Performance Analysis of a Solar Farm Capturing a Unique Real-time Performance Ratio through Data-driven Methodology

Ila Thakur[1,], Anjana Geetha Rajakumar[2], Remish Leonard Minz[2] , Nikhil Kulkarni[2], Prateek Mital[2] , Nilanjan Chakravortty[2] and Chandrasekar Venkatraman[3]*

[1,2]*Research & Development Centre, Hitachi India Pvt. Ltd., Bangalore-560001, India*

[3]*Hitachi America, Ltd., USA*

Abstract. Photovoltaic technology, a rising renewable energy source, relies heavily on irradiation and temperature for performance. Defects like shading, hot spots, and soiling can disrupt the current–voltage curve, challenging inverters to operate at maximum power point (MPP). Indirect measurement and real-time loss detection methods are essential due to the inaccessibility of PV cell circuits. The conventional methods of performance measurement consider the performance based on standard test conditions. The main focus of the present work is to develop a novel performance ratio, which captures real-time efficiency involved in Solar farm operations using data-driven methodologies. The major contributions of this work are: Monitoring health and performance of solar farms by collecting 8 months of SCADA data using IEC standard 61724; Developing a data-driven model to predict ideal I-V curve and MPP at each transient state in the absence of panel-level data; Developing a new performance ratio which provides insights into the transient operational state of the inverter and the deviation of real-time data from the MPP. The results reveal the interplay of different inverters and the evolving overall performance of solar farm over time. This underscores necessity for maintenance and highlights the potential for enhancing solar farm's output.

1 Introduction

Renewable energy adoption has been steadily increasing, comprising 30% of global final energy consumption as of 2023. The accelerated growth of the PV market can be attributed to several advantages offered by PV farms, including their stationary operation, low maintenance requirements, modular design, and high reliability. This expansion underscores the importance of continued research and analysis to effectively locate and operate PV solar farms optimally, ensuring sustainable and reliable renewable energy sources for the future.

There are various defects, like, shading, hot spots, soiling, cracks, damage etc., which can alter the current-voltage (I-V) and power–voltage (P–V) characteristics, posing challenges for the inverter to work at maximum power point (MPP) [1], [2], [3]. Given the inaccessibility of the circuit inside PV cells, there is a need for indirect measurement and loss detection methods at the real-time operating points to address the actual real-time loss [4], [5], [6]. The conventional methods of performance measurement only consider performance based on standard test conditions [7], [8]. Therefore, the main focus of the present work is to study the performance using SCADA data and IEC standard (61724) parameters and to develop a distinct performance-measuring-parameter which captures real-time efficiency and losses involved in Solar farm operation using data-driven methodologies. The study focuses on the analysis of a 20 MW PV grid-connected solar farm (expanded in around 100 acres area) installed in India in 2017. Unlike standalone systems, grid systems, which typically don't require storage, ensure continuous full-capacity generation. The SCADA data is published over MQTT (Message Queuing Telemetry Transport) using multiple sensor devices, which is received by the MQTT brokers in the cloud.

2 Standard performance parameters

This section introduces the system performance parameters from the guidelines for the assessment of photovoltaic (PV) system from the IEC standard 61724. These parameters are crucial for assessing and optimizing the performance of solar PV systems, helping stakeholders make informed decisions about system design, maintenance, and efficiency improvements.

- Array Efficiency: Array Efficiency is a measure of the ability of the solar PV array in converting sunlight into usable electrical energy. It is calculated by dividing the actual electrical output of the array by the solar energy incident on the array.
- Inverter Efficiency: Inverter Efficiency represents the ability of the inverter in converting direct current (DC) electricity generated by the solar panels into alternating current (AC) electricity for use in the grid. It is calculated by dividing the AC power output by the DC power input.
- System Efficiency: System Efficiency is an overall measure of the efficiency of the entire PV system, encompassing both the array and the inverter. It is calculated by multiplying the array efficiency and the inverter efficiency.

---

* Corresponding author: ila.thakur@hitachi.co.in

Fig. 1. Real-time visualization using Grafana dashboard for 10 days data from 30-12-2023 to 10-01-2024

- Array Yield/Efficiency: The array yield is the ratio of daily, monthly, or yearly direct current (DC) energy output from a PV array to the rated PV array power. Present work talks about the monthly yield for the duration of specific months.
- Reference Yield: The reference yield is the ratio of total in-plane solar radiation to the reference irradiance at standard test conditions (STC). It represents the total in-plane solar radiation or an equivalent number of hours at the reference irradiance.
- Final Yield: The final yield can be defined as the total AC energy during a given period divided by the rated PV array power.
- Capture Losses: Capture Losses represent the energy losses that occur due to factors such as shading, soiling, and other inefficiencies that reduce the panels' ability to capture available sunlight.
- System Losses: System Losses encompass all losses in the PV system, including losses in the array, inverter losses, and other inefficiencies. It is expressed as a percentage of the reference yield and provides an overall measure of system performance.
- Performance Ratio (PR): Performance Ratio is a key indicator of the overall performance of the photovoltaic system.

$$PR = \frac{\text{Actual Array efficiency}}{\text{Rated Array efficiency}}$$
$$= \frac{\text{Energy produced/Actual irradiance}}{\text{Rated energy at STC/Rated irriadiance}}$$

- Capacity Factor: Capacity Factor is a ratio that represents the actual energy output of PV system over a specific period compared to its maximum peak output if it operated at

$$CF= \text{Energy produced/ Rated energy at STC}$$

3 Results and discussion

3.1. Real-time visualization through Grafana dashboard

The Grafana dashboard provides real-time visualization of critical data fed through InfluxDB. Utilizing various queries, the dashboard fetches data for selected timelines, with a focus on essential fields such as 'AC power' and 'irradiance', etc. The time variation of these outputs is plotted, with minute-level granularity. The dashboard also allows users to generate plots for daily or monthly data using different commands. Fig. 1 illustrates the time variation of fields for different inverters, showcasing real-time and daily average variations of ac power at the inverter and farm levels over a 10-day span. The plots in 1(a&b) depict the real-time and daily average variations

---
* Corresponding author: ila.thakur@hitachi.co.in

of ac power at the inverter level. Plots 1(c) showcase real-time ac power generation for the entire farm, aggregating data from all 15 inverters. Irradiance, measuring incident radiation per square meter, follows the sun's cycle, peaking around 1000 W/m² in the afternoon and returning to zero at sunset (Fig. 1(d)). This variation aligns with the one-on-one link between incident irradiance on solar panels and power generation. Cloudy days exhibit lower irradiance, leading to reduced dc power generation. Higher afternoon temperatures impact dc power

generation rates, as solar panel efficiency decreases with temperature rise.

Monitoring efficiency and temperature shows that efficiency is higher for the days when the temperature is lower and vice versa. The bar chart 1(e) quantifies the daily average incident energy versus dc power energy generated, emphasizing the conversion of solar energy to electricity. Fig. 1(f) compares ac and dc power generation, highlighting the role of inverters in converting dc power for grid transmission. This way, the Grafana dashboard serves as a powerful tool for real-time data visualization and monitoring.

Table 1. Performance assessment parameters for first three inverters

Inverter-1														
month	dc_power	ac_power	dc_energy	ac_energy	Array_effi	INV_effi	Sys_effi	A_yield	R_yield	F_yield	C_Loss	S_Loss	F_Loss	PR	CF
September	656.6	642.8	236392.3	231410.9	0.139	0.979	0.136	124.5	150.2	121.9	0.171	0.021	0.188	0.812	0.169
October	725.7	713.2	269968.9	265299.3	0.135	0.983	0.133	142.2	176.3	139.8	0.193	0.017	0.207	0.793	0.188
November	682.0	669.3	245531.3	240965.8	0.137	0.981	0.134	129.4	158.1	127.0	0.182	0.019	0.197	0.803	0.176
December	694.3	682.3	258278.1	253804.8	0.139	0.983	0.136	136.1	164.1	133.7	0.171	0.017	0.185	0.815	0.180
January	645.4	634.8	240079.2	236162.1	0.138	0.984	0.136	126.5	152.8	124.4	0.172	0.016	0.186	0.814	0.167
February	909.4	901.6	316473.6	313744.3	0.120	0.991	0.119	166.7	232.6	165.3	0.283	0.009	0.289	0.711	0.238
March	820.1	819.0	305083.6	304683.8	0.120	0.999	0.120	160.7	224.0	160.5	0.282	0.001	0.283	0.717	0.216
April	820.0	819.9	295204.9	295174.1	0.114	1.000	0.114	155.5	228.0	155.5	0.318	0.000	0.318	0.682	0.216
Inverter-2														
month	dc_power	ac_power	dc_energy	ac_energy	Array_effi	INV_effi	Sys_effi	A_yield	R_yield	F_yield	C_Loss	S_Loss	F_Loss	PR	CF
September	661.9	649.9	238298.0	233953.7	0.129	0.982	0.127	116.0	150.2	113.9	0.227	0.018	0.241	0.759	0.158
October	733.0	722.2	272668.6	268656.3	0.126	0.985	0.124	132.8	176.3	130.8	0.247	0.015	0.258	0.742	0.176
November	690.2	678.4	248458.7	244240.2	0.128	0.983	0.126	121.0	158.1	118.9	0.235	0.017	0.248	0.752	0.165
December	705.0	693.4	262244.3	257939.6	0.130	0.984	0.128	127.7	164.1	125.6	0.222	0.016	0.235	0.765	0.169
January	649.4	639.4	241584.5	237869.1	0.129	0.985	0.127	117.6	152.8	115.8	0.230	0.015	0.242	0.758	0.156
February	909.0	899.9	316347.7	313149.9	0.111	0.990	0.110	154.0	232.6	152.5	0.338	0.010	0.345	0.655	0.219
March	836.3	833.9	311114.2	310192.5	0.113	0.997	0.113	151.5	224.0	151.0	0.324	0.003	0.326	0.674	0.203
April	839.9	842.1	302381.6	303148.3	0.108	1.003	0.108	147.2	228.0	147.6	0.354	-0.003	0.353	0.647	0.205
Inverter-3														
month	dc_power	ac_power	dc_energy	ac_energy	Array_effi	INV_effi	Sys_effi	A_yield	R_yield	F_yield	C_Loss	S_Loss	F_Loss	PR	CF
September	665.2	652.3	239484.0	234819.4	0.131	0.981	0.129	118.1	150.2	115.8	0.214	0.019	0.229	0.771	0.161
October	745.6	731.4	277357.3	272079.7	0.130	0.981	0.127	136.8	176.3	134.2	0.224	0.019	0.239	0.761	0.180
November	699.9	686.6	251954.3	247193.2	0.131	0.981	0.129	124.2	158.1	121.9	0.214	0.019	0.229	0.771	0.169
December	712.0	699.1	264881.7	260052.1	0.133	0.982	0.131	130.6	164.1	128.2	0.204	0.018	0.219	0.781	0.172
January	657.9	647.2	244750.7	240770.6	0.132	0.984	0.130	120.7	152.8	118.7	0.210	0.016	0.223	0.777	0.160
February	909.0	893.0	316324.8	310758.9	0.112	0.982	0.110	156.0	232.6	153.2	0.330	0.018	0.341	0.659	0.220
March	859.0	845.2	319545.4	314419.5	0.118	0.984	0.116	157.6	224.0	155.0	0.296	0.016	0.308	0.692	0.208
April	839.0	823.8	302029.0	296580.7	0.109	0.982	0.107	148.9	228.0	146.2	0.347	0.018	0.359	0.641	0.203

Table 2. Monthly average performance of whole farm consisting of 15 inverters

The whole farm														
month	dc_power	ac_power	dc_energy	ac_energy	Array_effi	INV_effi	Sys_effi	A_yield	R_yield	F_yield	C_Loss	S_Loss	F_Loss	PR	CF
September	665.3	653.3	239523.0	235183.3	0.133	0.982	0.131	120.3	150.2	118.1	0.199	0.018	0.213	0.787	0.164
October	736.9	725.3	274142.0	269827.2	0.130	0.984	0.128	137.7	176.3	135.5	0.219	0.016	0.231	0.769	0.182
November	696.6	684.8	250771.1	246537.4	0.133	0.983	0.130	126.0	158.1	123.8	0.203	0.017	0.217	0.783	0.172
December	705.0	693.2	262242.1	257858.8	0.134	0.983	0.131	131.7	164.1	129.5	0.198	0.017	0.211	0.789	0.174
January	654.4	643.9	243427.5	239517.8	0.133	0.984	0.131	122.3	152.8	120.3	0.200	0.016	0.213	0.787	0.162
February	917.4	906.1	319271.0	315332.9	0.115	0.988	0.113	160.4	232.6	158.4	0.311	0.012	0.319	0.681	0.228
March	850.3	843.0	316308.3	313598.9	0.118	0.991	0.117	158.8	224.0	157.5	0.291	0.009	0.297	0.703	0.212
April	828.9	824.1	298396.7	296692.0	0.110	0.994	0.109	150.0	228.0	149.1	0.342	0.006	0.346	0.654	0.207

3.2. Monitoring the health of solar farm using IEC standard and SCADA data

As described in Section 2, various performance assessment parameters were computed to assess the overall operational performance of different inverters within the solar farm. Table 1 presents the calculated parameters for all three inverters in the master inverter room every month, allowing for a fair comparison. The

monthly performance analysis reveals that Inverter-2 exhibits lower performance compared to the other two inverters. Metrics such as energy generated, array efficiency, yields, performance ratios, and capacity factors are observed to be the lowest for Inverter-2. Conversely, Inverter-1 demonstrates superior performance among the three. It is important to note that capture losses, system losses, etc. are higher for Inverter 2, responsible for lower performance.

Similar to Table 1, Table 2 illustrates the monthly average performance of the entire farm, representing the

cumulative data from all 15 inverters. Fig. 2 visualizes the monthly average energy generation from dc power for all 15 inverters. February and March record the highest energy generation, while September achieves lower energy generation. Across different inverters, the plots reveal that Inverter 7, 12 and 14 exhibit the higher energy generation and Inverters 4 demonstrate low energy generation. To understand the performance, it was important to plot monthly average performance ratios for different months across all the inverters (Fig. 3). A monthly comparison highlights February, March, April as the lower performing months, with December, January as the best performing month, which is not depicted clearly through energy plots. Across the inverters, Inverter 5 shows the best performance and Inverter-4 exhibits the least performance. A large discrepancy can be seen for inverter-9 in the month of April. It was understood that it was because of some open circuit issue. These daily and monthly averaged performance comparisons serve as valuable indicators for early maintenance requirements.
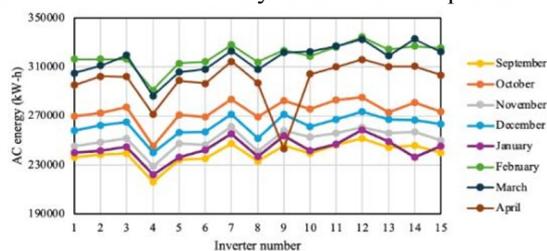

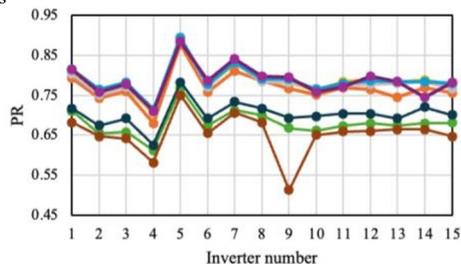**Fig. 2.** Monthly average energy generated for all inverters across 8 months

Fig. 3. Monthly average performance ratio for all inverters across 8 months (refer the legends from Fig.2)

3.3. Model development to obtain ideal I-V curves, MPPs and novel performance ratio

The PR in section 3.2 was obtained using conventional approach. A new method was developed to calculate the ideal (in the absence of any defects) I-V characteristics to obtain the difference between ideal MPP and real operating points at different irradiance and temperatures. The model development involved the following steps:

- Acquire the data from the I-V plots provided by the manufacturer's datasheets corresponding to various solar irradiance and temperatures (at the solar panel). Comprehend the inverter architecture. For INV-1: 20 solar panels are connected in series to form 1 string, 18-32 strings are connected in parallel to create 1 string combiner box (SCB), and 8 SCBs are connected in parallel to form 1 inverter.

- One significant limitation of the current study is the absence of panel-level data, necessitating the extrapolation of data to generate plots at the inverter level. Extend the data to obtain accurate

data and ideal plots at inverter level.

$$V_{String} = 20 \times V_{panel}; V_{SCB} = \max(V_{string}); V_{inverter} = \max(V_{SCB})$$

$$I_{String} = I_{panel} \ ; \ I_{SCB} = N_{string} \times I_{string} \ ; \ I_{inverter} = N_{SCB} \times I_{SCB}$$

An example of such dataset is shown in Table 3, it shows how the panel-level information has been extrapolated to get the inverter-level information.

Table 3. Extrapolation of panel-level data (voltage, current and power) to inverter-level data.

Irradiance	Module temp	Panel voltage	Panel current	Panel power	INV dc voltage	INV dc current	INV Rated power
400W/m2	25 ℃	0.01994 V	3.3026 A	330W	0.3989 V	753.009 A	1504800

- The present study utilized a support vector regression model to predict the ideal I-V characteristics and identify the ideal maximum power point. The model considers rated power, irradiance, module temperature, and DC voltage as input variables, with current as the output variable. The input array, X, consists of data[['Rated_power', 'Irradiance', 'module_temp', 'dc_voltage']], while the output array, y, comprises data['dc_current']. Presently, the errors for the test data stand at a Test R-squared value of 0.999, and a Test percentage RMSE of 0.791. Using this, an ideal I-V chart was obtained for corresponding rated power, irradiance and temperature.

- Now, delving into the actual dataset through influxdb, we possess data points encompassing Rated Power, Irradiance, Temperature, Voltage, and Current at each time stamp at the inverter level. For every timestamp, corresponding to each row of data, our objective is to generate an I-V chart by calculating the current across the entire voltage range (0-1000 V). The depicted curve in Fig. 4 is a forecast for a single dataset: [Rated Power = 1504800, Irradiance = 1000, Module Temperature = 25, Voltage Range = 0 to 1000 V]. The voltage range spans from 0 to 1000 V with a step size of 0.5 V. The model forecasts the current for the complete voltage range, and the plot is subsequently generated. Power for each combination of current and voltage is calculated as predicted power = dc voltage × predicted current. The task is to identify the combination where power reaches its maximum value. Once the enhanced model at the inverter level is established, the Maximum Power Point (MPP) value can be obtained for actual irradiance and temperature at each timestamp.
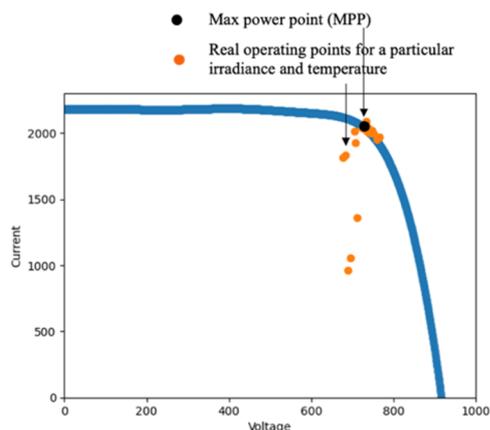
Fig. 4. The accurate state of real-operating points on the I-V curve, which depicts how far they exist from MPP at 1000 W/m² irradiance and 25°C temperature. The blue curve shows the ideal predicted I-V curve for corresponding irradiance and temperature

- The MPP values are crucial for deriving precise ideal power values for real-time data under actual irradiance and temperature conditions. Ultimately, the new PR is computed as the ratio of Actual Power to Ideal Power: (New PR= Energy produced/ Maximum energy it can produce at instantaneous irradiance and temperature)

The average value of the new performance ratio for Inverter-1 is determined to be 0.77. Traditional performance ratios (PR$_{Conventional}$ = Actual efficiency/Rated efficiency) consider rated efficiency in denominator. The tradiational way does not take denominator as the maximum efficiency it can produce at that instantaneous irradiance and temperature. The conventional method of real-time performance measure does not consider the variation of irradiance and temperature, hence do not capture the accurate performance and losses. Consequently, this ratio fails to offer an accurate assessment of the farm's performance capabilities at different timestamps, as it doesn't consider the dynamic nature of irradiance and temperature changes. The updated performance ratio provides precise insights into the operational efficiency of the inverter, offering an accurate gauge of its current performance and highlighting potential areas for improvement. Table 4 illustrates the monthly (for September) average performance ratio of three inverters using both conventional methods and the current approach. To calculate the monthly average PRs, the instantaneous PRs at each timestamp were averaged out over a month. The current approach considers transient variations in irradiance and temperature, calculating I-V curves and MPP at each timestamp for accurate performance assessment. The new performance ratio (nPR) example for inverter-1 is 0.77, hence 1-0.77 accounts accurately for all losses, including shading, hotspots, soiling, and imperfect MPP algorithms.

Table 4. Monthly average performance ratio of three inverters using both conventional methods and the current approach

	PR$_{Conventional}$	New PR
Inverter-1	0.81	0.77
Inverter-2	0.75	0.70
Inverter-3	0.77	0.72
Inverter-4	0.70	0.66
Inverter-5	0.89	0.82
Inverter-6	0.77	0.73
Inverter-7	0.82	0.76
Inverter-8	0.79	0.75
Inverter-9	0.78	0.73
Inverter-10	0.76	0.72
Inverter-11	0.78	0.73
Inverter-12	0.78	0.73
Inverter-13	0.78	0.73
Inverter-14	0.78	0.73
Inverter-15	0.77	0.72

4 Conclusion

This study spans eight months, from September 2023 to April 2024, with the primary goal of evaluating different aspects of solar farm's performance. The motivation behind this assessment is the need to enhance the farm's output, considering the limited access to alternative energy resources and the potential for increasing national photovoltaic (PV) penetration. Achieving optimal farm performance requires continuous monitoring and proactive maintenance to address variabilities promptly. Monitoring the solar farm using SCADA data and analytical analysis, the critical performance indicator, array efficiency, is observed to be approximately 13-14%, notably lower than the rated efficiency. Variations are identified across different inverters and months. The accurate modeling of losses is deemed crucial for estimating potential improvements. March and April emerge as low-performing months, contrasting with the higher performance observed in December and January. This trend can be understood with the fact that March and April are higher temperature month due to summer season, while December, January are colder months. The dc energy generated is increased in summer weather due to increase in irradiance, while performance is degraded, which can be attributed to the inverse proportionality to higher temperature. Among the inverters, Inverter 4 exhibits lower performance, while Inverter 5 stands out as the best performer. The discrepancies in inverter-9 in april month is identified as the open circuit failure/defect. The new performance ratio serves as a valuable parameter, providing accurate insights into inverter performance and the potential for improvement. By predicting the ideal I-V curve and determining the Maximum Power Point (MPP) value at each timestamp (for the real-time values

of irradiance and temperature), which provides a transparent insight into the transient operational state of the inverter and the deviation of real-time data from the MPP. The developed performance ratio offers precise details on the inverter's efficiency and signifies the potential for improvement.

References

1. H. Chen, H. Yi, B. Jiang, K. Zhang, and Z. Chen, "Data-Driven Detection of Hot Spots in Photovoltaic Energy Systems," *IEEE Trans Syst Man Cybern Syst*, vol. 49, no. 8, pp. 1731–1738, Aug. 2019, doi: 10.1109/TSMC.2019.2896922.

2. E. E. Van Dyk, A. R. Gxasheka, and E. L. Meyer, "Monitoring current-voltage characteristics and energy output of silicon photovoltaic modules," *Renew Energy*, vol. 30, no. 3, pp. 399–411, Mar. 2005, doi: 10.1016/j.renene.2004.04.016.

3. A. Gopi, P. Sharma, K. Sudhakar, W. K. Ngui, I. Kirpichnikova, and E. Cuce, "Weather Impact on Solar Farm Performance: A Comparative Analysis of Machine Learning Techniques," *Sustainability (Switzerland)*, vol. 15, no. 1, Jan. 2023, doi: 10.3390/su15010439.

4. C. Buerhop, L. Bommes, J. Schlipf, T. Pickel, A. Fladung, and I. M. Peters, "Infrared imaging of photovoltaic modules: a review of the state of the art and future challenges facing gigawatt photovoltaic power stations," *Progress in Energy*, vol. 4, no. 4. Institute of Physics, Oct. 01, 2022. doi: 10.1088/2516-1083/ac890b.

5. M. Dey, S. P. Rana, C. V. Simmons, and S. Dudley, "Solar farm voltage anomaly detection using high-resolution µPMU data-driven unsupervised machine learning," *Appl Energy*, vol. 303, Dec. 2021, doi: 10.1016/j.apenergy.2021.117656.

6. N. Mansouri, A. Lashab, D. Sera, J. M. Guerrero, and A. Cherif, "Large photovoltaic power plants integration: A review of challenges and solutions," *Energies*, vol. 12, no. 19. MDPI AG, Oct. 08, 2019. doi: 10.3390/en12193798.

7. S. Bentouba, M. Bourouis, N. Zioui, A. Pirashanthan, and D. Velauthapillai, "Performance assessment of a 20 MW photovoltaic power plant in a hot climate using real data and simulation tools," *Energy Reports*, vol. 7, pp. 7297–7314, Nov. 2021, doi: 10.1016/j.egyr.2021.10.082.

8. M. Mpholo, T. Nchaba, and M. Monese, "Yield and performance analysis of the first grid-connected solar farm at Moshoeshoe I International Airport, Lesotho," *Renew Energy*, vol. 81, pp. 845–852, Sep. 2015, doi: 10.1016/j.renene.2015.04.001.