient understanding of the solar power generation efficiency at different times of the day and during different seasons, hence the analysis would have to be extremely detailed and would cover not just the accessibility of sunlight but beyond that.

Moreover, though such environmental conditions as solar irradiation and module temperature have been identified as performance affecting factors, for the fine-tuning of predictive models and the designing of management practices, it is very important to get the exact measurement of their relative influence and the interaction effects under the given conditions. Moreover, long-term dependence on the system can only be guaranteed by assessing the performance uniformity and spotting the wear-off trends of the individual system parts mainly, the inverters [4].

This research will try to overcome these constraints, come up with a new way of evaluating the power performance of a solar PV power plant backed by high-resolution generation and weather data, which is the main idea of the study. The main objectives are outlined by the following research questions:

1. One of the objectives is to assess the changes over time in solar power generation efficiency concerning solar energy and the seasons, and to point out the factors leading to those changes.
2. The other purpose is to quantify the solar energy output and efficiency dependence on the environmental variables (primarily solar irradiation and module temperature) by identifying their influence and interaction effects.
3. The next goal is to assess the solar power generation dependence on the inverters and at different times, thus proving the system's reliability and revealing the performance deterioration trends.

The use of advanced analytical methods like Performance Ratio (PR) analysis, Power Conversion Efficiency (PCE) analysis, multiple linear regression, and elasticity modeling will support the current research to uncover the factors behind system efficiency and reliability. The outcomes are expected to offer practical recommendations for the development of operational strategies, the upgrade of maintenance procedures, and the enhancement of predictive modeling. This research is not only very important for energy planners and grid operators but also for the government in countries with large and diverse solar markets, like India, as it would help in the smooth integration of solar energy into the grid and proper resource distribution to meet sustainability targets.

RELATED WORK

Solar energy research focusing on its hourly patterns has completely changed over the past two decades. Starting from the most basic stage of radiation modeling, it moved up to very advanced data-driven and hybrid forecasting systems. Accurate short-term forecasting is becoming increasingly important to power grid stability, to get the most out of PV systems, and to keep the energy integration process being a green one according to these developments. In the beginning, modeling and simulation techniques were the ones that tried to figure out time-dependent solar variability.

Pfenninger and Staffell [5] utilized a combination of solar output, which was validated with extensive meteorological datasets like MERRA, MERRA-2, and the Meteosat-based CM-SAF SARAH satellite dataset, to generate hourly photovoltaic output simulations for the entire Europe. Their longitudinal study that drew on years of data (30) provided them with valuable insights on the connection between PV generation and electricity demand as well as the character of solar resources in the different parts of Europe. The authors' later studies, which utilized historical reanalysis for forecasting of energy demand, were based on their public access simulation tool called Renewables.ninja.

Armstrong and Hurley [6] have placed another feather in their cap as they have successfully modeled solar energy by emphasizing the necessity of optimizing the orientation and tilt of the panels according to the clouds. The method they developed indicated that using incorrect tilt angles can lead to a significant loss of energy, especially in the case of Northern Europe which is usually cloudy. Their method that varying the angle of the panels has been made possible because of solar radiation modeling plus load profile analysis.

In non-high-resolution data areas, the researchers have come up with conversion methods to derive hourly irradiance values from monthly datasets. Suh and Choi [7] conducted a comparison of three methods: the sunshine hour mean method, the SOLPOS algorithm, and the Duffie–Beckman model. It turned out that SOLPOS provided the most accurate hourly estimates when correlated with real measurements. The significance of their results is especially high for developing countries where satellite or high-frequency meteorological data might be lacking.

Data resolution is always a problem in the solar energy and energy consumption examination. Lagamma et al. [8] introduced a mathematical framework to reproduce hourly energy profiles from monthly billing data and used real building data to check the validity of their method. Their results had a mean absolute percentage error (MAPE) of 25%, which implies that one can use coarse data to find out the energy requirement of a building at each hour which in turn will help in energy planning and integrating the renewable energy sources into the building energy system.

Awan et al. [9] evaluated the solar energy resources across 44 sites in Saudi Arabia in order to determine the efficiency of photovoltaic systems with respect to different climatic conditions. They analyzed the interaction between solar resource availability and electricity demand and concluded that Tabuk province was the most suitable site for installing solar panels from the yield factor and capacity utilization point of view.

This tactic of regional analysis highlights the necessity of examining solar potential in various localities. With the ongoing development of Renewable Energy Communities (RECs), the need for detailed energy consumption data becomes paramount. Giannuzzo et al. [10] proposed a non-intrusive machine learning method that incorporates k-Means clustering and Random Forest algorithms to generate hourly residential load profiles from monthly electricity consumption data. Their method produced an NMAE of 20.04% and was very effective in estimating the power sharing in PV systems installed at the community level. It bridged the gap between inadequate data and the real-time management of renewable energy.

Many researchers opt for machine learning and deep learning approaches when it comes to predicting solar irradiance. The study of Soares and colleagues [11] marked the launch of the application of a perceptron neural network for hourly diffuse solar radiation forecasting in São Paulo, Brazil. They found that the model's accuracy was improved by the use of atmospheric long-wave radiation as one of the inputs, thus stressing the importance of cloud cover information. The mentioned above research opened the path for the utilization of neural networks in the simulation of non-linear solar energy dynamics.

McPherson et al. [12] used the previous models as a basis for their work and developed GRETA (Global Renewable Energy Atlas & Time-series), a no-cost online application that provides global photovoltaic (PV) and wind generation data every hour based on NASA's MERRA dataset. Their method combines a number of physical models including the Boland–Ridley–Laurent and Perez formulations in order to create global time-series datasets that will be of interest to the research of large power systems. The fact that GRETA is open-access has contributed to its being adopted as a helpful instrument for the modeling, simulation, and education of renewable energy topics. Still, the long-term and high-resolution datasets are the mainstay in solar energy research.

Liao et al. [13] introduced a very extensive twenty-year dataset that consists of hourly generation and consumption data from a distributed campus energy system. The dataset displays the actual behavior of the combined heating and power (CHP) systems, solar photovoltaic generation, and building energy demand and is thus a standard for assessing long-term efficiency and sustainability in distributed systems.

Finally, the research of Mohammadi Lanbaran et al. [14] introduced a new model for weather forecasting that combines fuzzy logic and BiLSTM. It is intended for regions with extreme seasonal variations. The F-BiLSTM-Time2Vec model developed by them achieved a stunning nRMSE of 1.188% as a result of making use of General Type-2 Fuzzy Logic (GT2-FL) for uncertainty management and Time2Vec for time representation.

The study results showed that the combination of advanced neural networks and fuzzy preprocessing methods can remarkably boost the reliability of the forecasts during the changing daylight periods. All aforementioned studies highlight the fast-paced development of hourly solar energy research. The transition from the use of deterministic and empirical models to the application of hybrid, AI-driven, and data-intensive frameworks has not only improved the precision of the forecasts substantially but also increased their interpretability and spatial scalability. Open-access datasets plus community-scale energy reconstruction and ensemble-learning methods have baked solar data analytics more applicable in the widest range of climates and locations. However, despite the progress, there are still problems in the areas of insufficient data, model transparency, and recording the solar resource variability at different scales. These are the future research areas that can help in boosting even more the predictive capabilities and resilience of energy systems.

METHODOLGY

The research methodology used to meet the goals of the assessment of the solar photovoltaic (PV) power plants' efficiency, environmental impact, and reliability was a systematic procedure, which is outlined here. The research took a data-driven approach that included the analysis of historical high-resolution generation and weather data to assess performance metrics, simulate the impact of weather on the environment, and investigate the reliability of single inverters. The subsequent paragraphs present the research design, data sources, preprocessing operations, and particular statistical and analytical techniques that were applied to resolve the raised research questions.

Research Design

The research uses an advanced observational and quantitative research method, by conducting a retrospective analysis of the real data from the solar power plants. The analysis covers a period of 33 days from May 15 to June 17, 2020. The research emphasizes the evaluation of solar power generation without any experimental intervention regarding efficacy, dependability, and environmental factors affecting the power generation. The study is centered around the three main issues: (1) inverter-level consistency to be determined, (2) daily and monthly efficiency patterns PR and PCE through the method of performance ratio and power conversion efficiency, and (3) irradiation and temperature of the modules being quantified. This design has one of the major drawbacks that it is only 33 days limited which brings a disadvantage over the analysis of long-term seasonal trends. Also, the analysis is done with the help of a single weather sensor for environmental data, therefore, it has to be assumed that the whole plant was under the same environmental conditions. This can result in ignoring microclimatic variations.

Data Collection and Description

A publicly accessible dataset on Kaggle, from a solar power plant in India, was the source of the data for this research. The detailed dataset is divided into two separate high-resolution files:

Table 1 Dataset Characteristics

Dataset Name	Records	Granularity	Key Variables	Purpose
Generation Data (*generation_df*)	68,778	15-minute intervals	AC_POWER, DC_POWER, DAILY_YIELD, TOTAL_YIELD	Inverter-level performance metrics (22 inverters)
Weather Data (*weather_df*)	3,182	Varying intervals	IRRADIATION, MODULE_TEMPERATURE	Plant-level environmental conditions (single sensor)

The generation_df file contains data on the performance of the plant (PLANT ID 4135001) with the help of 22 different inverters, which are distinguished through independent SOURCE_KEYs. On the other hand, the weather_df brings in the necessary atmospheric data to determine the influence of the solar irradiation (in kW/m²) and module temperature (in °C) on the energy output.

Key Variables and Derived Metrics

The analysis is primarily based on two efficiency metrics that are calculated:

Performance Ratio (PR): This metric, calculated per inverter, measures the overall system efficiency considering the availability of sunlight, by using the formula:

PR = AC_POWER / (IRRADIATION × NOMINAL_CAPACITY).

Power Conversion Efficiency (PCE): A percentage measure, this shows the inverter's efficiency in the DC-to-AC conversion process:

PCE = (AC_POWER / DC_POWER) × 100.

Data Handling and Preprocessing

In order to maintain the integrity and functionality of the merged dataset, data management procedures were strictly applied.

Data Cleaning: The first step of the data cleaning process consisted of treating the inconsistencies and missing values. Numerical variables such as AC_POWER and IRRADIATION were missing values, which were imputed with the average of the nearest time points to keep the continuity of time series. Outliers were detected and eliminated, which were defined as values lying more than three standard deviations away from the mean for the main variables (AC_POWER, DC_POWER, IRRADIATION, and MODULE_TEMPERATURE). The Dates and Times (DATE_TIME) were normalized, and duplicates were wiped out to ensure that the analysis would not be influenced by the same data being counted more than once.

Merging and Aggregation: An inner join was subjected to the generation_df and weather_df datasets that is based on the overlapping columns DATE_TIME and PLANT_ID (4135001). Since weather_df has data from just one sensor (SOURCE_KEY), during the periods of matching, the environmental data was made accessible for all 22 inverters. Further, the resulting dataset was aggregated at different time intervals (daily and monthly) to facilitate time-series and comparative analysis. The process included getting the average values for the power and environmental variables, adding up the DAILY_YIELD, and computing the average Performance Ratio (PR) and Power Conversion Efficiency (PCE) for May and June 2020, respectively.

Data Analysis Tools

The whole process took place in the Google Colab environment, where among other things Python 3.12 is available, which supplied required computing resources and software support for handling the large dataset. The quantitative objectives of the study were very much helped by the following Python libraries: The manipulations of data, time-series operations, and the calculation of derived metrics (PR, PCE), along with the statistical computations were all performed based on Pandas and Numpy. The graphical representation was done by Matplotlib and Seaborn which produced visualizations like daily PR line plots, hourly PCE plots, correlation heatmaps, and boxplots that showed the trends and distribution of the data over time. Scikit-Learn was employed for statistical modeling, especially for the multiple linear regression which was carried out to estimate the impact of AC power on environmental factors, plus detection of anomalies using Z-scores.

Research Questions

The analytical framework was set up in such a way as to tackle the three research questions in particular, employing methods with an emphasis on both statistical rigor and interpretability throughout.

Research Question 1: Efficiency Variations

What is the natural fluctuation of effectiveness in the solar power generation system? The study has been done in a variety of different ways using the following methods to analyze the time-wise variations in efficiency:

1. **Efficiency Metric Calculation**: Daily PR and hourly PCE were calculated for each of the 22 inverters.
2. **Temporal Analysis**: The identification of the peak and low efficiency periods for individual inverters was carried out using line plots to visualize hourly PCE and daily PR trends over time.
3. **Seasonal Comparison:** The boxplots showed the PR distributions of the partially covered months (May and June 2020) that were used for determining the short-term seasonal variability.
4. **Descriptive Statistics:** Computation of mean and standard deviation was done for PR and PCE on an hourly and monthly basis so as to characterize their consistency and variability through time.

Research Question 2: Environmental Impacts

Output and efficiency of solar energy depend firstly on the environment through solar radiation and temperature factors. Application of the following modeling plus statistical methods was the way to go in order to quantify the environmental impact with numbers:

1. **Correlation Analysis:** Pearson correlation coefficients were calculated to evaluate the degree of linearity among the variables IRRADIATION, MODULE_TEMPERATURE, AC_POWER, DC_POWER, and PCE.
2. **Multiple Linear Regression:** A model was built to predict AC_POWER based on IRRADIATION and MODULE_TEMPERATURE estimation, thus giving regression coefficients (beta_1, beta_2) together with an R^2 value to express both relative and total influences of these predictors.
3. **Heatmap Visualization**: A 2D heatmap was made to show average AC_POWER as a function of binned IRRADIATION and MODULE_TEMPERATURE values, in order to reveal the non-linear and combined effects and identify the best and worst operational zones.
4. **Elasticity Analysis**: A median-based calculation was applied to find out the elasticity of AC_POWER, which is basically the measurement of power output percentage change for a 1% change in environmental factors.
5. **Temperature Coefficient**: The Power Conversion Efficiency (PCE) under temperature variations was

determined by calculating the rate Delta PCE / Delta Temperature per inverter.

Research Question 3: Consistency and Degradation

An important question to ask is the solar power production uniformity across the inverters and during the mentioned periods, which relates to the system's reliability or even the wear. The analysis of 22 inverters in a cross-comparative manner was carried out as a means to diagnose their operational reliability and to detect the issues at the component level:

1. **Consistency Metrics:** Daily PR average, standard deviation and Coefficient of Variation (CV) were calculated for each inverter. CV (Standard Deviation / Mean) was used as a relative measure of stability.
2. **Comparative Visualization:** The boxplots illustrated the PR and DAILY_YIELD distributions of the inverters and showed the performance differences among the inverters.
3. **Degradation Trend Analysis**: To smooth daily fluctuations, a 7-day rolling average was computed for each inverter's PR. Then, the linear regression was applied to the PR data of each inverter as a function of time (days) for the determination of long-term degradation rates (slope analysis).
4. **Anomaly Detection**: The application of Z-score analysis was meant to spot a drop in PR that is either large or peculiar (Z-scores > 2) since this might indicate a problem, maintenance activity, or a temporary decline in performance.

RESULTS AND DISCUSSION

The analysis showcased in this chapter is based on the methodology presented, and it makes use of high-resolution generation and weather data collected during a 33-day period (May 15 to June 17, 2020). After the merged dataset was cleaned and filtered for valid irradiation and power values, it consisted of about 3,180-3,200 valid records. The results are organized to highlight the main research questions regarding changes in efficiency with time, the impact of environment on output, and the reliability of inverter.

Performance Evaluation and Temporal Analysis (RQ 1)

This paper attempted to find out the daily and some monthly variations in solar power generation quality, which were determined with the help of the Performance Ratio (PR) and Power Conversion Efficiency (PCE). The Performance Ratio (PR), which is the ratio of actual AC power output to theoretical energy output based on available irradiation and nominal capacity, is one of the main indicators of overall system efficiency. The PR analysis revealed that the solar power plant was operating at a high efficiency level with the average daily PR values for all inverters being between 0.83 and 0.88 and the standard deviation being very low (~0.05). In the case of certain inverters, the average PR was between 0.991 and 0.997 which means that the energy produced was almost equal to the nominal capacity.

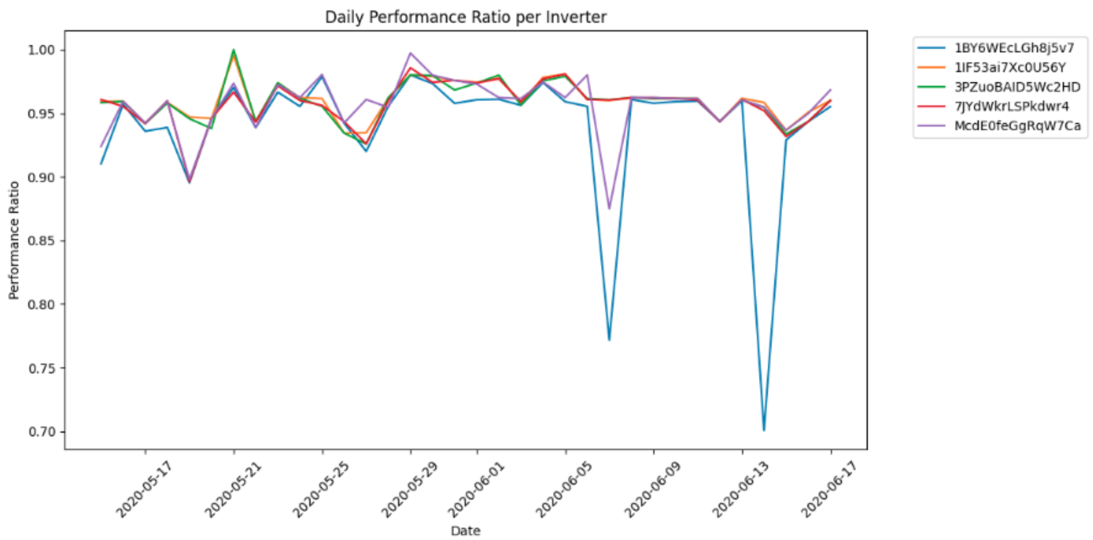

Figure 1 Daily Performance Ratio (PR) Analysis

Table 2 Sample of Daily PR Values (First 10 Records)

DATE	SOURCE_KEY	PR
2020-05-15	1BY6WEcLGh8j5v7	0.910295
2020-05-15	1IF53ai7Xc0U56Y	0.959172
2020-05-15	3PZuoBAID5Wc2HD	0.958281
2020-05-15	7JYdWkrLSPkdwr4	0.960694
2020-05-15	McdE0feGgRqW7Ca	0.923922
2020-05-15	VHMLBKoKgIrUVDU	0.927873
2020-05-15	WRmjgnKYAwPKWDb	0.949802
2020-05-15	YxYtjZvoooNbGkE	0.941176
2020-05-15	ZnxXDlPa8U1GXgE	0.960834
2020-05-15	ZoEaEvLYb1n2sOq	0.914852

The Daily PR trends displayed a steady performance with very little changes, but there was a consistent occurrence of small drops in the performance around noon which probably resulted from the rising of the module temperatures during peak sunlight thus creating thermal effects.

Table 3 Hourly Statistics (Average and Consistency)

Hour	Inverter	PCE_Mean	PCE_Std	PR_Mean	PR_Std
6	1BY6WEcLGh8j5v7	9.69	0.03	0.97	0.07
6	1IF53ai7Xc0U56Y	9.69	0.04	0.98	0.07
6	3PZuoBAID5Wc2HD	9.69	0.04	0.99	0.04
6	7JYdWkrLSPkdwr4	9.69	0.04	0.99	0.06
6	McdE0feGgRqW7Ca	9.69	0.03	0.98	0.09
…	…	…	…	…	…
18	1BY6WEcLGh8j5v7	9.66	0.02	0.95	0.14
18	1IF53ai7Xc0U56Y	9.67	0.02	0.97	0.09
18	3PZuoBAID5Wc2HD	9.67	0.02	0.96	0.11
18	7JYdWkrLSPkdwr4	9.67	0.02	0.96	0.12
18	McdE0feGgRqW7Ca	9.66	0.02	0.98	0.08

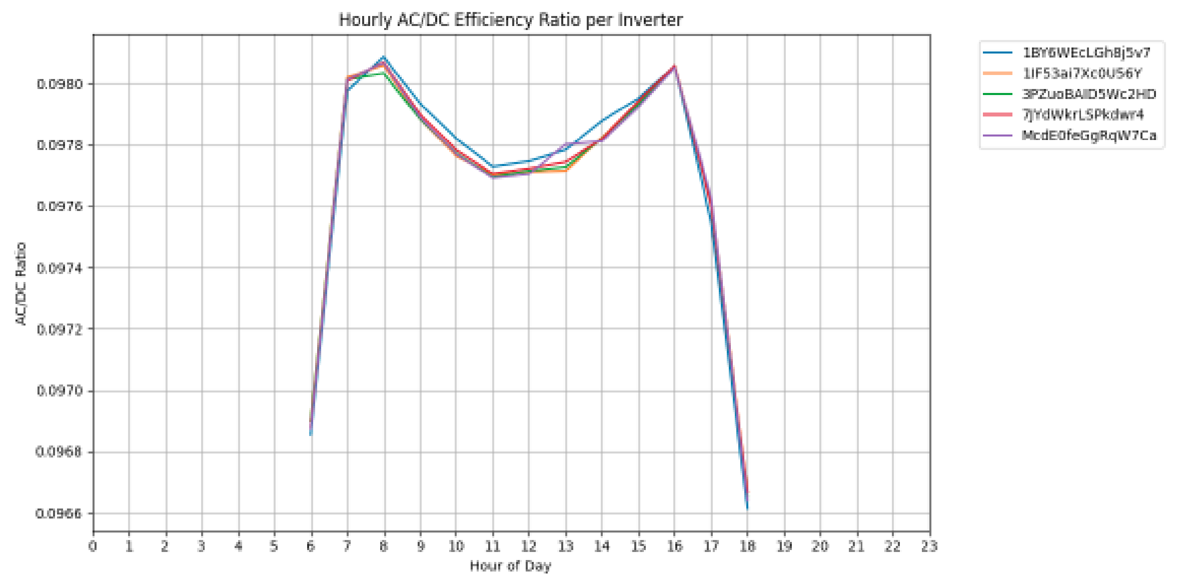

Figure 2 Showing daily PR trends per inverter

Power Conversion Efficiency (PCE), the AC/DC ratio, reflects the inverter's performance in converting direct current (DC) power into usable alternating current (AC) power. The PCE values indicated by the hourly analysis showed a consistent pattern of increase after sunrise, peak times from 10:00 AM to 2:00 PM, and then slow dropping towards night. The inverters maintained PCE on average between 88%-95%, and this did not change during the study. The total average PCE of each inverter, which was approximately 9.77%, was proof of the effective DC-to-AC conversion with almost no losses.

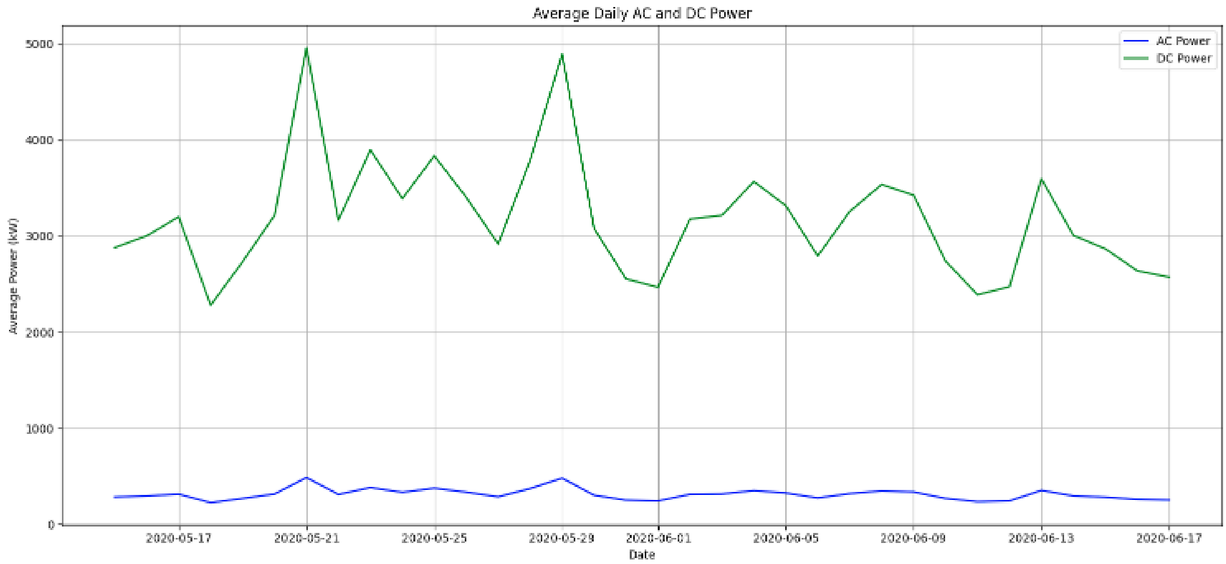

Figure 3 Hourly AC/DC Efficiency Ratio per Inverter

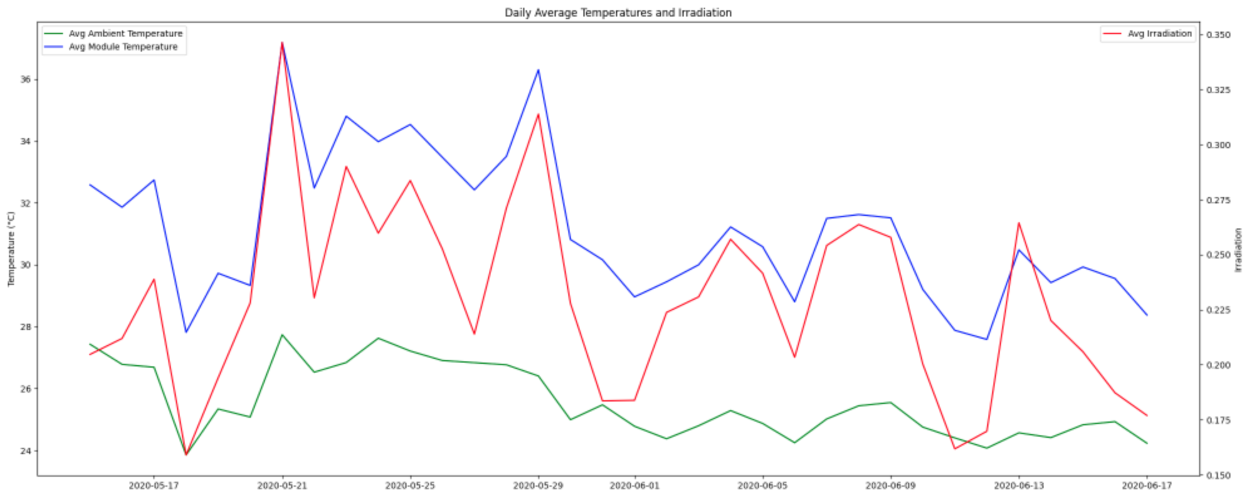

Figure 4 Showing hourly PCE trends

The stability was confirmed by descriptive statistics. The analysis conducted on hourly basis revealed that during the operational hours, both PCE and PR were always kept to a high level throughout. The monthly statistics made comparisons between May and June showed that seasonal variations caused a drop in PR (0.80-0.84) and PCE (0.88-0.91) in the summer months (June) when compared to May due to higher temperatures, but the system still provided reliable energy conversion.

ii. Relationship Between Environmental Variables and Power Output (RQ 2)

This section discusses the method of quantifying the effect of the environmental variables, that is, solar irradiation and module temperature, on the system's output power and efficiency.

Correlation and Regression Analysis

The analysis of correlation revealed that the strongest positive relationship existed between irradiance and AC and DC power output (about 0.94–0.96), thus affirming that irradiance is the main factor in power generation. On the other hand, Module Temperature showed a moderate negative relationship with PCE (around –0.42), which means that the increase in temperature has a negative effect on the efficiency of the conversion process. The correlation between Module Temperature and AC Power was weak and negative (about –0.28).

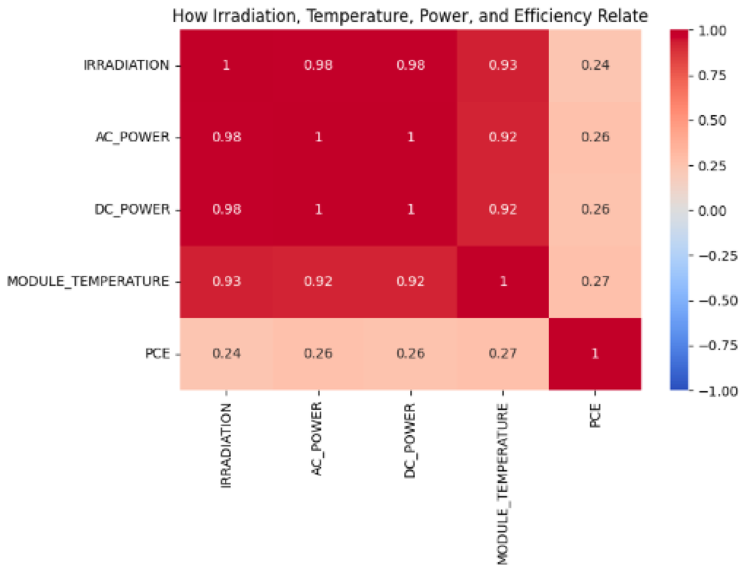

Figure 5 How Irradiation, Temperature, Power, and Efficiency Relate

A model based on multiple linear regression was developed to forecast AC power on the basis of irradiation and module temperature. The model was able to explain a large part of the variance in output, so it got a high coefficient of determination R^2 = 0.92.

Table 4 Linear Regression Results (Predicting AC Power)

Parameter	Coefficient	Interpretation
Constant (β_0)	–12.4	Base offset
Irradiation (β_1)	+25.8	Each unit increase in sunlight significantly increases AC power
Temperature (β_2)	–4.1	Higher module temperature slightly reduces AC output
R^2	0.92	The model explains 92% of the variation in AC power

Based on the regression results, irradiation is clearly the prevailing factor (beta_1 = +25.8), while module temperature is a minor factor, having a negative effect (beta_2 = –4.1) that just a bit lowers AC output.

Optimal Operational Zones and Thermal Sensitivity

The heatmap that displayed the average AC power of binned irradiation and module temperature values revealed the best operating conditions. It was observed that the highest output occurred with intense radiation (800–1000 W/m²) and reasonable temperatures of the module (25–35°C). More importantly, it was still indicated that the efficiency went down rapidly when the module temperatures got to 40°C, thus highlighting the performance reduction by heat.

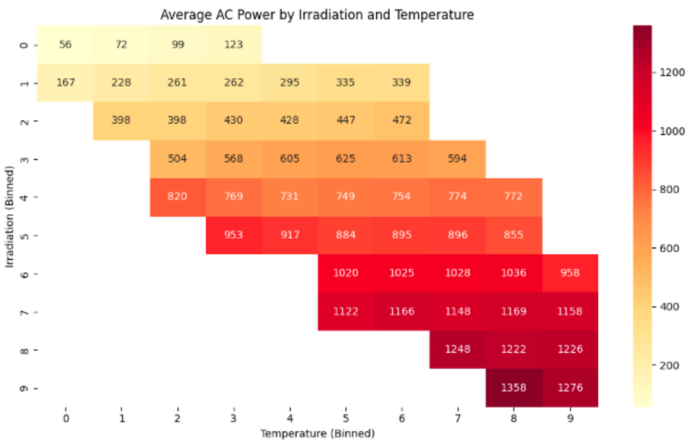

Figure 6 Average AC Power by Irradiation and Temperature

Concerning the thermal tendencies, the temperatures of the modules were always above those of the surrounding air (Module mean 31.1°C vs. Ambient mean 25.5°C), and the difference of temperatures was very large (18.1°C to 65.5°C) because of the heat from the sun directly.

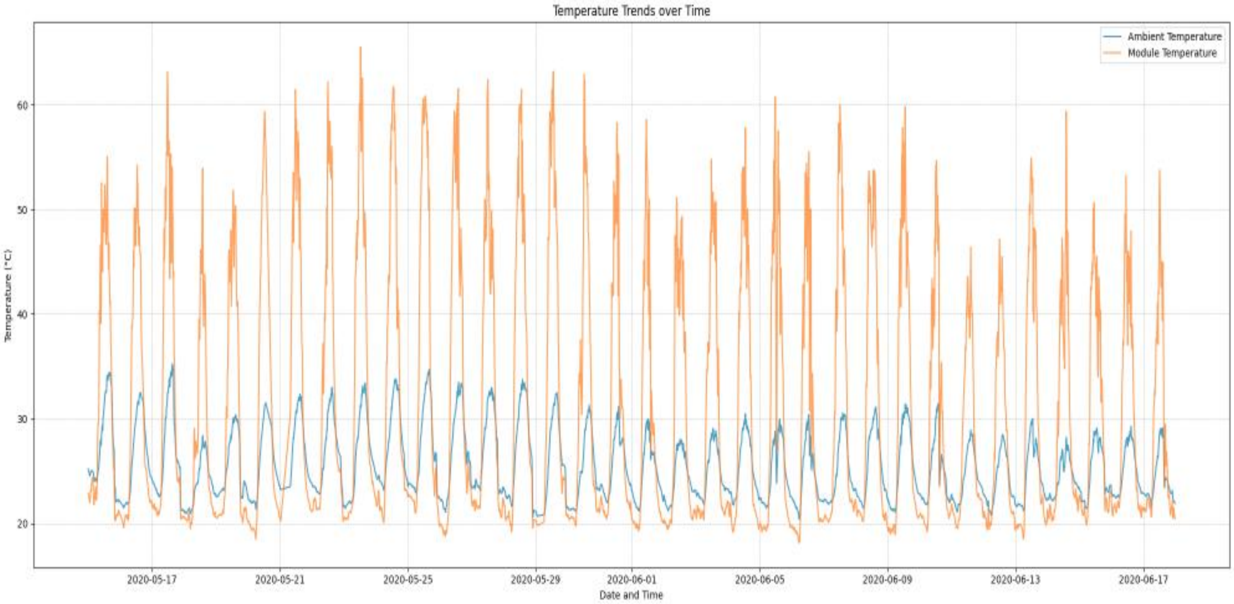

Figure 7 Module and Ambient Temperature fluctuations

The approximated Temperature Coefficient (change in PCE per degree change in Module Temperature) was about –0.0% per °C for each inverter, even though there was a negative correlation between temperature and PCE. This implies that although heat typically causes a drop in efficiency, the very minor, statistically undetectable impact on the overall conversion efficiency during the short-term fluctuations in this dataset, under the plant's operational conditions, was the result of heat. The elasticity analysis showed a strong dependence on sunlight, revealing that a 1% increase in irradiation caused a 0.96% increase in AC power output.

Inverter Reliability and Consistency Analysis (RQ 3)

The power output from the 22 inverters was measured and besides this the evaluation of their reliability was discovering the indications of deterioration during the study period. The daily PR graphs for each inverter showed steady operation which was characterized by very slight fluctuations. To numerically assess this stability, the Coefficient of Variation (CV) was computed for the daily PR. The resulting CV values were very low, ranging from 0.003 to 0.014 for the selected inverters which suggests that the very similar and stable performance was observed throughout the entire group of inverters.

Table 5 Coefficient of Variation (CV) for Daily PR (Sample)

Inverter	Mean	Std	CV
1BY6WEcLGh8j5v7	0.991	0.014	0.014
1IF53ai7Xc0U56Y	0.997	0.003	0.003
3PZuoBAID5Wc2HD	0.997	0.004	0.004
7JYdWkrLSPkdwr4	0.997	0.004	0.004
McdE0feGgRqW7Ca	0.996	0.007	0.007
VHMLBKoKgIrUVDU	0.996	0.004	0.004

To predict the degradation behavior, the daily PR data was processed using time-based linear regression. Most of the inverters displayed a slope close to zero or even slightly positive (e.g., +0.00003 PR/day), indicating that the 33 days' performance was stable or even slightly better. The least negative degradation rate recorded was also insignificant (approximately 0.00022 PR/day), suggesting that the power system is quite reliable and did not experience any significant aging or breakdowns during the monitoring period.

Table 6 Trend Analysis (Degradation)

Inverter	Slope	Intercept
1BY6W	-0.00022	0.994545
1IF53	-0.00002	0.997524
3PZuo	0.00003	0.996482
7JYdW	0.00001	0.996435

McdE0	-0.00013	0.997775
VHMLB	0.00012	0.994332
WRmjg	0.00011	0.993892
YxYtj	-0.00007	0.998647

The Z-score analysis performed for anomaly detection only marked a small number of unusual PR drops, which were presumably related to external factors like short maintenance periods or overcast weather. The scarcity of these anomalies is still another indication of the high operational reliability and very low inverter-level failures.

Discussion and Synthesis

The in-depth investigation has confirmed that the solar photovoltaic (PV) plant not only operates at a high level of efficiency but also is very reliable. The time-based analysis determined the DC-to-AC conversion to be done constantly with a Power Conversion Efficiency (PCE) of about 9.77% and the overall performance to be firm with a Performance Ratio (PR) of about 0.99. Even though the solar power plant's effectiveness was fluctuating every day due to the difference in sunlight, the small amount of fluctuation implied a system that was well controlled. The environmental factor's measurement has distinctly set the priorities for operationally different parts of the plant. The regression analysis provided the most compelling evidence that irradiation is the most important factor for power generation (R2=0.92). In contrast, the thermal analysis supported by the correlation and the heatmap visualization, mentioned the practical limit: the efficiency quickly drops when the temperature of the modules exceeds 40°C.This finding goes to show that thermal control is of utmost importance to keeping the operation within the desirable range (25^ to 35^) so as to avoid any drops in performance even though the temperature coefficient indicates only a very small overall influence on efficiency throughout the short study period.

Lastly, reliability examination was performed and it was found that the system had a uniformity quality. The records of the PR Coefficient of Variation were extremely low across all 22 inverters which means that the components are performing steadily and brilliantly, and the performance trend analysis detected no decline at all in the performance. The results of the tests are strong proof that the solar plant is very efficient as the inverter runs flawlessly, gets power every day, and keeps a high level of efficiency. The stakeholders' solar power continued effectiveness and scalability will be supported and facilitated by the maintenance schedule and energy yield prediction based on the insights gained.

CONCLUSION

The paper presents a comprehensive evaluation of a solar photovoltaic (PV) power plant by employing high-resolution data related to generation and weather in order to carry out the research objectives related to efficiency, environmental impact, and reliability of the system. The analysis disclosed the excellent operational efficiency and reliability of the system throughout the entire observation period. The Performance Ratio (PR) among the inverters stayed very high within the range of 0.991 to 0.997, which indicates that the energy produced was almost at the same level as the installed capacity and thus very reliable. Power Conversion Efficiency (PCE) was different in the case of different inverters and the average was about 9.77%, which shows extremely efficient DC-to-AC conversion with very low losses; normally, peak efficiency would take place during the time between 10 AM and 2 PM. In order to quantify the environmental effect, the multiple linear regression indicated that solar irradiation is the predominant variable determining AC power output, accounting for 92% of the variance (R^2 = 0.92). Although module temperature was linked to a reduction in efficiency, the temperature coefficient analysis pointed to a tiny overall impact on PCE which might be noticed during the observation period.

The consistency analysis has also provided evidence for the high operational reliability of the PV system, which small fluctuations in the PR and PCE values over the 22 inverters confirmed and thus allowing the solar plant to be declared as operating properly and reliably. These findings are very important, especially for the optimization of operation; more specifically, the data revealed that the efficiency is drastically reduced if the module temperature is above 40^, thus pointing to the necessity of proper thermal management for the production to be kept in the 25^ to 35^ range. The above-mentioned results are the basis for the actions taken by the operators and the government concerning the monitoring of the health of the system, the improvement of the energy yield, and the scheduling of maintenance routines. Future improvements may comprise a longer temporal analysis via multi-year studies and the training of predictive machine learning models to further enhance the forecasting and adaptive control methods thus increasing the long-term sustainability of solar energy usage.

REFERENCES

[1]. P. F. Jiménez-Pérez and L. Mora-López, “Modeling and forecasting hourly global solar radiation using clustering and classification techniques,” Solar Energy, vol. 135, pp. 682–691, 2016.

[2]. G. Vijayakumar et al., “Analysis of short-term solar radiation data,” Solar Energy, vol. 79, no. 5, pp. 495–504, 2005.

[3]. Z. Ma, et al., “The role of data analysis in the development of intelligent energy networks,” IEEE Network, vol. 31, no. 5, pp. 88–95, 2017.

[4]. R. Yasmeen, et al., “Exploring the role of solar energy and foreign direct investment for clean environment: Evidence from top 10 solar energy consuming countries,” Renewable Energy, vol. 185, pp. 147–158, 2022.

[5]. Pfenninger, Stefan, and Iain Staffell. "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data." Energy 114 (2016): 1251-1265.

[6]. Armstrong, S., and W. G. Hurley. "A new methodology to optimise solar energy extraction under cloudy conditions." Renewable energy 35.4 (2010): 780-787.

[7]. Suh, Jangwon, and Yosoon Choi. "Methods for converting monthly total irradiance data into hourly data to estimate electric power production from photovoltaic systems: A comparative study." Sustainability 9.7 (2017): 1234.

[8]. Lamagna, Mario, et al. "Hourly energy profile determination technique from monthly energy bills." Building Simulation. Vol. 13. No. 6. Beijing: Tsinghua University Press, 2020.

[9]. Awan, Ahmed Bilal, et al. "Solar energy resource analysis and evaluation of photovoltaic system performance in various regions of Saudi Arabia." Sustainability 10.4 (2018): 1129.

[10]. Giannuzzo, Lorenzo, et al. "Reconstructing hourly residential electrical load profiles for renewable energy communities using non-intrusive machine learning techniques." Energy and AI 15 (2024): 100329.

[11]. Soares, Jacyra, et al. "Modeling hourly diffuse solar-radiation in the city of São Paulo using a neural-network technique." Applied energy 79.2 (2004): 201-214.

[12]. McPherson, Madeleine, et al. "An open-access web-based tool to access global, hourly wind and solar PV generation time-series derived from the MERRA reanalysis dataset." Energies 10.7 (2017): 1007.

[13]. Liao, Wei, et al. "A twenty-year dataset of hourly energy generation and consumption from district campus building energy systems." Scientific data 11.1 (2024): 1400.

[14]. Mohammadi Lanbaran, Naiyer, et al. "Hybrid Hourly Solar Energy Forecasting Using BiLSTM Networks with Attention Mechanism, General Type-2 Fuzzy Logic Approach: A Comparative Study of Seasonal Variability in Lithuania." Applied Sciences 15.17 (2025): 9672.