



Universidad de Jaén

**PONTIFICIA UNIVERSIDAD
CATÓLICA DEL PERÚ**

Escuela de Posgrado

Contribution to the characterization and modeling of
photovoltaic generators

Tesis para obtener el grado académico de Doctor en Física que
presenta:

Jose Ruben Angulo Abanto

Asesor:

Dr. Jan Amaru Palomino Töfflinger

Dr. Juan de la Casa Higuera

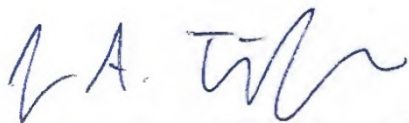
Lima, 2022

Declaración jurada de autenticidad


Yo, Jan Amaru Palomino Tofflinger, docente de la Escuela de Posgrado de la Pontificia Universidad Católica del Perú, y yo, Juan de la Casa Higuera, docente de la Universidad de Jaén, asesores de la tesis titulada “*Contribution to the characterization and modeling of photovoltaic generators*” del autor Jose Ruben Angulo Abanto, para la obtención del Doble Título de Doctor en Régimen de Cotutela Internacional, dejamos constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de 23%, pero se tiene que omitir las referencias 1 y 3 del informe dado que estas son artículos elaborados por el mismo alumno producto de la tesis doctoral. El índice de similitud, omitiendo las referencias señaladas, es de 6 %, lo que está dentro del límite establecido. Así lo consigna el reporte de similitud emitido por el software Turnitin el 1/12/2022.
- Hemos revisado con detalle dicho reporte y confirmamos que cada una de las coincidencias detectadas no constituyen plagio alguno.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

Lugar y fecha: Lima, 05 de diciembre de 2022

Apellidos y nombres de asesor 1: Palomino Tofflinger, Jan Amaru	
DNI: 48866095	Firma: 
ORCID: https://orcid.org/0000-0002-2076-4046	

Lugar y fecha: Jaén, 05 de diciembre de 2022

Apellidos y nombres de asesor 2: de la Casa Higuera, Juan	
DNI: 26001227A	Firma: 
ORCID: https://orcid.org/0000-0002-1781-3417	

Acknowledgments

I want to express my gratitude to my supervisors for their effort in establishing a successful collaboration between PUCP and the University of Jaén, Spain. First, I would like to thank Dr. Jan Amaru Palomino Töfflinger for all his support, patience, and confidence during my thesis completion. Thanks to his scientific knowledge and advice, he helped me mature scientifically and personally. Second, I would like to thank Dr. Juan de la Casa Higuera, my second director and friend, for all his patience and support. To the professors who lead the MatER-PUCP research group: Dr. Andres Guerra and Dr. Rolf Grieseler.

I am grateful to the National Council of Science, Technology and Technological Innovation (CONCYTEC) for funding and stipend (236-2015-FONDECYT). I recognize the support provided by the Vicerectory of the PUCP through research activities PO0088, PI0664 and PI0737. Finally, I am grateful for the mobilization funds for stays at the University of Jaen - Spain (161-2019-FONDECYT) and the mobility aid for doctoral students (UNIA-2021) aimed at members of La Rábida 2021 Ibero-American universities. The PV-system data from Peru was obtained thanks to the financial support by PROCENCIA through contract N°045-2018-FONDECYT-BM-IADT-MU and to the company GERION INGENIERIA for giving access to the facilities and permission for monitoring the PV plant in Granada, Spain.

Thanks to everyone who has helped me develop this thesis, especially my family, for always supporting me.

Abstract

A crucial aspect of evaluating and maintaining a photovoltaic (PV) installation connected to the grid is the availability of models that describe its operation reliably in real operating conditions. The nominal power of the PV generator (P_M^*) is considered an essential input parameter, and several models have been proposed to estimate P_M^* for characterizing the PV system.

In the case of PV generators in outdoor conditions, the American Society for Testing and Materials, the International Electrotechnical Commission, and others have proposed procedures to determine the P_M^* of the generator. As part of these procedures, monitoring days with ideal conditions is mandatory, notably days with a clear sky, high irradiance values, and low wind speeds. Such restrictions can limit the number of suitable monitoring days, especially in places where clouds frequently form.

This thesis proposes a new approach that allows estimating the P_M^* with data even from non-ideal, partially cloudy days. Based on non-parametric statistics, this procedure identifies and filters out noise as well as deviations from ideal conditions of irradiance, allowing for an estimation of P_M^* with similar accuracy as for a clear-sky day. This new procedure enables the characterization of a PV generator on a daily basis without the requirement to meet ideal conditions, thus, considerably enhancing the number of suitable monitoring days. To overcome the limitation in the P_M^* estimation and considerably extend the number of monitoring days, the new procedure can be applied to ideal and non-ideal conditions, such as partially cloudy days. This procedure determines the most probable nominal power value within one monitoring day using non-parametric statistics.

In order to test the new procedure, a 109.44 kW photovoltaic plant in Granada, Spain, was monitored for six months. A referential procedure reported in the literature for large PV plants under ideal climatic conditions is first applied to estimate its nominal power. The results indicate that the nominal power can be estimated reliably in non-ideal conditions, maintaining the same precision as in ideal conditions.

Then validating the procedure for a smaller PV generator and under different conditions, two small grid-connected 1.5 kW PV arrays were used. The PV systems in question are located in two different cities in Peru: Chachapoyas (tropical highland) and Lima (coastal desert). The objective of this study in Chachapoyas was to validate the methodology in a tropical climate with a high presence of clouds but at the same time with high irradiance values above 800 W/m². According to the results obtained, under these conditions, the nominal power of the system can be calculated with reasonable certainty. As a precaution, monitoring for more than one day is recommended to obtain more data (at least 3 hours with high irradiance) to reduce uncertainties. Lima, Peru's second location under study, has a particular climate. Since the capital is located in a desert with high relative humidity values, dust deposition increases and power output decreases due to these conditions. For this purpose, the nominal power was used as a parameter to determine the maintenance schedule. Since keeping the system in optimal performance, considering this in future installations for operation and maintenance costs, is essential.

The new procedure developed in this work can be applied to facilitate technical due diligence and quality control processes for PV generators of different sizes and under different operating conditions that are being re-purchased or have been recently installed. The possibility of daily monitoring of the P_M^* also enables long-term monitoring of a PV generator to ensure the correct operation or identify possible degradation effects.

Resumen

Un aspecto crucial a la hora de evaluar y mantener una instalación fotovoltaica (FV) conectada a la red es la disponibilidad de modelos que describan su funcionamiento de forma fiable en condiciones reales de funcionamiento. La potencia nominal del generador fotovoltaico (P_M^*) se considera un parámetro de entrada esencial y se han propuesto varios modelos para estimar P_M^* para caracterizar el sistema fotovoltaico.

En el caso de generadores fotovoltaicos en condiciones exteriores, la Sociedad Estadounidense de Pruebas y Materiales (abreviatura del inglés ASTM), la Comisión Electrotécnica Internacional (abreviatura del inglés IEC) y otros han propuesto procedimientos para determinar la P_M^* del generador. Como parte de estos procedimientos, es obligatorio monitorear los días con condiciones ideales, en particular los días con cielo despejado, valores de irradiancia altos y velocidades de viento bajas. Tales restricciones pueden limitar la cantidad de días de monitoreo adecuados, especialmente en lugares donde se forman nubes con frecuencia.

Esta tesis propone un nuevo enfoque que permite estimar la P_M^* con datos incluso de días parcialmente nublados no ideales. Basado en estadística no paramétricas, este procedimiento identifica y filtra el ruido, así como las desviaciones de las condiciones ideales de irradiancia, lo que permite una estimación de P_M^* con una precisión similar a la de un día de cielo despejado. Este nuevo procedimiento permite la caracterización diaria de un generador fotovoltaico sin el requisito de cumplir con las condiciones ideales, lo que aumenta considerablemente el número de días de monitoreo adecuados. Para superar la limitación en la estimación de P_M^* y extender considerablemente el número de días de monitoreo, el nuevo procedimiento se puede aplicar a condiciones ideales y no ideales,

como días parcialmente nublados. Este procedimiento determina el valor de potencia nominal más probable dentro de un día de monitoreo utilizando estadísticas no paramétricas.

Para probar el nuevo procedimiento, se monitorizó durante seis meses una planta fotovoltaica de 109,44 kW en Granada, España. Primero se aplica un procedimiento referencial reportado en la literatura para grandes plantas fotovoltaicas en condiciones climáticas ideales para estimar su potencia nominal. Los resultados indican que la potencia nominal se puede estimar de forma fiable en condiciones no ideales, manteniendo la misma precisión que en condiciones ideales.

Luego, para validar el procedimiento para un generador fotovoltaico más pequeño y en diferentes condiciones, se utilizaron dos pequeños generador fotovoltaicos de 1,5 kW conectados a la red. Los sistemas fotovoltaicos en cuestión están ubicados en dos ciudades diferentes de Perú: Chachapoyas (altiplano tropical) y Lima (desierto costero). El objetivo de este estudio en Chachapoyas fue validar la metodología en un clima tropical con alta presencia de nubes pero al mismo tiempo con altos valores de irradiancia por encima de 800 W/m^2 . De acuerdo con los resultados obtenidos, en estas condiciones se puede calcular con razonable certeza la potencia nominal del sistema. Como precaución, se recomienda monitorear durante más de un día para obtener más datos (al menos 3 horas con alta irradiación) para reducir las incertidumbres. Lima, la segunda localidad del Perú bajo estudio, tiene un clima particular. Dado que la capital está ubicada en un desierto con altos valores de humedad relativa, la deposición de polvo aumenta y la producción de energía disminuye debido a estas condiciones. Para ello, se utilizó la potencia nominal como parámetro para determinar el programa de mantenimiento. Ya que mantener el

sistema en un desempeño óptimo, considerando esto en futuras instalaciones para costos de operación y mantenimiento, es fundamental.

El nuevo procedimiento desarrollado en este trabajo se puede aplicar para facilitar los procesos de diligencia debida técnica y control de calidad para generadores fotovoltaicos de diferentes tamaños y en diferentes condiciones de funcionamiento que se están recomprando o que se han instalado recientemente. La posibilidad de monitorear diariamente la P_M^* también permite monitorear a largo plazo un generador fotovoltaico para asegurar el correcto funcionamiento o identificar posibles efectos de degradación.



Contents

Acknowledgments	I
Abstract	II
Resumen	IV
List of Figures	IX
List of Tables	XI
Chapter I: Introduction	1
1.1 International Standards for PV Outdoor Characterization	2
1.2 Reference Procedure for Nominal Power Estimation.....	4
1.3 Hypotheses and Objectives.....	8
Chapter II: Novel procedure for the nominal power of PV generators	10
2.1 Probability Density Function.....	11
2.2 Non-Parametric Statistics	13
2.3 Applications of Non-Parametric statistics in PV	17
2.4 Proposed Procedure	21
Chapter III: Nominal Power of a PV Generator	25
3.1 Experimental Details	25
3.2 Identified Challenges in PV Generator Monitoring	28
3.3 Nominal Power following Reference Procedure	33
3.4 Nominal power following NPKDE Procedure	34
Chapter IV: Nominal Power of Small PV Generators	40
4.1 Location 1: Chachapoyas.....	41
4.2 Location 2: Lima, Dust Effect	46
Conclusions	50
Future Lines	53

Reference	55
Publications	67
Scientific publications indexed with relative quality index.....	67
Scientific publications without relative quality index	68
Communications in congresses with paper indexed in scopus.....	69



List of Figures

1.1	For five days, irradiance and DC power output were measured every 30 seconds. The yellow-marked day (16.05.2018) is the only one that meets the ASTM E2848-13 requirements for estimating nominal power.....	3
1.2	An example of a day (18.04.2018) that satisfies the test conditions required by Martinez-Moreno et al.. [29]. The DC power temperature-corrected to 25 °C versus irradiance is shown	6
2.1	Gaussian or normal distribution representation with levels of uncertainty	11
2.2	Left is the kurtosis test and right is the skewness test to identify deviations from normal distribution.....	12
2.3	Using the PDF as a guide to determine the upper and lower boundaries Image taken from [66]	20
2.4	(a) Temperature-corrected DC power versus irradiance for a sunny day with ideal conditions (18.04.2018). The hysteresis effect is observable. Data for probability density analysis is marked in red. (b) Corresponding histogram of the instantaneous nominal power values P_M^i calculated for irradiances $> 800 \text{ W/m}^2$. The black arrows indicate the three local maxima of this histogram.....	21
2.5	For irradiances above 800 W/m^2 , PDF probability density functions with three different bandwidths were applied to the same data set. PDF with optimal bandwidth $h = 0.115 \text{ kW}$ is generated using the ISJ algorithm. The red arrow indicates the mode of the photovoltaic generator for that particular day, which represents the nominal power...	23
3.1	The photovoltaic plant under study in Granada, Spain, has a datasheet nominal power of 109.44 kW. Top: Photo taken at ground level. Bottom: Satellite Image (Latitude: 37.287, Longitude: -3.057).....	26
3.2	For a sunny day with ideal conditions (18.04.2018), (a) Plane-of-array irradiance, module temperature, and DC power. (b) DC power versus irradiance shows a hysteresis effect in the morning and afternoon due to linear and non-linear behavior.....	29
3.3	PVLib simulation, (a) Irradiance, and (b) DC power simulated for four days with different azimuths..	32
3.4	Values of daily nominal power according to Reference Procedure [29]. For the 39 days analyzed, the mean value is shown together with the standard deviation.....	33
3.5	An exemplary case of a partially cloudy, non-ideal day (30.04.2018). In (a), the plane-of-array irradiation, module temperature, and DC power are shown for the entire day. (b) The temperature-corrected DC power versus irradiance, red indicates data sets with irradiances $> 800 \text{ W/m}^2$ for PDF analysis.....	34
3.6	The probability density function (PDF) for a clear-sky day (18.04.2018) compared to a partially cloudy day (30.04.2018). A mode marked by an arrow represents the nominal power with the highest	

	probability. For ideal and non-ideal conditions, in (b) daily nominal power values with the highest probability (modes).....	36
3.7	For each of the three cases (1), (2), and (3) under analysis, box plots depict the nominal power values for every single day.....	37
4.1	Top view, at the bottom left is the PERC array under analysis at Chachapoyas.....	41
4.2	An example of two days with irradiance and DC power output measured every minute. There is predominantly cloudy weather during the day	42
4.3	Boxplot analysis when setting the minimum data required in one day..	43
4.4	Estimation of the daily nominal power for the year under consideration	45
4.5	Array in the upper row is the PERC generator in Lima.....	48
4.6	A pyranometer and a calibrated module are used to estimate nominal power. Dust effects are taken into account.....	49



List of Tables

3.1	Main characteristics of the PV generator with DC parameters at STC, according to the module datasheet.....	26
3.2	Analysis of the nominal power estimation procedures from a statistical perspective	38
4.1	Main characteristics of the PV array at STC	40
4.2	Analysis of the nominal power estimation procedures from a statistical perspective	44



Chapter I:

Introduction

As the electricity demand continues to grow, power companies are increasingly turning to sustainable and clean energy sources to meet this demand. Solar and wind are becoming more cost-effective as production costs continue to decline [1], the cost of photovoltaic installations have dropped below 0.4 USD/W_p in 2021 and also have a long useful life of more than 30 years, making them more viable alternatives to traditional energy sources [2] and one of the most promising energy sources. Therefore, large-scale utility investment in photovoltaic (PV) systems has been installed rapidly [3] [4]. There is no doubt that photovoltaic power generation will become one of the primary electrical power sources in most countries where grid parity is achieved [4]. Moreover, small PV systems are installed in different geographical locations in European countries based on the amount of solar irradiation they receive and the local weather conditions [5].

PV systems perform differently due to factors [6] such as solar radiation, temperature, solar spectrum, humidity, moisture, thermal cycling, UV light exposure, etc. PV modules in an array can degrade due to one or a combination of these environmental stresses. Any reliable PV analysis or characterization requires a solid understanding and validation of data quality regardless of this purpose [7]. In this scenario, determining the status of the PV system is crucial to assess its proper operation; for an owner or investor, it is helpful to evaluate whether the projected profitability is within the expected margins. By doing so, they can make informed decisions about their investment.

Based on international standards, the **Performance Ratio** is one parameter widely used to determine the general status of a PV system [8] or large photovoltaic plants [9]. As part of the performance ratio calculation of PV plants, partial shading of strings or shading of irradiance sensors or PV plants is taken into consideration, PV plants can be evaluated

incorrectly when shadows are present [10]. Additionally, it is important to measure irradiance, ambient temperature, or module temperature [11] and know the nominal power of the PV generator.

The **nominal power** of a PV module is the maximum power that the module can produce under standard test conditions (STC). These conditions consider an irradiance of 1000 W/m², a module temperature of 25 °C, and the AM1.5G spectrum according to the IEC 60904-3 standard. In the laboratory, measuring the power under STC conditions can be done panel by panel. However, measuring each panel would not be feasible for large, installed PV generators in order to estimate the total nominal power.

Therefore, accurately determining the nominal power (P_M^*) of a PV generator under outdoor conditions is necessary [12]. This P_M^* is a critical parameter in many applications, particularly those where commercial transactions are involved, e.g., product guarantees and return on investment calculations or quality control of photovoltaic modules or for identifying defective modules or arrays [13]. Additionally, comparing the built and planned state of PV plants by means of the P_M^* can help to optimize their design and improve their overall efficiency [14], as well as understand why a PV plant's expected performance may deviate from its actual performance [15].

1.1 International Standards for PV Outdoor Characterization

PV Characterization in outdoor conditions is intended to determine the characteristics of the PV generator and compare it with the manufacturer's specifications [16]. With the Current-Voltage curve, P_M^* is calculated in accordance with the **International Electrotechnical Commission** (IEC 61829). In addition to the nominal power value, this procedure also provides the essential electrical parameters from the manufacturer's data sheet [18]. Additionally, the shape of the curve I-V gives information about detecting possible anomalies, such as disconnected arrays/modules, broken cells, shading, or

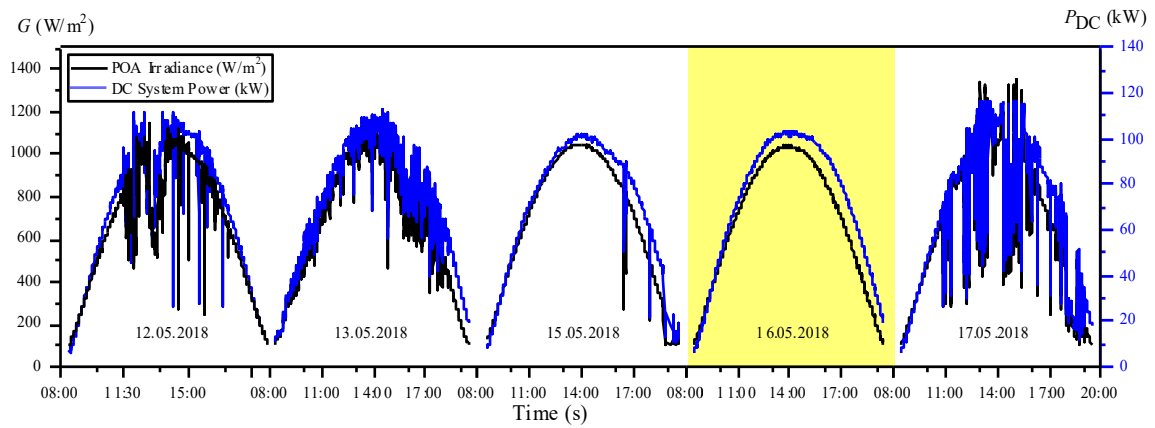


Fig. 1.1. For five days, irradiance and DC power output were measured every 30 seconds. The yellow-marked day (16.05.2018) is the only one that meets the ASTM E2848-13 requirements for estimating nominal power.

module mismatches [19] and maintenance alerts or cleaning schedules can be determined by the shape of this curve [20].

Tracing the current-voltage curve for the entirety of the generator is ideal, but that is not always possible [21][23]. As the size of a PV generator increases, so does the complexity of measuring. This is because more precautions must be taken into account, and adequate measurement equipment is required for high-power generators [22][23].

The **American Society for Testing and Materials (ASTM)** has developed a standard, E2848-13[24] to measure the capacity of photovoltaic systems in outdoor conditions. By setting so-called Reporting Conditions (RC), this procedure provides a way to analyze and determine the installed capacity of a PV system at RC. The RC are user-defined and agreed-upon values; for instance, RC could be an irradiance of 1000 W/m^2 , an ambient temperature of $20 \text{ }^\circ\text{C}$ and a wind velocity of 1 m/s [25]. In order to comply with standard E2848-13, a minimum of 10 hours of solar irradiance data must be collected. This data must show an average of 1000 W/m^2 or more solar irradiance [26]. The E2848-13 can be used to test and verify the performance of newly installed DC or AC systems and monitor the PV system's performance.

For accurate characterization in outdoor conditions, the standard requires having clear skies, low wind speeds, and appropriate measurement equipment. The ASTM E2848-13 specifies that these outdoor conditions may not typically be encountered during a monitoring period at some locations [19]. Exemplarily, five days of monitored DC power and irradiance are shown in Fig. 1.1, but only one day, highlighted in yellow, meets the necessary conditions to calculate P_M^* .

NREL introduced another methodology for evaluating PV system's nominal power under so-called Performance Test Conditions (PTC) [27]. It involves finding the best correlation between the power generated by the monitored system, the irradiance, the wind speed, and the ambient temperature. Once the parameter settings have been determined, the power is evaluated under Standard Test Conditions [28]. As a limitation, this procedure also requires high irradiance, close to 1000 W/m^2 , which is not always the case.

1.2 Reference Procedure for Nominal Power Estimation

Due to the lack of an international standard for evaluating the performance of PV systems with general-purpose instrumentation, Martinez et al. [29] proposed a procedure for calculating the P_M^* of a PV generator; and compared its results to those obtained by applying (IEC 61829).

In [29] they provide an experimental procedure for calculating the nominal power of a photovoltaic (PV) generator under outdoor conditions, which has the following considerations:

1. High-precision and calibrated instruments to measure power output and operating conditions

The **DC output power** monitoring requires instruments with a measurement uncertainty of around 0.2% in accordance with the IEC 61724-1 standard. PV modules have these requirements [32], but larger generators have similar considerations.

Usually, thermopile pyranometers or calibrated cells are used to measure the irradiance received by a PV generator. Compared to other sensors, pyranometers have better accuracy with errors in the range of 1-5% [30]; however, their spectral response is spectrally flat in the range of 300-3000 nm, whereas PV modules present different ranges in the spectral response. On the other hand, the calibrated cell can be chosen of similar technology as the modules of the PV generator, such as crystalline silicon. This way, the cell responds to the same spectral range as the installed technology in the PV generator. Therefore, there is no consensus regarding the appropriate instrument to measure irradiance as it depends on the purpose of the measurement [31][32]. For instance, one can use a pyranometer when the purpose is to precisely measure the incident irradiance independent of module technology and a calibrated cell when the purpose is to measure the captured irradiance by the module. For estimating the P_M^* , the use of a cell is recommended since the angular and spectral effects that are inherently included in the response of the cells contribute to better accuracy in the modeling of the PV generator.

The module temperature can be affected by multiple climatic factors, such as irradiance, ambient temperature, humidity, wind speed, and airflow [33][34]. Moreover, the modules can present significant temperature differences throughout the plant due to different climatic conditions, for example, differences in wind impact between the center and the border of the generator. Because the models do not consider temperature gradients, reproducibility of the results is difficult. Module temperature is typically measured with the PT100 sensor on the back or with a calibrated photovoltaic module that allows calculating the module's temperature from its open circuit voltage [29].

2. Weather conditions

A minimum of one day with clear skies and low wind speed is required during the data measurement time. As well as experimental data filtering, which considers high

irradiance values, values between 800 - 1000 W/m² [35], and eliminates anomalous values caused by shadows or inverter saturation, or non-linear PV module efficiency. The ASTM standard in Fig. 1.1 also considers these weather conditions.

3. Measurement time

The **plane-of-array irradiance (G)**, **module temperature (T_m)** and **DC power (P_{DC})** values have to be measured at least every minute or quicker for at least one full day.

Based on these three considerations, the nominal power in outdoor conditions can be calculated. The temperature dependence, the incident irradiance, and the power output can be expressed in empirical equations [36] or linear dependence on the operation conditions [37].

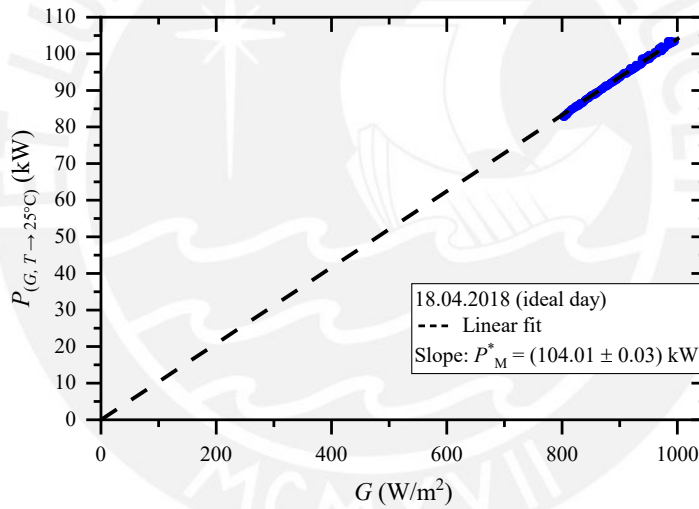


Fig. 1.2. An example of a day (18.04.2018) that satisfies the test conditions required by Martinez-Moreno et al. [29]. The DC power temperature-corrected to 25 °C versus irradiance is shown.

The Reference Procedure suggests translating the measured DC output power to the DC power temperature-corrected to 25 °C ($P_{DC} \rightarrow P_{(G,T \rightarrow 25^\circ C)}$), which is expressed as:

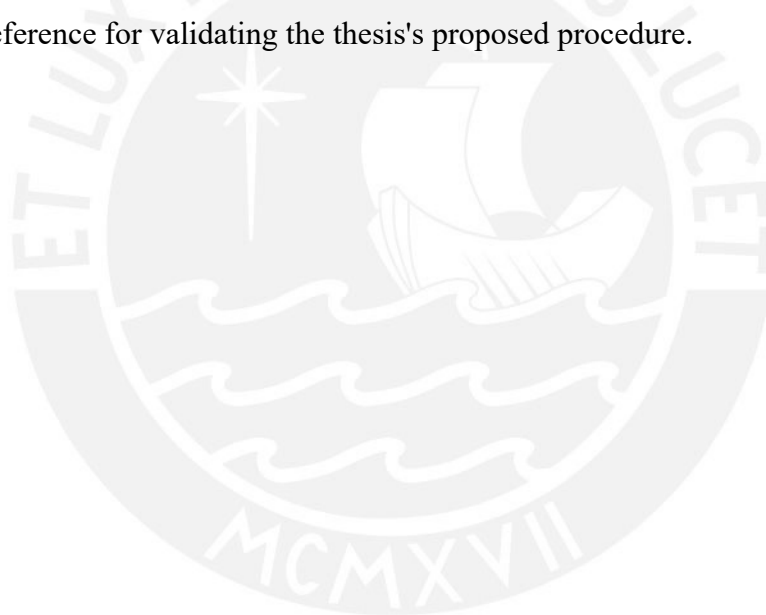
$$P_{(G,T \rightarrow 25^\circ C)} = \frac{P_{DC}}{(1 + \gamma(T_m - T_m^*))} \quad (1.1)$$

where γ is the power temperature coefficient provided by the manufacturer ($\%/^{\circ}\text{C}$), T_m^* is the module or cell temperature at STC (25°C).

Then, the slope of the linear regression ($800\text{-}1000\text{ W/m}^2$) gives the nominal power P_M^* , through:

$$P_{(G,T \rightarrow 25^{\circ}\text{C})} = P_M^* \times \frac{G}{G^*} \quad (1.2)$$

Where, G^* is the irradiance at STC (1000 W/m^2). Hence, by performing a linear global fit, one can obtain P_M^* with data from a single day. In Fig. 1.2, the nominal power is calculated using the reference procedure described above under clear skies. In this case, the slope gives an experimental nominal power of $P_M^* = (104.01 \pm 003)\text{kW}$. This value is used as a reference for validating the thesis's proposed procedure.



1.3 Hypotheses and Objectives

Photovoltaic generators installed on a site can be successfully characterized in terms of their nominal power using the **Reference Procedure**. However, it requires the specific climatic conditions mentioned above, and therefore days without these conditions during the monitoring process have to be excluded, such as days with high irradiance values but with clouds. For extending the reference procedure to estimate P_M^* , this thesis developed a method that can extend the characterization to days when monitoring conditions are not ideal. For instance, in Fig. 1.1, four days do not meet the requirements because they deviate considerably (12.05., 13.05. and 17.05.2018) or slightly (15.05.2018) from the highlighted ideal day. Applying the ASTM equations, IEC standards, or Reference Procedures lead to inaccurate estimates of P_M^* due to their high level of noise.

In summary, there is no dedicated standard or recognized guide on how to characterize a PV generator under any weather conditions, which motivates and highlights the need for further research. The hypothesis is based on the notion that non-parametric statistics can be applied to identify operating conditions outside a system's expected behavior and filter those out to estimate the nominal power of the PV generator using the remaining suitable data.

The novel procedure effectively filters out noise and deviations from ideal irradiance conditions and estimates the most probable P_M^* value with the remaining data for any day. Thus, daily characterization of a PV generator or array can be done without requiring that ideal conditions are fully met.

The **general objective** of the thesis is to focus on the correct estimation of the Nominal Power and uses statistical tools to test its validity. In the thesis, the following **specific objectives** were achieved:

- 1) The development of statistical processing procedure for estimating the nominal power under varying operating conditions, **see chapter 2**.

- 2) Validate the new proposed procedure against the reference procedure described above for a 109.4 kW photovoltaic generator located in Granada, Spain, **see chapter 3.**
- 3) Characterize the array of a generator in Chachapoyas, Peru, under tropical highland conditions, **see chapter 4.**
- 4) Determine how dust affects the nominal power calculation of a generator located in Lima, Peru, and determine the generator's maintenance schedule for better performance, **see chapter 4.**



Chapter II:

Novel procedure for the nominal power of PV generators

As a general basis for the interpretation of parametric statistics, sample data are assumed to come from a probability distribution that is adequately modeled with parameters that allow identifying the quality of a measurement. By taking a set of n measurements $(x_1, x_2 \dots x_n)$, the mean is defined by the following formula:

$$\mu = \frac{x_1 + x_2 \dots + x_n}{n} \quad (2.1)$$

Upon determining the mean μ , it will be necessary to determine the average deviation from the mean in order to determine the measurement precision, which is then calculated by the standard deviation. Standard deviation (σ) is the measurement of this deviation:

$$\sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{n - 1}} \quad (2.2)$$

The parameter X is then represented by μ and σ ($X = \mu \pm \sigma$). The calculations from eq. (2.1) and (2.2) are derived from a representation of the data in the form of a Gaussian distribution, also known as a Normal distribution. As a general rule, most experiment measurements will assume and follow a Normal distribution $N_{(x;\mu,\sigma)}$:

$$N_{(x;\mu,\sigma)} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (2.3)$$

An effective way to understand a normal distribution is to describe it graphically and then identify its most important parameters. An example of the typical shape of a Normal distribution, seen in Fig. 2.1, suggests that the data is centered on the main peak of the curve, which is defined by equation eq. (2.3). Furthermore, the mean, median, and mode are all the same and centered because most of the values are in the center of the distribution.

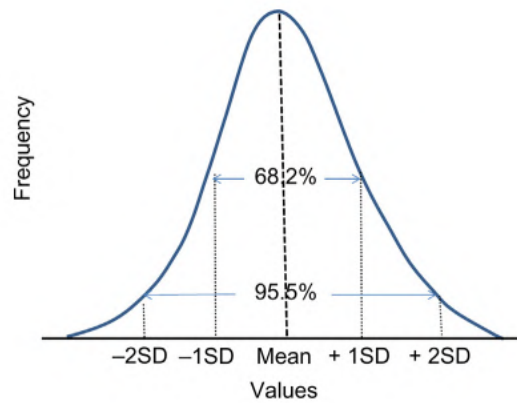


Fig. 2.1 Gaussian or normal distribution representation with levels of uncertainty.

There are a number of other important characteristics in normal distributions, such as the precision of the measurement represented by the Standard Deviation. The Standard Deviation of the data is shown with 68.2 % in the experiments. With a greater range, however, greater precision is achieved by representing 95.5% (2σ) and 99.7% (3σ) of the all-data values.

2.1 Probability Density Function

In order to standardize the frequency distribution, the probability density function (PDF) is a concept that must be defined to determine the frequency density [38]. An analysis of stochastic processes and probability theory uses this PDF to represent a random phenomenon. A random variable consists of values where some range may be assigned to it in order to represent the range of possible values. Hence, it has a close relationship with frequency distributions in this sense. Hence, the PDF is characterized by a distribution function ($f(x)$) of a random variable x .

There are two fundamental requirements that the PDF must meet. First, all values of the random variable must be nonnegative, and second, the integral of the PDF over all values of the random variable must be one. Therefore, the PDF is defined as follows:

$$\int_R f(x) dx = 1 \quad (2.4)$$

where R is the total region of the measurement space and $f(x)$ contains the experimental information. According to the parametric function, $f(x)$ is considered a member of the family of distributions that constitute a parameterized distribution.

The shape of a PDF is not always that of a normal curve or Gaussian. In some cases, the data are densely concentrated in the center and in other cases they are dispersed widely. To determine if there is a bias in the measurement that can cause an asymmetry in the normal curve, some tests are performed (Kurtosis and Skewness).

A kurtosis is shown to the left of Fig. 2.2. This measures how flat or dispersed a PDF is relative to a normal distribution for a sample or an indicator of concentration values at the center of a normal distribution. A positive kurtosis value indicates that the data is centered at the center of a normal distribution and is termed leptokurtic. For negative values of kurtosis, it indicates that the distribution is significantly dispersed and is referred to as platykurtic.

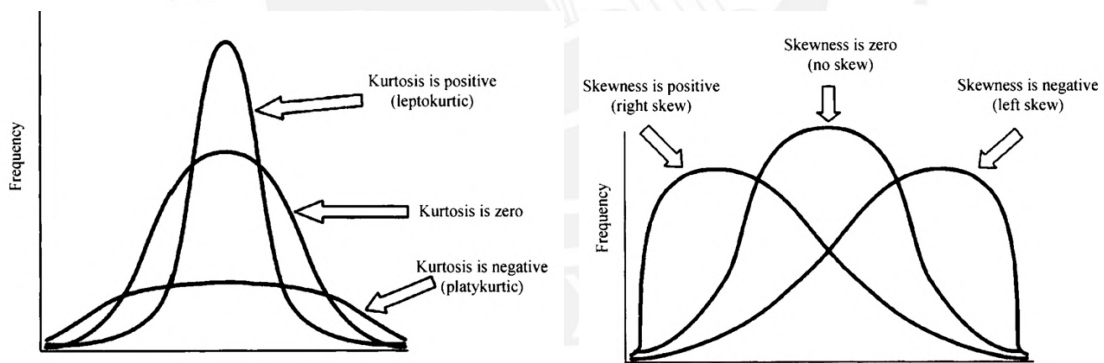


Fig. 2.2. Left is the kurtosis test and right is the skewness test to identify deviations from normal distribution.

Data collection bias is represented by a biased (non-symmetric) sample, which is calculated using the arithmetic mean, right of Fig. 2.2. A distribution is skewed to the left or has a negative skew when your data is concentrated on the right side of the curve or if your tail gets longer for values below the mean. The distribution data will be skewed to

the right when they are concentrated on the left side of the curve. Positive skewness is the result of a right-skewed distribution.

Unknown data often have a normal distribution, and this is one of the most commonly assumed assumptions. This assumption is generally made due to the normal distribution's versatility and good behavior. In most cases, the distribution of the measurements in a laboratory is normal because most uncertainty in measurements comes from the instruments used. In cases where experimental data does not approximate a known function, non-parametric statistics can be used.

2.2 Non-Parametric Statistics

Non-parametric statistics is a tool used in data science that is often referred to as the study of the distribution or the analysis of given data. Compared with parametric methods, non-parametric methods do not require assumptions such as observed independence and normal distributions for populations, which are required for parametric methods [39]. Therefore, this non-parametric method is more robust because it does not make assumptions before calculating the PDF. For instance, a histogram is the simplest form of non-parametric density estimation. The PDF in non-parametric estimation is divided into two categories: the Kernel Density and the bandwidth. The following provides the basis for understanding the new methodology. Further theoretical details can be found in section 3 in [40].

Kernel Density Estimation

The PDF of an unknown distribution can be calculated with the kernel density estimation (KDE) [41]. The KDE defined as the convolution of multiple functions (e.g. Gaussianas, triangular, delta, etc.) [42]. The KDE is expressed as

$$K(u) = (2\pi)^{-1/2} \exp\left(-\frac{1}{2}u^2\right) \quad (2.5)$$

$K(u)$ satisfies condition eq. (2.4) and to be symmetric, u is a general variable. Therefore, for a random parameter (x_1, x_2, \dots, x_n) from an unknown density, the Kernel Density Estimator looks like:

$$\hat{f}_{(x,h)} = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{\mu - x_i}{h}\right) \quad (2.6)$$

Where $\hat{f}_{(x)}$ is the estimated PDF from de KDE, n represents the sample size, h is the bandwidth and has the interpretation of the standard deviation σ , also called a smoothing parameter, $K_{(z)}$ is the kernel function of z , x_i is the random variable, and μ is the mean of the set of values in n .

The kernel function $K_{(z)}$ is the base function that keeps counts of the x_i . The simplest kernel, $K_{(z)} = \text{box function}$, is presented in the histogram. This central rectangular function's discontinuous or non-uniform nature makes identifying its peak probabilities difficult. The proper kernel that smooths the curve is the Gaussian function [43], which will be considered here. Therefore, derived from eq. (2.5) and (2.6), the KDE will be:

$$\hat{f}_{(x,h)} = \frac{1}{n} \sum_{i=1}^n \frac{1}{\sqrt{2\pi}h} \exp\left(-\frac{(x - \mu)^2}{2h^2}\right) \quad (2.7)$$

Performance Criteria

The accuracy of KDE depends strongly on the bandwidth value, and the performance criteria will help to estimate the h value [44]. Given a target PDF that will estimate $(f_{(x)})$ and a KDE function $\hat{f}_{(x,h)}$, the Mean Squared Error (MSE) is an appropriate measure of the closeness of $\hat{f}_{(x,h)}$ to $f_{(x)}$:

$$MSE(\hat{f}_{(x,h)}) = E(\hat{f}_{(x,h)} - f_{(x)})^2 \quad (2.8)$$

Variance and bias are rewritten as:

$$MSE(\hat{f}_{(x,h)}) = Var(\hat{f}_{(x,h)}) + (E\hat{f}_{(x,h)} - f_{(x)})^2 \quad (2.9)$$

In the case of a local error at x , a global error criterion is usually considered over the entire domain of x , that is, the Integrated Squared Error (ISE $\hat{f}_{(:,h)}$). Following the eq. (2.8), the expected value is analyzed by Mean Integrated Squared Error (MISE):

$$MISE(\hat{f}_{(:,h)}) = E(\text{ISE } \hat{f}_{(:,h)}) = E \int (\hat{f}_{(x,h)} - f(x))^2 dx \quad (2.10)$$

MISE is the average value of the global ISE eq. (2.10) and it can be rewritten as:

$$MISE(\hat{f}_{(:,h)}) = \int \text{Var}(\hat{f}_{(x,h)}) dx + \int \text{Bias}^2 \hat{f}_{(x,h)} dx \quad (2.11)$$

For an optimal h value the $MISE(\hat{f}_{(:,h)})$ will be minimum or:

$$h_{MISE} = \text{argmin} (MISE(\hat{f}_{(:,h)})) \quad (2.12)$$

Calculating eq. (2.12) is not trivial, and eq. (2.11) needs a Taylor expansion for variance and bias:

$$\text{Var}(\hat{f}_{(x,h)}) = (nh)^{-1} h^2 R(K) f(x) + o((nh)^{-1}) \quad (2.13)$$

$$E\hat{f}_{(x,h)} - f(x) = \frac{1}{2} h^2 u_2(K) f''(x) + o(h^2) \quad (2.14)$$

Where:

$$R(K) = \int K(x)^2 dx$$

$$u_2(K) = \int x^2 K(x)^2 dx \quad (2.15)$$

$$R(f'') = \int f''(x)^2 dx$$

By adding eq. (2.13) and (2.4), the possibility of obtaining:

$$MISE(\hat{f}_{(:,h)}) = (nh)^{-1} h^2 R(K) f(x) + \frac{1}{2} h^4 u_2(K)^2 f''(x)^2 + o((nh)^{-1} + h^4) \quad (2.16)$$

After that, perform the integration in dx

$$MISE(\hat{f}_{(:,h)}) = (nh)^{-1} R(K) + \frac{1}{4} h^4 u_2(K)^2 R(f'') + o((nh)^{-1} + h^4) \quad (2.17)$$

In the previous equation, the AMISE was defined as:

$$AMISE(\hat{f}_{(:,h)}) = (nh)^{-1}R(K) + \frac{1}{4}h^4u_2(K)^2R(f'') \quad (2.18)$$

As a result of differentiation with respect to h , the AMISE is able to find a solution by calculating the root of the derivative. As a result, the optimal bandwidth will be the following:

$$h = \left(\frac{R(K)}{u_2(K)^2R(f'')n} \right)^{1/5} \quad (2.19)$$

Bandwidth selectors for kernel density estimation

In the normal distribution, the bandwidth corresponds to the interpretation of the standard deviation. Choosing the optimal bandwidth is not easy, and there is not just one best method that can be applied universally to solve eq. (2.19) because this is not a trivial task [45]. Three types of selectors should be distinguished: rule-of-thumb, cross-validation (CV) and plug-in (PI) selectors.

Rule-of-thumb

The unknown density is assumed to belong to the normal distribution[46]. Specifically, for the $K(x)$ gaussian or normal kernel function for the case of a single dimension, eq. (2.15) is rewritten as:

$$\begin{aligned} R(K) &= (\sqrt{2\pi})^{-1} \\ u_2(K) &= 1 \\ R(f'') &= 3(\sqrt{8\pi})^{-1} \end{aligned} \quad (2.20)$$

Then, from eq. (2.19) the optimal bandwidth is as follows:

$$h \approx 1.06 \hat{\sigma}n^{-1/5} \quad (2.21)$$

$\hat{\sigma}$ represents the estimated standard deviation based on the input data. In the case of a PDF that does not match a normal distribution, the rule-of-thumb is easy to compute but widely inaccurate.

Cross-validation

This method uses a subset of a dataset to evaluate another subset of the same data group.

In other words, cross-validation is a method that uses a subset of the data to evaluate

another subset of the same data group, referred to as the validation data [47]. This process produces an unbiased estimate of the error in the model and relies on the dependency of h on the least squared. This method is also known as ‘unbiased cross validation’. Eq. (2.18) is observed in the sum of two terms, h^{-1} and h^4 . To obtain the minimum AMISE, h should not be too small nor too large.

Plug-in (PI) selectors: Improved Sheather-Jones algorithm

Introducing a bias in the eq. (2.19) for h allows to solve for the optimal value of h , since first- and second-order derivatives $R(f'')$ need to be calculated [48]. This bias contains a function that is derived using a recursive form. This method is applied when the distribution of measurements cannot be assumed normal but rather multimodal or other non-normal distributions.

Fast Fourier Transformer-based algorithms

Fast Fourier Transformer-based algorithms or fast Fourier transform algorithms is a set of mathematics that allows computational speedups in the discrete Fourier transform. This is done through the use of data structures or computation optimizations. In this algorithm, the sampled data is divided into discrete frequency bins for analysis. The basis of sampling is to reduce the data size to be analyzed by extracting representative data, which preserves most information from this experimental data.

2.3 Applications of Non-Parametric statistics in PV

In renewable energy studies, KDE has found a different application in probability distribution. For instance, the KDE is also used to study the probability distribution of wind power density and power output of wind turbines [43]. The main applications of non-parametric statistics in PV will be discussed:

PV Forecasting Error Distributions

As renewable energy becomes increasingly in demand, the need for accurate forecasting of PV energy production grows alongside. Precise and reliable solar energy forecasts are essential for economic viability [49].

The Bayesian bootstrapping method is a statistical technique that can be used to improve the accuracy of PV energy forecasting models. This method works by resampling data from the original dataset to create new, synthetic datasets. These synthetic datasets are then used to train and test the forecasting model [50], when capturing non-symmetric profiles of global hourly irradiation time, machine learning can be used as a non-parametric clustering technique [51].

Forecasting can be done by using the distribution functions generated by KDE, which uses recent data sets to create distribution curves. [52] used historical data from previous campaigns to determine what case was most similar to the current one and thus used that point as a prediction.

Forecasts are likely to have uncertainties, and additional information can provide those uncertainties. By using the standard error distribution of predictions, forecasts become more accurate by including a prediction interval [53]. In addition, [54] demonstrated that KDE represents probability distributions more accurately than parametric methods. [55] used the prediction error with KDE to evaluate their interval prediction method for solar generation.

As a feasible solution to reduce the uncertainty level of photovoltaic generator forecasts, deep learning methods, probability density functions, and meteorological parameters have been applied [56]. The study in [57] uses an extreme learning machine (ELM) to perform interval forecasting for generated energy, and the forecast error distribution is fitted by estimating the kernel density and compared with a normal distribution model, a

logistic distribution model, demonstrating that the non-parametric method is effective for describing the PV energy forecast error distribution.

Fault Detection

Identifying, detecting, and classifying faults from data is vital in grid-connected systems. A framework for fault diagnosis using data analysis has been developed [58]. Based on the fact that energy distributions are never perfectly Gaussian, [59] proposes a non-parametric statistical algorithm to monitor the energy performance of photovoltaic plants and detect anomalies. Through KDE, statistical-based fault detection attempts to identify outliers in measurement data [60]. Since voltage is not strongly affected by changes in irradiance [61], these tools rely mainly on the deviation between the observed and expected data. A comparison was conducted between normalized and non-parametric statistics for the distribution of k-nearest neighbor distances, with better results obtained when calculations were made with KDE for fault detection [62]. For shading or dust problems, [63] highlights non-parametric detection thresholds' superiority over conventional parametric methods.

Other techniques require threshold values for proper fault detection, as shown in Fig. 2.3, indicating voltage, current, or power failures by defining a threshold boundary. The probability distribution model is constructed using non-parametric statistics [64]. Despite this, [65] uses a rule-based method to detect and classify faults without setting a threshold.

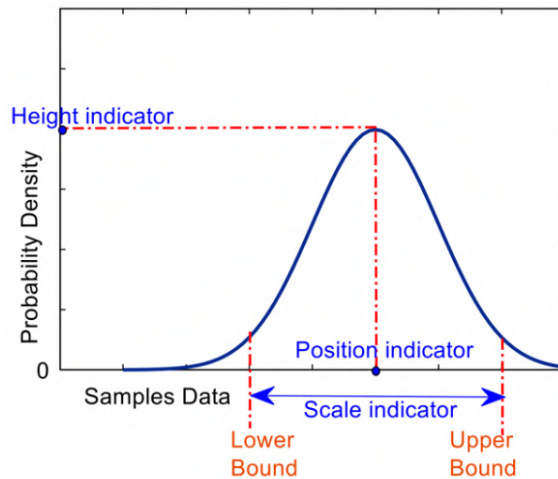


Fig. 2.3 Using the PDF as a guide to determine the upper and lower boundaries Image taken from [66]

Other Applications

[67] suggests that KDE is a powerful tool for dealing with skewed or unbalanced datasets since it generates new experimental samples naturally and statistically. For the calculation of probabilistic power flow for unbalanced power distribution, a non-parametric algorithm is used to analyze certain energy generation and load [41].

A concept known as the ‘Duck Curve’ is important in the operation and planning of electrical systems. It describes electricity's net load curve. The uncertainty and variability of electricity's net charge must therefore be accurately modeled, which is accomplished with the probabilistic Duck Curve. KDE produces the best estimate of probability distributions compared to beta, normal, Weibull function, and data histogram [68]. To identify generation patterns that vary seasonally and diurnally in [69], a KDE-based model was used, which allowed KDE to identify different PDFs based on the season and time of day.

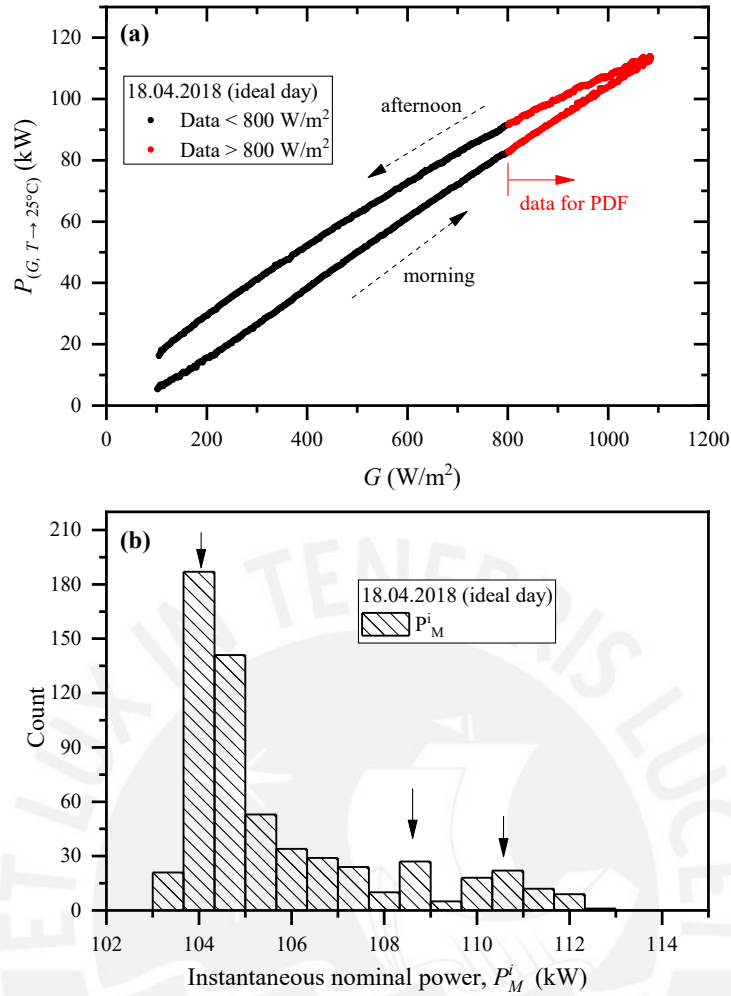


Fig. 2.4. (a) Temperature-corrected DC power versus irradiance for a sunny day with ideal conditions (18.04.2018). The hysteresis effect is observable. Data for probability density analysis is marked in red. (b) Corresponding histogram of the instantaneous nominal power values P_M^i calculated for irradiances $> 800 \text{ W/m}^2$. The black arrows indicate the three local maxima of this histogram.

2.4 Proposed Procedure

The computation of the nominal power with the reference procedure requires the indirect application of parametric statistics since the error distribution is assumed to be normal or Gaussian. In order to overcome the weather limitations of the standard method of calculating nominal power under ideal conditions, a new method based on non-parametric

statistics is proposed. Compared to previous applications in photovoltaics, this one offers a new approach. To begin with, let us define the instantaneous nominal power P_M^i for an i -th data set:

$$P_M^i = P_{(G,T \rightarrow 25^\circ\text{C}),i} \times \frac{G^*}{G_i} \quad (2.21)$$

In Fig. 2.4. (a), the temperature-corrected DC power $P_{(G,T \rightarrow 25^\circ\text{C})}$ over the POA irradiance G for the exemplary ideal day (18.04.2018) is depicted. There is one effect that will be discussed in detail chapter III, which is the hysteresis. According to eq. (2.21), local nominal power P_M^i for each set of data points i is calculated, but only for irradiances $> 800 \text{ W/m}^2$ where the most linear behavior is expected. Fig. 2.4 (b) presents the histogram of the resulting values P_M^i for this particular day. The histogram has multiple distribution nodes or peaks, indicated by black arrows. Thus, it becomes evident that a method is required to calculate unknown or non-parametric probability distributions based on the parameters discussed, such as the kernel, bandwidth, and sampling algorithm (FFT).

Using eq. (2.21) the respective local nominal power P_M^i for each set of data points i is calculated, but, only for irradiance values $> 800 \text{ W/m}^2$ where a linear power response is expected. From eq. (2.6) The KDE is expressed as:

$$\hat{f}_{(p)} = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{p-p_i}{h}\right) \quad (2.22)$$

Where p_i is the random variable, and p is the mean of the set of values in n . Then, the Gaussian kernel ($K(z)$) is considered due to its ability to adequately smooth the PDF curve and for h , when the distribution is multimodal, the previously described ISJ method is particularly useful. For a large data set [70], the Fast Fourier Transformer technique is used to perform the calculation.

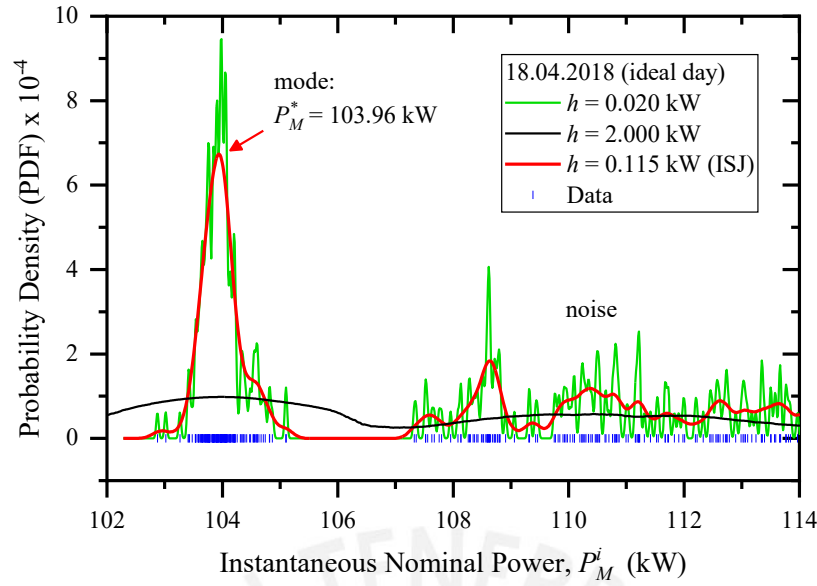


Fig. 2.5. For irradiances above 800 W/m^2 , PDF probability density functions with three different bandwidths were applied to the same data set. PDF with optimal bandwidth $h = 0.115 \text{ kW}$ is generated using the ISJ algorithm. The red arrow indicates the mode of the photovoltaic generator for that particular day, which represents the nominal power.

Based on the above, calculating the PDF and determining its main peak or mode can be done. As shown in Fig. 2.5, the nominal power P_M^* is the value with the highest probability density, highlighted with red arrow.

For this particular clear sky day (18.04.2018), three values are shown to illustrate the effect of bandwidth. A low bandwidth (e.g. $h = 0.02 \text{ kW}$) does not correctly represent the underlying data and generates many oscillations in the PDF. Using high bandwidths (e.g. $h = 2.00 \text{ kW}$) might remove peaks from the underlying data distribution, making it even more challenging to identify them. The ISJ algorithm finds a bandwidth ($h = 0.115 \text{ kW}$) that balances overestimating and underestimating, demonstrating higher reliability in the automated calculation of the PDF and, thus, the nominal power estimations.

The nominal power calculated with Reference Procedure ($P_M^* = 104.01 \text{ kW}$) for the same day (18.04.2018) is close to that calculated in fig. 2.4 ($P_M^* = 103.96 \text{ kW}$), making the proposed method reliable for the same set of evaluated data. Due to the process of calculating the nominal power of the entire generator, different sources of error can occur during measurement. Minor peaks are usually associated with noise, outliers, and/or uncertainty. Non-Parametric Kernel Density Estimation (NPKDE) can be referred to as the proposed methodology and a detailed discussion of days under non-ideal conditions follows in the next chapter.



Chapter III:

Nominal Power of a PV Generator

The purpose of this chapter is to present the results of calculating the nominal power of a utility-scale PV generator following the Non-Parametric Kernel Density Estimation (NPKDE) methodology described in section 2. Firstly, a description of the PV generator and the measurement equipment used for the measurement analysis is given. Challenges when measuring a large PV generator, giving rise to the phenomenon called ‘hysteresis,’ are identified. A detailed description and discussion of the hysteresis are given, evaluating the reliability of the methodology to calculate the nominal power in these experimental conditions. The results obtained using the NPKDE methodology are compared with those obtained using the standard reference method described in the introduction under ideal measurement conditions.

3.1 Experimental Details

A 109.44 kW PV generator, which has been in operation since 2008, is analyzed to calculate its nominal power at working conditions. Fig. 3.1 shows the ground view (top) and the satellite image (bottom) of the PV system under study. Table 3.1 summarizes the main electrical characteristics of the generator.

At the beginning of its operation, the photovoltaic arrays operated with a two-axis tracker for higher annual production, resulting in economic benefits [71]. However, due to internal decisions of the owner, tracking was stopped, and all arrays were oriented south and tilted at 30° degrees [72] during the experimental campaign. Most arrays are aligned and coplanar, as shown in the satellite image; however, some arrays are not aligned and coplanar. This misalignment is caused by uncertainties in the axes' mechanical positioning and uneven mounting ground.



Fig. 3.1. The photovoltaic plant under study in Granada, Spain, has a datasheet nominal power of 109.44 kW. Top: Photo taken at ground level. Bottom: Satellite Image (Latitude: 37.287, Longitude: -3.057).

Table 3.1. Main characteristics of the PV generator with DC parameters at STC, according to the module datasheet.

Characteristics of the PV generator	
Current at the maximum power point (A)	257.60
Voltage at the maximum power point (V)	574.20
Power at maximum power point (kW)	109.44
Power temperature coefficient (%/°C)	-0.43
Number of modules per string connected in series	18
Number of strings connected in parallel	32

The PV generator is monitored by measuring its operating conditions (irradiance and module temperature) and DC power at maximum power point (MPP) at the inverter. The

irradiance measurement was done with a calibrated PV module of the same technology and spectral response as the PV modules installed in the PV generator under study [32]. In addition, the PV module was installed at the same angle and orientation as the adjacent string. In Fig. 3.1, an arrow indicates the position of the calibrated module aligned with the adjacent string.

Despite cloudy or partially sunny irradiance days, the photovoltaic irradiance sensor can accurately measure solar irradiance for PV performance analysis [73]. Several factors affect the irradiance measurement quality in the PV generator, including the uncertainty related to the short-circuit current of the calibrated PV module (1.8%, according to the calibration certificate) and the uncertainty related to the shunt resistance (0.5%, according to manufacturers' specifications). The uncertainty introduced by the resolution of the datalogger (analog-digital conversion-16 bits, with a full scale of 100 mV) and the accuracy of this device (0.1% of a 100 mV-full scale) is negligible in this uncertainty budget. If a calibrated solar module is available, using a module is preferable to using a pyranometer.

Indirect measurements of module temperature can be obtained by measuring the open-circuit voltage of calibrated modules [74], or by sensors (PT100 or thermocouple) on the module's backside [75]. To measure T_m under operating conditions for the PV generator in this work, a J-type thermocouple was placed on the backside of a PV module [72].

T_m measurements are affected both by uncertainty associated with the accuracy of J-type thermocouples (± 1.5 °C, according to the manufacturer's specifications), and the uncertainty due to the accuracy of the datalogger (± 1.7 °C, according to the manufacturer's specifications).

Temperature and irradiance measurements were made with a photovoltaic module attached to the string adjacent to the irradiance sensor near the geometric center of the

photovoltaic generator, see photo in Fig. 3.1. Throughout the experiment, temperature and irradiance measurements were made with the same module under the same operating conditions.

A calibrated YOKOGAWA WT1600 wattmeter (uncertainty of less than 0.5%) was used to measure the DC power at the inverter input. From March 27th to September 30th, 2018, the irradiance, module temperature, and DC power were recorded every 30 seconds, as recommended in [29].

3.2 Identified Challenges in PV Generator Monitoring

Different challenges were identified in monitoring the operating conditions during the experimental campaign, which contributed to uncertainties in evaluating the PV generator's operating conditions and electrical parameters. One of the biggest challenges in PV monitoring is measuring the operating temperature. Module temperature analysis and modeling revealed a temperature difference of up to 2.5°C between the cell and the back-side of the module [76], giving uncertainty of the actual temperature of the PV system. Furthermore, a single module's temperature may not represent the operating temperature for the entire plant. When wind speed increases, the difference between cell and ambient temperature decreases [77]. It has been demonstrated that wind speeds and directions change randomly, resulting in operating temperature inhomogeneities or temperature gradients up to 10 °C in a PV plant [78]. Due to this, a single local temperature measurement in a PV generator rarely represents its average temperature, increasing the uncertainty of the measurement.

Further inhomogeneities in a PV generator derive from slightly different aligned strings. Such misalignments may arise from the uneven ground or slightly different mounting structures, as in most plants such as the one under investigation in Fig. 3.2. Consequently, different strings may capture distinct levels of plane-of-array irradiance (POA)

simultaneously. Therefore, the local irradiance measurements in a PV generator may not represent the irradiance captured by the entire generator, thus, introducing an additional source of uncertainty.

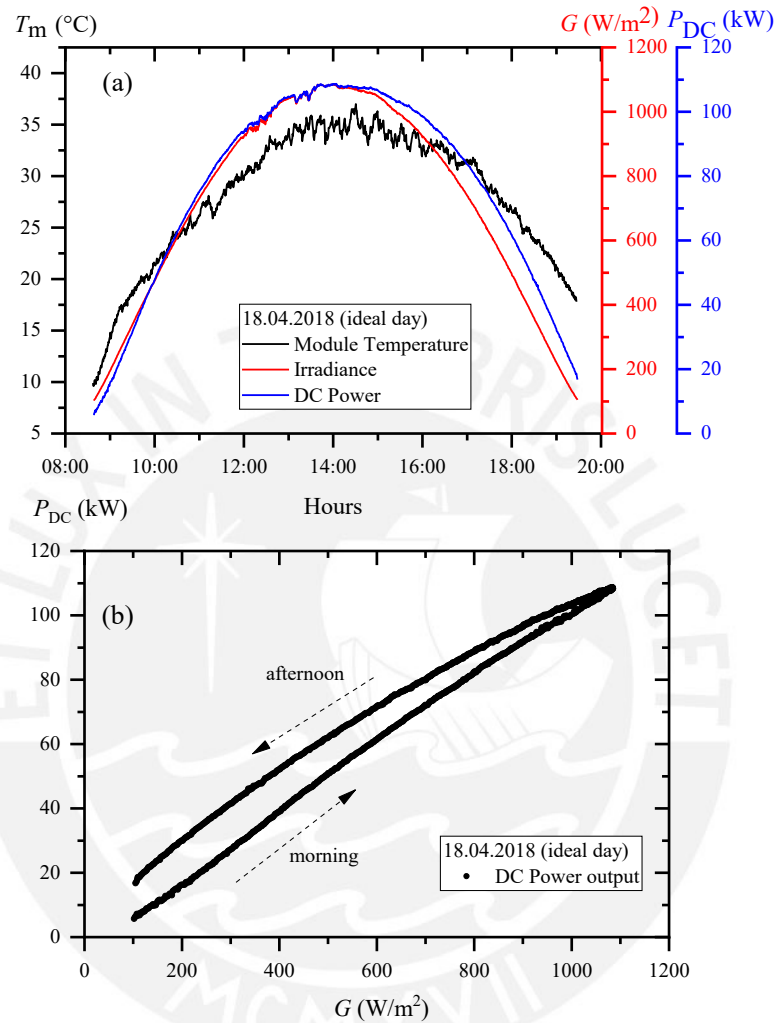


Fig. 3.2. For a sunny day with ideal conditions (18.04.2018), (a) Plane-of-array irradiance, module temperature, and DC power. (b) DC power versus irradiance shows a hysteresis effect in the morning and afternoon due to linear and non-linear behavior.

An exemplary day with ideal operating conditions is shown in Fig. 3.2 (POA irradiance, module temperature, and DC power). During morning hours, DC power and irradiance data appear to be aligned. During afternoon hours, there is a temporal delay between these two parameters. The delay is likely caused by the misalignment of the irradiance sensor

and some strings. Fig. 3.1 shows that the string used to align the calibrated module is slightly different from other PV strings. The irradiance sensor was placed in this mean orientation to minimize the impact of this misalignment. Nevertheless, this experimental campaign followed exactly the protocol proposed in [29] in order to investigate the validity of that protocol when applied to PV plants with non-ideal characteristics during an extended campaign that involved applications under very different weather conditions during an extended campaign. Additionally, this drawback allowed to test the robustness of the proposed experimental analysis method in chapter 2. Therefore, the string that occupies the geometric center of the PV generator was installed with both the irradiance sensor and the temperature sensor.

Fig. 3.2 (a) also shows a temporary delay in module temperature registering higher operating temperatures during the afternoon at the same irradiance levels as the morning. During the morning and afternoon. Different wind conditions and different heating and cooling speeds can cause this temporal asymmetry of the module's temperature. One irradiance sensor and one module temperature sensor may not provide a complete characterization of the PV generator's actual operating conditions. There is a possibility that this limitation may have an impact on the relationship between the PV plant's power and the measured operating conditions. Fig. 3.3 (b) shows, for example, a quasi-linear response to irradiance from morning to noon and a counterclockwise hysteresis response in the afternoon.

As a result of a hysteresis effect attributed to temporal delays in responding to temperature changes (cooling and heating) during the morning and afternoon, [79] demonstrated nonlinear behavior for averaged irradiance, ambient temperature, and power values on a small system level (3 kW). For the case under study, however, if temperature were the

dominant factor in hysteresis formation, DC power values at the same irradiances would be lower in the afternoon for higher temperature values.

As suggested in [29], one of the PV modules located in a strategic position in the plant could approximate the operating temperature of the entire PV generator. In this case, hysteresis would be clockwise. As shown in Fig. 3.2 (b), the observed hysteresis is counterclockwise. It seems reasonable to speculate that in this particular case, the hysteresis formation is primarily caused by the uncertainty in the irradiance measurement due to the misalignment of the PV strings. Consequently, there are different angles of incidence between some of the photovoltaic arrays and our irradiance sensor [80]. Since the values registered by the single sensor cannot take into account possible inhomogeneities in the temperatures of the modules that compose the entire PV generator, the hysteresis may be slightly affected by temperature as well. However, the hypothesis is that the main cause of hysteresis is the misalignment of the generator with the irradiance sensor. By simulating the output power, the hypothesis can be verified. A photovoltaic simulation program known as PVLlib was used. PVLlib is an online database for Python developed by Sandia National Laboratories. It contains a repository with weather data to model PV systems [81] and the quality of the package is comparable to that of commercial packages like PVsyst [82]. PVLlib's particular features make it an excellent package for modeling PV system performance [83]. In this sense, the software allows a number of azimuths to be simulated for irradiance and also calculate the DC output power for each of these azimuths based on the irradiance values.

Fig. 3.3 (a) depicts the RMSE calculated from experimental irradiance sensor and simulated irradiance as a function of the azimuth for four exemplary days. A minimum RMSE around 175° is observed, indicating that the PV panel for irradiance measurement

is misaligned by -5° to -7° . In Fig. 3.3 (b) the RMSE calculated from the experimentally measured DC power and the power calculated with the simulated irradiance in PVLlib.

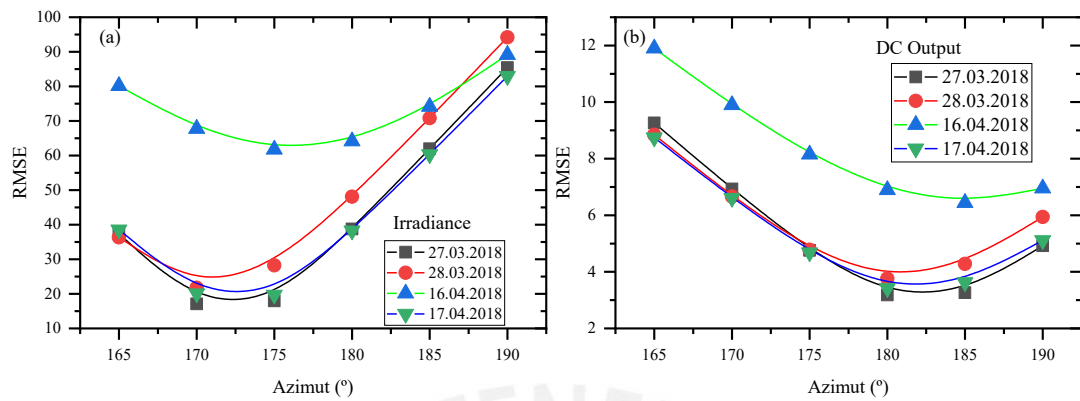


Fig. 3.3. PVLlib simulation, (a) Irradiance, and (b) DC power simulated for four days with different azimuths.

Here, the minimum can be found between 180° and 185° , indicating that the average orientation of the arrays is misaligned by 0° to 5° .

For irradiances above 800 W/m^2 , only the quasi-linear region of the output power is considered for calculating the nominal power according to the introduction section. Since the afternoon data may not represent the correct operation of the PV plant, hysteresis data must be omitted for this nominal power estimation procedure.

3.3 Nominal Power following Reference Procedure

According to the datasheet of the module manufacturer, the generator's nominal power is 109.44 kW at STC. In [84], each string of this PV generator was characterized through I-V curve measurements. Based on the IEC 61829 2015, the P_M^* of the entire plant was estimated at 103.90 kW, with a standard deviation of 1.2%. When compared to the manufacturer's value, the theoretical loss decreased by 4.9% after 10 years. [85], [86] proposed a number of explanations for this difference, including annual polycrystalline panel degradation of $\sim 0.5\%$, resistive wiring losses, and module/string mismatch [87]. A total of 39 out of 135 days with reliable monitoring data met the operational conditions for the reference procedure [29] during the six-month monitoring period. In Fig. 3.4, the actual daily nominal power estimations are shown along with their deviation in % from their mean value. For the 39 days, the mean nominal value was 103.97 kW with a standard deviation of 0.77 kW. As a result, the deviations from the mean value are less than 2%. This is in agreement with what was reported in [84].

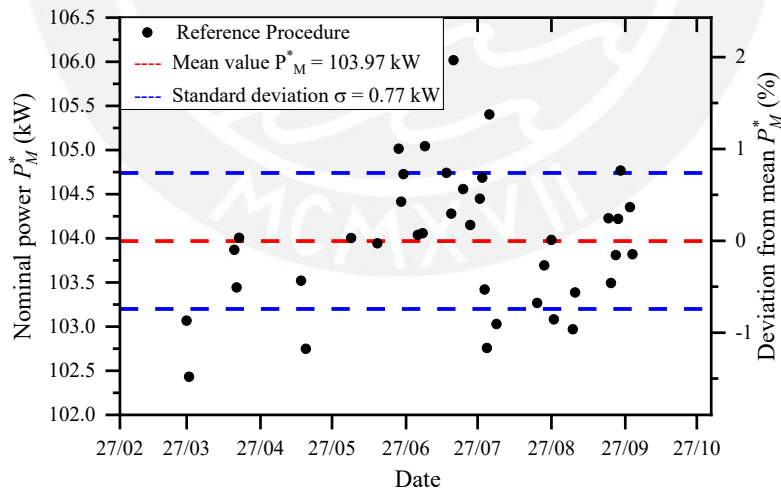


Fig. 3.4 Values of daily nominal power according to Reference Procedure [29]. For the 39 days analyzed, the mean value is shown together with the standard deviation.

3.4 Nominal power following NPKDE Procedure

A total of 96 days with reliable monitoring data did not meet the operating conditions required by the reference procedure in [29]. An exemplary partly cloudy day with relatively few instances of high irradiance was used in this work to demonstrate the applicability of the proposed procedure for these 96 days with non-ideal operational conditions.

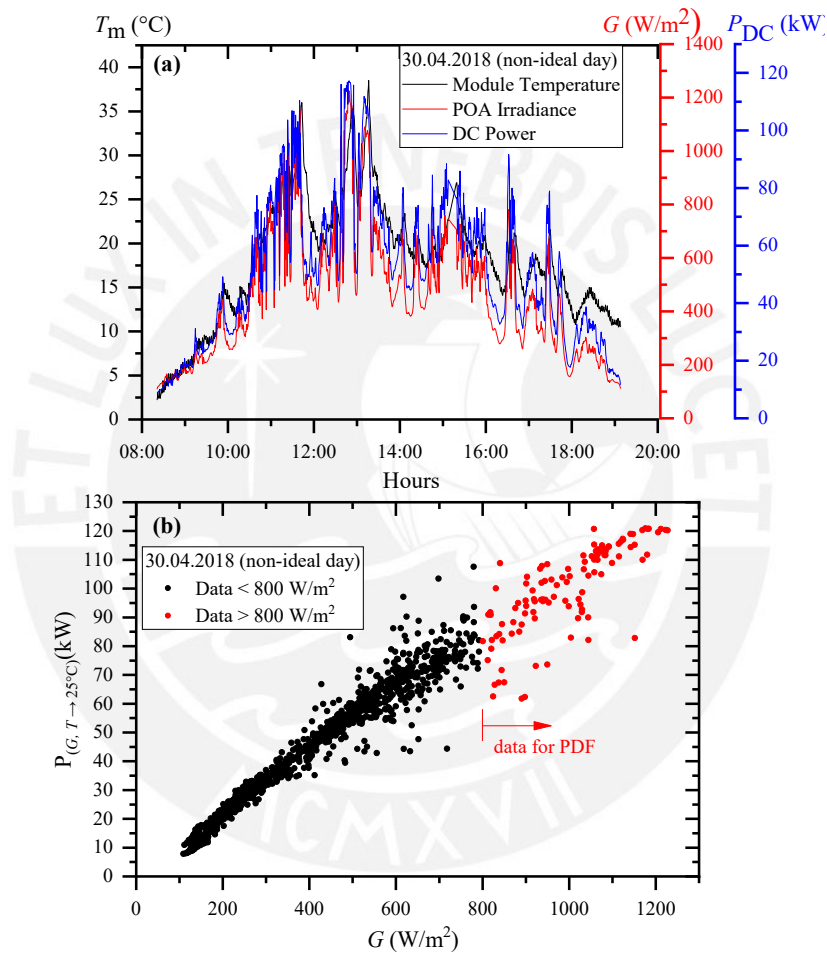


Fig. 3.5. An exemplary case of a partially cloudy, non-ideal day (30.04.2018). In (a), the plane-of-array irradiance, module temperature, and DC power are shown for the entire day. (b) The temperature-corrected DC power versus irradiance, red indicates data sets with irradiances $> 800 \text{ W}/\text{m}^2$ for PDF analysis.

Fig. 3.5(a) shows the plane-of-array irradiance, module temperature, and DC power for such a predominantly cloudy day. Occasionally, high irradiance values ($G > 800 \text{ W}/\text{m}^2$)

have been observed. As a result of the cloud crossing, irradiance and module temperatures are likely to vary, and there are likely to be additional variations in POA irradiance between strings and the irradiance sensor. In Fig. 3.5(b), there is considerable noise in the relation between $P_{(G,T \rightarrow 25^\circ\text{C})}$ and G . Clearly, this data cannot be represented by a single linear regression, so the procedure described by [29] will not work.

Fig. 3.6 (a) shows a non-parametric KDE for this exemplary day with non-ideal weather conditions (30.04.2018), which indicates that after completing the steps of the PDF calculation described above for this day, the resulting PDF shows a mode with a value of $P_M^* = 103.53 \text{ kW}$.

The calculated value is very similar to that for the clear-sky day (18.04.2018), $P_M^* = 103.96 \text{ kW}$. A day with non-ideal conditions has a much broader probability density distribution than a day with ideal conditions due to cloud noise. The difference between these two P_M^* values is 0.43 kW (0.41%), which is within the standard deviation calculated under clear sky conditions of 0.77 kW (0.74 %) and is less than the uncertainty of the power measurement of 0.5%. As can be seen from this slight difference, the proposed procedure can be used with great accuracy to estimate P_M^* on this particular non-ideal day.

Fig. 3.6 (b) shows the nominal power calculated according to the procedure proposed in this thesis for each of the 135 days. The P_M^* values are shown in black for non-ideal weather conditions and in red for ideal weather conditions. Note that both ideal and non-ideal cases have hysteresis data, the only restriction being $G > 800 \text{ W/m}^2$. A deviation in % from $P_M^* = 103.97 \text{ kW}$ resulting from [29] reference procedure is given. Based on the mean value of P_M^* , the values of P_M^* are centered around this value. According to the standard deviation of 0.77 kW (0.74%), most of the results are within the limits of the standard deviation. Some values can deviate by up to 2.5% from the mean of P_M^* ,

depending on the conditions on the day, for days with and without ideal conditions. Assuming that measurement uncertainties are taken into account, these deviations are within a reasonable range, and following the reference procedure for ideal days produces uncertainties of up to 2% in Fig. 3.4.

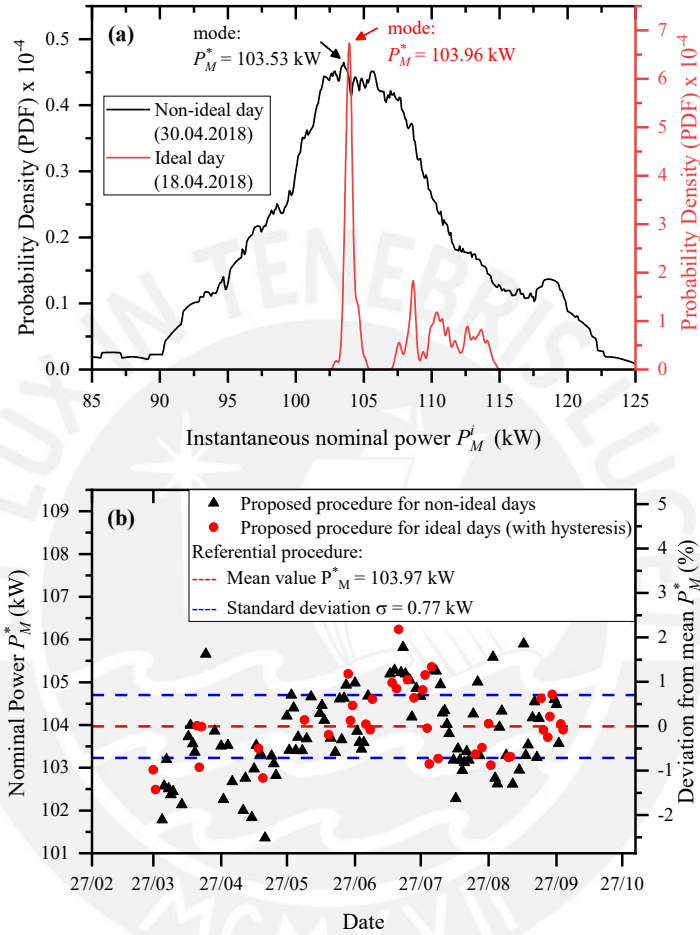


Fig. 3.6. The probability density function (PDF) for a clear-sky day (18.04.2018) compared to a partially cloudy day (30.04.2018). A mode marked by an arrow represents the nominal power with the highest probability. For ideal and non-ideal conditions, in (b) daily nominal power values with the highest probability (modes)

A statistical analysis of the daily nominal power values resulting from the photovoltaic plant can be seen in Fig. 3.7. Table 2 summarizes the main statistical parameters of the plant. The statistical analysis of three cases is presented.

For the reference procedure or case (1), the method already established by Martinez et al. [29] requires the omission of non-linear data from the hysteresis and can only be applied on clear sky days under ideal conditions. In case (2), the proposed procedure with the non-parametric filter is tested on the same data set for clear sky days. This case does not require omitting the nonlinear data related to the hysteresis effect. As case (3), the NPKDE proposed procedure is also statistically evaluated for partly cloudy days in non-ideal conditions.

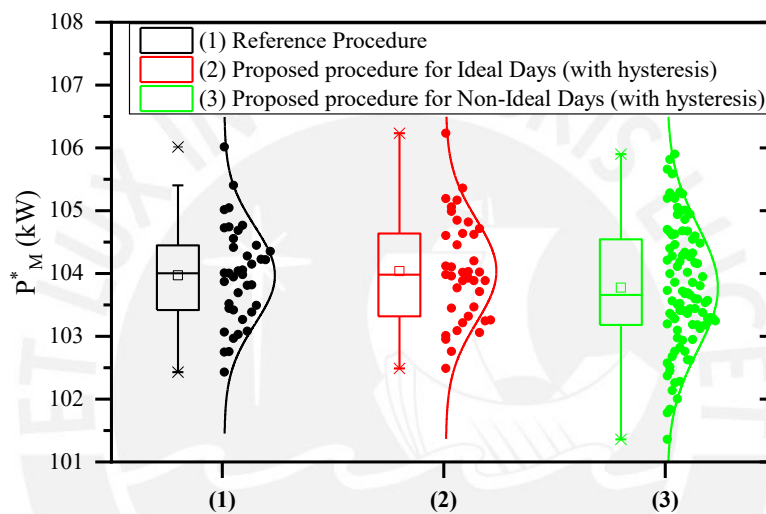


Fig. 3.7. For each of the three cases (1), (2), and (3) under analysis, box plots depict the nominal power values for every single day.

The black points in case (1) show the P_M^* values for the 39 days in ideal conditions. According to the black line, daily values follow a Gaussian or normal distribution, as shown by the similar mean and median P_M^* values in Table 3.2. As a result, the mean of P_M^* values for case (1) is (103.97 ± 0.77) kW. An estimate of the measurement's uncertainty is made by calculating the interquartile range (Q3-Q1). In this range, 50% of the values are distributed. The respective IQR is 1.03 kW which indicates a low level of uncertainty. In PV systems, 2.5% to 5% uncertainty has been reported [88]. Despite a standard deviation of 0.74%, the uncertainty presented in Table 3.2 does not exceed 0.5%.

Essentially, these low uncertainty values are the result of uncertainty in the measurement of irradiance and temperature, which can only be controlled in a laboratory setting [86]. As shown in Fig. 3.7, the daily P_M^* values for case (2) are distributed similarly to those in case (1). According to Table 3.2, the mean P_M^* value of (104.04 ± 0.82) kW and IQR are both very similar to those of case (1). Under these ideal conditions, both procedures have statistically equivalent results, validating the results of the non-parametric KDE procedure.

Table 3.2. Analysis of the nominal power estimation procedures from a statistical perspective

Case	Days	Median P_M^* (kW)	Mean P_M^* (kW)	Standard deviation (kW)	Q1 (kW)	Q3 (kW)	Q3-Q1 (kW)
Case (1): Martinez et al. [29], ideal conditions	39	104.00	103.97	0.77	103.42	104.45	1.03
Case (2): This work, ideal conditions with hysteresis data	39	103.98	104.04	0.82	103.32	104.64	1.32
Case (3): This work, non-ideal conditions	96	103.66	103.77	0.99	103.18	104.54	1.36

The dispersion of the daily P_M^* values is slightly higher for case (3) in Fig. 3.7 compared to the other two cases. As can be seen from Table 3.2, the standard deviation is slightly higher, and the IQR is 1.36 kW. The mode may be affected by noise, i.e., the probability density function's most probable P_M^* value. Nevertheless, the mean P_M^* value of (103.77 ± 0.99) kW for case (3) is still within the range of the other two cases where the nominal PV plant power would be expected.

The results obtained in both procedures have two important aspects to consider. Firstly, procedures get statistically indifferent mean values for nominal power values for days with clear skies under ideal conditions. Secondly, the Non-Parametric filter procedure estimates the nominal power without removing non-linear data from the hysteresis effect

manually. The NPKDE procedure inherently filters out the hysteresis effect on the F_M^* estimation.



Chapter IV:

Nominal Power of Small PV Generators

It was demonstrated in chapter III that, under partially cloudy sky conditions, the NPKDE method is valid for evaluating a utility-scale photovoltaic generator. For these non-ideal conditions, the P_M^* can be calculated with a low degree of uncertainty [89].

In order to test the validity of the NPKDE methodology on small-size PV generators and under different climatic conditions, two geographic locations (Chachapoyas and Lima) with grid-connected 1.5 kW PV systems were selected. The locations have entirely different climates, Chachapoyas with a tropical climate and Lima with a coastal desert-like climate. The systems' modules are based on Passivated Emitter and Rear Cell (PERC) technology. The motivation for studying PERC technology is due to the new generation of solar modules using PERC technology that has become the standard for cell technology representing the new mainstream of the market [90].

In both locations, the PERC array has a nominal power of 1675 W according to the datasheet and is installed on a fixed north-oriented structure with a tilt angle of 15°. Additionally, the array is connected to a 1.5 kW inverter (Sunny Boy 1.5). For more technical details, a brief description of the main characteristics of PV module at STC can be found in Table 4.1.

Table 4.1 Main characteristics of the PV array at STC

Characteristics of the PV generator	PERC
Current at the maximum power point (A)	8.96
Voltage at the maximum power point (V)	37.4
Power at maximum power point (W)	335
Power temperature coefficient (%/°C)	-0.37
Number of modules per string connected in series	5
Number of strings connected in parallel	1

4.1 Location 1: Chachapoyas

Chachapoyas, Peru (6.234632 S, 77.854394 W), is the first place under analysis. This city is the capital of the Amazonas Department, where the annual average temperature is 15.6 °C [91], and the lowest minimum temperatures are reported from June to August. The rainfall varies widely from 1,000 to 2,500 mm over the year [91]. Rainfall is reported every month [92], so this rain creates self-cleaning in the modules, which prevents dust from affecting them. According to a model proposed in [93], wind shear also helps remove particles with diameters between 0.1 and 100 μm .

In Fig. 4.1, the PERC photovoltaic system is shown on the bottom left array. G , T_m and P_{DC} values were monitored to apply the procedure in section 2. The values were obtained by using an EKO- MS-80 pyranometer, module temperature was measured using two PT100 sensor mounted on the back, one in the center, the other at the border of the PV module, and the DC power was obtained from the Sunny Boy 1.5 inverter. The parameters were measured every minute almost a year from 25.01.2021 to 31.12.2021.



Fig. 4.1 Top view, at the bottom left is the PERC array under analysis at Chachapoyas.

Chachapoyas's weather is rainy and cloudy, but it also has high irradiance. On-site, an was measured in one year of monitoring time. Due to the almost continuous presence of clouds, the nominal power can be challenging to estimate for due diligence purposes.

During the monitoring year, only four days were identified with ideal conditions, which would allow calculating the nominal power considering the standard procedure in the introduction part. Hence, it is evident that the nominal power calculation for this location can more suitably be estimated using the NPKDE methodology.

Based on the entire data set, a typical day in Chachapoyas can be observed in Fig. 4.2. Despite the clouds' impact, the KDE can calculate the PDF that identifies the nominal power value. Nevertheless, clouds significantly impact the amount of data with irradiances above 800 W/m^2 (high irradiances). High irradiance values are present for short periods only during the day. The nominal power estimation becomes more reliable when more accumulated data is available with the PV system exposed to high irradiance. High levels of irradiance must be sufficiently present during the day. For instance, the nominal power estimated from data collected over 30 accumulated minutes with high irradiance differs from that collected over 3 accumulated hours. To decrease the uncertainty in the nominal power estimation, the amount of time or data required to generate a reliable KDE should be determined.

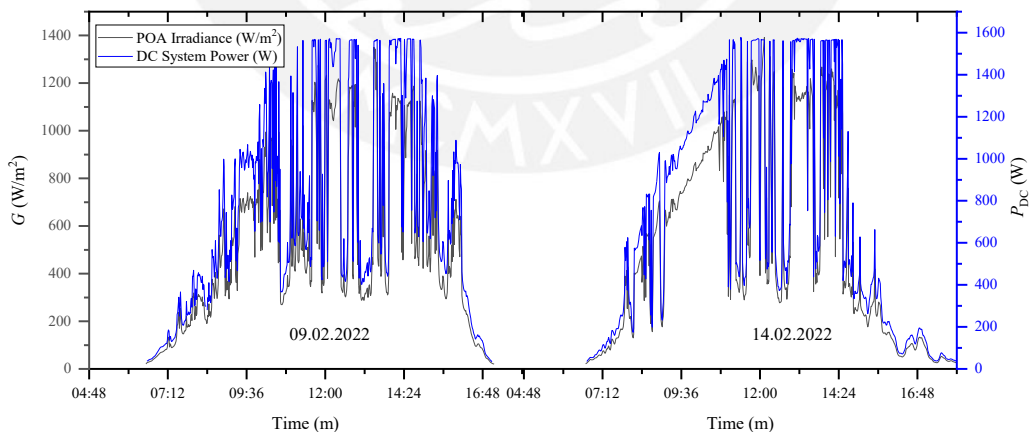


Fig. 4.2. An example of two days with irradiance and DC power output measured every minute. There is predominantly cloudy weather during the day.

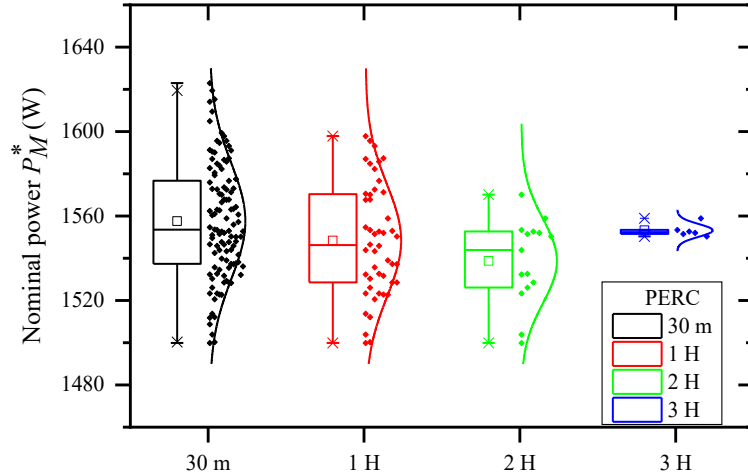


Fig. 4.3 Boxplot analysis when setting the minimum data required in one day

A minimum monitoring time can be set to ensure that the photovoltaic system receives enough accumulated time with high irradiance levels. Fig. 4.3 presents the boxplot analysis of the daily P_M^* for four different time periods. For instance, if the minimum monitoring time is set to 30 minutes, the photovoltaic system will be at high irradiance for at least 30 accumulated minutes daily. Similarly, the same limits would apply for one, two, and three hours of accumulated time with high irradiances with the corresponding box plot.

The nominal power at STC for PERC system is 1675 W. The first boxplot shows the case of 30 minutes. The nominal power values are scattered with a median of 1553.6 W and a mean of 1557.6 W. A difference of almost 7% to the nominal power at STC can be attributed to various factors, including mismatches between panels, wiring losses, irradiance sensor and arrays do not have the same spectral response [94], among others. The data dispersion is evidenced in the uncertainty up to 27.4 W (1.76%) or in the total interquartile range of 39.4 W (2.52%). The number of days when high irradiance was measured was 202. Of these days, only 110 meet the condition of presenting at least 30 accumulated minutes of data at high irradiance, as seen Table 4.2.

Setting the limit to one accumulated hour of high irradiances, the number of days that comply decreases to 48, but the standard deviation (26.4 W) remains similar to the one for 30 min. The mean and median values are also similar at 1546.2 W and 1548.4 W, respectively. Therefore, its nominal power distribution is normal.

By increasing the filter for an accumulated time of high irradiance to 2 hours, the standard deviation decreases to 20 W (1.29 %). As accumulated monitoring time increases, the standard deviation of the daily nominal power decreases. This tendency is most pronounced when the monitoring time is increased to 3 hours, where the standard deviation decreases to 3 W. However, this only occurs on 6 out of the 202 days under analysis. By restricting one day to three hours or 180 data at high irradiances, it becomes difficult to estimate the nominal power for this geographical location.

A comparison of all the mean values for each case of accumulated monitoring time shows no significant differences. However, a minimum of three hours of data collection could be recommended if a higher certainty for value for the nominal power is required. KDE provides more reliable PDF calculations for cloudy conditions the more days, or experimental data points are available with high irradiances.

Table 4.2. Analysis of the nominal power estimation procedures from a statistical perspective

Case	Days	Median P_M^* (W)	Mean P_M^* (W)	Standard deviation (W)	Q1 (W)	Q3 (W)	Q3-Q1 (W)
30 m	110	1553.6	1557.6	27.4	1537.4	1576.8	39.4
1 H	48	1546.2	1548.4	26.4	1528.6	1570.4	41.8
2 H	15	1544.0	1538.7	20.0	1526.1	1552.7	26.6
3 H	6	1552.4	1553.2	3.0	1551.5	1553.5	2.0

Fig. 4.4 below shows the nominal daily power for all twelve months under study. The ideal situation would be to have a large sample size to generate a more accurate KDE, but

10 minutes or 10 experimental data points have been considered sufficient for this calculation. The daily estimation of the nominal power may allow for identifying more quickly if the values are not within the expected range [95].

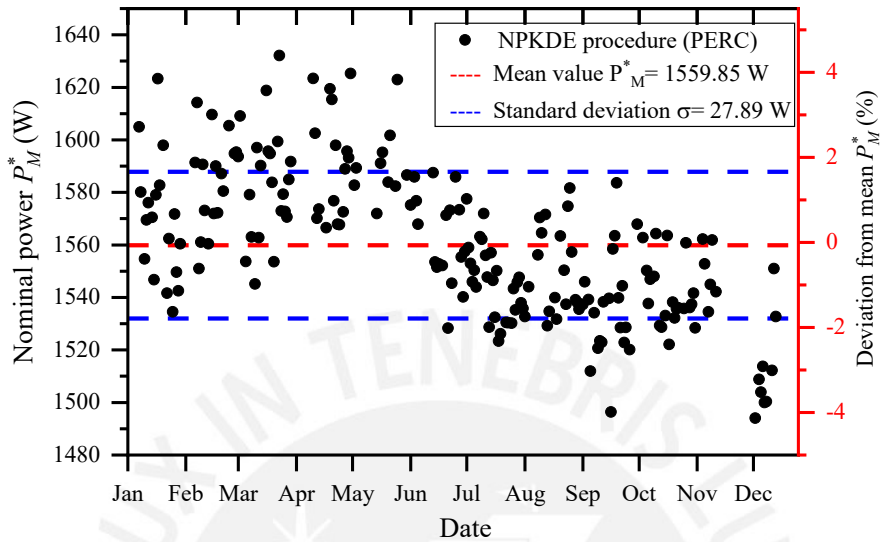


Fig. 4.4. Estimation of the daily nominal power for the year under consideration

The P_M^* calculated for the 202 days with high irradiances indicates a slight nominal power loss during the months of July to December. This loss may be due to the dry season in this region from June to August [92], when rainfall decreases causing a decrease of P_M^* due to dust deposition. In terms of maintenance costs, these otherwise self-cleaning effects are essential [96]. In Lima's case, dust deposition's effect will play an essential role in estimating the nominal power.

PV systems require continuous outdoor monitoring to assess their long-term reliability and performance. Mitigating the economic risk associated with installing new PV technologies [99] and accurately analyzing outdoor performance and degradation is essential. The results of this chapter demonstrated that the nominal power under these conditions could be used to assess long-term degradation.

4.2 Location 2: Lima, Dust Effect

Lima's urbanization, large consumer base, and growing industrial and commercial capacity present significant potential for distributed generation through small-medium PV systems [97]. Lima city has an annual reference yield greater than 1,541.5 kWh / kW or an average daily reference yield of 4.36 kWh / kW. This city is considered, after Cairo, the second largest city located in a desert [98] with rainfall of only 9 mm per year. August recorded a minimum average temperature of 14.7°C, while February recorded a maximum average of 27.3°C [99]. Moreover, this city presents microclimates that make estimating the PV power potential challenging. Despite its solar resources, its location in the desert makes the soiling or dust deposition a factor required to consider. Consequently, the PV system can considerably suffer from lower performance due to higher levels of dirt and pollutants.

Different factors can affect the deposition of dust on a PV module, such as the amount of dust in the atmosphere, wind, humidity, temperature levels, and the tilt angle of the module [100]. The site where the module is installed also plays a role, as does the elevation [101]. All these elements can affect the dust deposition on the surface of the PV module, which can impact its output power.

The dust effect on photovoltaic panels' electrical and thermal behavior is essential when operating these systems. Dust can accumulate on the bottom edging of solar panels and cause inhomogeneities on the panel surface and, thus, losses in power output [102]. The mechanism by which dust bonds to PV surfaces is by adhesion forces. Capillary, van der Waal, electrostatic, and gravitational forces are the fundamental adhesion forces [103].

The humidity in the air can significantly impact how much dust sticks to the surface of solar panels. The humidity of a location significantly affects the adhesion force between dust particles and the surface of a PV module [104]. An increase in relative humidity from 40% to 80% leads to an increase in adhesion by approximately 80% [105].

In Lima, dust accumulation is one of the main factors that cause losses in the efficiency of the PV system. This effect must be characterized in the installation location since environmental factors cannot be estimated or predicted (e.g., wind speed, relative humidity, degree of pollutants in the air, etc.).

According to [109], dust depositions depend on the city's degree of pollution, the soil's nature, the type of dust, and installation factors. In [102] it was found that dust deposition reduces performance in terms of power output by up to 50% over six months. Therefore, module cleaning schedules are also influenced by economic factors which are based on local costs [106]. Additionally, depending on environmental conditions, dust accumulation rates may vary between seasons [107].

In regions with arid or desert climates and low rainfall, there is much effort to optimize and determine cleaning frequency; since they cannot self-clean naturally due to low wind speed and lack of rain. In [108] the authors show that in controlled chambers, there is a linear relationship between the accumulated dust density and the normalized power loss (1.7% per g/m^2). Over six months, dirty solar PV modules can lose up to 50% of their power output in desert regions [109]. For instance, a super-hydrophobic film can reduce the dust deposition's impact on solar PV module efficiency [113]. This type of film creates a barrier that prevents dust from settling on the surface of PV modules, keeping them clean and efficient.

The PV installation under study in Lima is shown in Fig. 4.5. Both facilities are identical, and the same modules and inverters were installed as in Chachapoyas. Considering Lima's high relative humidity and pollution levels, calculating the nominal power is essential for determining cleaning time.



Fig. 4.5. Array in the upper row is the PERC generator in Lima.

Fig. 4.6 shows the calculation of the daily nominal power for six months, during the summer and autumn, when high irradiances ($>800 \text{ W/m}^2$) frequently occur in Lima. Two different irradiance sensors were used: a pyranometer and a calibrated mini-module. At the beginning of measurements in January, the P_M^* using both sensors are very similar. In both cases, there is a general drop in nominal power over the following months. The vertical lines represent dates when the PV modules were manually cleaned. After cleaning, the nominal power returned to its initial value. In two instances, at the end of February and the end of March, only the pyranometer was cleaned, causing a sudden drop of the P_M^* calculated with the irradiance from the pyranometer. Note that calculating P_M^* with irradiance data from the pyranometer and the mini-module results in different slopes of the temporal evolution of the P_M^* . The dome of the pyranometer accumulates less dust than the PV module and the mini-module, and, hence, measures higher irradiance values than effectively reach the solar cells in the PV modules. Hence, the slope of the P_M^* calculated using the pyranometer is larger than when using the mini-module. Three

regions with linear losses are highlighted with slopes m_1 , m_2 , and m_3 . Losses range from 2-3 W per day. The lower slope observed when calculating P_M^* with irradiance data from the calibrated mini-module indicates that the dust deposition rates on the surfaces of the PV modules and the mini-module differ, with a lower rate on the mini-module.

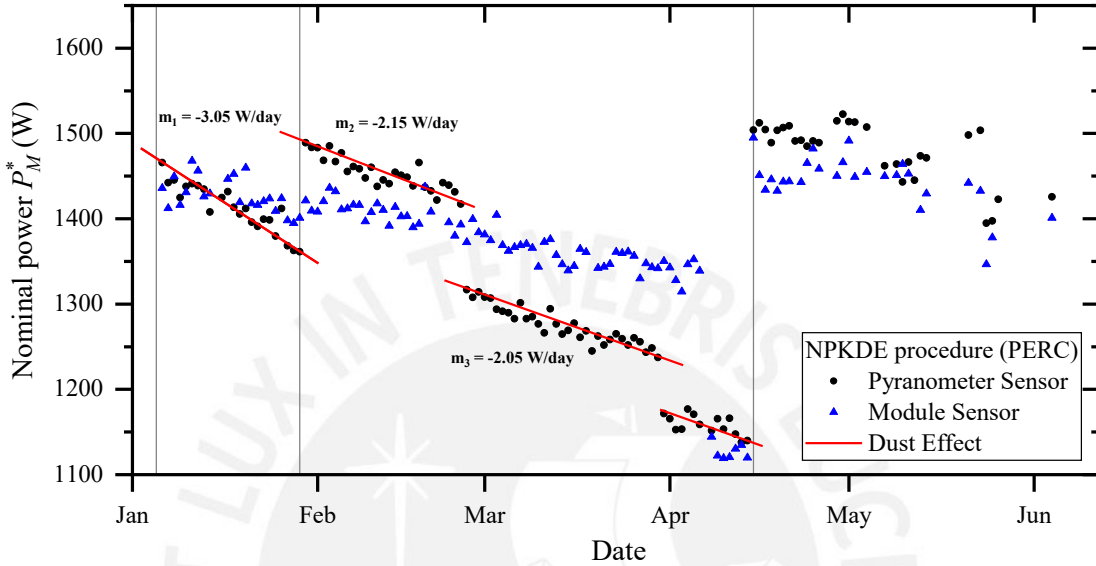


Fig. 4.6. A pyranometer and a calibrated module are used to estimate nominal power. Dust effects are taken into account.

These results allow for determining and monitoring the impacts of the maintenance on the PV system. During the six months, the most prolonged period without cleaning the PV modules was between February and mid-April. During these 2.5 months, the nominal power dropped from about 1500 W to 1150 W, a loss of about 23%. The findings are in agreement with those of other desert locations [109]. For instance, PV panels in desert regions (Bahawalpur, Pakistan) have been studied extensively to determine the best frequency of cleaning [110]. The smallest particle blocks the most sunlight [111] and when the power output reduction and particle concentration are equal to 5% and 100 $\mu\text{g}/\text{m}^3$, the frequency of PV module cleaning was fixed at approximately 20 days.

Conclusions

In the literature, different procedures exist for obtaining the nominal power of medium- and large-scale PV generators that are in outdoor operation. Applying these procedures requires ideal weather conditions with clear-sky days, which, depending on the locations and weather, can sometimes make these procedures non-applicable. By utilizing non-parametric data processing statistics, this Doctoral Thesis proposes new approaches to the characterization of PV generators under ideal and non-ideal conditions. This new approach can be applied by integrating intermediate steps into the reported procedures for the characterization of photovoltaic generators. Mainly in the nominal power of photovoltaic generators, reducing the restrictions imposed and associated uncertainty.

The new procedure estimates the nominal power of a photovoltaic generator under outdoor conditions regardless of its size, particularly monitoring in partly cloudy conditions. The procedure was validated for three different cases:

1. First, the procedure enabled the characterization of a utility-scale PV generator of 109.40 kW. This PV plant presented a hysteresis effect due to the linear and non-linear behavior of the DC power data versus irradiance. The most likely causes of this effect were spatial inhomogeneities of module temperatures and the plane-of-array irradiance. The latter was due to inevitable string misalignments evidencing the challenges of monitoring large PV plants.

Additionally, under non-ideal operating conditions on partially cloudy days caused noise in the DC power versus irradiance data, most likely due to partial shadowing in the PV generator. Standard characterization procedures for the PV nominal power are not applicable under such conditions. The new proposed procedure presents an advance in both regards, demonstrating robustness towards

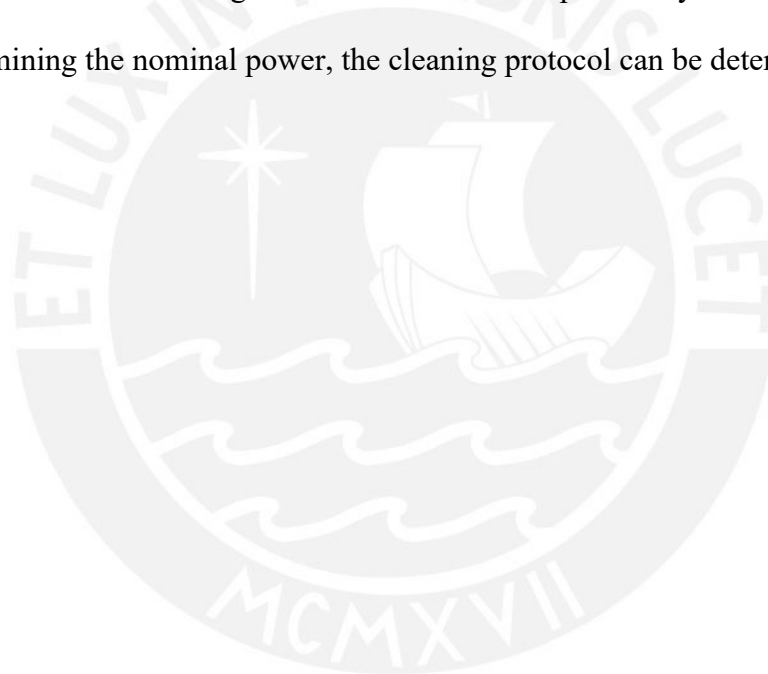
such monitoring challenges and non-ideal operating conditions. The procedure was validated through a 135-day experimental campaign with high-precision instrumentation for monitoring. The Reference Procedure was applied to estimate the PV plant's nominal power, resulting in a mean daily nominal power of (103.97 ± 0.77) kW. It was applicable only in 29% of the studied days, which offered ideal meteorological conditions (full clear sky). Furthermore, it required excluding non-linear data from the hysteresis effect. The proposed procedure applied to these same days with ideal operating conditions resulted in a mean nominal power of (104.04 ± 0.82) kW, with no statistically significant difference to the Reference. Unlike previous procedures, this new method did not require non-linear data to be excluded from the hysteresis effect.

Furthermore, applying the same procedure to the remaining 71% of monitoring days, which presented non-ideal conditions (partially cloudy sky), resulted in a mean nominal power of (103.77 ± 0.99) kW. This result is in good agreement with the previous results for ideal days. Therefore, the proposed non-parametric KDE procedure is suitable for ideal and non-ideal operating conditions and can reliably estimate the nominal power for a single monitoring day.

2. In addition to large generators, small DC power generators of 1675 W were also tested. Chachapoyas, the first location, is characterized by days with high cloud cover. Only 6 of the 202 monitoring days had ideal monitoring conditions. Hence, determining the minimum monitoring time is essential in order to obtain a reliable value under non-ideal conditions. For obtaining the nominal power with a low degree of uncertainty, three hours or 180 data points should be taken. The necessary amount of data can be collected over several days. Furthermore, the

panels are self-cleaning due to high precipitation, which minimizes the effect of dust.

3. As a final evaluation, the methodology was applied to the same PV system located in Lima, where soiling and humidity levels are high. High dust deposition rates were observed in the system. Since the dust deposition rate in the pyranometer and the array is different, it was possible to estimate the cleaning scheme by calculating the nominal power. Calculations indicate a cleaning interval of at least two weeks based on 2-5 W/day losses. Based on similar environmental conditions, these results are in agreement with those reported by other authors. By determining the nominal power, the cleaning protocol can be determined.



Future Lines

Three photovoltaic generators were characterized, resulting in results that will serve as a starting point for future work. Here are four research lines that can be pursued:

1. The new methodology can be used to characterize a larger sample of photovoltaic generators. In Peru there is currently a project: "Evaluación energética y técnico-económica de la generación de energía eléctrica renovable con nuevas tecnologías fotovoltaicas en diferentes zonas climáticas del Perú". A total of three photovoltaic technologies are being monitored at six different locations in Peru as part of this project. The analysis of this procedure under high irradiance conditions makes it very interesting. Furthermore, it would be beneficial to install a monitoring system for large solar plants of the MW order and to identify the challenges that may arise during the characterization process.
2. To improve the precision of the nominal power calculation, hysteresis in photovoltaic generators must be studied in more detail. DC power simulations can be used to determine the ideal position of the generator in order to minimize misalignment between the irradiance sensor and the generator. The temperature and its homogeneity throughout a large generator should also be studied more deeply to determine how it contributes to the hysteresis.
3. An alert system can be implemented through daily monitoring to identify when a system is not operating at its nominal power or to verify that the algorithm can be used effectively as an indicator of cleaning due to dust effects. Additionally, long-term monitoring can describe a technology's global degradation.
4. Although the methodology has excellent results, there is still an important restriction: the irradiance must be greater than 800 W/m^2 . As a result of this

limitation, only certain seasons can be monitored during the year at locations where there are frequent high irradiance occurrences. At low irradiances, a new characteristic parameter would be essential or define it.



Reference

- [1] Benda V and Černá L. 2020. PV cells and modules – State of the art, limits and trends. *Heliyon*. 6, 12.
- [2] Ghosh P *et al.* 2021. Scaling Sustainable Integrated PV Manufacturing Globally. *2021 IEEE 48th Photovoltaic Specialists Conference (PVSC)*. 0472–0476.
- [3] Jäger-Waldau A. 2019. *PV Status Report 2019, EUR 29938*.
- [4] Jäger-Waldau A. 2021. Snapshot of photovoltaics - March 2021. *EPJ Photovoltaics*. 12, 1–7.
- [5] Raghoebarsing A *et al.* 2022. The status of implementation of photovoltaic systems and its influencing factors in European countries. *Progress in Photovoltaics: Research and Applications*.
- [6] Sharma V, Kumar A, Sastry O S, and Chandel S S. 2013. Performance assessment of different solar photovoltaic technologies under similar outdoor conditions. *Energy*. 58, 511–518.
- [7] Lindig S, Louwen A, Moser D, and Topic M. 2020. Outdoor PV system monitoring—Input data quality, data imputation and filtering approaches. *Energies*. 13, (19) 1–18.
- [8] Kelly G, Spooner T, Volberg G, Ball G, and Bruckner J. 2014. Ensuring the reliability of PV systems through the selection of international standards for the IECRE conformity assessment system. *2014 IEEE 40th Photovoltaic Specialist Conference, PVSC 2014*. 914–918.
- [9] Martín-Martínez S, Cañas-Carretón M, Honrubia-Escribano A, and Gómez-Lázaro E. 2019. Performance evaluation of large solar photovoltaic power plants in Spain. *Energy Convers Manag*. 183, 515–528.

- [10] Lindig S, Theristis M, and Moser D. 2022. Best practices for photovoltaic performance loss rate calculations. *Progress in Energy*. 4, (2) 022003.
- [11] Shiva Kumar B and Sudhakar K. 2015. Performance evaluation of 10 MW grid connected solar photovoltaic power plant in India. *Energy Reports*. 1, 184–192.
- [12] Calcabrini A, Ziar H, Isabella O, and Zeman M. 2019. A simplified skyline-based method for estimating the annual solar energy potential in urban environments. *Nat Energy*. 4, (3) 206–215.
- [13] Coello J, Pérez L, Domínguez F, and Navarrete M. 2014. On-site quality control of photovoltaic modules with the PV MOBILE LAB. *Energy Procedia*. 57, 89–98.
- [14] Kiefer K, Reich N H, Dirnberger D, and Reise C. 2011. Quality assurance of large scale PV power plants. *Conference Record of the IEEE Photovoltaic Specialists Conference*. 001987–001992.
- [15] Reich N H, Mueller B, Armbruster A, Van Sark W G J H M, Kiefer K, and Reise C. 2012. Performance ratio revisited: is PR > 90 % realistic?. *Progress in Photovoltaics: Research and Applications*.
- [16] Rahman M M, Selvaraj J, Rahim N A, and Hasanuzzaman M. 2018. Global modern monitoring systems for PV based power generation: A review. *Renewable and Sustainable Energy Reviews*. 82, 4142–4158.
- [17] IEC 61829. 2015. Photovoltaic (PV) array - On-site measurement of current-voltage characteristics.
- [18] Navada H G, Singh S V, and Shubhanga K N. 2017. Modelling of a Solar Photovoltaic Power Plant for Power System Studies. *IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems*. 1–6.
- [19] Muñoz J and Lorenzo E. 2006. Capacitive load based on IGBTs for on-site characterization of PV arrays. *Solar Energy*. 80, (11) 1489–1497.

- [20] Joglekar A V and Hegde B. 2018. Online I-V Tracer for per string monitoring and maintenance of PV panels. *Proceedings: IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*. 1890–1894.
- [21] Quiroz J E, Stein J S, Carmignani C K, and Gillispie K. 2015. In-situ module-level I-V tracers for novel PV monitoring. *2015 IEEE 42nd Photovoltaic Specialist Conference, PVSC 2015*.
- [22] Muñoz J V, Nofuentes G, Fuentes M, de la Casa J, and Aguilera J. 2016. DC energy yield prediction in large monocrystalline and polycrystalline PV plants: Time-domain integration of Osterwald's model. *Energy*. 114, 951–960.
- [23] De la Parra I, Muñoz M, Lorenzo E, García M, Marcos J, and Martínez-Moreno F. 2017. PV performance modelling: A review in the light of quality assurance for large PV plants. *Renewable and Sustainable Energy Reviews*. 78, 780–797.
- [24] E2848 –13. 2013. Standard Test Method for Reporting Photovoltaic Non-Concentrator System. Reapproved 2018, 1–11.
- [25] Watts J L R, Singh R, and Dross F. 2017. Statistical analyses for capacity testing of photovoltaic systems. *2017 IEEE 44th Photovoltaic Specialist Conference, PVSC 2017*, 1–5.
- [26] Kimber A *et al.* 2009. Improved Test Method To Verify The Power Rating of a Photovoltaic (PV) Project. in *Conference Record of the IEEE Photovoltaic Specialists Conference*, 316–321.
- [27] Belluardo G, Ingenhoven P, Sparber W, Wagner J, Weihs P, and Moser D. 2015. Novel method for the improvement in the evaluation of outdoor performance loss rate in different PV technologies and comparison with two other methods. *Solar Energy*. 117, 139–152.

- [28] Kurtz S R *et al.* 2000. Outdoor rating conditions for photovoltaic modules and systems. *Solar Energy Materials and Solar Cells*. 62, (4) 379–391.
- [29] Martínez-Moreno F, Lorenzo E, Muñoz J, and Moretón R. 2012. On the testing of large PV arrays. *Progress in Photovoltaics: Research and Applications*. 20, (1) 100–105.
- [30] Vignola F, Grover C, and Lemon N. 2011. Building a bankable solar radiation dataset. *40th ASES National Solar Conference 2011, SOLAR 2011*. 1, (8) 705–712.
- [31] Riley E *et al.* 2012. Pyranometers and Reference Cells. Part 2: What Makes the Most Sense for PV Power Plants. *PV Magazine*.
- [32] Polo J, Fernandez-Neira W G, and Alonso-García M C. 2017. On the use of reference modules as irradiance sensor for monitoring and modelling rooftop PV systems. *Renew Energy*. 106, 186–191.
- [33] Kaldellis J K, Kapsali M, and Kavadias K A. 2014. Temperature and wind speed impact on the efficiency of PV installations. Experience obtained from outdoor measurements in Greece. *Renew Energy*. 66, 612–624.
- [34] Akhsassi M *et al.* 2018. Experimental investigation and modeling of the thermal behavior of a solar PV module. *Solar Energy Materials and Solar Cells*. 180, 271–279.
- [35] Whitfield K and Osterwald C R. 2001. Procedure for determining the uncertainty of photovoltaic module outdoor electrical performance. *Progress in Photovoltaics: Research and Applications*. 9, (2) 87–102.
- [36] Skoplaki E and Palyvos J A. 2009. On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations. *Solar Energy*. 83, (5) 614–624.
- [37] Osterwald C R. 1986. Translation of device performance measurements to reference conditions,” *Solar Cells*. 18, (3–4) 269–279.

- [38] Cirp P *et al.*. 2022. On the identification of the probability density function of a dimension from data with measurement errors. *Procedia CIRP*. 114, 60–66.
- [39] Wong T T. 2017. Parametric methods for comparing the performance of two classification algorithms evaluated by k-fold cross validation on multiple data sets. *Pattern Recognit.* 65, 97–107.
- [40] Gramacki A. 2018. *Bandwidth Selectors for Kernel Density Estimation*.
- [41] Nosratabadi H, Mohammadi M, and Kargarian A. 2019. Nonparametric Probabilistic Unbalanced Power Flow with Adaptive Kernel Density Estimator. *IEEE Trans Smart Grid*. 10, (3) 3292–3300.
- [42] Qin X, Li Y, Shen C, Zhang Z, and Zeng X. 2016. The Correlation Analysis of Clean Energy Output Based on Nonparametric Kernel Density Estimation Probability Models. *Atlantis Press*. 133, 24–28.
- [43] Han Q, Ma S, Wang T, and Chu F. 2019. Kernel density estimation model for wind speed probability distribution with applicability to wind energy assessment in China. *Renewable and Sustainable Energy Reviews*. 115, 109387.
- [44] Nakarmi J and Sang H. 2018. Central limit theorem for the variable bandwidth kernel density estimators. *J Korean Stat Soc*. 47, (2) 201–215.
- [45] Mugdadi A R and Ahmad I A. 2004. A bandwidth selection for kernel density estimation of functions of random variables. *Comput Stat Data Anal*. 47, (1) 49–62.
- [46] Lin J, Yang C, Liu Y, Luo S, and Wang L. 2022. Practical state estimation model and algorithm based on kernel density theory. *International Journal of Electrical Power and Energy Systems*. 134, 107442.
- [47] Dobrovidov A V and Rud’Ko I M. 2010. Bandwidth selection in nonparametric estimator of density derivative by smoothed cross-validation method. *Automation and Remote Control*. 71, (2) 209–224.

- [48] Sheather S J and Jones M C. 1991. A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*. 53, (3) 683–690.
- [49] Ahmed R, Sreeram V, Mishra Y, and Arif M D. 2020. A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. *Renewable and Sustainable Energy Reviews*. 124, 109792.
- [50] Bozorg M, Bracale A, Carpita M, de Falco P, Mottola F, and Proto D. 2021. Bayesian bootstrapping in real-time probabilistic photovoltaic power forecasting. *Solar Energy*. 225, 577–590.
- [51] Hassan M A, Abubakr M, and Khalil A. 2021. A profile-free non-parametric approach towards generation of synthetic hourly global solar irradiation data from daily totals. *Renew Energy*. 167, 613–628.
- [52] Lotfi M, Javadi M, Osório G J, Monteiro C, and Catalão J P S. 2020. A novel ensemble algorithm for solar power forecasting based on kernel density estimation. *Energies*. 13, 1.
- [53] Liu L *et al.* Prediction of short-term PV power output and uncertainty analysis. 2018. *Appl Energy*. 228, 700–711.
- [54] Gu B, Shen H, Lei X, Hu H, and Liu X. 2021. Forecasting and uncertainty analysis of day-ahead photovoltaic power using a novel forecasting method. *Appl Energy*. 299.
- [55] Pan C, Tan J, and Feng D. 2021. Prediction intervals estimation of solar generation based on gated recurrent unit and kernel density estimation. *Neurocomputing*. 453, 552–562,.
- [56] Rodríguez F, Galarza A, Vasquez J C, and Guerrero J M. 2022. Using deep learning and meteorological parameters to forecast the photovoltaic generators intra-hour output power interval for smart grid control. *Energy*. 239, 122116.

- [57] Han Y, Wang N, Ma M, Zhou H, Dai S, and Zhu H. 2019. A PV power interval forecasting based on seasonal model and nonparametric estimation algorithm. *Solar Energy*. 184, 515–526.
- [58] Livera A, Theristis M, Makrides G, and Georghiou G E. 2019. Recent advances in failure diagnosis techniques based on performance data analysis for grid-connected photovoltaic systems. *Renew Energy*. 133, 126–143.
- [59] Vergura S. 2018. Hypothesis tests-based analysis for anomaly detection in photovoltaic systems in the absence of environmental parameters. *Energies*. 11, (3) 25–27.
- [60] Taghezouit B, Harrou F, Sun Y, Arab A H, and Larbes C. 2021. A simple and effective detection strategy using double exponential scheme for photovoltaic systems monitoring. *Solar Energy*. 214, 337–354.
- [61] Ding H *et al.* 2018. Local outlier factor-based fault detection and evaluation of photovoltaic system. *Solar Energy*. 164, 139–148.
- [62] Harrou F, Taghezouit B, and Sun Y. 2019. Improved κ N-Based monitoring schemes for detecting faults in PV systems. *IEEE J Photovolt*. 9, (3) 811–821.
- [63] Taghezouit B, Harrou F, Sun Y, Arab A H, and Larbes C. 2020. Multivariate statistical monitoring of photovoltaic plant operation. *Energy Convers Manag*. 205, 112317.
- [64] Pan J, He W, Shi Y, Hou R, and Zhu H. 2021. Uncertainty analysis based on non-parametric statistical modelling method for photovoltaic array output and its application in fault diagnosis. *Solar Energy*. 225, 831–841.
- [65] Madeti S R and Singh S N. 2018. Modeling of PV system based on experimental data for fault detection using kNN method. *Solar Energy*. 173, 139–151.
- [66] Pan J, Zou Z, Sun S, Su Y, and Zhu H. 2022. Research on output distribution modeling of photovoltaic modules based on kernel density estimation method and its application in anomaly identification. *Solar Energy*. 235, 1–11.

- [67] Kamalov F. 2020. Kernel density estimation based sampling for imbalanced class distribution. *Inf Sci (N Y)*. 512, 1192–1201.
- [68] Hou Q, Zhang N, Du E, Miao M, Peng F, and Kang C. 2019. Probabilistic duck curve in high PV penetration power system: Concept, modeling, and empirical analysis in China. *Appl Energy*. 242, 205–215,.
- [69] Amarasinghe P A G M, Abeygunawardane S K, and Singh C. 2020. Kernel Density Estimation Based Time-Dependent Approach for Analyzing the Impact of Increasing Renewables on Generation System Adequacy. *IEEE Access*. 8, 138661–138672.
- [70] Chen Y C. 2017. A tutorial on kernel density estimation and recent advances. *Biostat Epidemiol*. 1, 161–187.
- [71] Hammad B, Al-Sardeah A, Al-Abed M, Nijmeh S, and Al-Ghandoor A. 2017. Performance and economic comparison of fixed and tracking photovoltaic systems in Jordan. *Renewable and Sustainable Energy Reviews*. 80, 827–839.
- [72] Muñoz-Cerón E, Lomas J C, Aguilera J, and De la Casa J. 2018. Influence of Operation and Maintenance expenditures in the feasibility of photovoltaic projects: The case of a tracking pv plant in Spain. *Energy Policy*. 121, 506–518.
- [73] Hishikawa Y *et al.* 2020. Improved precision of the outdoor performance measurements of photovoltaic modules by using the photovoltaic irradiance sensor. *Solar Energy*. 211, 82–89.
- [74] Huang B J *et al.* 2011. Solar cell junction temperature measurement of PV module. *Solar Energy*. 85, (2) 388–392.
- [75] Barykina E and Hammer A. 2017. Modeling of photovoltaic module temperature using Faiman model: Sensitivity analysis for different climates. *Solar Energy*. 146, 401–416.

- [76] Rao Golive Y, Kottantharayil A, and Shiradkar N. 2022. Improving the accuracy of temperature coefficient measurement of a PV module by accounting for the transient temperature difference between cell and backsheet. *Solar Energy*. 237, 203–212.
- [77] Kaldellis J K, Kapsali M, and Kavadias K A. 2014. Temperature and wind speed impact on the efficiency of PV installations. Experience obtained from outdoor measurements in Greece. *Renew Energy*. 66, 612–624.
- [78] Muñoz Escribano M, Garcia Solano M, De La Parra Laita I, Marcos Alvarez J, Marroyo L, and Lorenzo Pigueiras E. 2018. Module temperature dispersion within a large PV array: Observations at the amareleja PV plant. *IEEE J Photovol*. 8, (6) 1725–1731.
- [79] Filik U B, Filik T, and Gerek O N. 2018. A hysteresis model for fixed and sun tracking solar PV power generation systems. *Energies*. 11, 3.
- [80] Carrillo J M, Martínez-Moreno F, Lorenzo C, and Lorenzo E. 2017. Uncertainties on the outdoor characterization of PV modules and the calibration of reference modules. *Solar Energy*. 155, 880–892.
- [81] Stein J S, Holmgren W F, Forbess J, and Hansen C W. 2017. PVLIB: Open source photovoltaic performance modeling functions for Matlab and Python. *2017 IEEE 44th Photovoltaic Specialist Conference, PVSC 2017*. 1–6.
- [82] Gurupira T, Rix A J, Gurupira T, and Rix A J. 2017. PV Simulation Software Comparisons: Pvsyst, Nrel Sam and Pvlib.
- [83] Heusinger J, Broadbent A M, Sailor D J, and Georgescu M. 2020. Introduction, evaluation and application of an energy balance model for photovoltaic modules. *Solar Energy*. 195, 382–395.
- [84] Lomas J C. 2019. Propuesta de Una Nueva Metodología de Analisis de Plantas FV Atendiendo a los Cambios Normativos Motivados por el RD 661/2007. Universidad de Jaen.

- [85] Jordan D C and Kurtz S R. 2013. Photovoltaic degradation rates - An Analytical Review. *Progress in Photovoltaics: Research and Applications*. 21, (1) 12–29.
- [86] Carullo A, Castellana A, Vallan A, Ciocia A, and Spertino F. 2017. Uncertainty issues in the experimental assessment of degradation rate of power ratings in photovoltaic modules. *Measurement*. 111, 432–440.
- [87] Lorente D G, Pedrazzi S, Zini G, Dalla Rosa A, and Tartarini P. 2014. Mismatch losses in PV power plants. *Solar Energy*. 100, 42–49.
- [88] Dirnberger D, Bartke J, Steinhüser A, Kiefer K, and Neuberger F. 2010. Uncertainty of field IV Curve measurements in Large Scale PV systems. *25th EU PVSEC*. 6–10.
- [89] Angulo J R *et al.* 2022. Estimation of the effective nominal power of a photovoltaic generator under non-ideal operating conditions. *Solar Energy*. 231, 784–792.
- [90] Jia X, Zhou C, Tang Y, and Wang W. 2021. Life cycle assessment on PERC solar modules. *Solar Energy Materials and Solar Cells*. 227, 111112.
- [91] Newell F L, Ausprey I J, and Robinson S K. 2022. Spatiotemporal climate variability in the Andes of northern Peru: Evaluation of gridded datasets to describe cloud forest microclimate and local rainfall. *International Journal of Climatology*. 5892–5915.
- [92] Rascón J, Gosgot Angeles W, Quiñones Huatangari L, Oliva M, and Barrena Gurbillón M Á. 2021. Dry and Wet Events in Andean Populations of Northern Peru: A Case Study of Chachapoyas, Peru. *Front Environ Sci*. 9, 1–13.
- [93] Jiang Y, Lu L, Ferro A R, and Ahmadi G. 2018. Analyzing wind cleaning process on the accumulated dust on solar photovoltaic (PV) modules on flat surfaces. *Solar Energy*. 159, 1031–1036.
- [94] Riley E *et al.* 2012. Pyranometers and Reference Cells : Part 2 : What Makes the Most Sense for PV Power Plants. *PV Magazine*.

- [95] Velilla E, Cano J B, and Jaramillo F. 2019. Monitoring system to evaluate the outdoor performance of solar devices considering the power rating conditions. *Solar Energy*. 194, 79–85.
- [96] Tanesab J, Parlevliet D, Whale J, and Urmee T. 2016. Dust Effect and its Economic Analysis on PV Modules Deployed in a Temperate Climate Zone. *Energy Procedia*. 100, 65–68.
- [97] Espinoza R, Muñoz-Cerón E, Aguilera J, and De la Casa J. 2019. Feasibility evaluation of residential photovoltaic self-consumption projects in Peru. *Renew Energy*. 136, 414–427.
- [98] Hussein M. 2018. Sustainable Regeneration Of Urban Green Areas In Egypt ' S Desert Cities Sustainable Regeneration Of Urban Green Areas In Egypt's Desert Cities Adopting Green Infrastructure Strategies In New Borg El-Arab City. HafenCity University.
- [99] Giráldez L, Silva Y, Flores-Rojas J L, and Trasmonte G. 2022. Diagnosis of the Extreme Climate Events of Temperature and Precipitation in Metropolitan Lima during 1965–2013. *Climate*. 10, 8.
- [100] Chanchangi Y N, Ghosh A, Baig H, Sundaram S, and Mallick T K. 2021. Soiling on PV performance influenced by weather parameters in Northern Nigeria. *Renew Energy*, 180, 874–892.
- [101] Reise C, Müller B, Moser D, Belluardo G, and Ingehoven P. 2018. *Uncertainties in PV System Yield Predictions and Assessments*.
- [102] Abderrezek M and Fathi M. 2017. Experimental study of the dust effect on photovoltaic panels' energy yield. *Solar Energy*. 142, 308–320.
- [103] Isaifan R J, Johnson D, Ackermann L, Figgis B, and Ayoub M. 2019. Evaluation of the adhesion forces between dust particles and photovoltaic module surfaces. *Solar Energy Materials and Solar Cells*, 191, 413–421.

- [104] Mekhilef S, Saidur R, and Kamalisarvestani M. 2012. Effect of dust, humidity and air velocity on efficiency of photovoltaic cells. *Renewable and Sustainable Energy Reviews*. 16, (5), 2920–2925.
- [105] Gupta V, Sharma M, Kumar R, and Babu K N D. 2019. Comprehensive review on effect of dust on solar photovoltaic system and mitigation techniques. *Solar Energy*. 191, 596–622.
- [106] Tanesab J, Parlevliet D, Whale J, and Urmee T. 2018. Energy and economic losses caused by dust on residential photovoltaic (PV) systems deployed in different climate areas. *Renew Energy*. 120, 401–412.
- [107] Javed W, Wubulikasimu Y, Figgis B, and Guo B. 2017. Characterization of dust accumulated on photovoltaic panels in Doha, Qatar. *Solar Energy*. 142, 123–135.
- [108] Hachicha A A, Al-Sawafta I, and Said Z. 2019. Impact of dust on the performance of solar photovoltaic (PV) systems under United Arab Emirates weather conditions. *Renew Energy*. 141, 287–297.
- [109] Adinoyi M and Said S. 2013. Effect of dust accumulation on the power outputs of solar photovoltaic modules. *Renew Energy*. 60, 633–636.
- [110] Shah S F A, Khan I A, and Khan A H A. 2019. Performance evaluation of two similar 100MW solar PV plants located in environmentally homogeneous conditions. *IEEE Access*. 7, 161697–161707.
- [111] Hussain A, Batra A, and Pachauri R. 2017. An experimental study on effect of dust on power loss in solar photovoltaic module. *Renew Wind Water Sol*. 4, 1.

Publications

Scientific publications indexed with relative quality index

B. Calsi; L. A. Conde; **J.R. Angulo**; J. Montes-Romero; J.A. Guerra; J. de la Casa; J. A. Töfflinger, (2021). Monitoring versus prediction of the power of three different PV technologies in the coast of Lima-Peru. *Journal of Physics: Conference Series*, 1841 – 012001. **(Percentile 18th SCOPUS – SJR)**.

<https://doi.org/10.1088/1742/6596/1841/1/012001>

L. A. Conde; **J. R. Angulo**; M. A. Sevillano-Bendezu; G. Nofuentes; J. A. Töfflinger; J. de la Casa, (2021). Spectral effects on the energy yield of various photovoltaic technologies in Lima (Peru). *Energy*, 223, art. no. 120034. **(Q1 WOS-JCR)**

<https://doi.org/10.1016/j.energy.2021.120034>

J. R. Angulo; B. X. Calsi; L. A. Conde; J.A. Guerra; E. Muñoz-Cerón; J. de la Casa; J. A. Töfflinger (2022). Estimation of the effective nominal power of a photovoltaic generator under non-ideal operating conditions. *Solar Energy*; 231; pp. 784-792. **(Q1 WOS-JCR)**.

<https://doi.org/10.1016/j.solener.2021.12.015>

A. E. Berastain; L. A. Conde; **J. Angulo**; A. M. Carhuavilca; M. García; M. A. Sevillano-Bendezú; J. Montes-Romero; J. de la Casa; L. Chirinos; J. A. Töfflinger, (2022). Resolving challenges of monitoring PV systems: A case study for three 1.5 kW generators in Lima, Peru. *Journal of Physics: Conference Series*, 2180(1) – 012006. **(Percentile 18th SCOPUS – SJR)**.

<https://doi.org/10.1088/1742/6596/2180/1/012006>

J. Angulo; L. A. Conde; E. Muñoz-Cerón; J. de la Casa; J. A. Töfflinger, (2022). PV generator nominal power estimation using a ground sensor and the PVLIB online irradiance database. *Journal of Physics: Conference Series*, 2180(1) – 012005. **(Percentile 18th SCOPUS – SJR)**.

<https://doi.org/10.1088/1742/6596/2180/1/012005>

C.A. Espinosa-González; **J. Angulo**; L. A. Conde; A. M. Carhuavilca; E. F. Alfaro; R. Espinoza; J. A. Töfflinger; E. Muñoz-Cerón; J. de la Casa, (2022). Is the information provided by free satellite sources suitable for predicting or evaluating the performance of photovoltaic systems in Peru? (In memoriam to Heinrich Berg). *Journal of Physics: Conference Series*, 2180(1) – 012016. **(Percentile 18th SCOPUS – SJR)**.

<https://doi.org/10.1088/1742/6596/2180/1/012016>

Scientific publications without relative quality index

R. Perich-Ibañez; M. Sevillano-Bendezú; J. Montes-Romero; L. Conde-Mendoza; **J. R. Angulo**; J. de la Casa; J. A. Tofflinger (2020). Estudio de los métodos analíticos para la extracción de parámetros eléctricos de módulos fotovoltaicos de capas delgadas. *Revista Científica TECNIA*, 30, pp. 53-58.

<https://doi.org/10.21754/tecnia.v30i1.851>

J. R. Angulo; B. Calsi-Silva; E. F. Alfaro-Collazos; L. Conde-Mendoza; E. Muñoz-Cerón; R. Grieseler; J. A. Guerra; J. Palomino-Tofflinger; R. Espinoza-Paredes; **J. de la Casa**, (2020). Estudio del efecto del polvo y estimación de la potencia nominal en un string fotovoltaico. *Revista Científica TECNIA*, 30, pp. 27-33.

<https://doi.org/10.21754/tecnia.v30i1.832>

B. Calsi-Silva; **J.R. Angulo**; L. Conde-Mendoza; Emilio Muñoz Cerón; R. Grieseler; J.A. Guerra; J. de la Casa; J. Palomino-Tofflinger, (2020). Procedimiento del cálculo de la potencia nominal de un generador fotovoltaico. *Revista Científica TECNIA*, 30, pp. 22-26.

<https://doi.org/10.21754/tecnia.v30i1.834>

L. Conde-Mendoza; J. Montes-Romero; A. Carhuavilca-Vela; R. Perich-Ibañez; J.A. Guerra; M. Sevillano-Bendezú; B. Calsi-Silva; **J.R. Angulo**; J. de la Casa; J. A. Töfflinger, (2020). Puesta en marcha de un laboratorio para la caracterización de tecnologías fotovoltaicas a sol real bajo las condiciones climáticas de Lima, *Revista Científica TECNIA*, 30, pp. 80-89.

<https://doi.org/10.21754/tecnia.v30i1.835>

Communications in congresses with paper indexed in scopus

J. Angulo; B. Calsi; L. Conde-Mendoza; E. Muñoz-Ceron; R. Grieseler; J. A. Guerra; J. de la Casa; J. A. Töfflinger. MODELING OF THE NOMINAL POWER OF A PV GENERATOR UNDER CLEAR AND CLOUDY SKY CONDITIONS. **In:** *ISES Solar World Congress 2019*, Santiago de Chile, 2019, pp. 695-702. **Participation:** Poster and communication.

<http://dx.doi.org/10.18086/swc.2019.14.01>

L. A. Conde-Mendoza; J. Montes-Romero; A. Carhuavilca; R. Perich; J. A. Guerra; **J. Angulo**; E. Muñoz-Cerón; J. A. Töfflinger; J. de la Casa. PERFORMANCE EVALUATION AND CHARACTERIZATION OF DIFFERENT PHOTOVOLTAIC TECHNOLOGIES UNDER THE COASTAL DESERT CLIMATE CONDITIONS OF LIMA, PERÚ. **In:** *ISES Solar World Congress 2019*, Santiago de Chile, 2019, pp. 695-702. **Participation:** Poster and communication.

<http://dx.doi.org/10.18086/swc.2019.16.01>