2024 Fifth International Conference on Intelligent Data Science Technologies and Applications (IDSTA)

Dynamic Threshold-Based Anomaly Detection in Photovoltaic Generation Time Series Using Statistical Methods

Michelle Melo Cavalcante
FEEC/LESF-MV
University of Campinas (Unicamp)
Campinas, Brazil
michellemelo.c@gmail.com

João Lucas de Souza Silva
FEEC/LESF-MV
University of Campinas (Unicamp)
Campinas, Brazil
jlucas.souzasilva@gmail.com

Tárcio André dos Santos Barros
FEEC/LESF-MV
University of Campinas (Unicamp)
Campinas, Brazil
tarcio87@unicamp.br

***Abstract*—The efficient operation of photovoltaic (PV) plants requires continuous monitoring to identify and correct anomalies that may affect the performance and lifespan of the equipment. However, some challenges include defining which data can be collected and useful from the installation to identify anomalies, as well as which models can be applied. Therefore, this paper proposes an approach for anomaly detection in PV generation using dynamic threshold techniques based on descriptive statistics. The methodology involves monthly analysis of AC power characteristic curves and irradiance normalization, applying interquartile ranges and standard deviations to identify anomalies. AC power was chosen as it is the inverter output and, therefore, more easily obtained data. The methodology's implementation is validated using real PV inverter data to identify anomalous behaviors. Additionally, the data used are generally available in PV plants without the need for additional sensors. Therefore, this approach provides an effective tool for predictive maintenance and optimization of PV systems.**

***Index Terms*—Anomaly Detection, Photovoltaic Systems, Monitoring, Dynamic Thresholds, Descriptive Statistics.**

I. Introduction

The solar energy market has been showing growth. According to the 2023 report by the International Renewable Energy Agency (IRENA), the global installed capacity of solar photovoltaic (PV) energy exceeded 710 GW, reflecting an annual growth of about 22 % [1]. With the significant increase in installed solar energy capacity worldwide, the need for continuous monitoring becomes greater to ensure the reliability and performance of PV systems [2]–[4].

A monitoring framework aims to identify anomalies that can result in significant energy losses and additional maintenance costs [3], [5]. Many of the variables that can be monitored pertain to electrical quantities inherent to the system and data related to external factors, such as temperature and irradiance.

For grid-connected PV systems, electrical quantities are associated with the performance of various components aimed at converting sunlight into usable electricity and integrating it into the grid. The main components include PV modules, which capture solar energy and convert it into DC voltage and current, and inverters, which transform this DC voltage and current into alternating current (AC) compatible with the grid [2], [6], [7]. Power is related to voltage and current, and the converted electrical energy is the power accumulated over time. An example schematic is shown in Figure 1.

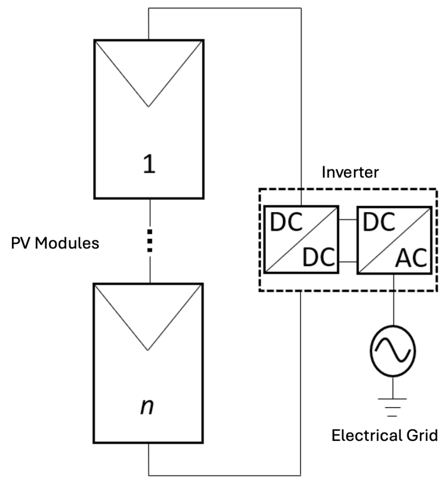

Fig. 1. Schematic of PV system connected to conventional grid [8].

In addition to electrical quantities, external factors such as ambient temperature, solar irradiance, and weather conditions also affect the performance of PV systems. Irradiance, for instance, is a measure of the power of the sun received per unit area and is crucial for calculating the amount of energy that PV modules can generate [9]. Ambient temperature also affects the efficiency of PV modules, as high temperatures can reduce the efficiency of converting solar energy into electricity [10]. Other climatic factors, such as the presence of clouds, rain, and dust, can influence the amount of sunlight reaching the modules, directly impacting energy generation [11]. Therefore, the collection and analysis of these external data can aid in identifying performance deviations and detecting system anomalies. Early fault detection is a key component for the efficient operation of PV plants [12]–[14].

Various approaches have been proposed for anomaly detection in PV systems, including machine learning-based methods and statistical techniques [15]. Some studies using machine learning algorithms, although effective, can be computationally intensive and depend on large volumes of labeled data, which may limit their applicability in scenarios with limited data or restricted computational resources [16]–[18].

On the other hand, methods based on descriptive statistics offer a simpler and more efficient solution for anomaly detection. Techniques such as regression analysis, interquartile range (IQR), and statistical control charts have been effective in detecting faults in PV systems by analyzing performance and identifying significant deviations from expected behavior. These methods are less dependent on large volumes of data and can be implemented with lower computational complexity [15], [19]–[21].

In this context, this article proposes an anomaly detection approach for PV generation using dynamic threshold techniques based on descriptive statistics. The presented methodology involves the monthly analysis of AC power characteristic curves and irradiance normalization, applying interquartile range and standard deviation limits to identify anomalies.

The implementation of the proposed methodology is evaluated using real data from inverters at a PV plant at the University of Campinas (UNICAMP) to identify anomalous behaviors, demonstrating that this approach provides a tool for predictive maintenance and optimization of PV systems.

II. Survey of Related Literature

In the literature, there are studies that seek to use statistical methods for the monitoring and detection of anomalies in PV systems. Among the methods, the use of hypothesis tests [22], descriptive statistics [20], [21], and multivariate statistical methods [17], [23], [24] are some examples, sometimes associated with other methods.

In [22], a methodology based on hypothesis testing is proposed to identify faults in PV systems. The approach involves collecting performance data, such as current and voltage, followed by defining null and alternative hypotheses. Test statistics are calculated and, based on the p-value and significance level, it is decided whether there is a fault in the system. This approach depends on accurate reference models, and noisy or incomplete data can lead to false detections.

In [20], the authors propose a method for detecting anomalies in solar power generation systems that involves normalizing the amount of electricity generated, allowing for the comparison of data between different systems and weather conditions. The method uses the mean and standard deviation of normalized generations to identify points outside the expected pattern, characterizing possible faults or anomalies. In the experiment, the number of panels per string was limited to four. Therefore, the effectiveness of the method was not applied when the number of panels in a sequence increased significantly.

In [21], linear regression analysis is applied to the Performance Ratio (PR) metric to estimate gradual degradation and performance reduction trends over time. For anomaly detection, statistical control charts, such as the Exponentially Weighted Moving Average (EWMA), are employed to monitor and identify sudden or anomalous changes in the system.

In [23], the method utilizes techniques like Principal Component Analysis (PCA) and Independent Component Analysis (ICA) to identify anomalous patterns and deviations in plant performance. This approach enables the early identification of operational problems, improving system efficiency and reliability. The authors also highlight that the implementation of these techniques can be complex, and the accuracy of the results can be affected by the quality and quantity of the available data, which may limit the effectiveness of monitoring under real-world conditions.

In [17], PCA is combined with genetic algorithms (GA) and artificial neural networks (ANN) to detect and diagnose faults in grid-connected PV systems. PCA is used to reduce data dimensionality while preserving the most relevant features, and GA selects the most important features. The ANN then classifies and diagnoses the faults. Despite the method's effectiveness, it has some limitations. The combination of GA and ANN can be computationally intensive, especially for large-scale systems. Moreover, the accuracy of the diagnosis heavily depends on the quality and quantity of available historical data. Another limitation is the model's generalization capability, which may not be effective under operational conditions not represented in the training data.

In [24], electrical and environmental variables are integrated to identify anomalies. Techniques such as PCA are used to reduce data dimensionality, and multivariate statistical models are applied to detect significant deviations from the system's normal behavior. Some limitations addressed include the necessity for high-quality and large-scale data to train the models, which can be challenging in some scenarios. Additionally, the approach can be sensitive to unmodeled variability in environmental conditions, leading to false detections. Another limitation is the computational complexity associated with processing large volumes of data and applying multivariate techniques such as PCA.

III. Case Study

The research used one of the installations from the Sustainable Campus project at the University of Campinas (UNICAMP) located in the city of Campinas, São Paulo (Brazil) as a case study. The PV plant, as shown in Fig. 2, is situated at the UNICAMP gymnasium, with a total capacity of 336.96 kWp and an estimated annual generation of 481.16 kWh [25].

AC Power and Irradiance data were collected over a total period of one year, recording measurements at 15-minute intervals. The decision was made to use the inverter's output power data as they are derived from the voltage and current relationships converted by the PV modules.

These data allow us to identify issues throughout the entire energy generation chain, including problems in the modules, the electrical grid, or the inverters themselves. The simplified schematic of the plant's operation with the DC-AC power component was presented in Fig. 1.

Fig. 2. PV plant installed at UNICAMP [26].

IV. Methodology

This study proposes a methodology to identify anomalies in the active power (P_{ac}) measurements of PV inverters using dynamic thresholds based on descriptive statistical methods and solar irradiance data.

A. Anomaly Detection Criteria

To identify anomalies in PV plants or systems, *(i)* monthly characteristic curves of AC power with dynamic thresholds were created, *(ii)* daily solar irradiance data were used, and *(iii)* the combination of both was also utilized. Using characteristic curves and irradiance data, anomaly detection focused on potential problems such as interruptions, failures or slowdowns associated with or absorbed by the inverter.

Creating monthly characteristic curves of active power (P_{ac}) with dynamic thresholds can contribute to identifying and monitoring the performance variations of PV inverters over time. By calculating monthly descriptive statistics, we can establish upper and lower limits that adapt to seasonal and daily variations in energy generation. Exceeding these dynamic limits allows for more precise anomaly detection, distinguishing between normal fluctuations and potentially problematic behaviors that may indicate system failures or inefficiencies.

Daily irradiance, which represents the total power of energy from the Sun per unit area, was compared with the inverter's output power (or P_{ac}). To facilitate the comparison between irradiance (*POA*) and P_{ac}, the *POA* data were normalized. Normalization allows for the identification of anomalies where active power decreases while irradiance increases, indicating possible system failures.

B. Monthly Characteristic Curve of AC Power

The monthly characteristic curves were defined by limits based on the active power (P_{ac}) values grouped monthly (*m*) and hourly (*h*). Two types of limits are calculated: Upper Limit ($U_{IQR_{mh}}$) and Lower Limit ($L_{STD_{mh}}$).

The $U_{IQR_{mh}}$ is calculated using the third quartile (*Q3*) and the interquartile range (*IQR*). The goal is to capture values that are above the typical variation range. The formula for the calculation is represented by Eq.1.

$$U_{IQR_{mh}} = min(Q3_{mh} + \alpha \times IQR_{mh}, max_{mh}) \tag{1}$$

The upper limit is adjusted by α. The value of α can vary depending on the context and the desired sensitivity for anomaly detection. In this work, an empirical value of 1.5 was used based on the interquartile method. To avoid unrealistic values, the upper limit was also constrained not to exceed the maximum power (max_{mh}).

$$max_{mh} = max(P_{ac_i}) \tag{2}$$

The *IQR* can be found by the difference between the third quartile (*Q3*) and the first quartile (*Q1*) according to Eq.3.

$$IQR_{mh} = Q3_{mh} - Q1_{mh} \tag{3}$$

The $L_{STD_{mh}}$ is based on the mean (μ) and the standard deviation (σ) of the P_{ac} measurements. Using the mean and standard deviation helps to avoid extrapolating to negative P_{ac} values, which are physically impossible and occur more frequently when using *IQR*. For these cases, which can happen at the beginning and the end of the curve due to being associated with the start and end of energy generation from the panels following sunrise and sunset, the values were adjusted to zero. The formula for the calculation is represented by the Eq.4.

$$L_{STD_{mh}} = max(\mu_{mh} - \alpha \times \sigma_{mh}, 0) \tag{4}$$

With μ and σ calculated according to Eq.5 and Eq.6.

$$\mu_{mh} = \frac{1}{n}\sum_{i=1}^{n}(P_{ac_i}) \tag{5}$$

$$\sigma_{mh} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_{ac_i} - \mu_{mh})^2} \tag{6}$$

C. Normalization of Irradiance Compared to Inverter Output Power

The irradiance (POA) data have the same distribution curve when compared to the P_{ac} curve under conditions of intense sunlight during the day. The difference lies in the scale of the values. For the region where the PV plant is located, POA ranges from 0 to 1000, while P_{ac} can range from 0 to over

55k. Normalizing the data facilitates the visual comparison of the curves.

The normalization of POA data (POA_{norm}) is used to scale the POA values to the same range as P_{ac} as per the Eq.7.

$$POA_{norm} = \left(\frac{POA - POA_{min}}{\Delta POA} \right) \times \Delta P_{ac} + P_{ac_{min}} \quad (7)$$

Where ΔPOA and ΔP_{ac} are calculated as per Eq.8 and Eq.9.

$$\Delta POA = POA_{max} - POA_{min} \quad (8)$$

$$\Delta P_{ac} = P_{ac_{max}} - P_{ac_{min}} \quad (9)$$

D. Summary of applied methodology

The logic for anomaly detection was based on two main criteria:

- Active Power outside the limits of the monthly characteristic curve: If the active power (P_{ac}) measured at a given moment was above the Upper limit ($U_{IQR_{mh}}$) or below the Lower limit ($L_{STD_{mh}}$).
- Drop in Active Power with an increase in Irradiance: An anomaly is also recorded if, in two consecutive moments, the active power (P_{ac}) decreases while the irradiance (POA) increases.

Both cases can also occur together at a given time of day. Thus, the anomaly detection focused on problems inherent to the equipment and not on anomalies caused by climatic variations.

V. Anomaly Detection in PV Plants

This section discusses the results of the case study. As reported in section IV, *(i)* the irradiance values were normalized to be compared with the power and to evaluate discrepancies between power and irradiance data that are directly proportional, and *(ii)* it can be observed the monthly characteristic curves of the inverter output power were created.

A. Normalization of Irradiance data

Under perfect weather conditions (sunny with no clouds) and without issues in the plant's equipment or the electrical grid, the inverter's output power and irradiance resemble a Gaussian curve. Fig. 3 shows the mentioned behavior and the climatic oscillations resulting in the variation of power for the 19th and 20th of May 2020. The characteristic curve for the month of May was added to Fig. 3

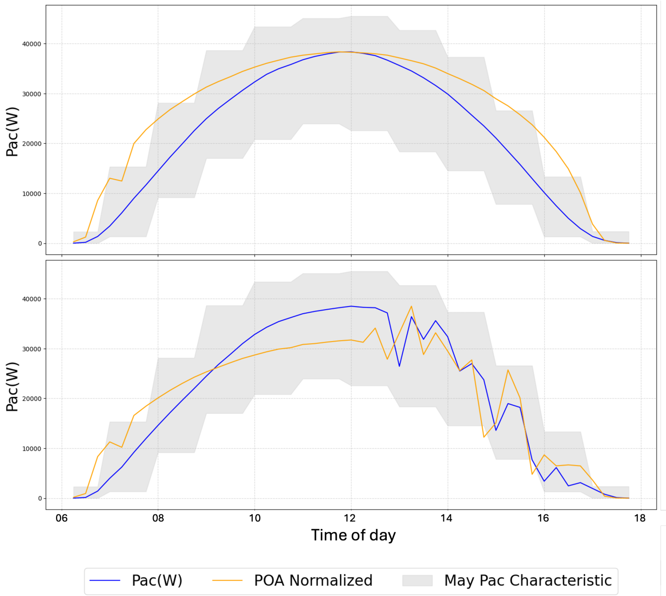

Fig. 3. Normalization of Irradiance data compared to AC Power.

B. Monthly Characteristic Curves of Power (P_{ac})

The Fig 4 below shows the monthly characteristic curves of active power (P_{ac}) for a one-year period. Each graph presents the mean, median, and the upper and lower limits based on the *IQR* and standard deviation.

Seasonal variability of the curves is noted, with higher P_{ac} values observed in the summer months and lower values in the winter months. Fig 4 includes information on minimum, mean, median, upper/lower limit calculated by *IQR* and lower limit calculated by standard deviation (STD) for P_{ac}.

It is observed that the mean and median values are approximately equal at the beginning and end of the day, indicating that, for this time period, most values are concentrated around the center, and the outliers are few and similarly distributed on both sides of the median.

In general, it is noted that the asymmetrical distribution increases closer to the peak of P_{ac}, where the median is always higher than the mean. This behavior suggests a left skew (skewed left), where there are some very low P_{ac} values pulling the mean down.

For the lower limit values by *IQR*, large oscillations are noted throughout the day, unlike when using the lower limit calculated using the standard deviation (L_{STD}), which becomes more similar to a normal distribution curve and can be used more generically.

C. Anomaly detection

The detected anomalies were identified in three ways: (I) anomalies captured by a drop in P_{ac} when there is an increase in POA, (II) anomalies captured when P_{ac} exceeds the upper or lower threshold of the characteristic curve, and (III) both methods captured the same anomaly, as shown in Fig. 5 which illustrates anomaly detection for the period from 02/05/2020 to 02/11/2020.

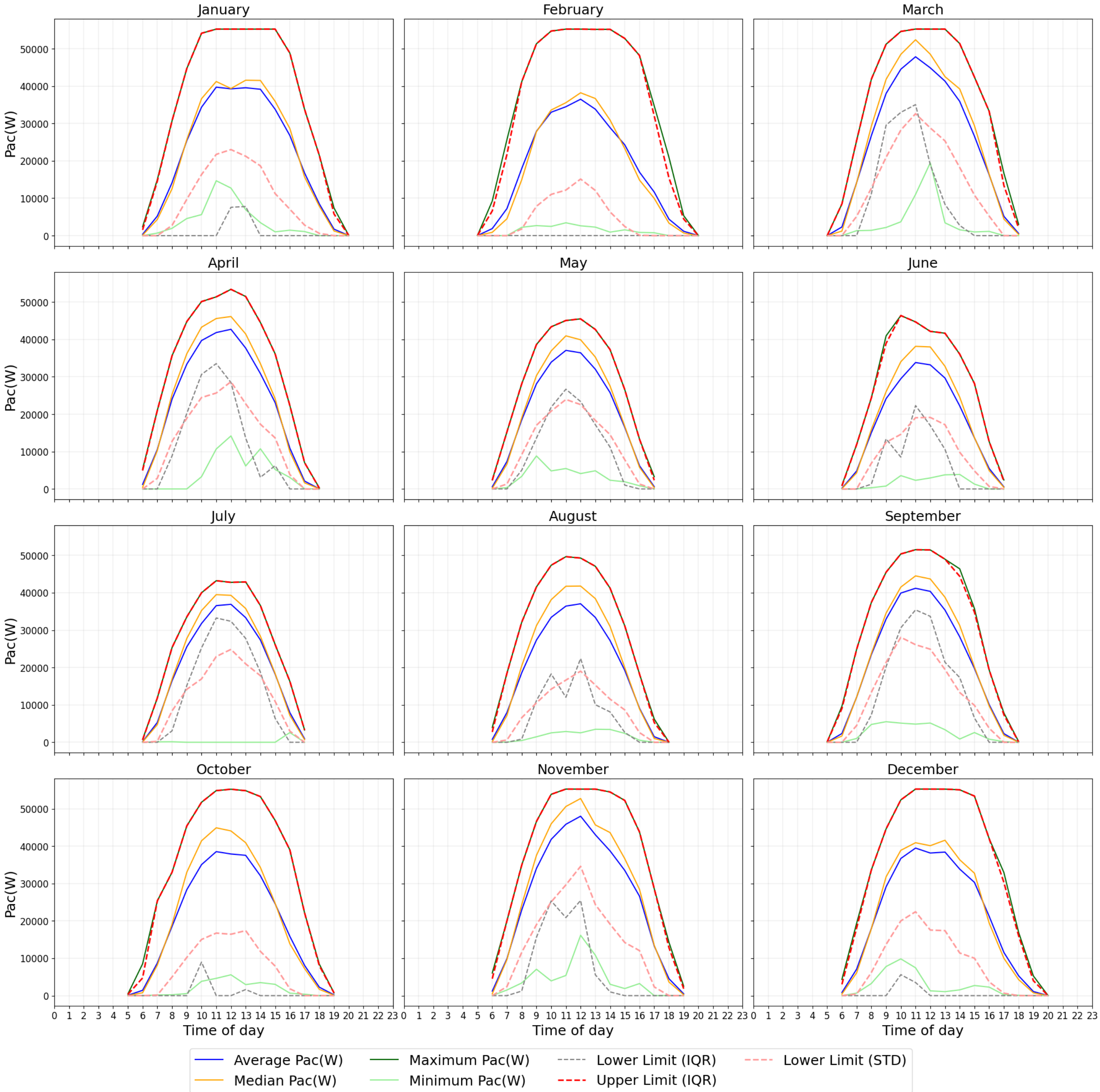

Fig. 4. Monthly Characteristic Curves of AC Power.

In Fig 5 we can observe that the variations in irradiance are likely a consequence of weather conditions (cloudy, rainy, etc.). On 11/02/2020, we can see that despite high irradiance, the AC power values are low, indicating issues with the inverter or the grid. For the period from 02/05/2020 to 02/11/2020, the number of identified anomalies is presented in Table I.

TABLE I
COMPARISON OF THE NUMBER OF ANOMALIES IDENTIFIED BY METHOD

Anomaly type	**Anomaly count**
(I) $P_{\mathrm{ac}}(W)$ drop while POA increasing	32
(II) $P_{\mathrm{ac}}(W)$ out of bounds	16
(III) Both	5

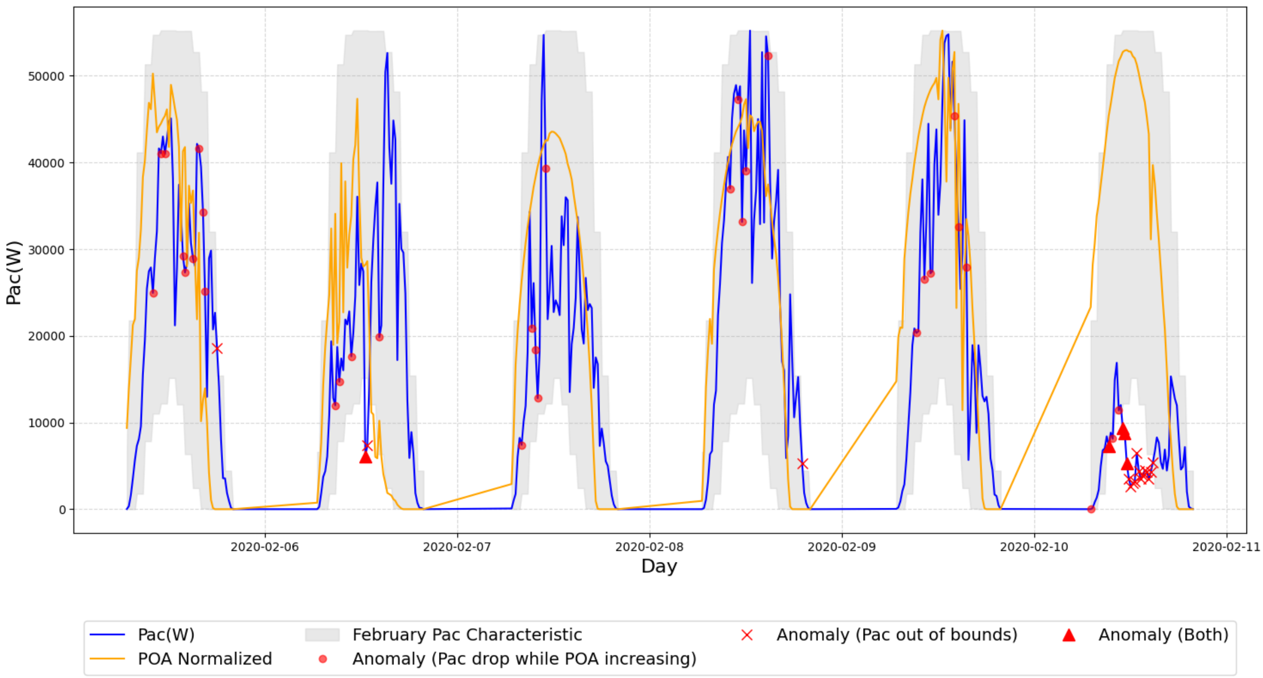

Fig. 5. P_{ac} and POA with Anomaly Detection (2020/2/5 - 2020/2/11).

TABLE II
NUMBER AND PERCENTAGE OF ANOMALIES DETECTED PER MONTH AND PER METHOD COMPARED TO THE TOTAL MEASUREMENTS (TOTAL DATA)

Moth	Anomaly count			Total Data
	I	II	III	
January	205 (12.0%)	79 (4.6%)	18 (1.1%)	1702
February	189 (12.2%)	70 (4.5%)	24 (1.5%)	1543
March	183 (11.5%)	100 (6.3%)	29 (1.8%)	1584
April	138 (09.4%)	89 (6.1%)	31 (2.1%)	1465
May	115 (07.9%)	103 (7.1%)	36 (2.5%)	1447
June	188 (13.9%)	78 (5.8%)	28 (2.1%)	1356
July	151 (10.6%)	63 (4.4%)	17 (1.2%)	1424
August	134 (08.9%)	115 (7.7%)	40 (2.7%)	1489
September	153 (10.2%)	95 (6.3%)	41 (2.7%)	1500
October	228 (14.0%)	94 (5.8%)	36 (2.2%)	1628
November	195 (11.8%)	87 (5.3%)	45 (2.7%)	1653
December	250 (14.5%)	76 (4.4%)	37 (2.1%)	1727

Considering the entire year of 2020, Table II shows the anomalies detected by month considering I) P_{ac} drop while POA increasing, II) P_{ac} out of bounds, and III) Both. It is noted that when considering only the increasing irradiance data with a drop in AC power (Case I), December has the highest number of anomalies both quantitatively and as a percentage of the total measurements taken. When considering only the characteristic curve (Case II) created, August has the highest number and percentage of anomalies. When considering the anomalies detected by both methods together (Case III), November is the month with the highest number of detected anomalies. Percentually, for this last case, November, September, and August have the same number of detected anomalies.

Overall, it is observed that the monthly characteristic curve can represent the behavior of AC power without considering small oscillations that may occur, making it a more conservative method.

VI. Conclusion

Anomaly detection in generated power $P_{\text{ac}}(W)$ over time can be a practical and efficient first step for minimal monitoring of PV plants. Given that power is a direct result of voltages and currents, monitoring these powers can be sufficient for systems that do not require extremely detailed monitoring.

This approach allows for the rapid identification of potential problems, serving as a basis for a more detailed analysis of other operational parameters. If applied consistently, it is possible to minimize interruptions, achieve greater longevity of PV systems, and maximize energy production by early capturing situations such as module degradation over time.

The inclusion of climatic data such as irradiance can provide greater accuracy and contribute to the detection of behaviors that may arise in the inverters that can occur more discreetly but reduce AC power generation and, consequently, the energy generated. However, there is a caution regarding the data recording process in PV plants, as techniques like moving averages can interfere with and cause errors in anomaly detection. Therefore, for future implementations, it is also advisable to use systems with a time step smaller than 15 minutes.

As future work, we intend to establish new criteria for defining anomalies based on the range of techniques presented; utilize datasets with a 1-minute time step; and, test new sensors and their correlation with PV power, such as adding humidity, temperature, and wind direction sensors.

ACKNOWLEDGMENT

We would like to thank TotalEnergies for the financial support. In addition, we are grateful to all collaborators from University of Campinas (UNICAMP). We acknowledge the support of ANP (Brazilian National Oil, Natural Gas and Biofuels Agency) through the R&D levy regulation. Acknowledgements are extended to the Center for Energy and Petroleum Studies (CEPETRO) and School of Electrical and Computer Engineering (FEEC). The authors would like to thank the Campus Sustainable Project Unicamp for the data.

REFERENCES

[1] Agência Internacional de Energia Renovável (IRENA), "Renewable Capacity Statistics 2023," 2023, acesso em: 05 jun. 2024. [Online]. Available: https://www.irena.org/publications/2023/Mar/Renewable-Capacity-Statistics-2023

[2] M. A. Eltawil and Z. Zhao, "Grid-connected photovoltaic power systems: Technical and potential problems - a review," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 1, pp. 112–129, 2010.

[3] T. Babasaki and Y. Higuchi, "Using pv string data to diagnose failure of solar panels in a solar power plant," in *2018 IEEE International Telecommunications Energy Conference (INTELEC)*, 2018, pp. 1–4.

[4] B. Khelifi, M. A. Zdiri, and F. B. Salem, "Machine learning for solar power systems - a short tour," in *2021 12th International Renewable Energy Congress (IREC)*, Hammamet, Tunisia, 2021, pp. 1–6.

[5] M. Ejgar and B. Momin, "Solar plant monitoring system: A review," in *Proceedings of the International Conference on Computing Methodologies and Communication, ICCMC 2017*, vol. 2018-January, 2018, pp. 1142–1144.

[6] B. Parida, S. Iniyan, and R. Goic, "A review of solar photovoltaic technologies," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 3, pp. 1625–1636, 2011.

[7] E. Kabir *et al.*, "Solar energy: Potential and future prospects," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 894–900, 2018.

[8] J. L. de Souza Silva, H. S. Moreira, M. V. G. dos Reis, T. A. dos Santos Barros, and M. G. Villalva, "Theoretical and behavioral analysis of power optimizers for grid-connected photovoltaic systems," *Energy Reports*, vol. 8, pp. 10 154–10 167, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352484722014196

[9] X. Han *et al.*, "Influence of solar irradiance on the performance of photovoltaic systems," *Journal of Solar Energy*, vol. 123, pp. 45–53, 2017.

[10] N. Gokmen *et al.*, "Temperature and its effect on the efficiency of solar photovoltaic systems," *Renewable Energy*, vol. 96, pp. 123–130, 2016.

[11] R. Gupta *et al.*, "Impact of environmental factors on the performance of solar photovoltaic systems," *Energy Reports*, vol. 5, pp. 822–830, 2019.

[12] S. Ayala *et al.*, "Comprehensive analysis of climatic factors affecting photovoltaic system performance," *Journal of Renewable and Sustainable Energy*, vol. 12, pp. 456–470, 2020.

[13] A. Lazaroiu, M. Gmal Osman, C.-V. Strejoiu, and G. Lazaroiu, "A comprehensive overview of photovoltaic technologies and their efficiency for climate neutrality," *Sustainability*, vol. 15, p. 16297, 2023.

[14] National Renewable Energy Laboratory, "Best practices for operation and maintenance of photovoltaic and energy storage systems," 2018, acesso em: 05 jun. 2024. [Online]. Available: https://www.nrel.gov/docs/fy19osti/73822.pdf

[15] M. M. Cavalcante, J. L. De Souza Silva, S. B. Martins, I. F. Silva Nunes, A. C. Ribeiro, and T. A. Dos Santos Barros, "Comparison and application of data science techniques for anomaly detection in photovoltaic systems," in *2023 IEEE 8th Southern Power Electronics Conference and 17th Brazilian Power Electronics Conference (SPEC/COBEP)*, 2023, pp. 1–5.

[16] A. F. Amiri, H. Oudira, A. Chouder, and S. Kichou, "Faults detection and diagnosis of pv systems based on machine learning approach using random forest classifier," *Energy Conversion and Management*, vol. 301, p. 118076, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0196890424000177

[17] A. Hichri, M. Hajji, M. Mansouri, M.-F. Harkat, A. Kouadri, H. Nounou, and M. Nounou, "Fault detection and diagnosis in grid-connected photovoltaic systems," in *2020 17th International Multi-Conference on Systems, Signals Devices (SSD)*, 2020, pp. 201–206.

[18] J. L. de Souza Silva, E. Mahmoudi, R. R. M. Carvalho, and T. A. dos Santos Barros, "Classification of anomalies in photovoltaic systems using supervised machine learning techniques and real data," *Energy Reports*, vol. 11, pp. 4642–4656, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352484724002488

[19] C. Hirschl *et al.*, "Statistical methods for degradation estimation and anomaly detection in photovoltaic plants," *Sensors*, vol. 21, no. 11, p. 3733, 2021.

[20] Y. Akiyama, Y. Kasai, M. Iwata, E. Takahashi, F. Sato, and M. Murakawa, "Anomaly detection of solar power generation systems based on the normalization of the amount of generated electricity," in *2015 IEEE 29th International Conference on Advanced Information Networking and Applications*, 2015, pp. 294–301.

[21] V. Dimitrievska, F. Pittino, W. Muehleisen, N. Diewald, M. Hilweg, A. Montvay, and C. Hirschl, "Statistical methods for degradation estimation and anomaly detection in photovoltaic plants," *Sensors*, vol. 21, no. 11, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/11/3733

[22] F. Harrou, B. Taghezouit, B. Bouyeddou, Y. Sun, and A. H. Arab, "Fault detection in solar pv systems using hypothesis testing," in *2021 IEEE 19th International Conference on Industrial Informatics (INDIN)*, 2021, pp. 1–6.

[23] B. Taghezouit, F. Harrou, Y. Sun, A. H. Arab, and C. Larbes, "Multivariate statistical monitoring of photovoltaic plant operation," *Energy Conversion and Management*, vol. 205, p. 112317, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0196890419313238

[24] G. G. Kim, W. Lee, B. G. Bhang, J. H. Choi, and H.-K. Ahn, "Fault detection for photovoltaic systems using multivariate analysis with electrical and environmental variables," *IEEE Journal of Photovoltaics*, vol. 11, no. 1, pp. 202–212, 2021.

[25] J. L. de Souza Silva, K. Barbosa de Melo, K. V. dos Santos, E. Yoiti Sakô, M. Kitayama da Silva, H. Soeiro Moreira, G. Bolognesi Archilli, J. G. Ito Cypriano, R. E. Campos, L. C. Pereira da Silva, and M. Gradella Villalva, "Case study of photovoltaic power plants in a model of sustainable university in brazil," *Renewable Energy*, vol. 196, pp. 247–260, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960148122009430

[26] UNICAMP, "About Campus Sustentável Project," http://www.campus-sustentavel.unicamp.br/en/about-campus-sustentavel/, 2021.