This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 953016

Ref. Ares(2021)2247061 - 06/04/2021

D3.1 Revision of state of the art of automatic PV performance supervision systems

T3.1 Review of the current status in monitoring, data assessment and data analytics

Grant Agreement nº:	**953016**
Call:	H2020-LC-SC3-2020-RES-IA-CSA / LC-SC3-RES-33-2020
Project title:	**Smooth, REliable aNd Dispatchable Integration of PV in EU Grids**
Project acronym:	SERENDI-PV
Type of Action:	Innovation Action
Granted by:	Innovation and Networks Executive Agency (INEA)
Project coordinator:	Fundación TECNALIA Research & Innovation
Project website address:	*(not yet available)*
Start date of the project:	October 2020
Duration:	48 months

Document Ref.:	SERENDI-PV_D3.1 SotA automatic PV performance supervision systems_v1.0.docx
Lead Beneficiary:	QPV
Doc. Dissemination Level:	**PU** – Public
Due Date for Deliverable:	31/03/2021
Actual Submission date:	06/04/2021
Version	V1.0

Summary

This document is a state of the art in the automated supervision of photovoltaic installations. This is the result of the study work carried out in task 3.1 and therefore its deliverable.

The document deals with different topics related to the automated supervision and analysis of PV plants: monitoring, data quality, operation analysis, digital twins, fault detection and diagnosis and integration in O&M. This document reviews the state of the art in recent years in this regard, providing references from both the academic world and industry regulation.

Document Information

Title	"Revision of state of the art of automatic PV performance supervision systems"
Lead Beneficiary	"QPV"
Contributors	"TEC, SGIS, AKUO, CNR, COB, MLS"
Distribution	"PU"
Report Name	"Revision of state of the art of automatic PV performance supervision systems"

Document History

Date	Version	Prepared by	Organisation	Approved by	Notes
30/03/2021	V1.0	QPV	QPV	TECNALIA	

Acknowledgements

The work described in this publication has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 953016.

Disclaimer

This document reflects only the authors' view and not those of the European Commission. This work may rely on data from sources external to the members of the SERENDI-PV project Consortium. Members of the Consortium do not accept liability for loss or damage suffered by any third party as a result of errors or inaccuracies in such data. The information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and neither the European Commission nor any member of the SERENDI-PV Consortium is liable for any use that may be made of the information.

© Members of the SERENDI-PV Consortium

Contents

Summary ii
Document Information ii
Document History ii
Acknowledgements iii
Disclaimer iii

1 EXECUTIVE SUMMARY 1
1.1 Description of the deliverable content and purpose 1
1.2 Reference material 3
1.3 Relation with other activities in the project 3
1.4 Abbreviation list 4
2 MONITORING AND DATA QUALITY 7
3 PV DATA ANALYTICS 13
3.1 Operation conditions 13
3.2 Generator 14
3.3 Inverter 16
3.4 Tracker 17
3.5 Peculiarities of technologies: 18
3.5.1 Bifacial 18
3.5.2 Floating PV 18
3.5.3 BIPV 20
4 COMPONENTS DIGITAL TWINS 22
4.1 PV inverter digital twin 23
4.2 Battery digital twin 26
4.2.1 Physical model of the battery cell 27
4.2.2 SOC and SOH estimation methods 31
4.2.3 Battery cell data driven model for life prediction 35
5 PV FAULT DIAGNOSIS 37
5.1 PV common techniques 37
5.2 Large PV plants 41
5.3 Small and medium PV plants 41
6 O&M integration 44
6.1 General O&M procedure for monitoring and controlling PV plants 44
6.1.1 Meteorological data 44
6.2 Drone IR inspection 45
7 REFERENCES 47

Tables

Table 1.1: Relation between current deliverable and other activities in the project ... 3
Table 1.2: Abbreviation list ... 4

Figures

Figure 1.1: Market potential for the next decades. Source [1]. ... 2
Figure 2.1: Example of asynchronous power measurement corrected by linear interpolation. Source [16]. ... 11
Figure 3.1: Configurations of grid-connected PV inverter [59] ... 16
Figure 3.2: Topology of the power system in a building with BIPV, storage, consumption and grid connection. Source [111]. ... 21
Figure 4.1: Optimization of a physical model to transform it into a hybrid model. ... 23
Figure 4.2: EPCOS/TDK capacitor lifetime as a function of temperature and current ripple. ... 24
Figure 4.3: ECM. Simple [162] ... 27
Figure 4.4: ECM. RC model [162] ... 28
Figure 4.5: ECM. ESC model [162] ... 28
Figure 4.6: Test schedule flow chart carried out in [164] ... 29
Figure 4.7: Zoom into discharge pulse of the LiNMC cell at around 90% SOC at 22°C [163] ... 29
Figure 4.8: Degradation mechanisms in a battery [168] ... 31
Figure 4.9: Comparison table between SOH and SOC estimation algorithms [163] ... 33
Figure 4.10: Flowchart of methodology for lifetime prognosis by [160] ... 35
Figure 5.1: Classification of Failure detection methods according to available monitoring system, proposed methods, and required input data. ... 40
Figure 6.1: Images obtained from an IR inspection for cells anomalies and hotspots identification respectively ... 46

1 EXECUTIVE SUMMARY

1.1 Description of the deliverable content and purpose

The objective of this document is to present a state of the art of the topics covered in WP3. This will serve as a starting point and framework for the WP developments, helping their integration into the current sector.

The topic covered in this document is the automated supervision in the photovoltaic sector. In general, Automation is a reality that affects multiple aspects of society and is especially relevant to industry. All industrial sectors are experiencing exponential growth in automated processes. This automation has had in turn a great growth in the supervision processes, where more and more amounts of data are analyzed and more complex analyzes are carried out. Although communications, computing and artificial intelligence have existed for decades, in recent years there has been massive growth together with the development of new concepts such as Big Data, Machine Learning, IoT, cloud computing, etc. These bring a revolution in data analysis and automated monitoring.

The case of the energy industry is not an exception and therefore the photovoltaic sector. Photovoltaic technology reaches maturity at the same time that the market grows exponentially annually, and this has created new challenges that both the academic and business world must face:

- The size of the plants is increasing, which implies an increase in the elements to be supervised and an increase in the difficulty of their management.
- PV plant portfolios are widely distributed geographically throughout the world.
- The amounts of data to handle from a portfolio of PV plants are massive.
- The growth of PV plant portfolios is very fast, making the management model of these new assets difficult.
- New technologies have appeared in the sector that must be quickly integrated into supervision systems.

There is therefore an immediate need for a paradigm shift in the supervision of photovoltaic power plants, a need for process automation and the introduction of distributed architectures of intelligent models that ensure flexibility, scalability and adaptation to emerging technologies. The sector needs a revolution based on massive data management using massive processing techniques through cloud computing, essential factors for the viability of future photovoltaic solar plants.

The latest reports from the global photovoltaic market show the continuation of the upward trend of the last decade. Forecasts by IRENA (International Renewable Energy Agency) place the market potential for the next decade at close to 3,000 GW, which will mean a five-fold increase in current power [1]. In the long term, it is expected to exceed 8,000 GW in 2050, which will mean a growth rate of close to 9% per year for the next thirty years.

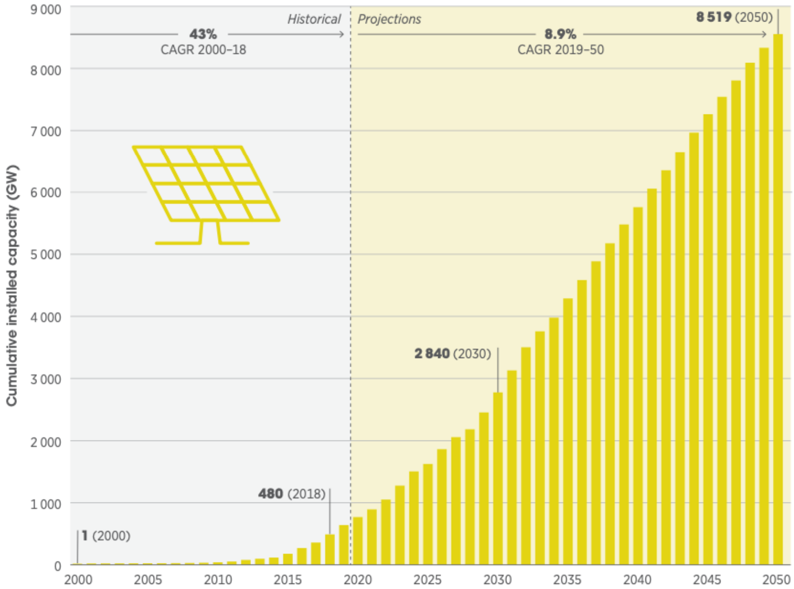


Figure 1.1: Market potential for the next decades. Source [1]**.**

The design and implementation of new photovoltaic plant supervision systems is not limited to new installations but encompasses the total installed power worldwide since old plants also benefit from optimized supervision architectures with the latest improvements developed.

The growth of the photovoltaic market is linked to the increase in the diversity of facilities, both in geographical locations and in the technologies used by the installed equipment. In addition, related to the supervision of these facilities we also find factors that increase their complexity: PV installation size (e.g., BIPV, large PV plant), geographic distribution with different types of sites (e.g., floating), the electrical grid integration, new generation technology (e.g. bifacial) and other new technologies (e.g. drones or cleaning robots).

This document is a state of the art related to the monitoring, analysis and diagnosis of different types of PV installations. The structure of the document is described below:

- Section 2 - Monitoring and Data Quality: for a correct supervision and analysis of PV installations it is necessary to have a good monitoring system. In addition, the quality of the data must be guaranteed, which will be basic for the quality of the subsequent analyzes. This section discusses the issues associated with monitoring data quality and how they are handled.
- Section 3 – PV Data Analytics: The plant operation analysis is used by all the stakeholders of the sector. There are different indicators or KPIs depending on the objective of the analysis or the element analyzed. Also, the different technologies or types of plants may have specific indicators, or some common ones will charge different values.
- Section 4 – Components Digital Twins: The simulation of the operation of the different elements of the plant allows to know their expected behavior. These simulations are based on different models and techniques from parametric models to the use of artificial intelligence, which are widely studied in the literature. This section summarizes the related studies and places special emphasis on the interaction of storage systems with PV.
- Section 5 – PV Faults Diagnosis: In a plant there can be multiple malfunctions. Its detection, diagnosis and analysis of its effects is key for proper supervision. In this section, the state of these techniques in the sector will be reviewed, dealing with general techniques and those aimed at different plant sizes.

- Section 6 – O&M Integration: The automation of the supervision must serve the correct O&M of the plants. Its integration together with new analysis techniques in the field is key.

1.2 Reference material

This document is not based itself on any other document, apart from the information obtained from the references cited in the bibliography.

1.3 Relation with other activities in the project

Table 1.1 depicts the main links of this deliverable to other activities (work packages, tasks, deliverables, etc.) within SERENDI-PV project. The table should be considered along with the current document for further understanding of the deliverable contents and purpose.

Table 1.1: Relation between current deliverable and other activities in the project

Project activity	Relation with current deliverable
T3.2	The specific data analytics for new technologies (bifacial, floating PV, BIPV) needs to consider the data that can actually be registered by the monitoring system (which is studied in section 3.5 of this document), as this will be the input data for the data analytics.
T3.3	The fault diagnosis toolbox for improved O&M in large PV plants and aggregations depends directly on the monitoring system and accordingly, the state-of-the-art of monitoring systems (which is reviewed in section 5 of this document) set the basis for this study.
T3.4	The fault diagnosis toolbox for improved O&M in medium size commercial, residential PV plants and aggregations depends directly on the monitoring system and accordingly, the state-of-the-art of monitoring systems (which is reviewed in section 5 of this document) set the basis for this study.
T3.5	The data analytics based on image acquisition with UAV/drones is always compared with data registered by the monitoring system to cross-check the results and to obtain more accurate and reliable results. This is analysed in section 6.2 of this document.
T3.6	The input data for the digital twins models will be that obtain by the monitoring system, which is assessed in section 4 of this document.
T3.7	The commercial tools developed need to be in line with the current requirements and needs of PV plant monitoring systems which are reviewed in section 6.1 of this document.

1.4 Abbreviation list

Table 1.2: Abbreviation list

Abbreviation	Meaning
ANN	Artificial Neural Networks
ANOVA	Analysis Of Variance
AoV	Angle of View
ARIMA	Autoregressive Integrated Moving Average
BBO	Biogeography-Based Optimization
BIPV	Building Integrated Photovoltics
BMS	Battery Management Systems
CSD	Classical Seasonal Decomposition
CUF	Capacity Utilization Factor
DoD	Depth of Discharge
DQR	Data Quality Routines
DST	Dynamic Stress Test
ECM	Electrical Circuit Models
EIS	Electrochemical Impedance Spectroscopy
EIS	Electrochemical Impedance Spectroscopy
EKF	Extended Kalman Filter
EMC	Electromagnetic Compatibility
EPC	Engineering, procurement and construction
EPI	Energy Performance Index
ESC	Enhanced Self Correcting
FOM	Full Order Models
GHI	Global Horizontal Irradiance
GTI	Global Tilted Irradiance
HPPC	Hybrid Pulse Power Characterization
HPPT	Hybrid Power Pulse Test
HSA	Harness Sub-Array
IEA	International Energy Agency
IEC	International Electrotechnical Commission
IRENA	International Renewable Energy Agency
IRT	Infrared Thermography

Abbreviation	Meaning
IT	Information and Telecommunication
I-V	Current-Voltage
KF	Kalman filter
KF	Kalman Filter
LCOE	Levelized Cost of Energy
LeTID	Light and Elevated Temperature Induced Degradation
LID	Light Induced Degradation
LO	Luenberger observer
LOCF	Last Observation Carried Forward
LOESS	Locally Weighted Scatterplot Smoothing
LOF	Local Outlier Factor
LR	Linear Regression
MAR	Missing at random
MCAR	Missing completely at random
MNAR	Missing not at random
MPPT	Maximum Power Point Tracking
NDS	Network Disturbance System
NOCB	Next Observation Carried Backward
NREL	National Renewable Laboratory
O&M	Operation and Maintenance
OCV	Open Circuit Voltage
P2P	Performance to Peers
PBM	Physics-Based Models
PDE	Partial Differential Equations
PI	Performance Index
PID	Potential Induced Degradation
PIO	Proportional Integral Observer
PLR	Performance Loss Rate
POA	Plane of Array
PR	Performance Ratio
PR_STC	Performance Ratio under STC conditions
PSC	Partial Shading Conditions

Abbreviation	Meaning
PV	Photovoltaic
PVPS	Photovoltaic Power Systems Programme
QC	Quality Controls
QCP	Quality Control Procedure
R&D	Research and Development
RLS	Recursive Least Square
RLS	Recursive Least Squares
RLS	Recursive Least Squares
ROI	Return on Investment
ROM	Reduced-Order Models
RUL	Remaining Useful Life
SCADA	Supervisory Control And Data Acquisition
SMO	Sliding Mode Observer
SOC	State of Charge
SOH	State of Health
SPD	Surge Protection Device
STC	Standard Test Conditions (1000 W/m^2 under spectrum AM1.5 and 25 °C)
STS	Solar Tracking Systems
SVR	Support Vector Regression
TF	Transfer Functions
THD	Total Harmonic Distortion
UAV	Unmanned Aerial Vehicle
UKF	Unscented Kalman Filter
UV	Ultraviolet
WMO	World Meteorological Organization
YOY	Year-on-Year

2 MONITORING AND DATA QUALITY

Monitoring PV systems is a fundamental part of the design, commissioning and maintenance of all types of PV installations. This process includes data acquisition, data transmission, data storage and initial processing for a proper adequacy of the data. The data resulting from this process are the basis of all subsequent analysis in the plant, from the calculation of the main KPIs to advanced diagnostics and the use of techniques based on artificial intelligence.

In addition, these data can be used to compare facilities, verify simulations or even contractual commitments that require guaranteeing established quality levels. This fact causes a need in the sector for research and standardization of this process. In this second line the IEC published an international standard in the PV sector: IEC 61724. Like any norm, this define general guidelines for the co-operation between agents of the sector, but it is these agents who must implement it.

In this section we will show a state of the art of monitoring and data quality, both from an academic and industrial point of view. Therefore, there is a need to define methodologies that allow to unify and optimize this process to face the frequent problems in PV installations, since currently both researchers and companies dedicated to data monitoring present disparate solutions adapted to specific situations.

In the initial phases of commissioning a photovoltaic installation, it is essential to verify the monitoring systems, with special emphasis on three stages:

1. Correct installation of monitoring systems and equipment.
2. Verification of the existence, availability and quality of the data.
3. Evaluation of the performance of the installation and comparison with expectations.

In order to assure the reliability of the data received there are a range of tests and validations needed to verify that the photovoltaic system has been installed correctly both from an installation and configuration point of view. The data itself should not contain extreme or anomalous data, devices have to be correctly monitored and continuously sending data, and also it is necessary to validate various metadata.

Regarding installation problems, photovoltaic modules and the accompanying production and consumption data monitoring systems are often unintentionally incorrectly installed. For example, a frequent error is the ‘upside-down’ connection of current measuring clips which can either give a negative reading for a positive current or for non-amperometric clips, a value of zero will be returned when negative values are not detected.

On the other hand, even if the installation of all hardware was successful, the setup process can also cause problems. The conversion factors of the sensors or their calibrations are usually a source of failures in the configuration of the monitoring system. In the case of BIPV, by identifying which appliances are consumer appliances, which are production appliances, and depending on the monitoring system installed, a variety of other parameters, mistakes can be made that lead to the data being received from an energy-consuming appliance, such as a heater of water, are identified as the energy production of the photovoltaic installation.

Detecting all these types of errors is further complicated by the accumulation of errors. Several different types of errors within the same installation can make individual errors virtually impossible to diagnose, or at least much more advanced and time-consuming techniques are needed (or even human intervention which for thousands of installations may be little realistic).

In terms of type and size of facilities, there are no notable differences in monitoring systems in the PV part, but there are in data storage and processing systems. However, there are differences depending on the level of detail and depth of monitoring. The higher the data quality and accuracy goal, the higher the cost of the complete monitoring system.

Therefore, there is a need to define methodologies that allow to unify and optimize this process to face the frequent problems in PV installations, since currently both researchers and companies dedicated to data monitoring present disparate solutions adapted to specific situations.

Some papers have allowed to study revision monitoring systems as a complete process [2], identifying the needs for monitoring and evaluating the progress of the last decade, allowing to define pending challenges such as:

- Improvement of systems to endure and keep quality measurements despite harsh environmental conditions.
- Reliability and latency requirements.
- Monitoring and prediction of efficiency degradation of PV modules.
- Resource constraints as energy efficiency, data storage and data processing.
- System calibration, focusing on remotely procedures to reduce necessary time and cost of calibration at the operation site.
- Future requirements of the systems as system automation, real time operation and visualization of operational, more efficient data logging, storage and transmission, system control for output maximization.

The level of compliance with these challenges depends on many factors such as the type of installation, size, geographic location, costs and the desired quality of monitoring. In order for there to be defined guidelines regarding monitoring, IEC published its specific regulations (IEC 61724-1) [3]. This International Standard defines classes of PV performance monitoring systems and serves as guidance for various monitoring system choices.

The monitoring system should be adapted to the PV system's size and user requirements. In general, larger and more expensive PV systems should have more monitoring points and higher accuracy sensors than smaller and lower-cost PV systems. The IEC standard [3] defines three classifications of monitoring system with differentiated requirements which are appropriate to a range of purposes.

This standard is a good reference for the classification of photovoltaic installations. In any case, it focuses a lot on the size of the installation, but we must also look for other factors related to technology such as bifacial, floating or BIPV.

Some works present the possibility of creating simple low-cost data acquisition solutions [4], to monitor in real time and under real conditions a small PV system, opening a great way to avoid the high cost of commercial data acquisition systems, which still are a barrier in emergent nations. On other hand, the ability to analyze lower quality data, with data gaps and monitoring failures, was deeply studied in large fleets of small PV installations [5], comparing each one of them with its environment to correct the quality of the data.

The quality of the data directly influences the performance of PV installations. In previous reviews of the measurement systems, results showed that that the quality and availability of data is considered essential for the early detection of anomalies and the automation of these processes, highlighting the importance of the design and discussion of new filters [6].

Data quality also affects to PV systems performance calculation, and special emphasis is placed on the importance of measurements of operating conditions since the linearity with respect to the power shown by the models constitutes a powerful indicator of anomaly detection. In this line the IEC standard also emphasizes it (IEC 61724-2 2016). This standard indicates the types of filtering for both operating conditions and production. It also recommends that these filters should be included in the operation reports, so that the quality of the data used is known.

A specific study where the problems associated with measurement of operating conditions of small PV installations based on a set of meteorological stations storing data every ten minutes of an area are examined [7], presented the data source as a proof of difficulties to measure correctly, and showed that the lack of data and wrong values often brings to incomplete analysis of PV systems.

Taking into account data or measurement problems, the IEC-61724-2 standard is studied as a first filtering approach [8], although it is considered that it causes a loss of information with low irradiance.

The problems associated with filtering data are common in the methodologies proposed by the different studies.

Some of the studies propose a filtering of outliers from normalized data, to try to eliminate measurement errors and noise [9] to variables such as irradiance, clearness index, temperature and site-specific factors.

However, other studies focused on accurately estimating the yield in PV systems present a double filter to get high accurate input data [10].

The proposed solution to data filtering, usually applied to reduce the uncertainty and offsets of the input data, uses the correlation between power output and irradiance to filter outliers of the Gaussian distribution of linear relationship. This method uses and additional step of smoothing curves based on 25th and 75th percentiles assuming that points outside this range are not useful to feed the models and becoming very effective way to remove unrealistic data.

Recently, more complex methods such as clustering and neural networks have also been studied.

The idea behind the use of clustering in anomaly detection is that the outliers do not belong to any other cluster or have their own clusters. The k-means method is one of the best-known and easy-to-implement clustering algorithms, despite the difficulty of choosing an appropriate k-value to clearly differentiate anomalous values from the rest of the data.

To solve this need, there are methods such as density-based spatial grouping of applications with noise (DBSCAN), which does not require any predefined number of clusters and only has two parameters: minimum number of points in a cluster and epsilon, which is the distance between clusters. DBSCAN is a high-performance method that is easy to tune in just a few iterations, constituting a very powerful option to perform anomaly detection.

Moreover, neural network-based approaches allow to detect anomalies even if the knowledge about the data range or data quality is limited, using techniques as autoencoder.

Autoencoder is an unsupervised type of neural networks, and mainly used for feature extraction and dimension reduction, and also it is a good option for anomaly detection problems. In encoding part, main features are extracted which represents the patterns in the data, and then each sample is reconstructed in the decoding part. The reconstruction error will be minimum for normal samples. On the other hand, the model is not able to reconstruct a sample that behaves abnormal, resulting a high reconstruction error. So, basically, the higher reconstruction error a sample has, the more likely it is to be an anomaly.

At the same time as data filtering, it is also common to aggregate data over longer periods (hourly or daily), allowing to reduce the detrimental effects of measurement quality and noise.

In the study of Performance Loss Rate (PRL) in long-term analysis of PV plants [9], a key idea of data quality is presented: the previous works of significance have used various data filtering and aggregation criteria, reaching very different results and making very difficult to compare between analysis. The proposed methodology is based on a process of data collection, normalization, filtering and aggregation.

Results show how the different levels of data filtering and data aggregation affect to variation of PRL and present the methodology itself as a first step to a consensus of PV data cleaning to enable benchmarking across plants.

In addition to data out of range or measurement errors, monitoring presents another major problem, data loss or data missing. The causes can be multiple, including measurement problems, communication failures, problems in the data acquisition system or outputs of the applied filters.

In [5] the author explores the value of the accuracy on PV systems data to evaluate and improve performance, where data quality analysis is performed in fleets of small PV installations (mainly residential PV systems) and methods of automatic detection of data value deviations are based in comparation of each installation with its environment. Research also details magnitude of lack of data in evaluating performance and at alarm activation frequency, establishing itself as a challenge in the balance between yield fault detection and monitoring fault detection. As a conclusion, this works shows how undetected wrong data mask fault detection and leads to a reduction of PV systems lifetime.

As in data filtering, data losses are especially relevant when they affect operating conditions. Related works had analyzed the influence of data loss in energy yield evaluation of PV installations and proposed a method of back-filling data gaps [11], [12]. Results are measured by mean square error, mean absolute error and mean bias error for three different cases: missing meteorological data, electrical monitoring system failure and failure of both systems. Conclusions show that data loss of meteorological data causes underestimation of in-plane irradiation (and overestimation of PR) while gaps of electrical data could be filled and still estimate with precision energy yield.

Detecting missing data is a simple process, but the categorization and labeling of this data influences the next stages of analysis. Missing values are generally represented as null values, to clearly differentiate them from other types of problems when analyzing data.

It is important to distinguish two metrics of missing data: the rate of missing data and its categorization.

The missing data rate is the fraction of missing data to the total number of data points. This rate, with the category of the missing data is important for determining which treatment method is appropriate.

Missing data can fall into three categories:

- Missing completely at random (MCAR). Values in a data set are missing completely at random if the events that lead to any particular data-item being missing are independent both of observable variables and of unobservable parameters of interest and occur entirely at random. When data is MCAR, the analysis performed on the data is unbiased; however, data is rarely MCAR.
- Missing at random (MAR). Occurs when the missingness is not random, but where missingness can be fully accounted for by variables where there is complete information. Since MAR is an assumption that is impossible to verify statistically, we must rely on its substantive reasonableness. Depending on the analysis method, these data can still induce parameter bias in analyses. However, if the parameter is estimated with Full Information Maximum Likelihood [11][13], MAR will provide asymptotically unbiased estimates.
- Missing not at random (MNAR). Data that is neither MAR nor MCAR (i.e., the value of the variable that's missing is related to the reason it's missing).

Some previous work has detailed complete methods for treating missing data [13]. Missing values were treated by data deletion (listwise deletion) for missing data rates lower than 10% and by data inference techniques (application of empirical models, multiple and univariate data imputation) for missing data rates higher than 10% (applicable only for the MCAR case). For the MAR and MNAR cases, the reason of missingness was further examined before determining an approach to handle the invalid values.

Many additional solutions have been proposed for filling data gaps such as:

- Approach based on Multiple Imputation by Chained Equations [7], to artificially generate missing and filtered compared to usual methods as Multiple Linear Regression. Results show the effectiveness of the proposal to recover missing data of meteorological stations with wrong or void values and allowing to analyze the complete time-series of solar irradiation even when some of the sensors fail.
- Classical Seasonal Decomposition (CSD) method [14], which shows the best performance to estimate degradation with data imputation for whole range of missing data percentages, although Autoregressive Integrated Moving Average (ARIMA) works better when data imputation is not applied.

- Data Quality Routines (DQRs) methodology [11], to identify data anomalies and reconstruct datasets with outliers and missing data, could improve PV monitoring systems permeance. Results show that missing or outlying data up to 40% can be reconstructed without significative accuracy loss on PV performance evaluation using Sandia Array Performance Model (SAPM).

Some examples of data imputation under these conditions are the following:

- Last observation carried forward (LOCF).
- Next observation carried backward (NOCB).
- Linear interpolation.
- Spline interpolation. These methods rely on the assumption that adjacent observations are similar to one another. These methods do not work well when this assumption is not valid, especially when the presence of strong seasonality.

However single data point imputation like this can introduce biases into the dataset. It is often preferable to perform multiple interpolation. In a multiple imputation, instead of substituting a single value for each missing data, the missing values are replaced with a set of plausible values which contain the natural variability and uncertainty of the right values. Any analyses are then performed on each 'different' data set. Subsequently, by combining these analysis results, a single overall analysis result is produced [15].

Finally, there is another additional process in the adequacy of the data, the synchronization between sources. When the data comes from different devices of different nature, these can be out of sync for two main reasons: 1) the time base is different (1 s., 1 min., 5 min., etc.) or 2) the data collection is out of sync. Most operating analysis, from basic to advanced, requires combining data from more than one source. Combining out-of-sync data causes analysis errors or directly the inability to perform these analyzes.

Data synchronization may require computationally heavy techniques, or usually through a combination of integrations, interpolations, regressions or transformations [16]. Figure 2.1 shows a conceptual example of a power measurement that is received with asynchronous measurements (black dots). If the rest of the monitoring had measurements at times t_1, t_2,..., t_n, it would be necessary have the power values at these times to be able to combine these measurements. In this case, it is solved by using a linear interpolation to obtain a synthetic value in the time base of the monitoring system.

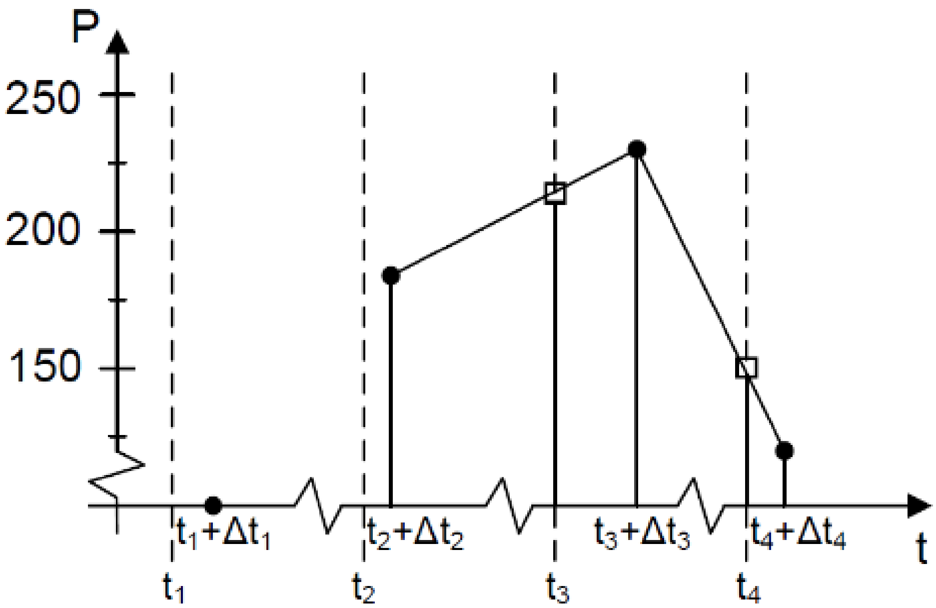


Figure 2.1: Example of asynchronous power measurement corrected by linear interpolation. Source [16]**.**

As we have seen, the process of monitoring and data adapting has different stages that can be approached with different techniques. All these stages are key for a subsequent analysis of the facilities and the complete procedure must be stipulated from the beginning of the operation of an installation. A recent work presents a basic proposal of methodology for PV data processing and data quality verification [13]. The method is composed of these steps:

- Identification of the recording interval and reporting period.
- Dataset examination for consistency (timestamp gaps, repetitive and duplicate records, and synchronization issues).
- Data filtering (daylight filter).
- Data anomalies (identification of outliers and missing values)
- Identification of missing data rate.
- Handling invalid values and dataset reconstruction.
- Data aggregation.
- Data statistics summary.

This sequentially process is a novel approach to a methodology to compare and reproduce PV performance and degradation analysis. Conclusion exposes that analyzing data quality and reconstructing low missing data rates could drastically improve PV systems monitoring and automated performance analysis.

The improvement of monitoring systems, including measurement, data acquisition, communication, storage and pre-processing (data filtering, data cleaning, data aggregation), has to maintain the balance between start-up and maintenance costs, and sufficient quality of the data as well as for performance analysis, degradation and failure detection. The pending challenges go through defining a methodology and specific techniques that serve as a common basis for the monitoring of PV system.

3 PV DATA ANALYTICS

3.1 Operation conditions

Throughout chapter two, the importance of the correct measurement of operating conditions for data analysis in photovoltaic installations has been highlighted. The measurement of operating conditions has two fundamental components: irradiance and temperature.

Regarding irradiance measurements, they are usually divided into global tilted irradiance (GTI) and global horizontal irradiance (GHI), although they can also be divided into its two main components: direct and diffuse. This latter division is not usually measured in most plants currently; however, the relationship between the two components tends to have considerable differences between the models and the reality due to the fact that the diffuse fraction is affected by many local components such as aerosols. This relationship can considerably affect some indices or the estimation of optimal values, such as the gain per tracking. In case a good evaluation of their behavior is required, it would be necessary to take measurements of these components. The common devices in almost every plant are mostly devoted to the differentiation between the GTI and GHI as aforementioned, such as pyranometers, calibrated cells or reference modules.

Lately, the measurement of the front and back radiation of the modules is taking special relevance, in the cases of using bifacial technology. The analysis of the operation of PV plants employing bifacial technology is more complex as KPIs are affected by these two radiation sources. The back radiation is highly influenced by the albedo, a component of radiation related to that reflected by the ground [17]. It is not common to find the measurement of albedo in plants, although the new bifacial plants are installing albedometers in their meteorological stations.

The operating conditions are the input measurement on which the performance of the PV plant will be evaluated. Most analysis and processes directly associated with operation conditions are usually focus on the generation of components from others or the estimation of the operating conditions the day before [18].

In relation to the models used to estimate irradiance (both GHI and GTI), different reviews have reached similar conclusions where the importance of time data aggregation is valued, as well as particular classic analysis of the anisotropic distribution of sky radiance [19], and the reflection component, which can reduce the bias in the GTI evaluation [20].

Another modeling topic studied throughout many previous works is the transformation from GHI to GTI [21] [22] [23], as well as vice versa [24] [25], allowing the performance of photovoltaic installations to be evaluated regardless of the sensor orientation. For this type of conversions, it is necessary to know or calculate the relationship between direct and diffuse radiation.

In addition to the precision of the sensors, measurement points, models and estimations, the effect of clouds is a fundamental factor to study in the analysis of irradiance data. This is usually evaluated by means of the clearness index, which is defined as the ratio of the horizontal global irradiance to the corresponding irradiance available out of the atmosphere (i.e., the extraterrestrial irradiance multiplied by the sinus of the sun height).

The clearness index and its effect have been extensively studied for its effect on the variation of irradiance [26] [27], and for its non-uniform distribution along different geographical locations [28] [29].

The prediction of this parameter has been studied both through statistical techniques [30], and more complex models based on neural networks [31], reaching the conclusion that the quality of the past data and the interval in which the index is to be calculated have a significant influence on the results, obtaining useful values for the modeling of irradiance on a daily basis.

Therefore, the analysis of the clearness index and the effect of the clouds allow different classifications to be made according to the type of day [32]. In a previous review, the classification of days was studied based on the clearness index and the probabilities of inter-day persistence from GHI [15], and in another review a classification of different types of sky was presented as a result of the application of machine learning techniques to identify patterns using irradiance and daylight illuminance data [33].

Another factor that affects the irradiance measurement is the soiling present in the PV plant, as both, the modules and the measurement equipment, are going to be affected by soil effects. To quantify the influence of soil on irradiance, a previous work defined an irradiance loss factor [34], that quantifies the relationship between irradiance, tilt angle and power output of a soiled panel with the soil particle size composition.

Another work related to this effect [35], presents the soil distortion factor (SDF), which establishes a relationship between particle size compositions of soil and irradiance received by a tilted soiled solar panel.

The growth of bifacial technology is also causing an increase in the study of effective incident radiation in the module with respect to different types of measurements carried out in the plant: combination of calibrated cells, albedometers or directly bifacial reference modules. Aalthough there is currently a great growth of this technology, there are studies of this type since the late 1980s [36], [37].

In addition to irradiance, the other fundamental measurement in operating conditions to know the performance of a photovoltaic installation is temperature. Previous works have thoroughly reviewed the relationship between ambient temperature, module operating temperature, and effects on the module behavior [38].

Different models have studied the evolution and influence of temperature on the behavior of the modules, fundamentally divided into predictive models and empirical models. The predictive models studied allow the temperatures to be accurately estimated by aggregating the data on a daily basis [39], although the precision is affected by thermal inertia. On the other hand, the empirical models show a higher precision [40], using wind speed and direction as additional input variables.

3.2 Generator

The analysis of the performance of photovoltaic generation systems has been studied both from a theoretical point of view through parametric models as well as from a practical perspective through empirical models.

Previous reviews have studied both, the operating conditions of photovoltaic modules and the effect on performance caused by the variation of these parameters [41] [42]. The influence on the expected performance of the installation location, the power, material, efficiency and materials of the photovoltaic cells has also been studied.

In relation to the operating conditions, as detailed in the previous section, it influences the photovoltaic performance: global tilt irradiance, cell temperature or other weather conditions as air temperature or wind speed.

Regarding the performance of the photovoltaic generator itself, the key characteristics to evaluate are the PV cell technology, module type and location.

These characteristics are related to each other, since for example the technology of the cells itself influences the behavior of the relationship between temperature and efficiency.

The relationship between the temperature of the photovoltaic modules and their electrical performance (understood as power generation or efficiency), has been widely studied throughout different reviews [42] [43], obtaining different types of correlation without a general criterion, since the most precise model is dependent on the type of installation and intrinsic characteristics of photovoltaic cells.

Numerous reviews have compared the results shown in previous works about the performance of different photovoltaic technologies in the market.

A very comprehensive review of all the technologies related to photovoltaic systems, and not only with the cell technology itself, was published comparing each of the cell types according to their application in small installations, BIPV, large plants and hybrid systems, depending on design criteria such as the purpose of the installation, needs and costs [44]. This work constitutes a very useful starting point for photovoltaic projects, for research or for the design of comparison and decision systems between different technologies.

In this line, a later work presents a decision-making approach model to evaluate the performance of the modules of different technologies even with insufficient data [45].

A previous work comparing different photovoltaic cell technologies presented variable performance results according to the season of the year [46]. According to it, each period of year benefits a series of specific technologies in particular, fundamentally caused by the relationship between temperature and performance presented above. On the other hand, other reviews have focused on the evolution of technologies comparing different PV generations [47], presenting as a conclusion the lines of research to follow to reduce costs and improve module efficiency.

In relation to the location of the generator, there is a key factor for the calculation of the performance, the shading.

Shading (total or partial) occurs when parts of the PV generators are shaded to some degree. This effect can be produced by natural conditions such as clouds, but also by design decisions of the installation, causing a lower energy output than expected.

In order to analyze the effect of shading on the energy performance of photovoltaic installations, previous works have shown methods to detect and quantify the effects through the analysis of the variation of the maximum power point tracking (MPPT) in partial shading conditions (PSC) [48] [49].

Other works have focused on the detection of this shading by combining models (shaded and unshaded) [50], and through more complex techniques such as neural networks [51], allowing not only the detection but also the precise quantification of the losses when comparing with the unshaded models.

To analyze the performance of photovoltaic generators, it is also necessary to consider the degradation of photovoltaic cells. Therefore, long-term performance is expressed as a function of the initial energy generation and the rate of degradation.

Some previous works have presented very in-depth reviews of long-term degradation and have reached the conclusion that this degradation rate is highly dependent on the cell technology and the methodology used to calculate the degradation. According to previous studies, the most recommended statistical techniques to analyze the degradation rate are: Linear Regression (LR), Classical Season Decomposition (CSD), Autoregressive Integrated Moving Average (ARIMA) and Locally Weighted Scatterplot Smoothing (LOESS) [52].

Although the degradation was considered initially as a linear function, some studies have shown the non-linearity of the degradation rate through Monte Carlo simulations with multiple scenarios and different measurement methodologies [53], reaching similar conclusions to other studies, where the rate is dependent on technology and methodology but showing that the linear approximation is insufficient for the necessary precision when designing photovoltaic projects in a large time interval.

These conclusions of non-linear rate degradation have been reaffirmed in other works that analyze degradation in real facilities over long periods [53] [54].

In relation to the variations in the degradation rate found according to the calculation methodology used, more recent studies have presented proposals for robust methodologies applying the year-on-year (YOY) method combined with a clear-sky model [55]. This method allows to avoid noise and other possible errors that may exist due to variations in operating temperature or in sensor calibrations, obtaining a significantly higher accuracy than previous methodologies, although it relies on high-quality data and clear-sky days.

3.3 Inverter

The performance of the inverters is directly linked to the performance of the complete photovoltaic installation. Grid-connected systems have a series of requirements that inverters must meet such as low total harmonic distortion of the currents injected into the grid, maximum power point tracking, high efficiency, and controlled power injected into the grid.

A previous review had deeply studied inverter topologies and configurations, and its control strategies for grid connected systems [56].

The problems associated with the grid-connected PV system are the grid disturbances if suitable and robust controllers are not designed and thus, it results in grid instability. The control strategies presented in [56] are divided in linear controllers, the non-linear controllers, the robust controllers, the adaptive controllers, the predictive controllers, and the intelligent controllers.

Inverter topologies had also reviewed and classified [56] [57] [58] dividing by the number of power processing stages, power decoupling, connection to transformers, multilevel inverters and soft-switching inverters.

The grid-connected inverters undergone various configurations can be categorized in to four types, the central inverters, the string inverters, the multi-string inverts and the ac module inverters [59], shown in the figure:

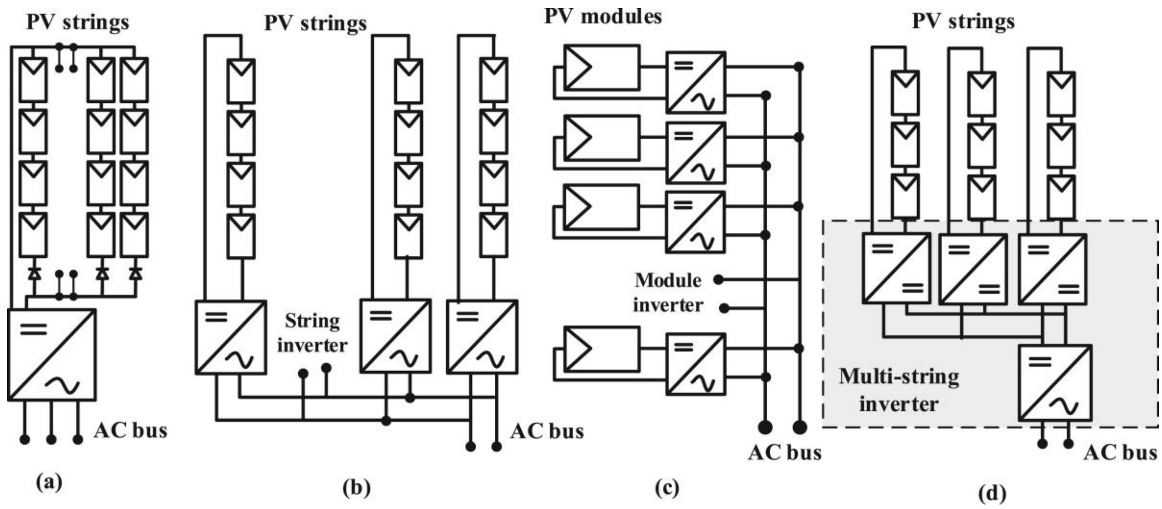


Figure 3.1: Configurations of grid-connected PV inverter [59]

In relation to the efficiency of inverters, both their modeling based on electrical parameters and their variations at an experimental level depending on the operating conditions have been deeply studied [60] [61]. Previous reviews also show that the inverter's topology and configuration are related to its efficiency [62].

Some methods of improving the efficiency of the inverters have been studied [63] [64], adjusting the power generation in situations of power limitations so that the inverters always work in optimal ranges of efficiency.

The factor that most affects the long-term performance of a photovoltaic installation is the failures of the inverter, since the losses produced by failures in this equipment represent very large losses with respect to the total installation (up to the total production in the case of installations with a single inverter).

The effect of failures in performance has studied in previous works, quantifying up to 10% of ROI reduction in residential solar PV systems [65], and about 5% in large PV installations [66].

3.4 Tracker

The analysis of data from Solar Tracking Systems (STS) is a fundamental component to determine the performance of photovoltaic installations where photovoltaic modules are mounted on trackers.

Among the previous reviews analyzed, there are works related to the type of mounting of the sun-tracking [67] [68] [69], analyzing and comparing the energy gain produced according to the size, distribution and number of axes and elevation, as well as the consumption associated with this equipment to be able to analyze the profitability of including these systems in photovoltaic installations.

On the other hand, other reviews focus on the control algorithms for the STS to track the optimal solar position [70], comparing both open-loop and closed-loop algorithms, determining that the choice of algorithm type is dependent on the type and the quality of the sensors used.

A recent review classifies active STS based on their technologies and driving methods [68], defining five categories:

- Sensor driver systems.
- Microprocessor driver systems.
- Open and closed loop driver systems.
- Intelligent driver systems.
- A combination of two or more of these driver systems.

Presenting the intelligent driver systems as the most promising among these tracking systems due to their capability to predict the exact position of the sun by using learning algorithms, like those presented in papers related to intelligent drivers for tracking based on adaptive neural fuzzy inference system [71].

The gain in yield obtained through STS has been extensively studied under different conditions.

A previous work presents the comparison in annual yield of a fixed PV system compared to a double-axis sun tracking system [72], with differences greater than 30% with the normalized data.

The increase in yield through tracking systems versus fixed systems has also been analyzed in relation to the geographical area, finding significant differences between the colder regions and hot regions [73], where the former has a higher tracking yield gain due to the fact that in hot areas there is overheating of the PV panels.

In relation to the gain obtained by STS according to environmental conditions, a previous work analyzed the results in different types of day according to the clearness index [74]. The best results were obtained on a totally clear day, being gradually less until totally cloudy days where there are no significant differences between the STS and the fixed systems because the tracking gain is compensated with the consumption of the motors and control electronics of the tracking equipment. As mentioned in section 3.1, the diffuse fraction also has great relevance in the gain per tracking.

A recent review of the advances in STS and future challenges [75], analyzes results obtained with the latest tracking technologies (both at the level of equipment and control systems) and quantifies the gain on clear days above 40%, although this value it is dependent on many factors that can cause failure to follow. It presents as possible errors and, at the same time, challenges in STS the misalignment of the tracking fixture, the level of pollution of the area, the shading of the sensors, the types of control schemes involved, the auxiliary units of the system, the lack of maintenance as well as the imperfection and power mismatch of connecting grids.

3.5 Peculiarities of technologies:

3.5.1 Bifacial

The rise of bifacial photovoltaic technology has also increased the interest of research in this field, from reviews of the operation of the technology itself [76], reviews about performance and efficiency analysis [77], to reviews about the characterization and simulation of this type of modules [78].

The review presented in [76] shows the evolution in the study of bifacial photovoltaic technology through its comparison with monofacial technology [79] [80], showing power outputs up to 30% higher, as well as its performance at different latitudes [81]. This point is analyzed in depth in another work, reaching the conclusion that this technology shows its full potential in latitudes higher than 40º or in lower latitudes with high albedo [82].

The electrical performance of bifacial modules is influenced by some common factors with monofacial modules such as tilt angle [81] or orientation [83]. However, the performance is largely influenced by the reflection of the earth (albedo) [84], and by the elevation of the modules [85], since both factors affect the irradiance received by the rear face of the module.

Related to the performance of the photovoltaic bifacial modules, a previous work explores the differences in performance with respect to the monofacial modules throughout a full year [86], showing small increases in performance on days with clear skies but very significant increases on cloudy days, mainly due to diffuse radiation.

Estimating photovoltaic bifacial modules performance and its behaviour requires a deep study of optical, electrical and thermal models [87].

A bifacial module can absorb a part of the sunlight to generate electricity with various optical losses. These optical losses occur at different interfaces on both sides, including reflection loss, absorption loss and transmittance loss [88].

The electrical model characterizes the bifacial gain, expressed as the relative yield of bifacial module in comparation to monofacial module under the same operating conditions [79] [89].

In relation to the thermal model, several approaches have been studied for the technology of bifacial modules such as Sandia [90], NOCT [85], or the PVsyst model [91], and all the models showing more precious results are the empirical ones such as Castillo’s model [92].

The electrical performance of the bifacial modules depends on the irradiance received by the rear face of the module. This radiation and its estimation have been extensively studied both at a theoretical [36] [93] and experimental level [94], and depending on the conditions of the installation and the operating conditions. The results have made it possible to create tools for calculating the rear irradiance [95].

3.5.2 Floating PV

Floating PV shows significant advantages in comparison to ground PV, such as the possibility of using large areas without competing for land use, the possibility of generating electricity very close to large energy consumers such as coastal areas with high population density, or the highest performance of PV modules due to the lower operating temperature in the marine environment [96] [97] [98].

However, all these advantages may be hampered by an increase in LCOE associated to the following factors:

1. The reduction in the energy production as consequence of a **higher rate of degradation** of the elements composing it [99]. The analysis of the state of art reveals a significant shortage of bibliography on the matter, since the technology has less than 5 years of experience in terms of large-scale projects. In this sense, it is expected that the failure modes that may originate specifically in floating PV technology will become known in existing plants in the coming years. However, on the

basis of the knowledge formed in ground PV, it is possible to anticipate what types of degradation could be predominant in floating PV. These would be linked to the main stressors in the marine environment: high relative humidity and high saline concentration (sea water), and high levels of mechanical stress (waves and wind). As a result, main non-transient failure modes that can be expected are:

a. The aquatic environment and high humidity conditions favour <u>potential-induced degradation (PID)</u> [100]. PID is a type of degradation in PV panels due to the presence of leakage currents caused by the presence of potential differences between the cells and the ground. These undesired currents can flow through the aluminium frame, glass or even other encapsulating materials (EVA, Tedlar) causing a panel power of up to 30%.

b. <u>Corrosion</u> of the different elements that make up the electrical circuit, both within the PV module and in other elements of the balance of the system, such as cables and interconnectors [101]. The delamination of the backsheet that protects against corrosion and electrically isolates the solar cells and the internal electrical circuit of the module [102]. Saltpetre can cause deterioration of both the mechanical elements such as the aluminium frame or the connections, especially if the module's junction box does not have the proper protections (IP 68), as well as the solar cells and encapsulants.

c. <u>The generation of breaks</u> or "cracks" in the component cells of the PV modules. The continuous movement of the panels due to the waves causes the formation of breaks or micro-breaks in the cells that make up the solar panel. Today it is still difficult to quantify the impact that these failures have on the efficiency of the modules during their useful life. Several studies on the matter [103] [104], conclude that the presence of micro-breaks in the panels has marginal effects on the power generated, as long as the different parts of the cell continue to be electrically connected. However, as it continues to operate, the module is subjected to stress or atmospheric agents of a different nature that cause the cracks to evolve into severe breaks, giving rise to inactive areas, hot spots and consequent physical deterioration [96].

Additionally, the marine environment conditions will induce other types of failure modes that could be defined as transient, since they produce a temporary loss of performance. These transient failure modes include:

a. In the case of the marine environment deposition of sea salt on the surface of PV panels is one of the factors causing <u>soiling</u> [105], along with the accumulation of bird droppings [106] and the formation of algae or other organisms (fungi, lichens). Losses due to soiling generally refer to the loss of electrical power generated resulting from the accumulation of these elements and which have the common effect of forming shadow areas in the surface area of the PV panels. The magnitude of this accumulation can be influenced by the wind, humidity and temperature of both the air and the modules themselves. Likewise, the type of dirt is specific to each environment and its accumulation is determined by structural factors such as the surface of the materials used for the PV modules or their orientation and inclination [103].

b. <u>Changes in the orientation and/or elevation of the PV array</u> as a consequence of the action of wind, waves or sea currents leading to mismatching effects and their corresponding performance losses [107].

2. **An increase in O&M costs** due to the intrinsic difficulties of accessibility and management of these kind of activities in a floating PV plant. It is important to remark that the aquatic environment introduces significant occupational risks for the operators in charge of the O&M tasks, such as falling into the water or electric shock risk if they are carried out with the floating PV system in operation [96]. All these factors can have a decisive influence on the final costs of O&M, which thus becomes a critical factor.

Thus, the O&M of floating PV systems must simultaneously address two confronted issues:

1. It must ensure that the floating PV plant is operating at high performance most of the time by detecting early potential failures.
2. It must reduce the tasks the operators must perform on site, considering the costs and technical difficulties associated with them.

The optimal solution to this apparent dilemma is through the use of digital techniques and the development of early failure detection and diagnosis (FDD) methods, also known as digital O&M. Digital O&M must enable the fast and efficient management of problems that may arise in a floating PV plant. Different FDD methods have been proposed depending on the type of data used and the way in which these data from the SCADA system of the PV plant are used [108]. The development of hybrid physical/statistical models are of special interest to the PV industry, as well as representing a significant challenge for digital O&M [109]. In the specific case of digital O&M applied to floating PV, the absence of an extensive historical data on the main existing failure modes and their impact on the performance makes it difficult to develop these hybrid models. In addition to this, most of the physical models used in ground PV must be readjusted or may even in some cases are not applicable for floating PV. Some examples of issues that are likely to need to be addressed could be:

- The irradiance on the Plane Of Array (POA) will suffer from oscillations due to the wind, waves and water currents that must be introduced in the optical model.
- The thermal performance of the PV panels may be not only influenced by the ambient temperature, incident irradiance and the wind speed, but also by the potential evaporation of the water deposited on the rear surface.
- The electrical performance of the floating PV array will be greatly impacted by mismatching effects, derived from the heterogeneous operating conditions such as the potentially different albedo or orientation and inclination of PV modules.

3.5.3 BIPV

BIPVs are usually integrated together with a broader concept that includes different elements of the building where they are integrated. These elements are of 3 different types: consumption, storage and connection to the electrical grid. The last element can evolve to smart grids as a convergence of ICT with power system engineering [110]. The Figure 3.2 shows a conceptual scheme of these elements [111].

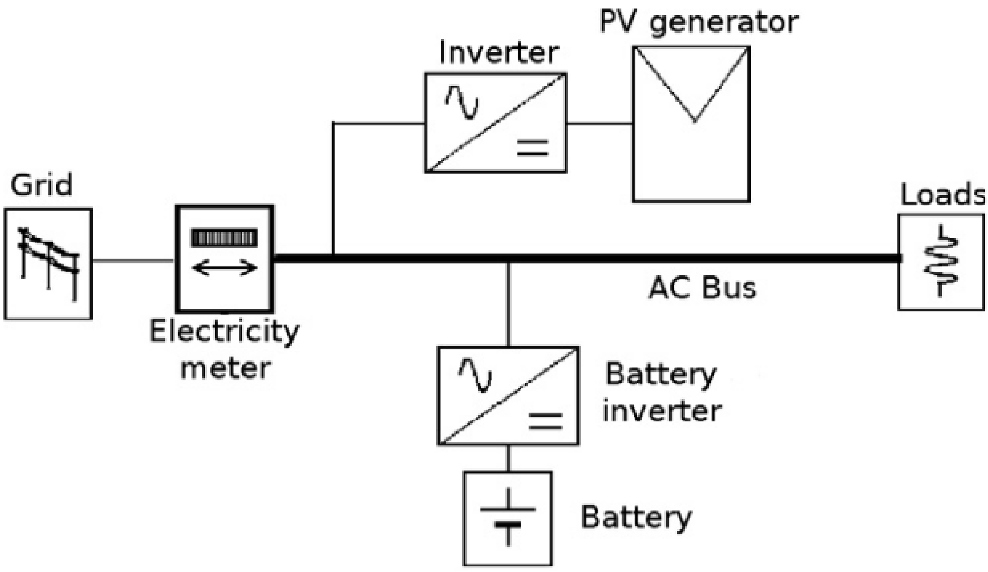


Figure 3.2: Topology of the power system in a building with BIPV, storage, consumption and grid connection. Source [111]**.**

The operation and its evaluation can be affected by the rest of the elements. This must be taken into account in the analysis, both when modifying indexes and methods already used and when creating new ones. For example, the use of solar energy to feed local consumption is widely studied, where are used indices such as Load Matching or Self-Consumption [111] [112] [113]. These indices are generally higher as the locally generated energy is used more and with them the associated benefits are increased. A concrete example of the use of these indices to evaluate the integration of BIPV with the rest of the elements is Solar Decathlon contest[1]. It is an international multidisciplinary competition in which around 20 university teams build and operate energy-efficient solar-powered houses. Each participating house is designed to optimize the use of its solar energy and the teams carry out different strategies to optimize this use, being a good example of how the evaluation of the BIPV operation can be strongly affected by different elements external to the PV [114].

1 https://solardecathlon.eu/

4 COMPONENTS DIGITAL TWINS

A Digital Twin -- the concept of digitally mirroring the behavior of a physical PV power plant in its operational conditions -- is one of the pressing challenges for the continued development of the industry. Due to the increasing penetration of solar energy in the grid, associated grid management issues are becoming commonplace. For solar energy's successful integration into the grid, the behavior of a powerplant has to be known to a greater degree of certainty, particularly from the monitoring and forecasting points of view.

One of the ways how to achieve it, and what some experts are calling for, is the parametrization of real PV output, which in turn can then be used to more accurately describe the inner workings of PV power plant components (inverters, PV arrays, etc.) and use it to refine the simulated output [115].

On the other side is the well-known single diode model [116], a cornerstone of today's simulation of the DC side of PV arrays, which is widely used in the industry to simulate the power output. However, once the power plant is built, it is often difficult to precisely match the real-world conditions of the site since input parameters (solar irradiation, meteorological parameters etc.) have inherent uncertainty. Moreover, a physical approach such as the single diode model, however, applies to ideal conditions, which unfortunately seldom occur in real life operations. Oftentimes detailed and accurate information (metadata) about PV plant is missing or are even incorrect. This is why there are numerous PV analytics regimes estimating this information from the production data [115] [117] [118]. Moreover, data quality procedures, such as those described above in Chapter 2, give more insights. Parametrization from production data is for example introduced in also in [119] the calculated parameters are then plugged back into the single diode model.

There are currently only a small number of research papers that discuss parameterization of the whole powerplant for construction of a digital twin in its full context and potential. One of them [120], uses the combination of various physical models, with parametrization and a genetic algorithm (GA), to construct a digital twin. While it addresses most of the aspects of a PV system's digital twin, there are many open questions where research has not yet begun. For example, the interconnection of the initial simulation using the well-known single diode model [116], with the real-world data monitoring and its refining over time, once more data become available.

Overall, this concept is at the cutting edge of current research and development. The importance of addressing the issues of PV power plant simulation and prediction of yield will continue to increase, and while the Digital Twin concept is still in the conceptual phase, the successful implementation of Digital Twins would address the shortcomings of other simulation and projection models and ease the grid integration challenges the industry is facing.

When modelling the real-world, there are three main different kind of models:

1. <u>Physical models</u> describing the known relationship between physic magnitudes. The main inconvenience is that are generally ideal, as degraded conditions are difficult to be estimated.
2. <u>Statistic or data driven models</u> based only on data and learning from previous events. They can accurately determine the existing correlation among the input and output variables of the model, but they are not able to predict the performance if input conditions have not happened before, nor carry out a diagnosis describing the reasons of deviations from the expected performance.
3. <u>Hybrid models</u> acquiring knowledge from the data but preserving the physical relationships. They allow a better characterization of the asset performance, optimizing the physical model with its specific design and operation characteristic, detecting anomalies and estimating its health. They can also provide a prediction of never-happened events.

A digital twin is a virtual representation of a real-world system, combining hybrid models of devices to reflect the most important asset behaviors to be explored. These digital twins can be used for design optimization or improvement of O&M activities, by means of characterizing and managing asset production, condition and degradation at the device and plant levels.

The precise characterization of an asset will allow to gain better understanding of failures that have already occurred and to develop early failure detection to improve predictive maintenance. However, monitoring data type and volume are essential to develop an effective digital twin. For instance, time resolution of available data will clearly determine which kind of characterization can be done. When dealing with power electronics, if the monitoring frequency is high (kHz), fast events like voltage peaks and grid unbalancing can be detected and their effects analyzed. But if the available monitoring data are every ten minutes, only long-term degradation effects can be examined.

Another of the most important benefits of a digital twin is the possibility of exploring potential scenarios in a safe and non-intrusive way. Thus, it is possible to simulate how the different actions that can be carried out affect the devices and system performance, even estimating the Remaining Useful Life (RUL).

As previously explained, a digital twin is a hybrid model, consisting of the optimization of a physical model with monitoring data to estimate the parameters in the model that minimize the residue. The residue is defined as the difference between the magnitudes measured in the real world and the ones estimated by the hybrid model, as shown in Figure 4.1. This optimization to minimize the residue can be carried out through genetic algorithms, gradient descent algorithm and/or recurrent neural networks, among other data analytics approaches.

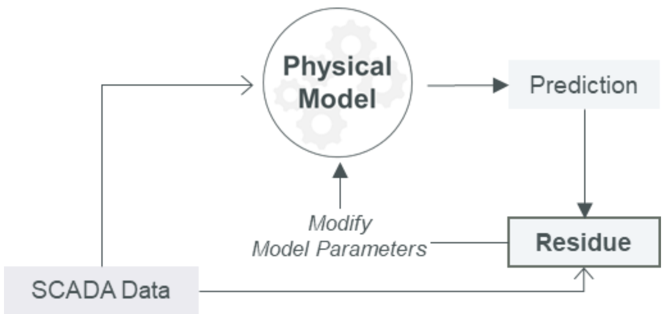


Figure 4.1: Optimization of a physical model to transform it into a hybrid model.

4.1 PV inverter digital twin

PV inverters are subject to frequent functional and environmental stress, which can induce failures. Recent studies show that PV inverter is responsible for almost a 30% of total energy losses due to failure modes in large PV plants [121]. Furthermore, in contrast to PV modules, unexpected failure of a single PV inverter means losing a significative ratio of the plant power production, which could even lead to regulatory penalties for failing to meet power production commitments.

Inverter failure modes may be both externally and internally caused. Among external causes we can find insulation failures on the PV array which according to inverter safety standards require the inverter to stop, or other inverter stops due to grid parameter values (voltage and frequency) out of grid code thresholds.

Main internal fault root causes are related to conversion power stack (semiconductors, DC-link capacitor, drivers, etc.), cooling system due to poor or lack of maintenance, fan failure, etc., and other less significant internal failures related to electronic boards, main AC and DC switches, etc.

There are two main categories of failure mechanisms in power converters: (1) abrupt failure due to singular events, manufacturing failures or unforeseen overstress conditions, almost independent from the application; and (2) degradation due to long-term operation under the particular working conditions of the application. It is very difficult to anticipate and work on the diagnosis of the first ones, so PV inverter digital twin aims to characterize the long-term degradation. After long-term operation, some components of PV inverter, mainly capacitors and IGBT/MOSFETs, become too fragile to cope with the electrical and thermal strains, leading to a collapse of the entire system. In fact, degradation process is primarily related to the thermal cycles of the components, as their dilation and compression during the heating and cooling process induce voltages that finally break the component. Thus, capacitor manufacturers provide information about the lifetime of their component depending on the ambient temperature and current ripple through it, as shown in Figure 4.2. IGBT and MOSFET manufacturers also estimate the lifetime of their switching devices as a function of their junction temperature, number and duration of thermal cycles. The main inconvenient in this case is that it is impossible to measure directly the junction temperature.

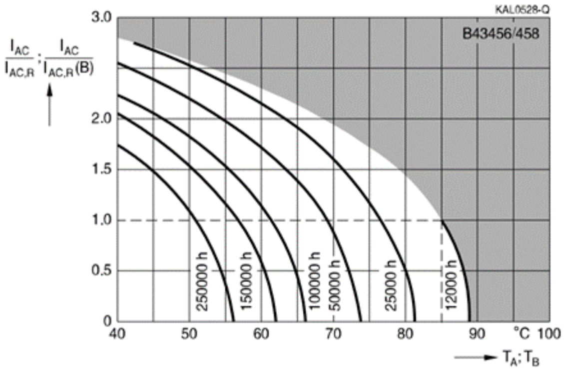


Figure 4.2: EPCOS/TDK capacitor lifetime as a function of temperature and current ripple.

The anticipation of the degradation progress of power converters is important and valuable, so failures can be early detected providing time for performing maintenance and, if necessary, for acquiring and installing a replacement device or the right parts.

It is observed that the degradation progress of the main components in power converters can be indicated by the alteration of their characteristic parameters, like the conduction resistance of switching semiconductors and capacitance of capacitor [122] [123]. Therefore, the measurement of those health indicators is critical and the first step of condition monitoring.

Component-level health indicators have been proposed for monitoring the degradation of power semiconductors and capacitors individually [124] [125] [126], which can be classified into two groups: electrical and thermal indicators. In [124] a review of past developments and recent advances in the area of condition monitoring and prognostics for IGBT is presented. In [125] reliability models for power electronics, including dominant failure mechanisms of devices are described first, followed by a description of recently proposed condition monitoring techniques. In [126] the capacitor condition monitoring methods are classified into three categories, then the respective technology evolution in the last two decades is summarized.

Regarding switching semiconductors, electrical indicators can be obtained from either the drain-source/collector-emitter terminals, such as the conductance voltage or resistance of power semiconductors, which can be measured by using measurement circuits [122] [127] [128] [129]. In [122] a custom designed accelerated ageing platform that can expose multiple discrete power MOSFETs to thermal stress simultaneously is introduced. In [127] required equipment and methodology for an advanced accelerated power cycling test of IGBT modules is presented. In [128] an in-situ diagnostic and prognostic technology to monitor the health condition of IGBTs with a focus on the IGBTs' solder layer fatigue is proposed. In [129] results of accelerated aging tests in IGBT are presented showing that the case temperature, the collector current, and the collector-emitter voltage are the failure precursor parameters that can be used for the development of a prognostic and health monitoring system for IGBTs and other medium-power switching supplies.

All these indicators show the highest sensitivity to the degradation of power semiconductors, but additional circuits are needed, which increases the implementation complexity. Existing gate-emitter related indicators includes the threshold voltage and miller plateau of power semiconductor, which can be obtained from the gate turn-on transient voltage waveform [130] [131] [132] [133]. In [130] gate oxide degradation mechanisms and effect are summarized, and the MOSFET turn-on process is analyzed. Then, a theoretical model is established to describe the relationship between miller platform voltage and two types of gate oxide defects, and miller platform voltage is identified as a new precursor. In [131] the duration of the Miller plateau during the IGBT turn-on transition is proposed as an online precursor indicating two dominant types of failures. In [132] a method for in situ high-bandwidth junction temperature estimation of insulated-gate bipolar transistors is introduced based on the acquisition of the gate voltage plateau during turn-on.

On the other hand, since junction temperature is strongly related to the health condition of power semiconductors, temperature sensitive electrical parameters can also be used as health indicator, such as pre-threshold voltage in [134], peak gate current in [135], kelvin-emitter voltage [136], and switching time in [137]. However, all of these gate related indicators are high frequency signals, demanding the high-speed data acquisition circuit with good noise-immune ability, which increases the complexity further. Moreover, the malfunction of the added circuit may induce the failure of the gate driver. Thermal signals are also proposed for condition monitoring, such as case temperature in [138] [139]and thermal resistance in [140]. However, case temperature may be easily interfered by other heat sources and shows a low degradation sensitivity. The measurement of thermal resistance strongly depends on the accuracy of measuring junction temperature, case temperature and power losses of interested module, which is complicated and difficult.

As for the capacitor, the condition monitoring can be achieved through two ways: 1) taking advantage of the discharge process of capacitor when the power converter is in off-line [141]; 2) obtaining the equivalent series resistance (ESR) and capacitance of the capacitor by measuring the ripple of capacitor voltage and current [123] [142]. However, in practice, obtaining the ripple requires both the data acquisition apparatus with higher sampling rate and higher resolution, and high frequency pass filter circuit [143] [144].

From the system-level view, various methods have been proposed to monitor the power semiconductors and capacitors, respectively. The frequency response of DC-DC converters is sensitive to the on-state resistance of power semiconductors in specific situations. In [145]. In addition, the output current harmonic of inverter is investigated to monitor the degradation of the solder layer of power semiconductor in [146], showing that low-order harmonics, caused by nonideal switching, are affected by the device junction temperature, which in turn depends upon module solder condition. Both methods require extra setups and show invasive to the system of interest. Moreover, they cannot distinguish the degradation of power semiconductor and capacitor.

Artificial neural network is also a potential way for monitoring the degradation of capacitor. In [147] [148]. However, both approaches require offline testing to obtain the enough training data, which is difficult to achieve in practice.

Conventional model-based parameter identification methods are used to modify the controller of power converters. It is known that the transfer function between the output voltage and duty cycle ratio is discretized when design controller. Thus, the coefficients of transfer function can be calculated by using different algorithms, such as recursive least square (RLS) in [149] and Kalman filter (KF) in [150], which is effective in tuning controller and improving the system performance. However, mapping the coefficients of transfer function into the internal parameters of the converter could cause transfer errors and even doesn't have feasible solutions when the number of unknown parameters is more than that of the known equations. In [151], a simplified model of boost converter is built, and a generalized gradient descent algorithm is applied to calculate the inductance and capacitance. A model for buck converter is developed in [152], where biogeography-based optimization (BBO) method is used to identify the internal parameters. The main issues with the above methods are that none of them focus on the degradation monitoring of the key components. Moreover, only the coefficients of the model (e.g., transfer function) or part of those physical parameters (e.g., inductance and capacitance) can be obtained. In addition, all of the above methods need to inject extra signal into the controller. It is worth to mentioning that these measured indicators need to be processed further to reduce noises and to indicate the health status of power converter numerically. Such as lowpass filter and gaussian process [153] [154]. Finally, Calibration with other impactors (e.g., temperature and current) is needed as well [127].

According to the analysis above, the challenges is to come up with a method non-invasive, calibration-free, without additional circuits and with the ability of monitoring both power semiconductor and capacitor.

Digital twin is a virtual representation of a physical system that virtually shares the same characteristics with its physical counterpart. It enables customers to better understand, optimize, predict, and monitor the performance of its installed systems [155]. The concept of digital twin has been applied in power converters for fault diagnosis recently [156] [157] [158], which is achieved by comparing the output signals of the digital twin and its physical counterpart in real-time. The digital twin technology includes two parts: the digital presentation of a physical system and an advanced algorithm for data analysis.

In [156] a digital twin of a power converter is defined as a real-time, probabilistic simulation model with stochastic variables, developed using generalized polynomial chaos expansion. The digital twin models are partitioned in perspective of control layers for power converter subsystems in the approach, with emphasis on the application and converter control layers.

In [157] presents the design methodology, mathematical analysis, simulation study, and experimental validation of a digital twin approach for fault diagnosis of a distributed PV system is presented. The fault diagnosis is performed by generating and evaluating an error residual vector, which is the difference between the estimated and measured outputs.

In [158] Particle Swarm Optimization algorithm is applied to carry out the data analysis and update the digital twin continuously according to the data coming from existing sensors in its physical counterpart, making the residue smaller than a pre-set threshold. Finally, a data-cluster concept is proposed to cover the estimated indicators at different possible operations.

4.2 Battery digital twin

Nowadays, lithium-ion is the fastest growing and most promising battery technology in the short term. It is widely used in several low power electronic devices but also in higher voltage applications, such as electric vehicles or electricity network support systems. The main advantages of this type of batteries are their high density, high efficiency and their relatively long cycle life [159]. Regarding storage systems for PV applications, also lithium-ion is the most widely used battery technology, in both distributed PV systems to increase self-consumption rate and in large PV plants to provide ancillary services to the grid.

The estimation of the State of Charge (SOC) and State of Health (SOH) of the battery is a challenging task due to the difficulty to assess the electrochemical phenomena inside the cells of a battery system, which depends on several factors including the characteristics of the battery, the operation conditions (temperature, current, Depth of Discharge - DoD) of the previous cycles, etc. This causes uncertainty on the expected performance and ageing of the storage, especially, towards its end of life.

Battery lifetime is highly dependent on application conditions. Usually, several ageing processes take place concurrently due to the complex operation conditions of most of the applications, i.e., combination of cycling, rest periods at different SOCs, partial charging, wide range of temperatures [160]. This happens in the case of independent cells and in the case of multicell systems. Conventional Battery Management Systems (BMS) try to achieve a uniform operation and ageing of cells based on the voltage control of the series elements within a battery pack. However, this does not prevent that, at the end of life of the module/s, all cells have different condition. This happens due to different temperatures within the package and because, even when cells are of the same type and manufacturer, their performance might not be identical due to manufacturing dispersion, which causes different ageing speed in battery elements.

As explained above, the digital twin concept is being developed in many engineering fields. It is basically a model, which main characteristic is that it is linked to a real system. The model will try to replicate the real system performance, and this will permit to have a broader understanding on the variables that have an influence on it, especially of those that cannot be directly measured with sensors or that are not precise enough due to noise in the process or in the measurement. The objective is to improve the operation and maintenance of the battery-based storage systems.

In the case of the battery, the digital twin proposed for improving O&M will be based on a hybrid model consisting of both a mathematical model, describing the electrical behavior of the cell, and a data driven model. This approach will permit to have a better estimation of the SOC and SOH, and a prediction of the Remaining Useful Life (RUL). However, these parameters are not directly measurable and, for this reason soft sensing techniques will be used to try to estimate the evolution of the characteristics of the system out of the available measurements.

4.2.1 Physical model of the battery cell

If we consider a physical model as a conceptual or mathematical model of a physical system, we could say that batteries can be modelled in accordance with their thermal characteristics, electrochemical nature or as electrical circuit [161].

4.2.1.1 Battery cell electrical circuit model

Cell equivalent Electrical Circuit Models (ECM) are behavioral models, which attempt to reproduce the cell electrical characteristics. As equivalent electrical circuit, battery cells may be defined through different configurations. The most common is to consider an internal resistance plus a number (*n*) of RC blocks. In [162], general types are considered, from simpler to more complex.

The simplest model consists of a voltage source, representing the Open Circuit Voltage (OCV), which is function of SOC and temperature plus a resistor in series.

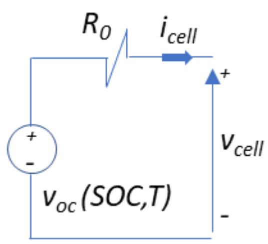


Figure 4.3: ECM. Simple [162]

If we add a RC block to this circuit (first order), we obtain the so called *Randles* model. In some cases, more than one RC block is added (higher order). According to [163] [164], the first order RC model is the best choice considering model complexity, accuracy and robustness for LiNMC cells. However, other cell types with more complex dynamics, e.g., lithium iron phosphate, may require more RC blocks, at least three according to [165], or one state plus hysteresis, according to [164]. In [161] [166], two RC blocks are considered, even if this is not justified on the top of a specific Li-ion chemistry.

The following figure shows the *Randles* or **RC model**:

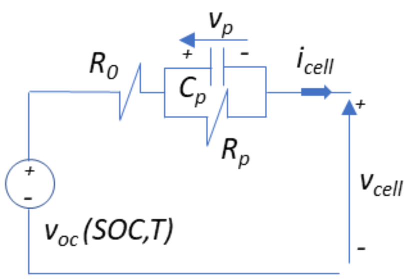


Figure 4.4: ECM. RC model [162]

The third model adds the hysteresis effect, v_h, in the cell's terminal voltage. This is called the **Enhanced Self Correcting** (ESC) model in [162].

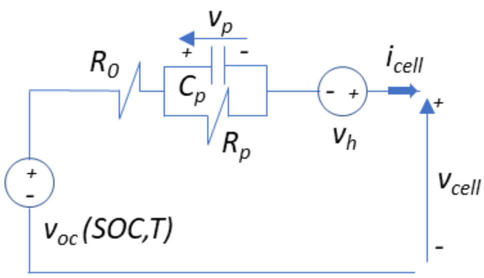


Figure 4.5: ECM. ESC model [162]

In [164], twelve battery models are compared, most of them variations of the three models presented above.

Once defined the model topology, parameter values must be determined based on the real performance of cells. In [163] [164], the following characterization tests are performed at three different temperatures (10°C, 22°C, 35°C):

- **Static capacity test**: the battery is tested extracting electrical energy from it over a time (normally one hour at 1C). This test is related with the discharge capacity, and the test would not return the actual capacity (total capacity), but the batteries capability to produce output equivalent to its ampere-hour rating as a minimum. Following this last reference, a battery is fit for use if its discharge capacity is more than 80% for one-hour discharge at 1C, which is equivalent to discharge lasting, at least, 48 minutes.
- **Hybrid Pulse Power Characterization (HPPC)**: the HPPC is used to determine the dynamic power capability over the usable voltage range. The voltages recorded are used to establish the cell's OCV behavior. From a 100% SOC situation, the battery is discharged in 10%DOD steps at constant current, each followed by one-hour rest period (to return to electrochemical and thermal equilibrium).
- **DC resistance test**: The resistance of modern lead acid and lithium-ion batteries stays flat through most of the service life. There are three methods for calculating the internal resistance/impedance of a battery:
    - **DC load method**: this method ignores the polarization capacitance of the equivalent circuit of a battery (C_p in **¡Error! No se encuentra el origen de la referencia.**) and blends both resistances (R_0 and R_p). A voltmeter measures the open circuit voltage (OCV) with no load, followed by the second reading with a load; Ohm's law calculates the resistance value. The load test is the preferred method for batteries that power DC loads.
    - **AC conductance**: The single-frequency method sees the components of the Randles model as one complex impedance called the *modulus* of Z (R_0, R_p and C_p cannot be distinguished).

The 1,000-hertz (Hz) ohm test is another common method (it can be used for lithium batteries). A 1,000Hz signal excites the battery and Ohm's law calculates the resistance.

  - **Electrochemical Impedance Spectroscopy** (EIS): it reads the three values in the Randles model; however, correlating the data into capacity estimations requires complex modelling.

- **Dynamic Stress Test (DST)**: it seeks to simulate dynamic discharging. Once battery is fully loaded, it is discharged by applying DST profiles repeatedly.
- **Ageing tests**: in each cycle, cells are charged and discharged at a constant rate until the cut-off voltage is reached.

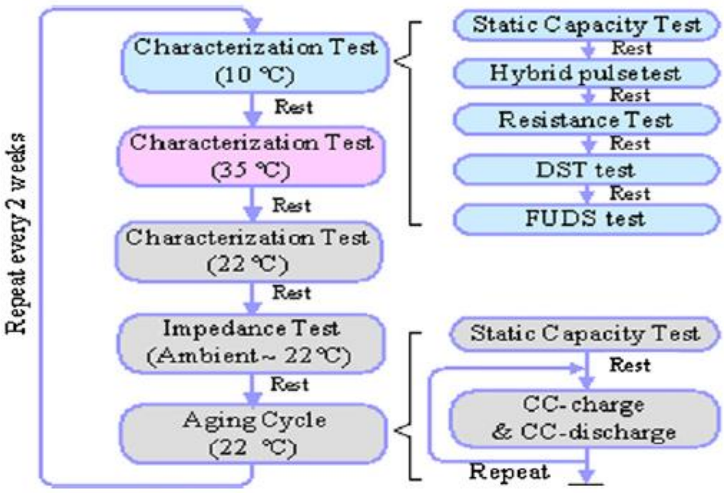


Figure 4.6: Test schedule flow chart carried out in [164]

The main specifications of the test should include rated values: nominal capacity, nominal voltage, upper cut-off voltage and lower cut-off voltage.

From the resulting voltage of the **hybrid pulse test**, the following information can be obtained for different SOC levels:

- The **open circuit voltage**, V_{oc}, value is obtained from the steady state voltage.
- The **DC resistance, R_0,** is proportional to the instantaneous voltage drop following a discharge pulse.
- The **polarization resistance and capacitance** (R_p, C_p), AC impedance, are related to the voltage amplitude change following the pulse start (see next figure).

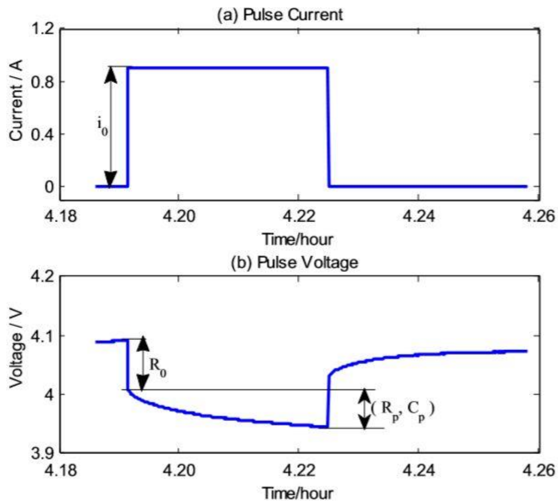


Figure 4.7: Zoom into discharge pulse of the LiNMC cell at around 90% SOC at 22°C [163]

The **Recursive Least Squares** (RLS) algorithm is used to calculate both parameters.

In [165], an equivalent electrical circuit with **more RC blocks** is evaluated. To determine the number of RC branches, the voltage evolution is examined during the relaxation phase after a current pulse is removed, during a pulse test. The adopted criterion for selecting the number of branches was that the curve needed to fit closely to the initial few data points in the transient. In this case, more than 3 exponential terms (RC blocks) where need to fit the curve obtained from the test.

4.2.1.2 Battery *cell electrochemical* model

While ideal-cell models enable prediction of non-degraded cell behavior and are useful for basic cell-level estimation and control (cell balancing, SOC estimation, power limit computations, etc.), advanced controls that aim to extend cell life (understand and control ageing), or improve, performance require Physics-Based Models (PBM) of cell degradation [167].

To understand degradation, additional state variables come into play, such as: potential in solid, potential in electrolyte, concentration of Li in solid, concentration of Li in electrolyte, and rate of lithium movement between phases [168].

PBMs can be developed at different length scales, form molecular to continuum. Continuum scale models can often run in PCs, but smaller length scales (more detailed) need supercomputers (high computational capability).

These Full Order Models (FOM) can adequately capture internal dynamics within the cell. However, the solution must be obtained numerically using finite element discretization methods and, as such is computationally burdensome, it is not suitable for near-real-time battery control [169]. To cope with this problem, Reduced-Order Models (ROM) are used [167] [168]:

- With the electrochemical richness, a model is developed from the control point of view. They capture important behaviors of FOM, but with lower computational costs.
- It predicts “ideal cell” electrochemical variables directly: potential, concentrations, rate of Li movement.
- These are inputs to follow-on degradation models, used by controls.
- It has linear discrete-time state and output equations. Linearization around an operating point, *f(SOC, T)*, is needed: one linear model for each operating point.
- Electrochemical variables and voltage are computed via nonlinear corrections to the terms computed by the output equation.
- ROM generation process involves the following steps:
    - Start with Partial Differential Equations (PDE).
    - Linearize PDEs, make transfer functions (TF).
    - Populate TFs with physical values.
    - Convert TFs to linearized state-space model.

It is very complex to model degradation. There are a lot of mechanisms that degrade a cell, some of them not well known.

In addition, the degradation mechanisms all interact:

- If we can model dominant degradation mechanism, and
- If there are controllable inputs to these mechanisms (current, SOC, temperature…).
- We can devise controls that minimize degradation while simultaneously maximizing cell life.

One of the main degradation mechanisms is lithium plating/deposition (formation of metallic lithium around the anode during charging) and that is the focus of the studies carried out in [168].

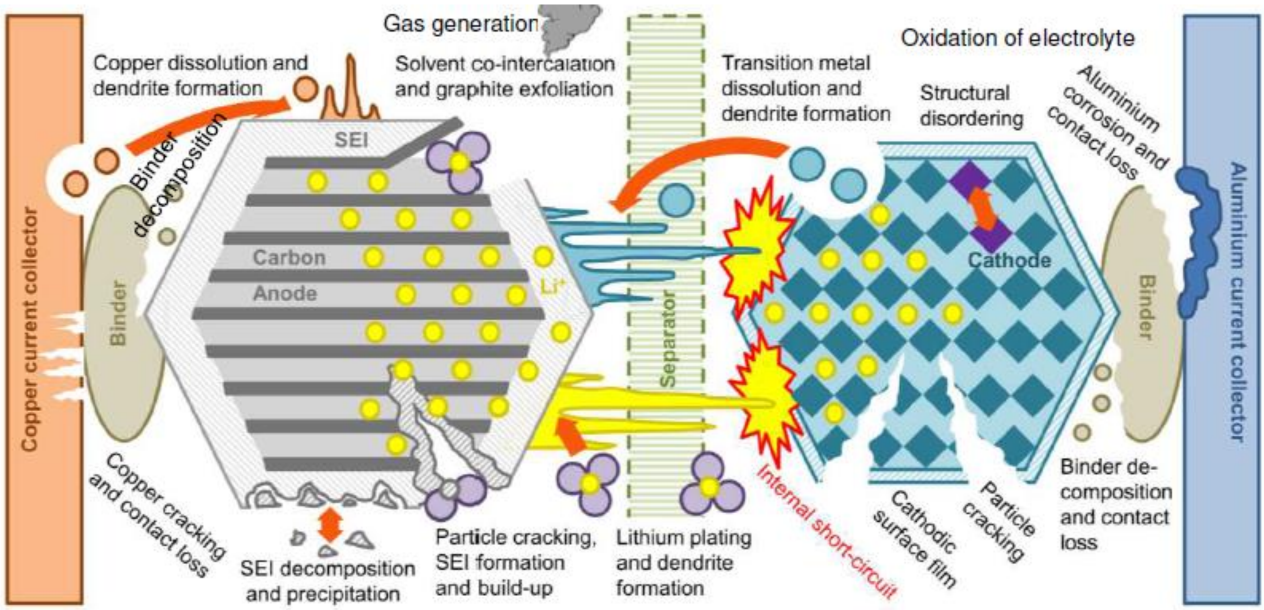


Figure 4.8: Degradation mechanisms in a battery [168]

4.2.2 SOC and SOH estimation methods

There are a lot of SOC estimation methods. According to [169], model-based methods are the best approach:

- *Observer methods*: Luenberger observer (LO), sliding mode observer (SMO), proportional integral observer (PIO).
- *Adaptive filter methods*: Kalman Filters, Recursive Least Squares (RLS).

A very common technique for **SOC** estimation is the **Coulomb counting**, where the number of Ah charged to and discharged from the battery are counted to estimate the remaining capacity or, even, life of the battery. The accuracy of this approach can be reduced if the initial SOC is not well known, and by other factors, such as self-discharge of the cells and leakage effects. The releasable charge is always less than the stored, due to the efficiency of the battery lower than 100%. These losses in addition to the self-discharging, cause cumulative errors due to the open-loop control characteristic of the measure. For more precise SOC estimation, these factors should be considered. In addition, the SOC should be recalibrated on a regular basis and the degradation of the battery capacity should also be considered [170]. In this work an **enhanced coulomb counting algorithm** is presented.

For example, in [166][171], the actual cell capacity is estimated through a **Support Vector Regression** (SVR) method given a set of training data, with couples of input and target values, a linear function is sought as flat as possible where a maximum deviation between both values is defined. Since SOC and SOH prediction will not be a linear regression problem, a method for the nonlinear case is required: the idea is to transform the input data nonlinearly in a higher dimensional feature space. In this case, the data for the SVR is obtained from a long-term cell ageing test: for one year a large test was performed on Li-ion cells: in them, regular check-ups were conducted to record the capacity fade in the cells; these check-ups consisted in capacity tests with 1C-rate and a pulse power characterization profile.

Another way to determine the SOC of a battery is to use the **open circuit voltage method**. It converts a reading of the battery voltage to the equivalent SOC value using the discharge curve (open circuit voltage vs.

SOC). The need for a stable voltage makes it difficult to implement, since a resting time for the battery is required prior to the measurement. It results also time consuming and, during the test, the system application is interrupted (offline method), contrarily to the coulomb counting (online method) [170]. Another issue is that the method is not effective for batteries with a flat OCV characteristic [163], such as $LiFeP0_4$ (Lithium Iron Phosphate).

Capacity decrease and power fading, related to the SOH, do not originate from one single cause, but from a number of various processes and their interactions: structural changes during cycling, chemical decomposition or dissolution reaction and surface film modification, loss of contact between the inactive components, metal dissolution, electrolyte decomposition [159].

A well know approach for SOH estimation is the measure of the battery AC impedance or DC resistance [159]. During its lifetime, the internal resistance of the battery tends to increase. Therefore, **DC resistance measurement** is used to evaluate battery degradation. This is carried out by applying a current pulse (at 1C) and measuring the voltage drop caused by it, in accordance with Ohm's law.

$$R_i(\Omega) = \frac{\Delta I}{\Delta \mathrm{V}} \qquad \textbf{(1)}$$

This relationship can be evaluated for different temperatures and SOC levels. In occasions different current rates (power) can be evaluated. Normally, the resistance changes significantly with temperature but not that much with the SOC according to [159], however, in our experience, this resistance changes significantly near the SOC limits (under 20% and over 80%).

Another technique to measure the internal resistance is through the **Joule effect**: the changes in temperature produced by losses are measured through a calorimeter.

The **dynamic power capability** of the battery, along its lifetime, can be determined using the Hybrid Power Pulse Test (HPPT), which consists of a test profile incorporating both charge and discharge pulses. With this method minimum and maximum cell voltages can be also determined.

The Electrochemical Impedance Spectroscopy (EIS) measures the **battery internal AC impedance** as a function of frequency. Different battery dynamics affect different frequency ranges and, therefore, this is an appropriate diagnostic tool: at high frequencies, inductive effects are prominent, while, as frequency decreases, impedance becomes purely ohmic and at lowest frequencies capacitive effects become more important [159]. In our case, frequencies below 1KHz are the most relevant. EIS equipment can also provide ECM parameters.

From measurements, a characteristic map can be drawn using **data fitting** techniques for calculating the internal resistance at every SOC and temperature. The main drawback is that one map must be parametrized for each cell type. Through this method, the influence of different parameters in the ageing of the battery can be assessed, such as DOD or temperature. In addition, apart from internal resistance, other effects in the battery can be evaluated: capacity loss, diffusion effects, etc.

There are also methods in which the capacity is estimated using a **probabilistic algorithm**. For example, by analyzing the charge and discharge data of batteries. A voltage is linked with a capacity through statistics from data of discharge curves of new and aged batteries. With this information a look up table is generated, where the capacity of the battery is related to a partial charge or discharge [159].

To estimate the SOH, other methods lie in measuring the mechanical fatigue or, in general, in **detecting the failures** of the cell [159].

One of the most used methods for SOH estimation is the **Kalman filter**, and related algorithms such as the **Extended Kalman Filter (EKF)** and **Unscented Kalman Filter (UKF)**. They permit to estimate non-measurable variables (inner state values) from a real system or variables that are affected by noise (due to measurement or process). It employs error correction mechanisms, and it can provide real time SOC and SOH estimations.

Because of this, it is a suitable method to develop digital twins, since the real system and the model can be continuously compared.

Fuzzy logic allows modelling nonlinear and complex system by processing measured data. It is a powerful method, but it requires a big amount of testing data, relatively large computations and good understanding of the batteries for an accurate SOH prediction.

The advantage of **Artificial Neural Networks** (ANN) for SOH estimation is that it is not necessary to know the details of the battery to perform the analysis. The main disadvantage is its high computational cost. To get good results out of it, it is necessary to train the ANN, which requires to have as many data available as possible.

In [163], a comparison between different SOC and SOH estimation methods is presented.

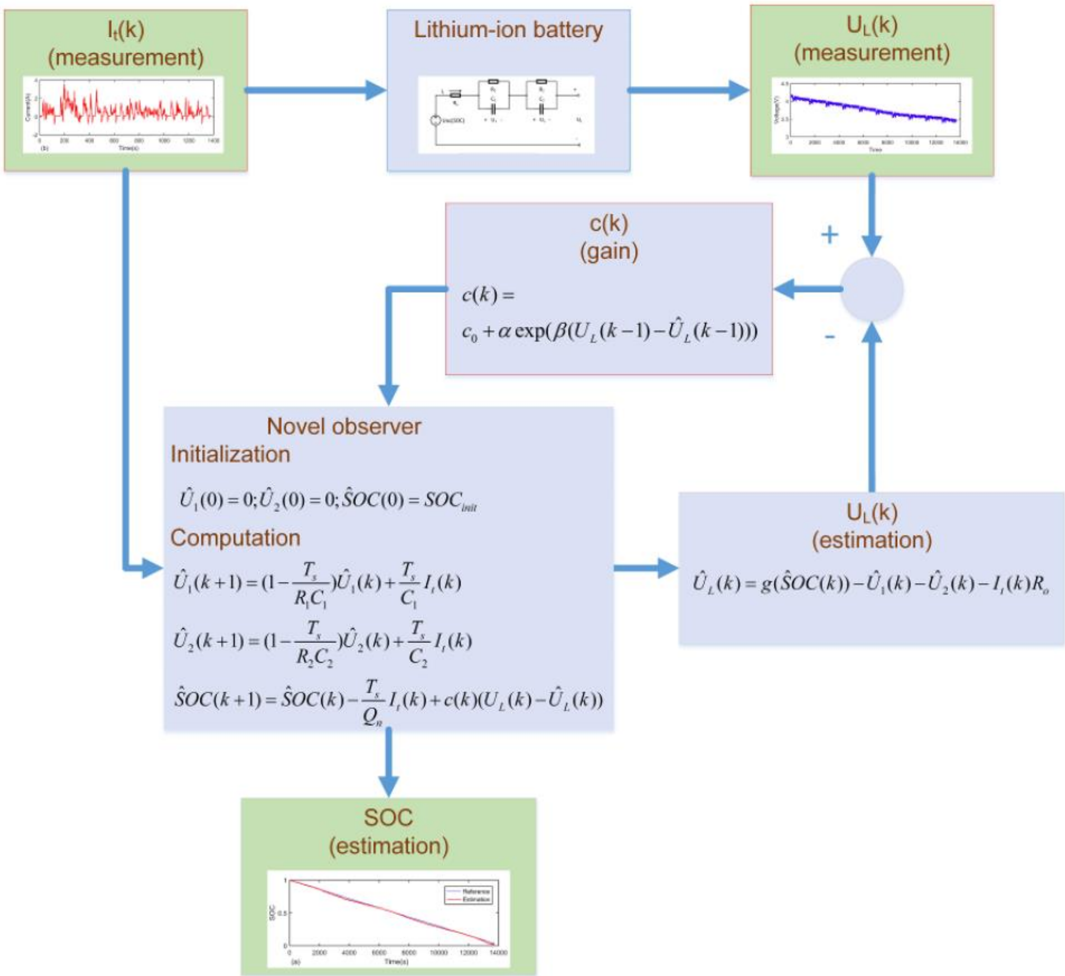
State of charge (SOC)			State of health (SOH)			
Method	Advantage	Disadvantage	Method		Advantage	Disadvantage
Coulomb counting [1–4]	Simple	Open-loop, sensitive to the current sensor precision, and uncertain to initial SOC	Durability model-based open-loop method	Durability mechanism [23,24]	Comprehensive understanding	Complex, need accurate input parameters
Open circuit voltage method [5]	Simple	Open-loop, sensitive to the voltage sensor precision, unsuitable for cells with flat OCV–SOC curves		Durability external characteristic [25–28]	Simple and easy to predict capacity fade and internal resistance increment	Based on a large number of experiments
Neural network [6]	Generic, good nonlinearity mapping approximation	Sensitive to the amount and quality of training data	Battery model-based parameter identification closed-loop method	DC resistance [21]	Simple	Not accuracy, sensitive to disturbances
Fuzzy logic [7]	Generic, good nonlinearity mapping approximation	Sensitive to the amount and quality of training data		AC impedance [22]	Accuracy	Complex
Support vector machine [8]	Generic, good nonlinearity mapping	Sensitive to the amount and quality of training data		Extend Kalman filter [11,29]	Quite easy to implement, accurate	Sensitive to modeling accuracy
Kalman filter [9–18]	Closed-loop, online, accuracy	More computationally expensive than non-feedback methods, and highly depend on the model accuracy.		Fuzzy logic [30]	Accuracy simple, accurate	Slow convergence
				Sample entropy [31–33]	Simple	Need large amount of data
Sliding mode observer [19,20]	Closed-loop, online, and accurate	More computationally expensive than non-feedback methods, and highly depend on the model accuracy.		Discharge voltage [30]	Easy	Not accurate
				Adaptive control system [31]	Online	Sensitive to modeling accuracy

Figure 4.9: Comparison table between SOH and SOC estimation algorithms [163]

According to this reference, most of the methods in the table were developed for either SOC or SOH estimation, and not for both. However, both parameters are related, and simultaneous estimation is beneficial. Compared to the SOC, battery SOH changes, typically, much slowly. Because of this, one option is to use **multi-scale Extended Kalman Filters** to estimate SOC and SOH, and the capacity estimation is periodically introduced in SOC update equation.

In [161], a **novel observer for SOC estimation** is proposed, using a two RC block battery model. According to this reference, most existing observers utilize the difference between the estimation output voltage and the measured voltage multiplied by coefficients to correct all of the estimated states. These observers have some disadvantages:

- The optimal gain coefficients are hard to determine.
- Some observers utilize a linearization method to approximate the nonlinear relationship between the OCV and the SOC, which causes linearization error.



As the battery cell ages, the estimation error increases due to the variation of model parameters. To keep the accuracy of SOC estimation under acceptable limits, it is possible to update the parameters through a SOH estimation. In [163], for offline parameters recalibration and SOH estimation, a fourth order EKF is used.

1. Results show [163] that it is undesirable to estimate the battery SOC without considering the capacity and resistance recalibration over the entire battery lifecycle. The fourth order EKF is computationally heavy, and it might suffer from stability issues, therefore, it is operated offline and the related capacity/resistance updates are instead used in the online (real-time) second order SOC estimator

In [166], a combination of Kalman filters is proposed to estimate the parameters of the battery model, consisting in two RC blocks:

1. Standard **Kalman Filter** (KF): it estimates the voltages through the two condensers in the RC branches and the R_0 resistance.
2. **Unscented Kalman Filter** (UKF): the output of the KF is fed into this filter that estimates the SOC and the resistances in the RC branches. The actual cell capacity, estimated through an SVR method (no C_{rated} capacity is used), can be used to predict the SOC.

The advantages of using this dual KF are the following [166]:

- Decoupling of estimations and, therefore, reduction of interactions and avoiding of building-ups of a filter.
- Separation of variables, which cannot be estimated by a single filter.

- Reduction of computation efforts, since two filters of lower dimensions are faster than one of higher dimension.

To estimate the SOC of a battery pack consisting of several cells, we could use one of the methods above, e.g., the Kalman filters, for each of the series elements of the pack (cells connected in parallel average their behaviour electrically, so they have a common SOC value). Even if this may work, the required level of computation might be prohibitive. Two methods are proposed that avoid this problem: the bar-delta filtering [169] and the mean cell model [172].

4.2.3 Battery cell data driven model for life prediction

There are different ways to predict the performance of a battery system, but they are all based in the availability of data.

In [160], an **ageing model** is presented for battery lifetime assessment that implies understanding calendar and cycling ageing, but also analysing their interaction for approaching real applications operation. The method is extensible to different operating scenarios and applications.

- It is based on on-board measurable parameters in real applications and, therefore, can be integrated into a Battery Management System (BMS) for SOH estimations.
- It achieves a compromise between the accuracy of the models and time-intensive experimental work.

The proposed methodology has an important testing component, and it is summarized in the following figure:

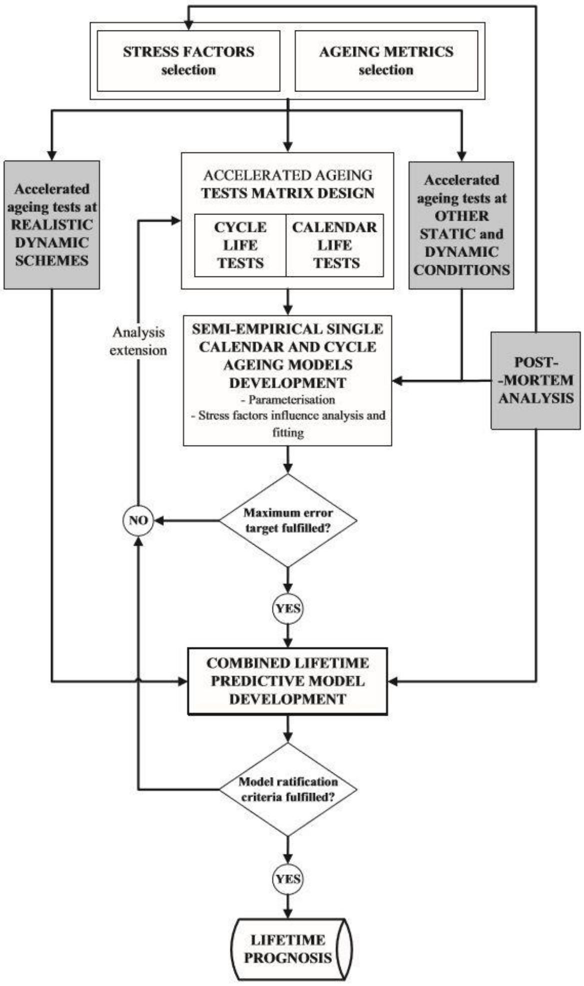


Figure 4.10: Flowchart of methodology for lifetime prognosis by [160]

The premise of the methodology is that the ageing model needs to be valid for any real condition. Ageing as a function of time (calendar) and as a function of usage (cycle) are studied individually. Calendar ageing is addressed by means of storage tests, at different temperatures and SOCs, which are defined as stress or impact factors (time domain):

$$A_{cal} = f(SOC, T, t) \quad \textbf{(2)}$$

In turn, cycle ageing is a function of C-rate and DOD, as stress factors (Ah domain):

$$A_{cyc} = f(DOD, C - rate, Ah) \quad \textbf{(3)}$$

The resulting ageing model is developed using a characterisation method based on the stress factors and by capturing the dominant effects on cell performance degradation to fit testing parameters (capacity and internal resistance increase are the metrics used to track the degradation: ageing meters). A weighting method is used to assess the different stress factor effects and it is assumed that calendar and cycle ageing effects can be superimposed.

$$A_{tot} = A_{cal} + A_{cyc} = f(SOC, T, t, DOD, C - rate, Ah) \quad \textbf{(4)}$$

Accumulated usage by former usage is considered to make future predictions at a different cycling or calendar event. The actual capacity loss is used as reference point for further predictions at different operating conditions.

This methodology has two main drawbacks: the proposed prediction method requires approximately 1-1.5 year of lifetime evaluation and results are applicable only to the reference cell considered in the work [160].

These disadvantages could be solved by the utilization of **machine learning** related techniques [173], which are more and more deployed in all engineering fields. The processing of large and complex data could predict the future behaviour of systems based on their past performance. For example, the data obtained from real battery systems could be used to predict the state of health of the system or to study where the most aggressive degradation comes from.

5 PV FAULT DIAGNOSIS

5.1 PV common techniques

The performance of the PV installations has increased over time [174] [175] [176], owing to technological advances and system optimizations. Over the last decades, numerous monitoring campaigns have allowed to report and analyze the performance of thousands of PV systems worldwide [177] [178] [179] [180] [181] [182] [183] [184] [185] [186] [187] [188] [189] [190] [191] [192] [193] [194] [195] [196] [197] [198]. These studies have shown a wide disparity in performance between the PV systems, with Performance Ratio (PR) values typically comprised between 60% and 90%. Some of the causes explaining the performance losses have been identified and quantified, but more remains to be done.

Thanks to the development of Information and Telecommunication (IT) technologies, these last years have witnessed an increase in the quantity and in the quality of the operating data that have been monitored, offering new possibilities to detect performance problems. Automatic fault detection procedures have experienced major developments in the field of engineering and their application to PV systems has undergone intense development, and it is now a trending Research and Development (R&D) topic [199] [200].

A fault is understood as a decrease in a performance indicator, generally abrupt, during a specific time period, and due to abnormal functioning. According to this definition, in certain cases, the overall performance of a PV system can be relatively low without the presence of any fault. For example, this can be the case of an old PV system whose generator nameplate power has decreased over time because of the aging of the PV modules [201]. This can also happen when the real nameplate power of the PV modules has been overrated, and it is lower than the expected nameplate power, according to datasheets [202]. These overall performance problems can be significant, but they are out of the scope of fault detection procedures strictly speaking.

Failures during the operation of photovoltaic installations are one of the factors that most affect the performance and maintenance. Unlike other phenomena such as shadowing, degradation or the effect of operating conditions, production losses due to malfunctions are not easily modellable or predictable, so they benefit significantly from advances in monitoring and control systems [203].

The process of analysis of faults produced in photovoltaic installations is divided fundamentally into three components: fault detection, fault diagnosis and quantification of the effect of these faults. These components have been studied both as individual processes and jointly, during the analysis of the performance of the installation.

Among the most frequent approaches to the detection of faults are techniques based on the modeling of PV systems that compare all the variables obtained in simulations (voltage, current, temperature) against the measurements [204]. There are also techniques based on the automatic detection of anomalies [200], in which deviations in different operation variables between different elements are compared. Both techniques can used together, especially to locate faults with a joint effect on the entire plant.

The classification of failure detection and diagnosis methods, as well as the classification of failures in different categories, has been widely studied in previous reviews [108] [205] [206] [200].

In relation to the classification of faults, a distinction is made between AC faults and DC faults [108]. On the DC side, the classification includes MPPT and PV array faults (mismatch, short circuit, open circuit, ground, bypass diode, asymmetrical, arc, bridging). On the AC side, the classification includes total black-out and grid outage.

Another work presents a very similar classification [205], but including hot spots, junction box failures and line-to-line faults. In the field of classifying fault detection and diagnosis techniques, this work presents also a first approximation divides the methods into statistical and signal processing approaches, I-V characteristics analysis, voltage and current measurement and artificial intelligence techniques.

In the classification of methods and failure analysis, another review also presents a classification of the same type [206], specifically dividing the analysis techniques for the detection of failures into electrical circuit simulation, statistical analysis, electrical signal approaches, artificial intelligence techniques, and predictive models and comparison with real models. In another review, in addition to statistical and computational methods, visual inspection and image processing techniques are presented as a tool [200].

Similar to the classification of fault types, diagnostic techniques are also classified with differentiation of DC and AC methods [207]. On the DC side, the classification of techniques includes electrical characterization, infrared thermal imaging, visual inspection, ultrasonic inspection, electroluminescence imaging and lock in thermography. On the AC side, the classification of techniques includes converter component fault and grid outage analysis.

Therefore, despite the absence of a specific consensus, fault detection and diagnosis techniques are generally divided into inspection-based methods, techniques based on statistical methods, artificial intelligence techniques, signal analysis, and comparison with simulations of a PV system model, although in general this comparison is applied in conjunction with another of the previous methods.

The PR is the most commonly used performance indicator [180] [208] [184] [175] [176] [207] [209] [210] [211] [212] [213]. It is based on the ratio between the energy output of a PV system and the solar irradiation that it receives. Several alternative fault detection procedures have been developed. The PR can be normalized in order to become more independent of the operating conditions [214].

Another method relies on comparisons between measured and simulated energy yields [215] [216] [217] [218]. Some procedures rely on the modeling of I-V curves [219] [220] [171] [221] [222], and others make use of abnormal variations in power, voltage or current as a fault indicator [223] [224] [225] [221] [226] [204] [227] [228] [229] [230]. Other methods make use of scatter plots, also called “stamp collections”, representing the outputs of the PV system (energy, AC or DC power, voltage, current) as a function of key input variables (solar irradiance, air temperature, cell temperature, solar elevation). These methods have shown promising capacities [231] [232] [233] [234] [235] [236] [237] [238] [239]. Another family of approaches relies on Artificial Neural Networks (ANN), fuzzy logic, or machine learning [240] [241] [242] [243] [244].

In terms of signal analysis for fault detection, there are several jobs related to automatic fault detection. The analysis is carried out from the voltage and currents measured, so that by analyzing the anomalies in both DC and AC, it is possible to accurately detect faults in a photovoltaic module, faults in a photovoltaic string, faults in an inverter, and a general fault that may include partial shading, PV aging, or MPPT error [227].

Another work focuses on this last point, presenting a procedure to automatically detect failures in systems with distributed maximum power point tracking at module level [245], monitoring the PV plant parameters at the module level (voltage and current at the operating power point) and being capable of detecting and diagnosing faults without the need for irradiance measurement.

A method widely used in the analysis of signals in photovoltaic systems is the Local Outlier Factor (LOF), which is utilized to identify fault conditions by comparing the parameters between fault and normal conditions, allowing to detect strings with lower energy performance than those of the environment.

The LOF technique has been widely studied throughout different datasets, reaching the conclusion that its precision requires correctly adjusting the thresholds and that it provides good results in large datasets, being necessary to adapt the algorithms in smaller samples [246].

The analysis of the local outliers bases their operation on the current signal characterized as non-stationary stochastic signal, therefore the analysis in time series improves the detection of failures using this system [247]. This is achieved by using a time series sliding window, calculating the LOF at each point and detecting those strings that deviate outside the threshold in much of the time window.

Statistical analysis methods have been studied through multiple implementations in all the variety of components of photovoltaic installations. Some works have analyzed the photovoltaic plant assembly using inferential statistics [248] [249], firstly analyzing the normal performance of the system to define thresholds and metrics, and secondly applying real-time analysis of variance (ANOVA) using previously calculated parameters to detect anomalies. This threshold-based real-time detection system has also been applied at the array level through outlier filters [250] [251].

Fault diagnosis techniques based on statistical approaches have also been specifically applied on the DC side, showing very precise results by means of exponentially weighted moving average of the differences between the theoretical power of the arrays and the real power in each measurement [252].

Classical statistical techniques have progressively evolved to more advanced models that have artificial intelligence, from machine learning algorithms to complex multilayer neural networks in deep learning models. This class of models have been used both in the detection of failures and in their diagnosis, directly by learning differentiating patterns in the failed devices, as well as in the simulation of photovoltaic systems for later comparison with the real system, allowing to detect failures on those more deviant devices.

Although dozens of theoretical models based on neural networks have been studied and applied [253], the most clarifying results have been obtained through multi-layer neural networks, artificial neural networks-based models, and neuro-fuzzy models, which combine artificial neural networks with fuzzy logic.

A recent review also shows convolutional neural networks as a promising tool since they are the most common type of deep neural networks, although they have not yet been widely applied in photovoltaic models with good results [254].

In relation to multi-layer neural networks, a previous work presented a neural network model trained in a hybrid way with both operation and simulation data for the detection of failures, so that in real-time operation it was also fed with these simulation and measurement data [255]. The results were far superior to the fault detection algorithms presented in previous reviews, with the additional advantage that the neural network took the voltage and current of the modules as input data, without the need for irradiance data or other operating conditions.

Multi-layer neural networks are also presented as a method with excellent results in the detection of short-circuit failures, especially in large PV plants [256], although for real-time operation prior training of neural networks with a large dataset is necessary [257].

In addition to fault detection, multi-layer neural networks have also been applied to the diagnosis of faults in photovoltaic installations [258], allowing the categorization of faults of a PV module in module mismatch fault, open circuit, short circuit or multiple faults under partial shading.

Artificial network-based modeling has been presented as a very powerful tool not only for fault detection, but for the study of degradation and shading of photovoltaic modules [191]. This method is based on modeling the photovoltaic system through multi-layer neural networks so that when comparing with real measurements, anomalies in operation can be easily distinguished and, in addition, the effect of shading and degradation is eliminated before diagnosing faults.

The neuro-fuzzy approach has been studied for the comparison of modules and the determination of their status as failed or not. This approach is based on the generation of fuzzy models under different operating conditions for their subsequent hybridization, allowing us to observe the evolution and detect anomalies that can be classified as failures [259]. The thresholds and parameters that need to be defined for these models are highly variable, but previous work has shown good results in the detection of failed photovoltaic modules using only the maximum power of the module and the open circuit voltage [260].

The results obtained through artificial neural networks have also been compared not only with classical models, but with complex machine learning methods. A previous work compared the methods of K-Nearest Nearest, Decision Trees, and Support Vector Machines against neural networks, the latter being much superior in terms of results [261]. In a similar study, neural networks were compared to decision trees, XGBoost models (based on multiple hybridized decision tree models) and random forest, again obtaining the most accurate results using neural networks [262].

In relation to their monitoring, PV systems can be classified into two main groups. The first group is constituted by large solar power plants, whose nameplate power stands between several hundreds of kW and several hundreds of MW. The second group is made of smaller installations, whose nameplate power typically lies between 1 kW and the hundreds of kW, generally mounted on the roof of a building, called distributed PV systems. The Operation and Maintenance (O&M) of PV plants benefits from a specific budget, and it is commonly assisted by Supervisory Control And Data Acquisition (SCADA) that registers sizeable quantities of data that are sometimes measured up to the PV module string level [263] [264], and a weather station that measures the operating conditions. Other tools and hardware specific to the O&M are also available, and they are used by trained professionals [265].

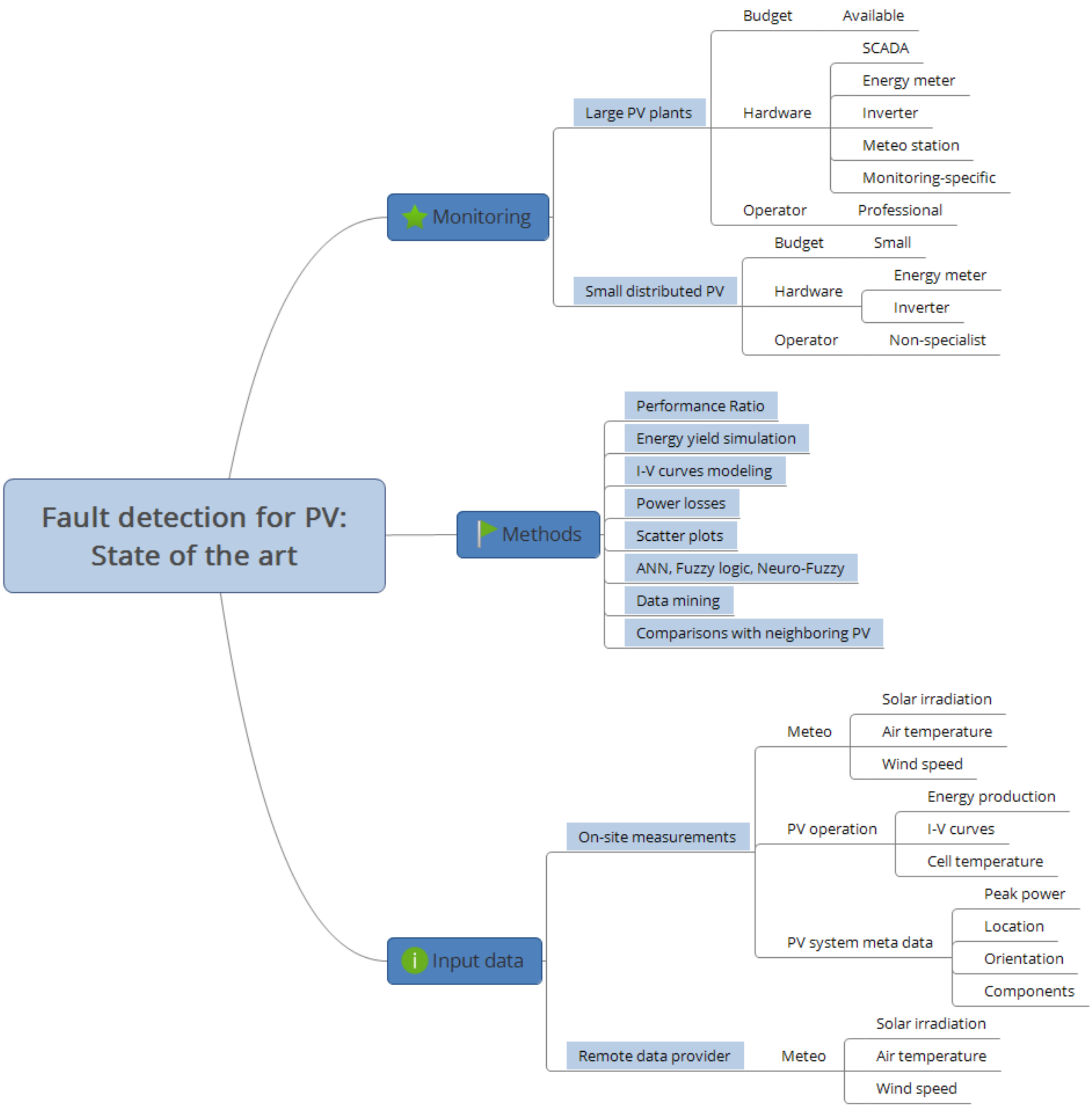


Figure 5.1: Classification of Failure detection methods according to available monitoring system, proposed methods, and required input data.

5.2 Large PV plants

Large plants require a higher level of precision than other installations. They are generally associated with large investments in assets with highly controlled profit margins. Any problem in the analysis or the supervision could imply considerable economic losses. This means it is absolutely necessary to have a good fault detection and diagnosis system that allows us to know them as soon as possible, measure their effects and know their origin.

The origin of failures in a plant are of a different nature, in most cases they are similar to other PV installations but with different implications. In [266] indicates three possible sources of failures in a PV plant. On the one hand, the reliability of the surveillance equipment is often lower than that of the equipment to be surveyed, this implies problems of the monitoring system. On the other hand, due to the multiple sensors and the size of the plant, we may have discrepancies between what is considered to be the operating conditions and the production, especially in areas far from the measurement point of these conditions. Finally, a PV plant is usually considered as one more generator of the electrical grid and therefore can be subject to production restrictions that modify its operation.

The large amount of information and the possible disparities in the operation of the different elements of the plant make the use of advanced statistical techniques or artificial intelligence to detect and diagnose anomalies absolutely necessary. It is usual to use the time evolution and the relationship with other components to detect faults, especially in the case of strings or stringboxes that are very numerous elements and do not have their own alarm systems. For example, the article [267] uses a system that uses space-time distributions of strings as inputs of a fuzzy system for the detection of anomalies in these devices. In the case of elements such as string or string boxes, noise and dispersion between elements has a great effect. In the article [268], where a fuzzy logic system is used to eliminate noise from different signals. In the latter case, the system used is very effective in reducing the number of false positives and in detecting different types of anomalies with very high precision.

However, most of the detection effort is focused on the generating part (the modules), leaving aside the equipment that is responsible for transforming the energy, the inverters. As can be seen in the article [269], the part of characterization of anomalies in photovoltaic modules is much more developed than in inverters, so the approach of this project is very interesting in the face of the current panorama.

5.3 Small and medium PV plants

Distributed PV installations consist of a large number of small or medium PV generation units spread out over the territory. In Europe, these distributed PV installations represent about 40% of the total installed PV capacity and more than 90% of the total number of PV installations. They differ from one another in their size, components, orientation and tilt angles, topology, and quality. They operate under weather conditions that vary from one installation to another, both in space and in time. Each one of them usually belongs to a different owner. Most of the time its operator is the owner, who has no expertise in IT or in photovoltaics [270]. The metadata on the technical characteristics of the PV system, such as its peak power, orientation, PV module technology, etc. are often reported with limited accuracy, or even unknown [271]. The budget allocated to the monitoring of the PV system is small, which makes cost-prohibitive the installation of monitoring hardware, although low-cost solutions exist [272]. Therefore, the energy output is often the only accurate information available. There is no on-site measurement of the operating conditions [273], which must be acquired through a remote data provider.

Unfortunately, most of the existing fault detection procedures rely on accurate input data on the PV system's characteristics and its operating conditions, and they have been developed and validated in large PV plants, or PV systems associated with research laboratories, whose technical characteristics and operating conditions were known with a high degree of accuracy. Therefore, they suffer from several limitations when the time comes to apply them to large groups of hundreds or thousands of PV installations, whose main technical characteristics and operating conditions are not known with accuracy. First, the inaccuracies in the metadata of the PV system, in particular, its peak power and its orientation, make it difficult to compare the energy production with the in-plane solar irradiation. Second, it is difficult to obtain reliable and low-cost solar irradiation data. The data measured by pyranometers are not available at a spatial resolution that is high enough, and the data retrieved from satellites convey inaccuracies that are due to their limited spatial and temporal resolution, cloudiness, snow cover, aerosols, and low solar elevation angles. Third, thermal losses affecting the PV systems depend on the weather conditions, notably including air temperature and wind speed. In fact, due to the uncertainties on solar irradiation at the hourly level, the PR can often go up to more than 100%, and it can show large variations. Under these conditions, it is difficult to build a reliable performance indicator that is stable enough to make it possible for accurate fault detections. As a result, many problems affecting distributed PV systems go undetected.

Several independent research teams have undertaken the task of analyzing the monitoring data from several hundred or thousands of PV systems. Their findings show that they faced common problems and reached similar conclusions on the inherent limitations from the lack of reliable input data and the need to cope with that reality [271][115] [274] [118] [275]. Several works have developed procedures that make it possible to correct for the inaccurate metadata, along with Quality Controls (QC) procedures and data inferences, making it possible to obtain more robust fault detections [276] [274] [118]. In parallel, other works have applied statistical techniques coming from other fields, such as data mining [277]or kriging [207], to construct a fault detector from the comparison of the performance of a group of PV systems, generally designed as PV system fleet [278] [279] [280] [281] [282] [283]. These studies have shown promising results, in particular in the presence of clear-sky conditions, and when the data come from several PV systems located in the same neighborhood and whose generators have similar orientations [276].

Appropriate information on solar resources is very important for a variety of technological areas, such as: agriculture, meteorology, forestry engineering, water resources and in particular in the designing and sizing of solar energy systems. Traditionally, solar radiation is observed by means of networks of meteorological stations. However, costs for installation and maintenance of such networks are very high and national networks comprise only a few stations. Consequently, the availability of solar radiation measurements has proven to be spatially and temporally inadequate for many applications.

Over the last decades, satellite-based retrieval of solar radiation at ground level has proven to be valuable for delivering a global coverage of the global solar irradiance distribution at the Earth's surface [215] . The recent deployment of PV systems offers an opportunity to extend the in-situ characterization of the incoming solar radiation at the Earth's surface through the reverse engineering conversion of the energy production of the PV systems to global solar irradiation data.

During the last years, thanks to the newest developments in telecommunication technologies, an increasing amount of electrical (power, frequency, current, voltage, etc.) data from the operation of PV systems has become available. These operational data are now typically available at time intervals of 10 min, opening the door to a totally new set of data analysis procedures that were difficult to imagine just a few years ago. Smart energy meters are projected to become widespread in a couple of years, which will make it possible to monitor the energy production and other electrical outputs of millions of PV systems in the world. Therefore, in the near future, it will be possible to know the energy production of PV systems located at millions of points spread over the whole globe. This conveys the potential that solar irradiation data can also be retrieved at all these points.

The information reported by the data providers is not always very accurate. Several works have reported these inaccuracies in the information on some key parameters such as the azimuth angle, the tilt angle, or the peak power. Inaccuracies and inconsistencies have also been observed in the energy production data and in the solar irradiation data. This leads to some important uncertainties in the results, although it is possible to minimize their impact, by applying filtering procedures, and by using analyses methods that are robust against erroneous data and outliers.

Mainly in the case of BIPV, the analysis of the effects of shadows is also of special relevance. At the level of analysis and diagnosis it is a widely studied problem, also through advanced statistics and artificial intelligence. For example, there are articles such as [264], which propose a model for the prevention of shadows in the modules by means of a statistical model; or as in [284], where a vector support machine is used in order to detect losses due to this type of anomaly.

New technologies have recently appeared in the market, and it would be interesting to analyze them in order to update the studies that have already been published. The uncertainties around many of the existing analyses are important, and further development may enable us to reduce them in order to obtain more accurate and more in-depth conclusions. We still have many gaps to cover to achieve a comprehensive understanding of the state-of-the-art of PV systems in Europe. By definition, a review of the state-of-the-art of PV systems requires regular iterations in order to continue to incorporate the latest advances.

Several performance indicators have been developed to carry out fault detections on PV systems. These are for example the Capacity Utilization Factor (CUF), the Performance Ratio (PR), the Performance Ratio under STC conditions (PR_STC), The Performance Index (PI) [285] or the Performance to Peers (P2P) [286].

The most novel performance indication is the P2P. It has been applied in recent studies to use correlations between neighboring PV systems for the fault detection applied to large PV system fleets. The first results show promising potential, but much remains to be done. Several pathways for future investigation work seem promising. This P2P method could be combined with some metadata correction procedures [115] [287] [239] for the translation of the PV output from different orientations to generate virtual peers and improve the whole process [288]. These procedures also revealed to be essential prior to the application of the generation of solar irradation data from PV energy output data. The P2P and other performance indicators and methods are not mutually exclusive but rather complementary. Analogously, there is a trend towards merging ground and satellite data to achieve better overall results [289] [290]. The integration of this novel P2P approach into the already existent fault detection procedures could provide added value to them.

Finally, much remains to be done in terms of proper fault diagnosis. In this regard, it would be convenient to combine the PR, PI, and P2P with scatter plots, sometimes also called stamp collections, to apply these performance indicators to the pattern recognition process that is particular to each type of fault.

6 O&M integration

O&M Contractor must provide the services in accordance with all laws, good industry practice, planning consents, manufacturer's warranties and operating manuals. In this way, the data monitoring is a core within the O&M good practice on daily basis. The PV monitoring system is fundamental to ensure a reliable and stable operation of the plant, as keeping track of electric variables (i.e., voltage, current, energy) and meteorological is essential to avoid major issues on the plant production and performance. The objective sought with monitoring is to provide enough information to accomplish an energy balance, considering the amount of solar resource available and losses that take place due to different circumstances [262]. Thus, data analysis could add a significant value to the understanding of the PV installation performance. There is a variety of variables which should be monitored in order to keep the operating performance analysis correctly and detect any failure on the performance. Although the variables monitored on PV plants varies as per the contractor and requirements in each case.

6.1 General O&M procedure for monitoring and controlling PV plants

The monitoring devices send real time data to a cloud or locally deployed database, which performs as gateway. Thus, the SCADA system of the plant collects all these data, showing it in the control board of the plant. Furthermore, the SCADA keeps a continuous communication with the plant operators. In general, there are two variants when it comes to monitoring and controlling PV plants and communications with plant operators: either the control system is centralized in a single control center, or each plant has its own communication system.

In any case, unavailability is considered to exist either when it is reported by the SCADA system or when an anomaly is detected in the on-site inspections made by the O&M personnel. In both cases, the plant manager shall be notified (via SMS or e-mail) so that he can solve the issue detected. In case the incident occurs outside the shift of the plant manager, an attempt will be made to remotely restart the plant. Additionally, if the resolution of the incident requires the replacement of parts/equipment, a purchase request must be launch and the warehouse stock must be updated.

6.1.1 Meteorological data

Meteorological stations are normally deployed throughout the plants, recording a range of variables of valuable interest. The data transmission is made wirelessly and read by the SCADA of the plant. These data are used by the contractor to assess important KPIs, such as the performance ratio, or analyzing operating conditions of the plant (i.e., during windy periods the trackers may not work as usual). Moreover, other data such as the wind speed is also used to fix the trackers in stow position for a certain wind threshold.

Among the records, the following variables could be normally found:

- Ambient temperature
- Atmospheric pressure
- Humidity
- Rainfall
- Wind speed and direction
- Solar radiation

For solar radiation, typically measured variables are global horizontal irradiation and global tilted (plane of array) irradiation. Sometimes also direct normal and diffuse horizontal irradiation components are measured.

The equipment is usually connected to a main datalogger which is the responsible of storing the data and communicating the devices with the SCADA of the plant. This data is then used by the plant personal to evaluate the performance considering the climatological conditions and assess any potential loss or fault related to the tracking. Additionally, different production trend during the different annual seasons could be obtained through implementing data analytics and using historical data of the plant production and climatological records.

Therefore, the maintenance of the meteorological stations is fundamental for keeping track of the good performance of the plant and estimating eventual trends. The main activities identified are:

- Pyranometer: Among others, the standard maintenance activities carried out for the pyranometer are: i) dome glass must be cleaned, avoiding any external event that may alter the readings, at least several times a week with a soft wipe to prevent it from any scratch; ii) difference physical components have to be checked on a weekly basis, such as the levelling, external appearance or fixing of the equipment; iii) in the long-term basis, the pyranometer has to be calibrated every year. Solar measurements are very prone to errors, therefore, very important part is also quality control of measured data, to avoid use of incorrect data records in further analysis.
- Anemometer: The main maintenance procedures are developed on a monthly basis and consist of i) visual and auditory inspection at low wind speed to verify it spins correctly; ii) replace the bearings when they become noisy; iii) replace the potentiometer after 50 million revolutions.
- Cells: They do not need of any specific maintenance, rather than keep the glass clean and an annual calibration.
- Datalogger: Annual inspection to discard any damage (i.e., corrosion, damaged wires), adopting the corresponding actions required. Calibration required every three years.

6.2 Drone IR inspection

This is a novel technique, which may not be implemented in most of the PV plants operating worldwide. The main aim is to collect thermal images of the plant solar PV panels and analyze them computationally in order to detect potential overheating of defective cells in PV modules, string boxes and transformers. For a correct deployment of the IR inspection, the monitoring campaign will be done when the entire plant is operating. A set of pre-requisites should be defined before the campaign is carry out: i.e., the irradiance has to be greater than a pre-set value. Additional sensors must be used during this inspection, such as temperature, wind or radiation sensors.

The monitoring campaign consists of an inspection flight with a pre-defined route covering the entire plant, from which an image dataset will be obtained. The size of the set must contain images from all the modules, string boxes and transformer. Before starting the inspection, it needs to be ensured that the radiation at the time of the flight is at least equal to the minimum value set in the protocol [291]. Moreover, the images resolution must be at least the one defined in the protocol. After analyzing the images, with the results extrapolated to a radiation for a nominal value (i.e., 1000 W/m2), potential hotspots in modules should be identified, as well as any other anomaly related to a significant temperature difference between modules. From the analysis of the images, different criteria are applied used to determine if there is potential failure or risk of failure in the different elements. The list below can give a clear idea of how these criteria could work:

- High-relevance – Nonacceptable: The temperature obtained cells or module hotspots is greater than the mean temperature obtained for the entire dataset by a pre-defined value (i.e., >20º).
- Low relevance: The difference between the temperature of the cells or hotspots in modules and the mean temperature of the entire set is lower than a pre-defined value.

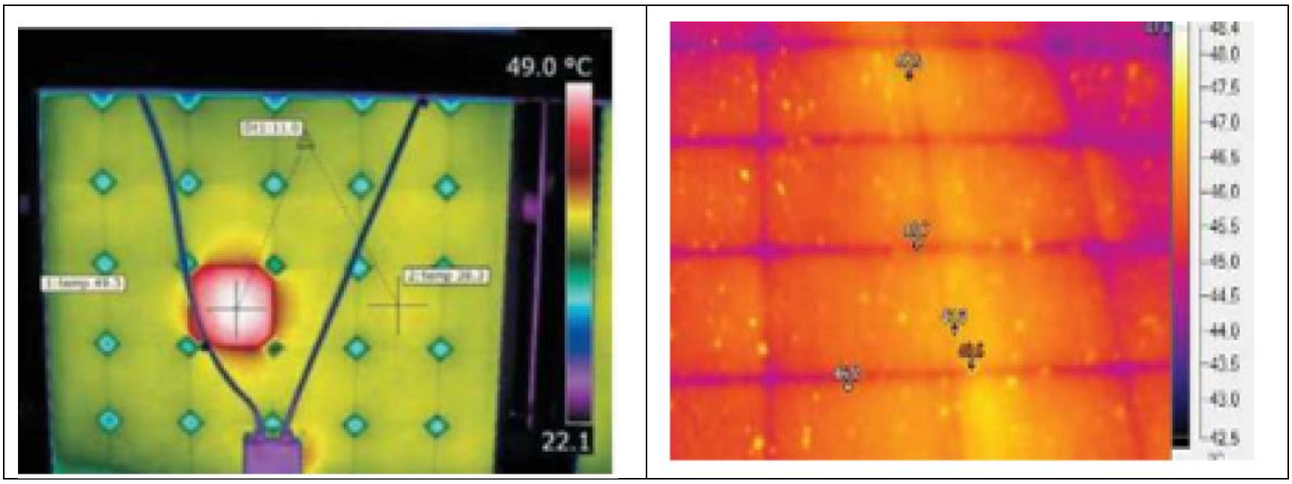


Figure 6.1: Images obtained from an IR inspection for cells anomalies and hotspots identification respectively

When a high-relevance alert is obtained for a certain module, the contractor must either replace it or repair it as per the O&M contract conditions. If a high-relevance alert is obtained in string boxes or transformers, a maintenance activity will be carried out.

The main anomalies detected during these inspections are below:

- Temperature difference between cells within the same module.
- Temporal shades, which may affect to the cells' temperature within the same module.
- Soiling, which may affect to the cells' temperature within the same module.
- Electric connection overheating.
- Defective connections.

7 REFERENCES

[1] IRENA, “Global energy transformation: A roadmap to 2050 (2019 edition),” 2019.

[2] S. R. Madeti and S. N. Singh, “Monitoring system for photovoltaic plants: A review,” *Renewable and Sustainable Energy Reviews*, vol. 67. 2017, doi: 10.1016/j.rser.2016.09.088.

[3] IEC, “IEC 61724-1: PHOTOVOLTAIC SYSTEM PERFORMANCE – Part 1: Monitoring.” IEC technical committee 82: Solar photovoltaic energy systems, 2016.

[4] H. Rezk, I. Tyukhov, M. Al-Dhaifallah, and A. Tikhonov, “Performance of data acquisition system for monitoring PV system parameters,” *Measurement: Journal of the International Measurement Confederation*, vol. 104, pp. 204–211, 2017, doi: 10.1016/j.measurement.2017.02.050.

[5] E. Koubli, “Impact of data quality on photovoltaic (PV) performance assessment,” 2017.

[6] S. Lindig, A. Louwen, D. Moser, and M. Topic, “Outdoor PV system monitoring—Input data quality, data imputation and filtering approaches,” *Energies*, vol. 13, no. 19, 2020, doi: 10.3390/en13195099.

[7] C. C. Turrado, M. del C. M. López, F. S. Lasheras, B. A. R. Gómez, J. L. C. Rollé, and F. J. de C. Juez, “Missing data imputation of solar radiation data under different atmospheric conditions,” *Sensors (Switzerland)*, vol. 14, no. 11, 2014, doi: 10.3390/s141120382.

[8] IEC, “IEC 61724-2: PHOTOVOLTAIC SYSTEM PERFORMANCE – Part 2: Capacity evaluation method.” IEC technical committee 82: Solar photovoltaic energy systems, 2016.

[9] B. Paudyal, M. Bolen, and D. Fregosi, “PV Plant Performance Loss Rate Assessment: Significance of Data Filtering and Aggregation,” 2019, doi: 10.1109/PVSC40753.2019.8981247.

[10] J. Ascencio-Vásquez, J. Bevc, K. Reba, K. Brecl, M. Jankovec, and M. Topič, “Advanced PV performance modelling based on different levels of irradiance data accuracy,” *Energies*, vol. 13, no. 9, 2020, doi: 10.3390/en13092166.

[11] A. Livera, M. Theristis, E. Koubli, G. Makrides, J. Stein, and G. Georghiou, “Guidelines for ensuring data quality for photovoltaic system performance assessment and monitoring,” 2020.

[12] E. Koubli, D. Palmer, P. Rowley, T. R. Betts, and R. Gottschalg, “Assessment of PV System Performance with Incomplete Monitoring Data,” in *31st European Photovoltaic Solar Energy Conference and Exhibition*, 2015, pp. 1594–1597.

[13] A. Livera *et al.*, “Data processing and quality verification for improved photovoltaic performance and reliability analytics,” *Progress in Photovoltaics: Research and Applications*, 2020, doi: 10.1002/pip.3349.

[14] A. Livera, A. Phinikarides, G. Makrides, and G. E. Georghiou, “Impact of Missing Data on the Estimation of Photovoltaic System Degradation Rate,” 2018, doi: 10.1109/pvsc.2017.8366442.

[15] B. O. Kang and K. S. Tam, “A new characterization and classification method for daily sky conditions based on ground-based solar irradiance measurement data,” *Solar Energy*, vol. 94, 2013, doi: 10.1016/j.solener.2013.04.007.

[16] M. Castillo-Cagigal *et al.*, “Heterogeneous collaborative sensor network for electrical management of an automated house with PV energy,” *Sensors*, vol. 11, no. 12, 2011, doi: 10.3390/s111211544.

[17] C. A. Gueymard, V. Lara-Fanego, M. Sengupta, and Y. Xie, “Surface albedo and reflectance: Review of definitions, angular and spectral effects, and intercomparison of major data sources in support of advanced solar irradiance modeling over the Americas,” *Solar Energy*, vol. 182. 2019, doi: 10.1016/j.solener.2019.02.040.

[18] B. O. Kang and K. S. Tam, “New and improved methods to estimate day-ahead quantity and quality of solar irradiance,” *Applied Energy*, vol. 137, 2015, doi: 10.1016/j.apenergy.2014.10.021.

[19] D. C. Mckay, "Estimating Solar Irradiance On Inclined Surfaces: A Review And Assessment Of Methodologies," *International Journal of Solar Energy*, vol. 3, no. 4–5, 1985, doi: 10.1080/01425918508914395.

[20] P. Ineichen, "Global irradiance on tilted and oriented planes: model validations," *University of Geneva*, no. February, 2011.

[21] H. Y. Cheng, C. C. Yu, K. C. Hsu, C. C. Chan, M. H. Tseng, and C. L. Lin, "Estimating solar irradiance on tilted surface with arbitrary orientations and tilt angles," *Energies*, vol. 12, no. 8, 2019, doi: 10.3390/en12081427.

[22] M. Lave, W. Hayes, A. Pohl, and C. W. Hansen, "Evaluation of global horizontal irradiance to plane-of-array irradiance models at locations across the United States," *IEEE Journal of Photovoltaics*, vol. 5, no. 2, 2015, doi: 10.1109/JPHOTOV.2015.2392938.

[23] R. Moretón, E. Lorenzo, A. Pinto, J. Muñoz, and L. Narvarte, "From broadband horizontal to effective in-plane irradiation: A review of modelling and derived uncertainty for PV yield prediction," *Renewable and Sustainable Energy Reviews*, vol. 78. 2017, doi: 10.1016/j.rser.2017.05.020.

[24] S. Halilovic, J. M. Bright, W. Herzberg, and S. Killinger, "An analytical approach for estimating the global horizontal from the global tilted irradiance," *Solar Energy*, vol. 188, 2019, doi: 10.1016/j.solener.2019.06.027.

[25] M. Gostein, B. Stueve, K. Passow, and A. Panchula, "Evaluating a model to estimate GHI, DNI, & DHI from POA irradiance," 2017, doi: 10.1109/PVSC.2017.8366160.

[26] A. Woyte, R. Belmans, and J. Nijs, "Fluctuations in instantaneous clearness index: Analysis and statistics," *Solar Energy*, vol. 81, no. 2, pp. 195–206, Feb. 2007, doi: 10.1016/j.solener.2006.03.001.

[27] C. S. Lai, X. Li, L. L. Lai, and M. D. McCulloch, "Daily clearness index profiles and weather conditions studies for photovoltaic systems," in *Energy Procedia*, 2017, vol. 142, pp. 77–82, doi: 10.1016/j.egypro.2017.12.013.

[28] S. Ravichandran and J. D. Rathnaraj, "Analysis of ratio of global to extra-terrestrial radiation (clearness index) at some tropical locations in India," *Thermal Science*, vol. 21, no. 3, pp. 1379–1387, 2017, doi: 10.2298/TSCI140530151R.

[29] O. O. Soneye, "Evaluation of clearness index and cloudiness index using measured global solar radiation data: A case study for a tropical climatic region of Nigeria," *Atmosfera*, vol. 34, no. 1, pp. 25–39, 2021, doi: 10.20937/ATM.52796.

[30] T. R. Ayodele and A. S. O. Ogunjuyigbe, "Prediction of monthly average global solar radiation based on statistical distribution of clearness index," *Energy*, vol. 90, pp. 1733–1742, 2015, doi: 10.1016/j.energy.2015.06.137.

[31] S. Kheradmand, O. Nematollahi, and A. R. Ayoobi, "Clearness index predicting using an integrated artificial neural network (ANN) approach," *Renewable and Sustainable Energy Reviews*, vol. 58. Elsevier Ltd, pp. 1357–1365, May 01, 2016, doi: 10.1016/j.rser.2015.12.240.

[32] F. Vignola, "GHI Correlations With Dhi And Dni And The Effects Of Cloudiness On One-Minute Data Frank Vignola Department of Physics 1274 – University of Oregon," *World Energy Forum, American Solar Energy Society, Colorado Renewable Energy Society*, no. June 2012, 2012.

[33] D. H. W. Li and S. Lou, "Review of solar irradiance and daylight illuminance modeling and sky classification," *Renewable Energy*, vol. 126. 2018, doi: 10.1016/j.renene.2018.03.063.

[34] S. Pulipaka and R. Kumar, "Analysis of irradiance losses on a soiled photovoltaic panel using contours," *Energy Conversion and Management*, vol. 115, pp. 327–336, May 2016, doi: 10.1016/j.enconman.2016.02.068.

[35] S. Pulipaka and R. Kumar, “Analysis of soil distortion factor for photovoltaic modules using particle size composition,” *Solar Energy*, vol. 161, pp. 90–99, Feb. 2018, doi: 10.1016/j.solener.2017.11.041.

[36] A. Krenzinger and E. Lorenzo, “Estimation Of Radiation Incident On Bifacial Albedo-Collecting Panels,” *International Journal of Solar Energy*, vol. 4, no. 5, 1986, doi: 10.1080/01425918608909865.

[37] P. Ineichen, O. Guisan, and R. Perez, “Ground-reflected radiation and albedo,” *Solar Energy*, vol. 44, no. 4, 1990, doi: 10.1016/0038-092X(90)90149-7.

[38] E. Skoplaki and J. A. Palyvos, “Operating temperature of photovoltaic modules: A survey of pertinent correlations,” *Renewable Energy*, vol. 34, no. 1, 2009, doi: 10.1016/j.renene.2008.04.009.

[39] P. Mora Segado, J. Carretero, and M. Sidrach-de-Cardona, “Models to predict the operating temperature of different photovoltaic modules in outdoor conditions,” *Progress in Photovoltaics: Research and Applications*, vol. 23, no. 10, pp. 1267–1282, Oct. 2015, doi: 10.1002/pip.2549.

[40] P. Mora, M. Piliougine, J. Carretero, G. Nofuentes, and M. Sidrach-De-Cardona, “Empirical model to predict the operating temperature of the modules of a photovoltaic system,” *International Journal of Smart Grid and Clean Energy*, vol. 6, no. 1, pp. 40–46, Jan. 2017, doi: 10.12720/sgce.6.1.40-46.

[41] A. S. Joshi, I. Dincer, and B. v. Reddy, “Performance analysis of photovoltaic systems: A review,” *Renewable and Sustainable Energy Reviews*, vol. 13, no. 8. pp. 1884–1897, Oct. 2009, doi: 10.1016/j.rser.2009.01.009.

[42] E. Skoplaki and J. A. Palyvos, “On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations,” *Solar Energy*, vol. 83, no. 5, pp. 614–624, May 2009, doi: 10.1016/j.solener.2008.10.008.

[43] S. Dubey, J. N. Sarvaiya, and B. Seshadri, “Temperature dependent photovoltaic (PV) efficiency and its effect on PV production in the world - A review,” in *Energy Procedia*, 2013, vol. 33, pp. 311–321, doi: 10.1016/j.egypro.2013.05.072.

[44] B. Parida, S. Iniyan, and R. Goic, “A review of solar photovoltaic technologies,” *Renewable and Sustainable Energy Reviews*, vol. 15, no. 3. pp. 1625–1636, Apr. 2011, doi: 10.1016/j.rser.2010.11.032.

[45] Mc. Guerrero Delgado, J. Sánchez Ramos, E. A. Rodríguez Jara, J. L. Molina Félix, and S. Álvarez Domínguez, “Decision-making approach: A simplified model for energy performance evaluation of photovoltaic modules,” *Energy Conversion and Management*, vol. 177, pp. 350–362, Dec. 2018, doi: 10.1016/j.enconman.2018.09.080.

[46] C. Cañete, J. Carretero, and M. Sidrach-de-Cardona, “Energy performance of different photovoltaic module technologies under outdoor conditions,” *Energy*, vol. 65, pp. 295–302, Feb. 2014, doi: 10.1016/j.energy.2013.12.013.

[47] L. el Chaar, L. A. Lamont, and N. el Zein, “Review of photovoltaic technologies,” *Renewable and Sustainable Energy Reviews*, vol. 15, no. 5. pp. 2165–2175, Jun. 2011, doi: 10.1016/j.rser.2011.01.004.

[48] J. Ahmed and Z. Salam, “A critical evaluation on maximum power point tracking methods for partial shading in PV systems,” *Renewable and Sustainable Energy Reviews*, vol. 47. Elsevier Ltd, pp. 933–953, Jul. 01, 2015, doi: 10.1016/j.rser.2015.03.080.

[49] M. A. Ghasemi, H. Mohammadian Forushani, and M. Parniani, “Partial shading detection and smooth maximum power point tracking of PV arrays under PSC,” *IEEE Transactions on Power Electronics*, vol. 31, no. 9, pp. 6281–6292, Sep. 2016, doi: 10.1109/TPEL.2015.2504515.

[50] B. A. Alsayid, S. Y. Alsadi, J. S. Jallad, and M. H. Dradi, “Partial Shading of PV System Simulation with Experimental Results,” *Smart Grid and Renewable Energy*, vol. 04, no. 06, pp. 429–435, 2013, doi: 10.4236/sgre.2013.46049.

[51] F. Salem and M. A. Awadallah, "Detection and assessment of partial shading in photovoltaic arrays," *Journal of Electrical Systems and Information Technology*, vol. 3, no. 1, pp. 23–32, May 2016, doi: 10.1016/j.jesit.2015.10.003.

[52] A. Phinikarides, N. Kindyni, G. Makrides, and G. E. Georghiou, "Review of photovoltaic degradation rate methodologies," *Renewable and Sustainable Energy Reviews*, vol. 40. Elsevier Ltd, pp. 143–152, 2014, doi: 10.1016/j.rser.2014.07.155.

[53] D. C. Jordan, T. J. Silverman, B. Sekulic, and S. R. Kurtz, "PV degradation curves: non-linearities and failure modes," *Progress in Photovoltaics: Research and Applications*, vol. 25, no. 7, pp. 583–591, Jul. 2017, doi: 10.1002/pip.2835.

[54] J. Y. Ye, T. Reindl, A. G. Aberle, and T. M. Walsh, "Performance degradation of various PV module technologies in tropical Singapore," *IEEE Journal of Photovoltaics*, vol. 4, no. 5, pp. 1288–1294, 2014, doi: 10.1109/JPHOTOV.2014.2338051.

[55] T. S. Liang *et al.*, "A review of crystalline silicon bifacial photovoltaic performance characterisation and simulation," *Energy and Environmental Science*, vol. 12, no. 1. Royal Society of Chemistry, pp. 116–148, Jan. 01, 2019, doi: 10.1039/c8ee02184h.

[56] K. Zeb *et al.*, "A comprehensive review on inverter topologies and control strategies for grid connected photovoltaic system," *Renewable and Sustainable Energy Reviews*, vol. 94. Elsevier Ltd, pp. 1120–1141, Oct. 01, 2018, doi: 10.1016/j.rser.2018.06.053.

[57] J. Jana, H. Saha, and K. das Bhattacharya, "A review of inverter topologies for single-phase grid-connected photovoltaic systems," *Renewable and Sustainable Energy Reviews*, vol. 72. Elsevier Ltd, pp. 1256–1270, 2017, doi: 10.1016/j.rser.2016.10.049.

[58] S. Kouro, J. I. Leon, D. Vinnikov, and L. G. Franquelo, "Grid-connected photovoltaic systems: An overview of recent research and emerging PV converter technology," *IEEE Industrial Electronics Magazine*, vol. 9, no. 1, pp. 47–61, Mar. 2015, doi: 10.1109/MIE.2014.2376976.

[59] F. Blaabjerg, R. Teodorescu, M. Liserre, and A. v. Timbus, "Overview of control and grid synchronization for distributed power generation systems," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 5. 2006, doi: 10.1109/TIE.2006.881997.

[60] G. A. Rampinelli, A. Krenzinger, and F. Chenlo Romero, "Mathematical models for efficiency of inverters used in grid connected photovoltaic systems," *Renewable and Sustainable Energy Reviews*, vol. 34. pp. 578–587, Jun. 2014, doi: 10.1016/j.rser.2014.03.047.

[61] A. A. Rahman, Z. Salam, S. Shaari, and M. Z. Ramli, "Methodology to determine photovoltaic inverter conversion efficiency for the equatorial region," *Applied Sciences (Switzerland)*, vol. 10, no. 1, Jan. 2020, doi: 10.3390/app10010201.

[62] S. B. Kjaer, J. K. Pedersen, and F. Blaabjerg, "A review of single-phase grid-connected inverters for photovoltaic modules," *IEEE Transactions on Industry Applications*, vol. 41, no. 5. pp. 1292–1306, Sep. 2005, doi: 10.1109/TIA.2005.853371.

[63] M. Liberos, R. González-Medina, G. Garcerá, and E. Figueres, "A method to enhance the global efficiency of high-power photovoltaic inverters connected in parallel," *Energies*, vol. 12, no. 11, Jun. 2019, doi: 10.3390/en12112219.

[64] L. Zhang, K. Sun, H. Hu, and Y. Xing, "A system-level control strategy of photovoltaic grid-tied generation systems for european efficiency enhancement," *IEEE Transactions on Power Electronics*, vol. 29, no. 7, pp. 3445–3453, Jul. 2014, doi: 10.1109/TPEL.2013.2279334.

[65] T. J. Formica, H. A. Khan, and M. G. Pecht, "The Effect of Inverter Failures on the Return on Investment of Solar Photovoltaic Systems," *IEEE Access*, vol. 5, pp. 21336–21343, Sep. 2017, doi: 10.1109/ACCESS.2017.2753246.

[66] M. S. Tariq, S. A. Butt, and H. A. Khan, "Impact of module and inverter failures on the performance of central-, string-, and micro-inverter PV systems," *Microelectronics Reliability*, vol. 88–90, pp. 1042–1046, Sep. 2018, doi: 10.1016/j.microrel.2018.06.096.

[67] H. Mousazadeh, A. Keyhani, A. Javadi, H. Mobli, K. Abrinia, and A. Sharifi, "A review of principle and sun-tracking methods for maximizing solar systems output," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 8. pp. 1800–1818, Oct. 2009, doi: 10.1016/j.rser.2009.01.022.

[68] A. Z. Hafez, A. M. Yousef, and N. M. Harag, "Solar tracking systems: Technologies and trackers drive types – A review," *Renewable and Sustainable Energy Reviews*, vol. 91. Elsevier Ltd, pp. 754–782, Aug. 01, 2018, doi: 10.1016/j.rser.2018.03.094.

[69] C. Y. Lee, P. C. Chou, C. M. Chiang, and C. F. Lin, "Sun tracking systems: A review," *Sensors*, vol. 9, no. 5. pp. 3875–3890, Apr. 27, 2009, doi: 10.3390/s90503875.

[70] N. AL-Rousan, N. A. M. Isa, and M. K. M. Desa, "Advances in solar photovoltaic tracking systems: A review," *Renewable and Sustainable Energy Reviews*, vol. 82. Elsevier Ltd, pp. 2548–2569, Feb. 01, 2018, doi: 10.1016/j.rser.2017.09.077.

[71] N. AL-Rousan, N. A. Mat Isa, and M. K. Mat Desa, "Efficient single and dual axis solar tracking system controllers based on adaptive neural fuzzy inference system," *Journal of King Saud University - Engineering Sciences*, vol. 32, no. 7, pp. 459–469, Nov. 2020, doi: 10.1016/j.jksues.2020.04.004.

[72] R. Eke and A. Senturk, "Performance comparison of a double-axis sun tracking versus fixed PV system," *Solar Energy*, vol. 86, no. 9, pp. 2665–2672, Sep. 2012, doi: 10.1016/j.solener.2012.06.006.

[73] S. A. S. Eldin, M. S. Abd-Elhady, and H. A. Kandil, "Feasibility of solar tracking systems for PV panels in hot and cold regions," *Renewable Energy*, vol. 85, pp. 228–233, Jan. 2016, doi: 10.1016/j.renene.2015.06.051.

[74] M. Koussa, A. Cheknane, S. Hadji, M. Haddadi, and S. Noureddine, "Measured and modelled improvement in solar energy yield from flat plate photovoltaic systems utilizing different tracking systems and under a range of environmental conditions," *Applied Energy*, vol. 88, no. 5, pp. 1756–1771, 2011, doi: 10.1016/j.apenergy.2010.12.002.

[75] W. Nsengiyumva, S. G. Chen, L. Hu, and X. Chen, "Recent advancements and challenges in Solar Tracking Systems (STS): A review," *Renewable and Sustainable Energy Reviews*, vol. 81. Elsevier Ltd, pp. 250–279, 2018, doi: 10.1016/j.rser.2017.06.085.

[76] W. Gu, T. Ma, S. Ahmed, Y. Zhang, and J. Peng, "A comprehensive review and outlook of bifacial photovoltaic (bPV) technology," *Energy Conversion and Management*, vol. 223. Elsevier Ltd, Nov. 01, 2020, doi: 10.1016/j.enconman.2020.113283.

[77] R. Guerrero-Lemus, R. Vega, T. Kim, A. Kimm, and L. E. Shephard, "Bifacial solar photovoltaics - A technology review," *Renewable and Sustainable Energy Reviews*, vol. 60. Elsevier Ltd, pp. 1533–1549, Jul. 01, 2016, doi: 10.1016/j.rser.2016.03.041.

[78] T. S. Liang *et al.*, "A review of crystalline silicon bifacial photovoltaic performance characterisation and simulation," *Energy and Environmental Science*, vol. 12, no. 1. Royal Society of Chemistry, pp. 116–148, Jan. 01, 2019, doi: 10.1039/c8ee02184h.

[79] X. Sun, M. R. Khan, C. Deline, and M. A. Alam, "Optimization and performance of bifacial solar modules: A global perspective," *Applied Energy*, vol. 212, 2018, doi: 10.1016/j.apenergy.2017.12.041.

[80] J. S. Stein, D. Riley, M. Lave, C. Hansen, C. Deline, and F. Toor, "Outdoor Field Performance from Bifacial Photovoltaic Modules and Systems," 2018, doi: 10.1109/pvsc.2017.8366042.

[81] M. T. Patel, M. R. Khan, X. Sun, and M. A. Alam, "A worldwide cost-based design and optimization of tilted bifacial solar farms," *Applied Energy*, vol. 247, 2019, doi: 10.1016/j.apenergy.2019.03.150.

[82] C. D. Rodríguez-Gallegos, M. Bieri, O. Gandhi, J. P. Singh, T. Reindl, and S. K. Panda, “Monofacial vs bifacial Si-based PV modules: Which one is more cost-effective?,” *Solar Energy*, vol. 176, 2018, doi: 10.1016/j.solener.2018.10.012.

[83] L. Sun, L. Lu, and H. Yang, “Optimum design of shading-type building-integrated photovoltaic claddings with different surface azimuth angles,” *Applied Energy*, vol. 90, no. 1, 2012, doi: 10.1016/j.apenergy.2011.01.062.

[84] S. Wang *et al.*, “Bifacial Photovoltaic Systems Energy Yield Modelling,” in *Energy Procedia*, 2015, vol. 77, doi: 10.1016/j.egypro.2015.07.060.

[85] U. A. Yusufoglu, T. M. Pletzer, L. J. Koduvelikulathu, C. Comparotto, R. Kopecek, and H. Kurz, “Analysis of the annual performance of bifacial modules and optimization methods,” *IEEE Journal of Photovoltaics*, vol. 5, no. 1, 2015, doi: 10.1109/JPHOTOV.2014.2364406.

[86] P. Sánchez-Friera *et al.*, “Development and characterization of industrial bifacial PV modules with ultra-thin screen-printed solar cells,” *22nd EUPVSEC Milan*, 2007.

[87] W. Gu, T. Ma, M. Li, L. Shen, and Y. Zhang, “A coupled optical-electrical-thermal model of the bifacial photovoltaic module,” *Applied Energy*, vol. 258, 2020, doi: 10.1016/j.apenergy.2019.114075.

[88] M. H. Saw, Y. S. Khoo, J. P. Singh, and Y. Wang, “Enhancing optical performance of bifacial PV modules,” in *Energy Procedia*, 2017, vol. 124, doi: 10.1016/j.egypro.2017.09.285.

[89] I. Shoukry, J. Libal, R. Kopecek, E. Wefringhaus, and J. Werner, “Modelling of Bifacial Gain for Stand-alone and in-field Installed Bifacial PV Modules,” in *Energy Procedia*, 2016, vol. 92, doi: 10.1016/j.egypro.2016.07.025.

[90] K. Keeratimahat, A. Bruce, J. K. Copper, and R. Evans, “Temperature Estimation of an Unconventional PV Array using the Sandia Module Temperature Model,” *2015 Asia-Pacific Solar Research Conference*, 2015.

[91] D. Faiman, “Assessing the outdoor operating temperature of photovoltaic modules,” *Progress in Photovoltaics: Research and Applications*, vol. 16, no. 4, 2008, doi: 10.1002/pip.813.

[92] J. E. Castillo-Aguilella and P. S. Hauser, “Multi-Variable Bifacial Photovoltaic Module Test Results and Best-Fit Annual Bifacial Energy Yield Model,” *IEEE Access*, vol. 4, 2016, doi: 10.1109/ACCESS.2016.2518399.

[93] B. Durusoy, T. Ozden, and B. G. Akinoglu, “Solar irradiation on the rear surface of bifacial solar modules: a modeling approach,” *Scientific Reports*, vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-70235-3.

[94] W. Gu, S. Li, X. Liu, Z. Chen, X. Zhang, and T. Ma, “Experimental investigation of the bifacial photovoltaic module under real conditions,” *Renewable Energy*, 2020, doi: 10.1016/j.renene.2020.12.024.

[95] J. R. Ledesma *et al.*, “A simulation model of the irradiation and energy yield of large bifacial photovoltaic plants,” *Solar Energy*, vol. 206, pp. 522–538, Aug. 2020, doi: 10.1016/j.solener.2020.05.108.

[96] World Bank, “Where Sun Meets Water: Floating Solar Market,” *World Bank Group, ESMAP and SERIS*, vol. 1, no. 2, 2019.

[97] M. Acharya and S. Devraj, “Floating Solar Photovoltaic (FSPV): A Third Pillar to Solar PV Sector ?,” *TERI Discussion Paper:Output of the ETC India Project (New Delhi: The Energy and Resources Institute)*, 2019.

[98] S. Oliveira-Pinto and J. Stokkermans, “Assessment of the potential of different floating solar technologies – Overview and analysis of different case studies,” *Energy Conversion and Management*, vol. 211, 2020, doi: 10.1016/j.enconman.2020.112747.

[99] A. Goswami and P. K. Sadhu, “Degradation analysis and the impacts on feasibility study of floating solar photovoltaic systems,” *Sustainable Energy, Grids and Networks*, vol. 26, 2021, doi: 10.1016/j.segan.2020.100425.

[100] J. wei Zhang, D. kun Cao, S. Diaham, X. Zhang, X. qian Yin, and Q. Wang, “Research on potential induced degradation (PID) of polymeric backsheet in PV modules after salt-mist exposure,” *Solar Energy*, vol. 188, 2019, doi: 10.1016/j.solener.2019.06.019.

[101] M. Kumar and A. Kumar, “Experimental validation of performance and degradation study of canal-top photovoltaic system,” *Applied Energy*, vol. 243, 2019, doi: 10.1016/j.apenergy.2019.03.168.

[102] G. Illya, V. Handara, M. Siahandan, A. Nathania, and A. S. Budiman, “Mechanical Studies of Solar Photovoltaics (PV) Backsheets Under Salt Damp Heat Environments,” *Procedia Engineering*, vol. 215, 2017, doi: 10.1016/j.proeng.2017.12.144.

[103] M. Köntges *et al.*, *Review of Failures of Photovoltaic Modules*. 2014.

[104] M. Paggi, S. Kajari-Schröder, and U. Eitner, “Thermomechanical deformations in photovoltaic laminates,” *Journal of Strain Analysis for Engineering Design*, vol. 46, no. 8, 2011, doi: 10.1177/0309324711421722.

[105] F. Setiawan, T. Dewi, and S. Yusi, “Sea Salt Deposition Effect on Output and Efficiency Losses of the Photovoltaic System; A case study in Palembang, Indonesia,” in *Journal of Physics: Conference Series*, 2019, vol. 1167, no. 1, doi: 10.1088/1742-6596/1167/1/012028.

[106] H. Liu, V. Krishna, J. Lun Leung, T. Reindl, and L. Zhao, “Field experience and performance analysis of floating PV technologies in the tropics,” *Progress in Photovoltaics: Research and Applications*, vol. 26, no. 12, 2018, doi: 10.1002/pip.3039.

[107] S. Z. Golroodbari and W. van Sark, “Simulation of performance differences between offshore and land-based photovoltaic systems,” *Progress in Photovoltaics: Research and Applications*, vol. 28, no. 9, 2020, doi: 10.1002/pip.3276.

[108] A. Mellit, G. M. Tina, and S. A. Kalogirou, “Fault detection and diagnosis methods for photovoltaic systems: A review,” *Renewable and Sustainable Energy Reviews*, vol. 91. Elsevier Ltd, pp. 1–17, Aug. 01, 2018, doi: 10.1016/j.rser.2018.03.062.

[109] X. Sun, R. V. K. Chavali, and M. A. Alam, “Real-time monitoring and diagnosis of photovoltaic system degradation only using maximum power point—the Suns-Vmp method,” *Progress in Photovoltaics: Research and Applications*, vol. 27, no. 1, 2019, doi: 10.1002/pip.3043.

[110] H. Farhangi, “The path of the smart grid,” *IEEE Power and Energy Magazine*, vol. 8, no. 1, 2010, doi: 10.1109/MPE.2009.934876.

[111] M. Castillo-Cagigal, A. Gutiérrez, F. Monasterio-Huelin, E. Caamaño-Martín, D. Masa, and J. Jimenéz-Leube, “A semi-distributed electric demand-side management system with PV generation for self-consumption enhancement,” *Energy Conversion and Management*, vol. 52, no. 7, 2011, doi: 10.1016/j.enconman.2011.01.017.

[112] M. Castillo-Cagigal *et al.*, “PV self-consumption optimization with storage and Active DSM for the residential sector,” *Solar Energy*, vol. 85, no. 9, 2011, doi: 10.1016/j.solener.2011.06.028.

[113] G. A. Dávi, M. Castillo-Cagigal, E. Caamaño-Martín, and J. Solano, “Evaluation Of Load Matching And Grid Interaction Indexes Of A Net Plus-Energy House In Brazil With A Hybrid PV System And Demand-Side Management,” 2016.

[114] E. Matallanas *et al.*, “Electrical energy balance contest in Solar Decathlon Europe 2012,” *Energy and Buildings*, vol. 83, 2014, doi: 10.1016/j.enbuild.2014.03.076.

[115] Y. M. Saint-Drenan, S. Bofinger, R. Fritz, S. Vogt, G. H. Good, and J. Dobschinski, “An empirical approach to parameterizing photovoltaic plants for power forecasting and simulation,” *Solar Energy*, vol. 120, 2015, doi: 10.1016/j.solener.2015.07.024.

[116] A. Luque and S. Hegedus, *Handbook of Photovoltaic Science and Engineering*. 2011.

[117] B. Meng, R. C. G. M. Loonen, and J. L. M. Hensen, “Data-driven inference of unknown tilt and azimuth of distributed PV systems,” *Solar Energy*, vol. 211, 2020, doi: 10.1016/j.solener.2020.09.077.

[118] S. Killinger, N. Engerer, and B. Müller, “QCPV: A quality control algorithm for distributed photovoltaic array power output,” *Solar Energy*, vol. 143, 2017, doi: 10.1016/j.solener.2016.12.053.

[119] J. D. Bastidas-Rodriguez, G. Petrone, C. A. Ramos-Paja, and G. Spagnuolo, “A genetic algorithm for identifying the single diode model parameters of a photovoltaic panel,” *Mathematics and Computers in Simulation*, vol. 131, 2017, doi: 10.1016/j.matcom.2015.10.008.

[120] D. E. Guzman Razo, B. Müller, H. Madsen, and C. Wittwer, “A Genetic Algorithm Approach as a Self-Learning and Optimization Tool for PV Power Simulation and Digital Twinning,” *Energies*, vol. 13, no. 24, 2020, doi: 10.3390/en13246712.

[121] I. Lillo-Bravo, P. González-Martínez, M. Larrañeta, and J. Guasumba-Codena, “Impact of energy losses due to failures on photovoltaic plant energy balance,” *Energies*, vol. 11, no. 2, 2018, doi: 10.3390/en11020363.

[122] S. Dusmez and B. Akin, “An accelerated thermal aging platform to monitor fault precursor on-state resistance,” 2016, doi: 10.1109/IEMDC.2015.7409238.

[123] M. A. Vogelsberger, T. Wiesinger, and H. Ertl, “Life-cycle monitoring and voltage-managing unit for DC-link electrolytic capacitors in PWM converters,” *IEEE Transactions on Power Electronics*, vol. 26, no. 2, 2011, doi: 10.1109/TPEL.2010.2059713.

[124] H. Oh, B. Han, P. McCluskey, C. Han, and B. D. Youn, “Physics-of-failure, condition monitoring, and prognostics of insulated gate bipolar transistor modules: A review,” *IEEE Transactions on Power Electronics*, vol. 30, no. 5, 2015, doi: 10.1109/TPEL.2014.2346485.

[125] S. Yang, D. Xiang, A. Bryant, P. Mawby, L. Ran, and P. Tavner, “Condition monitoring for device reliability in power electronic converters: A review,” *IEEE Transactions on Power Electronics*, vol. 25, no. 11. 2010, doi: 10.1109/TPEL.2010.2049377.

[126] H. Soliman, H. Wang, and F. Blaabjerg, “A Review of the Condition Monitoring of Capacitors in Power Electronic Converters,” *IEEE Transactions on Industry Applications*, vol. 52, no. 6, 2016, doi: 10.1109/TIA.2016.2591906.

[127] U. M. Choi, S. Jorgensen, and F. Blaabjerg, “Advanced Accelerated Power Cycling Test for Reliability Investigation of Power Device Modules,” *IEEE Transactions on Power Electronics*, vol. 31, no. 12, 2016, doi: 10.1109/TPEL.2016.2521899.

[128] S. Bęczkowski, P. Ghimre, A. R. de Vega, S. Munk-Nielsen, B. Rannestad, and P. Thøgersen, “Online Vce measurement method for wear-out monitoring of high power IGBT modules,” 2013, doi: 10.1109/EPE.2013.6634390.

[129] B. Ji *et al.*, “In situ diagnostics and prognostics of solder fatigue in IGBT modules for electric vehicle drives,” *IEEE Transactions on Power Electronics*, vol. 30, no. 3, 2015, doi: 10.1109/TPEL.2014.2318991.

[130] D. Astigarraga *et al.*, “Analysis of the Results of Accelerated Aging Tests in Insulated Gate Bipolar Transistors,” *IEEE Transactions on Power Electronics*, vol. 31, no. 11, 2016, doi: 10.1109/TPEL.2015.2512923.

[131] X. Ye, C. Chen, Y. Wang, G. Zhai, and G. J. Vachtsevanos, "Online Condition Monitoring of Power MOSFET Gate Oxide Degradation Based on Miller Platform Voltage," *IEEE Transactions on Power Electronics*, vol. 32, no. 6, 2017, doi: 10.1109/TPEL.2016.2602323.

[132] J. Liu, G. Zhang, Q. Chen, L. Qi, Y. Geng, and J. Wang, "In situ Condition Monitoring of IGBTs Based on the Miller Plateau Duration," *IEEE Transactions on Power Electronics*, vol. 34, no. 1, 2018, doi: 10.1109/TPEL.2018.2820700.

[133] C. H. van der Broeck, A. Gospodinov, and R. W. de Doncker, "IGBT Junction Temperature Estimation via Gate Voltage Plateau Sensing," *IEEE Transactions on Industry Applications*, vol. 54, no. 5, 2018, doi: 10.1109/TIA.2018.2836362.

[134] R. Mandeya, C. Chen, V. Pickert, and R. T. Naayagi, "Prethreshold Voltage as a Low-Component Count Temperature Sensitive Electrical Parameter Without Self-Heating," *IEEE Transactions on Power Electronics*, vol. 33, no. 4, 2018, doi: 10.1109/TPEL.2017.2749179.

[135] N. Baker, S. Munk-Nielsen, F. Iannuzzo, and M. Liserre, "IGBT Junction Temperature Measurement via Peak Gate Current," *IEEE Transactions on Power Electronics*, vol. 31, no. 5, 2016, doi: 10.1109/TPEL.2015.2464714.

[136] H. Luo, Y. Chen, P. Sun, W. Li, and X. He, "Junction Temperature Extraction Approach with Turn-Off Delay Time for High-Voltage High-Power IGBT Modules," *IEEE Transactions on Power Electronics*, vol. 31, no. 7, 2016, doi: 10.1109/TPEL.2015.2481465.

[137] H. Niu and R. D. Lorenz, "Sensing Power MOSFET Junction Temperature Using Gate Drive Turn-On Current Transient Properties," *IEEE Transactions on Industry Applications*, vol. 52, no. 2, 2016, doi: 10.1109/TIA.2015.2497202.

[138] Z. Wang, B. Tian, W. Qiao, and L. Qu, "Real-Time Aging Monitoring for IGBT Modules Using Case Temperature," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 2, 2016, doi: 10.1109/TIE.2015.2497665.

[139] D. Xiang, L. Ran, P. Tavner, A. Bryant, S. Yang, and P. Mawby, "Monitoring solder fatigue in a power module using case-above-ambient temperature rise," *IEEE Transactions on Industry Applications*, vol. 47, no. 6, 2011, doi: 10.1109/TIA.2011.2168556.

[140] W. Lai, M. Chen, L. Ran, O. Alatise, S. Xu, and P. Mawby, "Low Δtj Stress Cycle Effect in IGBT Power Module Die-Attach Lifetime Modeling," *IEEE Transactions on Power Electronics*, vol. 31, no. 9, 2016, doi: 10.1109/TPEL.2015.2501540.

[141] Y. Wu and X. Du, "A VEN Condition Monitoring Method of DC-Link Capacitors for Power Converters," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 2, 2019, doi: 10.1109/TIE.2018.2835393.

[142] N. Agarwal, M. W. Ahmad, and S. Anand, "Quasi-Online Technique for Health Monitoring of Capacitor in Single-Phase Solar Inverter," *IEEE Transactions on Power Electronics*, vol. 33, no. 6, 2018, doi: 10.1109/TPEL.2017.2736162.

[143] A. Wechsler, B. C. Mecrow, D. J. Atkinson, J. W. Bennett, and M. Benarous, "Condition monitoring of DC-link capacitors in aerospace drives," *IEEE Transactions on Industry Applications*, vol. 48, no. 6, 2012, doi: 10.1109/TIA.2012.2222333.

[144] Y. J. Jo, T. H. Nguyen, and D. C. Lee, "Condition monitoring of submodule capacitors in modular multilevel converters," 2014, doi: 10.1109/ECCE.2014.6953683.

[145] S. Dusmez, M. Bhardwaj, L. Sun, and B. Akin, "In Situ Condition Monitoring of High-Voltage Discrete Power MOSFET in Boost Converter Through Software Frequency Response Analysis," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 12, 2016, doi: 10.1109/TIE.2016.2595482.

[146] D. Xiang, L. Ran, P. Tavner, S. Yang, A. Bryant, and P. Mawby, "Condition monitoring power module solder fatigue using inverter harmonic identification," *IEEE Transactions on Power Electronics*, vol. 27, no. 1, 2012, doi: 10.1109/TPEL.2011.2160988.

[147] A. Soualhi *et al.*, "Heath Monitoring of Capacitors and Supercapacitors Using the Neo-Fuzzy Neural Approach," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 1, 2018, doi: 10.1109/TII.2017.2701823.

[148] H. Soliman, H. Wang, B. Gadalla, and F. Blaabjerg, "Condition monitoring for DC-link capacitors based on artificial neural network algorithm," in *International Conference on Power Engineering, Energy and Electrical Drives*, 2015, vol. 2015-September, doi: 10.1109/PowerEng.2015.7266382.

[149] M. Algreer, M. Armstrong, and D. Giaouris, "Active online system identification of switch mode DC-DC power converter based on efficient recursive DCD-IIR adaptive filter," *IEEE Transactions on Power Electronics*, vol. 27, no. 11, 2012, doi: 10.1109/TPEL.2012.2190754.

[150] M. Ahmeid, M. Armstrong, S. Gadoue, M. Al-Greer, and P. Missailidis, "Real-time parameter estimation of DC-DC converters using a self-tuned kalman filter," *IEEE Transactions on Power Electronics*, vol. 32, no. 7, 2017, doi: 10.1109/TPEL.2016.2606417.

[151] J. Poon, P. Jain, C. Spanos, S. K. Panda, and S. R. Sanders, "Fault Prognosis for Power Electronics Systems Using Adaptive Parameter Identification," *IEEE Transactions on Industry Applications*, vol. 53, no. 3, 2017, doi: 10.1109/TIA.2017.2664052.

[152] B. X. Li and K. S. Low, "Low sampling rate online parameters monitoring of DC-DC converters for predictive-maintenance using biogeography-based optimization," *IEEE Transactions on Power Electronics*, vol. 31, no. 4, 2016, doi: 10.1109/TPEL.2015.2472459.

[153] A. Alghassi, S. Perinpanayagam, M. Samie, and T. Sreenuch, "Computationally efficient, real-time, and embeddable prognostic techniques for power electronics," *IEEE Transactions on Power Electronics*, vol. 30, no. 5, 2015, doi: 10.1109/TPEL.2014.2360662.

[154] S. H. Ali, M. Heydarzadeh, S. Dusmez, X. Li, A. S. Kamath, and B. Akin, "Lifetime Estimation of Discrete IGBT Devices Based on Gaussian Process," in *IEEE Transactions on Industry Applications*, 2018, vol. 54, no. 1, doi: 10.1109/TIA.2017.2753722.

[155] A. Parrott and L. Warshaw, "Industry 4.0 and the digital twin," *Deloitte University Press*, 2017.

[156] M. Milton, C. O. de La, H. L. Ginn, and A. Benigni, "Controller-Embeddable Probabilistic Real-Time Digital Twins for Power Electronic Converter Diagnostics," *IEEE Transactions on Power Electronics*, vol. 35, no. 9, 2020, doi: 10.1109/TPEL.2020.2971775.

[157] P. Jain, J. Poon, J. P. Singh, C. Spanos, S. R. Sanders, and S. K. Panda, "A digital twin approach for fault diagnosis in distributed photovoltaic systems," *IEEE Transactions on Power Electronics*, vol. 35, no. 1, 2020, doi: 10.1109/TPEL.2019.2911594.

[158] Y. Peng, S. Zhao, and H. Wang, "A Digital Twin Based Estimation Method for Health Indicators of DC-DC Converters," *IEEE Transactions on Power Electronics*, vol. 36, no. 2, 2021, doi: 10.1109/TPEL.2020.3009600.

[159] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. van Mierlo, and P. van den Bossche, "Critical review of state of health estimation methods of Li-ion batteries for real applications," *Renewable and Sustainable Energy Reviews*, vol. 56. 2016, doi: 10.1016/j.rser.2015.11.042.

[160] E. Sarasketa-Zabala, E. Martinez-Laserna, M. Berecibar, I. Gandiaga, L. M. Rodriguez-Martinez, and I. Villarreal, "Realistic lifetime prediction approach for Li-ion batteries," *Applied Energy*, vol. 162, 2016, doi: 10.1016/j.apenergy.2015.10.115.

[161] B. Xia, W. Zheng, R. Zhang, Z. Lao, and Z. Sun, “A novel observer for lithium-ion battery state of charge estimation in electric vehicles based on a second-order equivalent circuit model,” *Energies*, vol. 10, no. 8, 2017, doi: 10.3390/en10081150.

[162] F. Zhang *et al.*, “State-of-charge estimation based on microcontroller-implemented sigma-point Kalman filter in a modular cell balancing system for Lithium-Ion battery packs,” 2015, doi: 10.1109/COMPEL.2015.7236525.

[163] Y. Zou, X. Hu, H. Ma, and S. E. Li, “Combined State of Charge and State of Health estimation over lithium-ion battery cell cycle lifespan for electric vehicles,” *Journal of Power Sources*, vol. 273, 2015, doi: 10.1016/j.jpowsour.2014.09.146.

[164] X. Hu, S. Li, and H. Peng, “A comparative study of equivalent circuit models for Li-ion batteries,” *Journal of Power Sources*, vol. 198, 2012, doi: 10.1016/j.jpowsour.2011.10.013.

[165] R. Jackey, M. Saginaw, P. Sanghvi, J. Gazzarri, T. Huria, and M. Ceraolo, “Battery model parameter estimation using a layered technique: An example using a lithium iron phosphate cell,” in *SAE Technical Papers*, 2013, vol. 2, doi: 10.4271/2013-01-1547.

[166] D. Andre, C. Appel, T. Soczka-Guth, and D. U. Sauer, “Advanced mathematical methods of SOC and SOH estimation for lithium-ion batteries,” *Journal of Power Sources*, vol. 224, 2013, doi: 10.1016/j.jpowsour.2012.10.001.

[167] M. Xavier, G. L. Plett, and M. S. Trimboli, “Power-limit estimation for lithium ion batteries using physics-based models and predictive methods,” in *35th Annual International Battery Seminar and Exhibit 2018*, 2018, vol. 2.

[168] A. K. de Souza, G. Plett, and M. S. Trimboli, “Lithium-Ion Battery Charging Control Using a Coupled Electro-Thermal Model and Model Predictive Control,” in *Conference Proceedings - IEEE Applied Power Electronics Conference and Exposition - APEC*, 2020, vol. 2020-March, doi: 10.1109/APEC39645.2020.9124431.

[169] G. L. Plett, *Battery Management Systems, Battery Modeling*, vol. 1, no. 8. 2015.

[170] M. Murnane and A. Ghazel, “A Closer Look at State of Charge (SOC) and State of Health (SOH) Estimation Techniques for Batteries,” *Analog devices*, 2017.

[171] K. H. Chao, S. H. Ho, and M. H. Wang, “Modeling and fault diagnosis of a photovoltaic system,” *Electric Power Systems Research*, vol. 78, no. 1, pp. 97–105, Jan. 2008, doi: 10.1016/j.epsr.2006.12.012.

[172] Y. Zheng *et al.*, “Cell state-of-charge inconsistency estimation for LiFePO4 battery pack in hybrid electric vehicles using mean-difference model,” *Applied Energy*, vol. 111, 2013, doi: 10.1016/j.apenergy.2013.05.048.

[173] A. K. Jain, “Data clustering: 50 years beyond K-means,” *Pattern Recognition Letters*, vol. 31, no. 8, 2010, doi: 10.1016/j.patrec.2009.09.011.

[174] T. Oozeki, T. Yamada, K. Otani, T. Takashima, and K. Kato, “Performance trends in grid-connected photovoltaic systems for public and industrial use in Japan,” *Progress in Photovoltaics: Research and Applications*, vol. 18, no. 8, 2010, doi: 10.1002/pip.992.

[175] N. H. Reich, B. Mueller, A. Armbruster, W. G. J. H. M. van Sark, K. Kiefer, and C. Reise, “Performance ratio revisited: Is PR > 90% realistic?,” in *Progress in Photovoltaics: Research and Applications*, 2012, vol. 20, no. 6, doi: 10.1002/pip.1219.

[176] W. G. J. H. M. van Sark, N. H. Reich, B. Müller, A. Armbruster, K. Kiefer, and C. Reise, “Review of PV performance ratio development,” in *World Renewable Energy Forum, WREF 2012, Including World Renewable Energy Congress XII and Colorado Renewable Energy Society (CRES) Annual Conferen*, 2012, vol. 6.

[177] W. R. Anis and M. A. S. Nour, “Energy losses in photovoltaic systems,” *Energy Conversion and Management*, vol. 36, no. 11, 1995, doi: 10.1016/0196-8904(94)00085-E.

[178] C. Baltus, J. Eikelboom, and R. van Zolingen, “Analytical monitoring of losses in PV systems.,” *14th European Photovoltaic Solar Energy Conference*, no. July. 1997.

[179] G. Blaesser, “PV system measurements and monitoring the European experience,” *Solar Energy Materials and Solar Cells*, vol. 47, no. 1–4, 1997, doi: 10.1016/S0927-0248(97)80008-6.

[180] B. Decker and U. Jahn, “Performance of 170 grid connected PV plants in northern Germany - Analysis of yields and optimization potentials,” in *Solar Energy*, 1997, vol. 59, no. 4-6–6 pt 4, doi: 10.1016/S0038-092X(96)00132-6.

[181] T. Erge, V. U. Hoffmann, and K. Kiefer, “The German experience with grid-connected PV-systems,” *Solar Energy*, vol. 70, no. 6, 2001, doi: 10.1016/S0038-092X(00)00143-2.

[182] S. Pelland and I. Abboud, “Comparing photovoltaic capacity value metrics: A case study for the city of Toronto,” *Progress in Photovoltaics: Research and Applications*, vol. 16, no. 8, 2008, doi: 10.1002/pip.864.

[183] Y. Ueda, K. Kurokawa, K. Kitamura, M. Yokota, K. Akanuma, and H. Sugihara, “Performance analysis of various system configurations on grid-connected residential PV systems,” *Solar Energy Materials and Solar Cells*, vol. 93, no. 6–7, 2009, doi: 10.1016/j.solmat.2008.11.021.

[184] M. A. Eltawil and Z. Zhao, “Grid-connected photovoltaic power systems: Technical and potential problems-A review,” *Renewable and Sustainable Energy Reviews*, vol. 14, no. 1. 2010, doi: 10.1016/j.rser.2009.07.015.

[185] M. Z. Hussin, N. Hasliza, A. Yaacob, Z. M. Zain, A. M. Omar, and S. Shaari, “A development and challenges of grid-connected photovoltaic system in Malaysia,” 2012, doi: 10.1109/ICSGRC.2012.6287160.

[186] J. Leloux, L. Narvarte, and D. Trebosc, “Review of the performance of residential PV systems in Belgium,” *Renewable and Sustainable Energy Reviews*, vol. 16, no. 1. 2012, doi: 10.1016/j.rser.2011.07.145.

[187] J. Leloux, L. Narvarte, and D. Trebosc, “Review of the performance of residential PV systems in France,” *Renewable and Sustainable Energy Reviews*, vol. 16, no. 2. 2012, doi: 10.1016/j.rser.2011.10.018.

[188] S. Wittkopf, S. Valliappan, L. Liu, K. S. Ang, and S. C. J. Cheng, “Analytical performance monitoring of a 142.5kW p grid-connected rooftop BIPV system in Singapore,” *Renewable Energy*, vol. 47, 2012, doi: 10.1016/j.renene.2012.03.034.

[189] M. Díez-Mediavilla, M. I. Dieste-Velasco, M. C. Rodríguez-Amigo, T. García-Calderón, and C. Alonso-Tristán, “Performance of grid-tied PV facilities: A case study based on real data,” *Energy Conversion and Management*, vol. 76, 2013, doi: 10.1016/j.enconman.2013.08.035.

[190] A. Woyte *et al.*, *Analytical Monitoring of Grid-connected Photovoltaic Systems Good practices for monitoring and performance analysis*, no. February. 2014.

[191] L. Bonsignore, M. Davarifar, A. Rabhi, G. M. Tina, and A. Elhajjaji, “Neuro-Fuzzy fault detection method for photovoltaic systems,” in *Energy Procedia*, 2014, vol. 62, pp. 431–441, doi: 10.1016/j.egypro.2014.12.405.

[192] G. Colantuono, A. Everard, L. M. H. Hall, and A. R. Buckley, “Monitoring nationwide ensembles of PV generators: Limitations and uncertainties. The case of the UK,” *Solar Energy*, vol. 108, 2014, doi: 10.1016/j.solener.2014.06.030.

[193] J. Leloux, J. Taylor, R. Moretón, L. Narvarte, D. Trebosc, and A. Desportes, “Monitoring 30,000 PV systems in Europe: Performance, Faults, and State of the Art,” 2015.

[194] S. Y. Myong, Y. C. Park, and S. W. Jeon, “Performance of Si-based PV rooftop systems operated under distinct four seasons,” *Renewable Energy*, vol. 81, 2015, doi: 10.1016/j.renene.2015.03.055.

[195] J. Taylor, J. Leloux, A. M. Everard, J. Briggs, and A. Buckley, “Monitoring thousands of distributed PV systems in the UK : Energy production and performance Sheffield Solar,” *11th Photovoltaic Science Application and Technology (PVSAT-11)*, 2015.

[196] O. Tsafarakis, K. Sinapis, and W. G. J. H. M. van Sark, “PV system performance evaluation by clustering production data to normal and non-normal operation,” *Energies*, vol. 11, no. 4, 2018, doi: 10.3390/en11040977.

[197] B. B. Kausika, P. Moraitis, and W. G. J. H. M. van Sark, “Visualization of operational performance of grid-connected PV systems in selected European countries,” *Energies*, vol. 11, no. 6, 2018, doi: 10.3390/en11061330.

[198] S. Seme, K. Sredenšek, B. Štumberger, and M. Hadžiselimović, “Analysis of the performance of photovoltaic systems in Slovenia,” *Solar Energy*, vol. 180, 2019, doi: 10.1016/j.solener.2019.01.062.

[199] S. Daliento *et al.*, “Monitoring, diagnosis, and power forecasting for photovoltaic fields: A review,” *International Journal of Photoenergy*, vol. 2017. 2017, doi: 10.1155/2017/1356851.

[200] S. R. Madeti and S. N. Singh, “A comprehensive study on different types of faults and detection techniques for solar photovoltaic system,” *Solar Energy*, vol. 158. Elsevier Ltd, pp. 161–185, 2017, doi: 10.1016/j.solener.2017.08.069.

[201] D. C. Jordan, S. R. Kurtz, K. VanSant, and J. Newmiller, “Compendium of photovoltaic degradation rates,” *Progress in Photovoltaics: Research and Applications*, vol. 24, no. 7, 2016, doi: 10.1002/pip.2744.

[202] F. Martínez-Moreno, E. Lorenzo, J. Muñoz, and R. Moretón, “On the testing of large PV arrays,” *Progress in Photovoltaics: Research and Applications*, vol. 20, no. 1, 2012, doi: 10.1002/pip.1102.

[203] J. D. Bastidas-Rodriguez, G. Petrone, C. A. Ramos-Paja, and G. Spagnuolo, “Photovoltaic modules diagnostic: An overview,” 2013, doi: 10.1109/IECON.2013.6699117.

[204] S. Silvestre, A. Chouder, and E. Karatepe, “Automatic fault detection in grid connected PV systems,” *Solar Energy*, vol. 94, pp. 119–127, Aug. 2013, doi: 10.1016/j.solener.2013.05.001.

[205] A. Triki-Lahiani, A. Bennani-Ben Abdelghani, and I. Slama-Belkhodja, “Fault detection and monitoring systems for photovoltaic installations: A review,” *Renewable and Sustainable Energy Reviews*, vol. 82. Elsevier Ltd, pp. 2680–2692, Feb. 01, 2018, doi: 10.1016/j.rser.2017.09.101.

[206] Syafaruddin and D. S. Zinger, “Review on methods of fault diagnosis in photovoltaic system applications,” *Journal of Engineering Science and Technology Review*, vol. 12, no. 5. Eastern Macedonia and Thrace Institute of Technology, pp. 53–66, 2019, doi: 10.25103/jestr.125.07.

[207] W. Chine, A. Mellit, A. M. Pavan, and S. A. Kalogirou, “Fault detection method for grid-connected photovoltaic plants,” *Renewable Energy*, vol. 66, pp. 99–110, Jun. 2014, doi: 10.1016/j.renene.2013.11.073.

[208] U. Jahn and W. Nasse, “Operational performance of grid-connected PV systems on buildings in Germany,” *Progress in Photovoltaics: Research and Applications*, vol. 12, no. 6, 2004, doi: 10.1002/pip.550.

[209] D. C. Jordan and S. R. Kurtz, “Reliability and geographic trends of 50,000 photovoltaic systems in the USA,” *29th European Photovoltaic Solar Energy Conference and Exhibition (EU PVSEC 2014)*, no. September, 2014.

[210] A. Woyte, M. Richter, D. Moser, S. Mau, N. Reich, and U. Jahn, “Monitoring of Photovoltaic Systems: Good Practices and Systematic Analysis,” *Journal of Chemical Information and Modeling*, vol. 53, no. 9, 2013.

[211] E. A. Essah, A. R. Arguelles, and N. Glover, “Assessing the performance of a building integrated BP c-Si PV system,” *Renewable Energy*, vol. 73, 2015, doi: 10.1016/j.renene.2014.04.002.

[212] K. Brecl and M. Topič, “Apparent performance ratio of photovoltaic systems - A methodology for evaluation of photovoltaic systems across a region,” *Journal of Renewable and Sustainable Energy*, vol. 8, no. 4, 2016, doi: 10.1063/1.4955088.

[213] M. Dhimish and V. Holmes, “Fault detection algorithm for grid-connected photovoltaic plants,” *Solar Energy*, vol. 137, 2016, doi: 10.1016/j.solener.2016.08.021.

[214] B. Herteleer, B. Huyck, F. Catthoor, J. Driesen, and J. Cappelle, “Normalised efficiency of photovoltaic systems: Going beyond the performance ratio,” *Solar Energy*, vol. 157, 2017, doi: 10.1016/j.solener.2017.08.037.

[215] R. Perez, M. Kmiecik, C. Herig, and D. Renné, “Remote monitoring of PV performance using geostationary satellites,” *Solar Energy*, vol. 71, no. 4, 2001, doi: 10.1016/S0038-092X(01)00050-0.

[216] A. Drews, H. G. Beyer, and U. Rindelhardt, “Quality of performance assessment of PV plants based on irradiation maps,” *Solar Energy*, vol. 82, no. 11, 2008, doi: 10.1016/j.solener.2008.04.009.

[217] T. Huld, R. Gottschalg, H. G. Beyer, and M. Topič, “Mapping the performance of PV modules, effects of module type and data averaging,” *Solar Energy*, vol. 84, no. 2, 2010, doi: 10.1016/j.solener.2009.12.002.

[218] A. Phinikarides, C. Shimitra, R. Bourgeon, I. Koumparou, G. Makrides, and G. E. Georghiou, “Development of a novel web application for automatic photovoltaic system performance analysis and fault identification,” 2017, doi: 10.1109/PVSC.2017.8366379.

[219] J. P. Charles *et al.*, “Faulty PV panel identification using the Design of Experiments (DoE) method,” *International Journal of Electrical Power and Energy Systems*, vol. 57, 2014, doi: 10.1016/j.ijepes.2013.11.037.

[220] N. Forero, J. Hernández, and G. Gordillo, “Development of a monitoring system for a PV solar plant,” *Energy Conversion and Management*, vol. 47, no. 15–16, 2006, doi: 10.1016/j.enconman.2005.11.012.

[221] N. Gokmen, E. Karatepe, B. Celik, and S. Silvestre, “Simple diagnostic approach for determining of faulted PV modules in string based PV arrays,” *Solar Energy*, vol. 86, no. 11, 2012, doi: 10.1016/j.solener.2012.09.007.

[222] Y. Hu, B. Gao, X. Song, G. Y. Tian, K. Li, and X. He, “Photovoltaic fault detection using a parameter based model,” *Solar Energy*, vol. 96, 2013, doi: 10.1016/j.solener.2013.07.004.

[223] S. Sopitpan, P. Changmuang, and S. Panyakeow, “Monitoring and data analysis of a PV system connected to a grid for home applications,” *Solar Energy Materials and Solar Cells*, vol. 67, no. 1–4, 2001, doi: 10.1016/S0927-0248(00)00317-2.

[224] T. Takashima, J. Yamaguchi, K. Otani, T. Oozeki, K. Kato, and M. Ishida, “Experimental studies of fault location in PV module strings,” *Solar Energy Materials and Solar Cells*, vol. 93, no. 6–7, 2009, doi: 10.1016/j.solmat.2008.11.060.

[225] A. Chouder and S. Silvestre, “Automatic supervision and fault detection of PV systems based on power losses analysis,” *Energy Conversion and Management*, vol. 51, no. 10, 2010, doi: 10.1016/j.enconman.2010.02.025.

[226] N. Gokmen, E. Karatepe, S. Silvestre, B. Celik, and P. Ortega, “An efficient fault diagnosis method for PV systems based on operating voltage-window,” *Energy Conversion and Management*, vol. 73, 2013, doi: 10.1016/j.enconman.2013.05.015.

[227] J. Solórzano and M. A. Egido, “Automatic fault diagnosis in PV systems with distributed MPPT,” *Energy Conversion and Management*, vol. 76, pp. 925–934, 2013, doi: 10.1016/j.enconman.2013.08.055.

[228] F. Spertino and F. Corona, “Monitoring and checking of performance in photovoltaic plants: A tool for design, installation and maintenance of grid-connected systems,” *Renewable Energy*, vol. 60. 2013, doi: 10.1016/j.renene.2013.06.011.

[229] S. Silvestre, M. A. da Silva, A. Chouder, D. Guasch, and E. Karatepe, “New procedure for fault detection in grid connected PV systems based on the evaluation of current and voltage indicators,” *Energy Conversion and Management*, vol. 86, 2014, doi: 10.1016/j.enconman.2014.05.008.

[230] S. Silvestre, L. Mora-López, S. Kichou, F. Sánchez-Pacheco, and M. Dominguez-Pumar, “Remote supervision and fault detection on OPC monitored PV systems,” *Solar Energy*, vol. 137, 2016, doi: 10.1016/j.solener.2016.08.030.

[231] A. Drews *et al.*, “Monitoring and remote failure detection of grid-connected PV systems based on satellite observations,” *Solar Energy*, vol. 81, no. 4, 2007, doi: 10.1016/j.solener.2006.06.019.

[232] S. K. Firth, K. J. Lomas, and S. J. Rees, “A simple model of PV system performance and its use in fault detection,” *Solar Energy*, vol. 84, no. 4, 2010, doi: 10.1016/j.solener.2009.08.004.

[233] A. Chouder, S. Silvestre, N. Sadaoui, and L. Rahmani, “Modeling and simulation of a grid connected PV system based on the evaluation of main PV module parameters,” *Simulation Modelling Practice and Theory*, vol. 20, no. 1, 2012, doi: 10.1016/j.simpat.2011.08.011.

[234] A. Chouder, S. Silvestre, B. Taghezouit, and E. Karatepe, “Monitoring, modelling and simulation of PV systems using LabVIEW,” *Solar Energy*, vol. 91, 2013, doi: 10.1016/j.solener.2012.09.016.

[235] A. Firman, V. Toranzos, L. Vera, A. Busso, and J. de La Casa, “Passive monitoring of the power generated in grid connected PV systems,” in *Energy Procedia*, 2014, vol. 57, doi: 10.1016/j.egypro.2014.10.028.

[236] A. Elkholy, F. H. Fahmy, A. A. Abou El-Ela, A. E.-S. A. Nafeh, and S. R. Spea, “Experimental evaluation of 8kW grid-connected photovoltaic system in Egypt,” *Journal of Electrical Systems and Information Technology*, vol. 3, no. 2, 2016, doi: 10.1016/j.jesit.2015.10.004.

[237] O. Hachana, G. M. Tina, and K. E. Hemsas, “PV array fault DiagnosticTechnique for BIPV systems,” *Energy and Buildings*, vol. 126, 2016, doi: 10.1016/j.enbuild.2016.05.031.

[238] M. Hosseinzadeh and F. Rajaei Salmasi, “Determination of maximum solar power under shading and converter faults - A prerequisite for failure-tolerant power management systems,” *Simulation Modelling Practice and Theory*, vol. 62, 2016, doi: 10.1016/j.simpat.2016.01.011.

[239] S. Killinger, B. Muller, Y. M. Saint-Drenan, and R. McKenna, “Towards an improved nowcasting method by evaluating power profiles of PV systems to detect apparently atypical behavior,” 2017, doi: 10.1109/PVSC.2017.8366169.

[240] Y. Yagi *et al.*, “Diagnostic technology and an expert system for photovoltaic systems using the learning method,” *Solar Energy Materials and Solar Cells*, vol. 75, no. 3–4, 2003, doi: 10.1016/S0927-0248(02)00149-6.

[241] A. Mellit and S. A. Kalogirou, “Artificial intelligence techniques for photovoltaic applications: A review,” *Progress in Energy and Combustion Science*, vol. 34, no. 5. pp. 574–632, Oct. 2008, doi: 10.1016/j.pecs.2008.01.001.

[242] A. Mellit and A. M. Pavan, “Performance prediction of 20 kWp grid-connected photovoltaic plant at Trieste (Italy) using artificial neural network,” *Energy Conversion and Management*, vol. 51, no. 12, 2010, doi: 10.1016/j.enconman.2010.05.007.

[243] F. Almonacid, C. Rus, P. Pérez-Higueras, and L. Hontoria, “Calculation of the energy provided by a PV generator. Comparative study: Conventional methods vs. artificial neural networks,” *Energy*, vol. 36, no. 1, 2011, doi: 10.1016/j.energy.2010.10.028.

[244] E. Spiliotis, N. Z. Legaki, V. Assimakopoulos, H. Doukas, and M. S. el Moursi, “Tracking the performance of photovoltaic systems: a tool for minimising the risk of malfunctions and deterioration,” *IET Renewable Power Generation*, vol. 12, no. 7, 2018, doi: 10.1049/iet-rpg.2017.0596.

[245] H. Ding *et al.*, “Local outlier factor-based fault detection and evaluation of photovoltaic system,” *Solar Energy*, vol. 164, pp. 139–148, Apr. 2018, doi: 10.1016/j.solener.2018.01.049.

[246] G. Chen, P. Lin, Y. Lai, Z. Chen, L. Wu, and S. Cheng, “Location for fault string of photovoltaic array based on current time series change detection,” in *Energy Procedia*, 2018, vol. 145, pp. 406–412, doi: 10.1016/j.egypro.2018.04.067.

[247] S. Vergura, G. Acciani, V. Amoruso, G. E. Patrono, and F. Vacca, “Descriptive and inferential statistics for supervising and monitoring the operation of PV plants,” *IEEE Transactions on Industrial Electronics*, vol. 56, no. 11, 2009, doi: 10.1109/TIE.2008.927404.

[248] S. Vergura, G. Acciani, V. Amoruso, and G. Patrono, “Inferential statistics for monitoring and fault forecasting of PV plants,” 2008, doi: 10.1109/ISIE.2008.4677264.

[249] Y. Zhao, B. Lehman, R. Ball, J. Mosesian, and J. F. de Palma, “Outlier detection rules for fault detection in solar photovoltaic arrays,” 2013, doi: 10.1109/APEC.2013.6520712.

[250] M. Davarifar, A. Rabhi, A. El-Hajjaji, and M. Dahmane, “Real-time model base fault diagnosis of PV panels using statistical signal processing,” 2013, doi: 10.1109/ICRERA.2013.6749826.

[251] F. Harrou, Y. Sun, B. Taghezouit, A. Saidi, and M. E. Hamlati, “Reliable fault detection and diagnosis of photovoltaic systems based on statistical monitoring approaches,” *Renewable Energy*, vol. 116, pp. 22–37, 2018, doi: 10.1016/j.renene.2017.09.048.

[252] S. A. Kalogirou, “Artificial neural networks in renewable energy systems applications: A review,” *Renewable and Sustainable Energy Reviews*, vol. 5, no. 4. pp. 373–401, Dec. 2000, doi: 10.1016/S1364-0321(01)00006-5.

[253] B. Li, C. Delpha, D. Diallo, and A. Migan-Dubois, “Application of Artificial Neural Networks to photovoltaic fault detection and diagnosis: A review,” *Renewable and Sustainable Energy Reviews*, vol. 138. Elsevier Ltd, Mar. 01, 2021, doi: 10.1016/j.rser.2020.110512.

[254] N. Sabri, A. Tlemcani, and A. Chouder, “Faults diagnosis in stand-alone photovoltaic system using artificial neural network,” Oct. 2018, doi: 10.1109/CEIT.2018.8751924.

[255] Syafaruddin, E. Karatepe, and T. Hiyama, “Controlling of artificial neural network for fault diagnosis of photovoltaic array,” 2011, doi: 10.1109/ISAP.2011.6082219.

[256] S. Laamami, M. Benhamed, and L. Sbita, “Artificial neural network-based fault detection and classification for photovoltaic system,” Oct. 2017, doi: 10.1109/GECS.2017.8066211.

[257] A. Ul-Haq, H. F. Sindi, S. Gul, and M. Jalal, “Modeling and Fault Categorization in Thin-Film and Crystalline PV Arrays through Multilayer Neural Network Algorithm,” *IEEE Access*, vol. 8, pp. 102235–102255, 2020, doi: 10.1109/ACCESS.2020.2996969.

[258] H. Mekki, A. Mellit, and H. Salhi, “Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules,” *Simulation Modelling Practice and Theory*, vol. 67, pp. 1–13, Sep. 2016, doi: 10.1016/j.simpat.2016.05.005.

[259] A. Belaout, F. Krim, and A. Mellit, “Neuro-fuzzy classifier for fault detection and classification in photovoltaic module,” in *Proceedings of 2016 8th International Conference on Modelling, Identification and Control, ICMIC 2016*, Jan. 2017, pp. 144–149, doi: 10.1109/ICMIC.2016.7804289.

[260] C. H. da Costa *et al.*, “A Comparison of Machine Learning-Based Methods for Fault Classification in Photovoltaic Systems,” Sep. 2019, doi: 10.1109/ISGT-LA.2019.8895279.

[261] K. Pahwa, M. Sharma, M. S. Saggu, and A. K. Mandpura, “Performance evaluation of machine learning techniques for fault detection and classification in PV array systems,” in *2020 7th International Conference on Signal Processing and Integrated Networks, SPIN 2020*, Feb. 2020, pp. 791–796, doi: 10.1109/SPIN48934.2020.9071223.

[262] N. Renewable Energy Laboratory, S. National Laboratory, S. Alliance, and S. National Laboratory Multiyear Partnership, “Best Practices for Operation and Maintenance of Photovoltaic and Energy Storage Systems; 3rd Edition,” 2018.

[263] J. P. Vargas, B. Goss, and R. Gottschalg, “Large scale PV systems under non-uniform and fault conditions,” *Solar Energy*, vol. 116, 2015, doi: 10.1016/j.solener.2015.03.041.

[264] E. Garoudja, F. Harrou, Y. Sun, K. Kara, A. Chouder, and S. Silvestre, “Statistical fault detection in photovoltaic systems,” *Solar Energy*, vol. 150, 2017, doi: 10.1016/j.solener.2017.04.043.

[265] N. Tyutyundzhiev, M. Petrov, F. Martínez Moreno, J. Leloux, and L. Narvarte Fernández, “Advanced testing of PV plants using portable SCADA tool,” *31st European Photovoltaic Solar Energy Conference and Exhibition*, no. 1, 2015.

[266] Rodrigo Moreton, “Contributions to uncertainty reduction in the estimation of PV plants performance,” 2016.

[267] H. Zhu *et al.*, “Study of joint temporal-spatial distribution of array output for large-scale photovoltaic plant and its fault diagnosis application,” *Solar Energy*, vol. 181, 2019, doi: 10.1016/j.solener.2019.01.083.

[268] M. Ammiche, A. Kouadri, L. M. Halabi, A. Guichi, and S. Mekhilef, “Fault detection in a grid-connected photovoltaic system using adaptive thresholding method,” *Solar Energy*, vol. 174, 2018, doi: 10.1016/j.solener.2018.09.024.

[269] A. Livera, M. Theristis, G. Makrides, and G. E. Georghiou, “Recent advances in failure diagnosis techniques based on performance data analysis for grid-connected photovoltaic systems,” *Renewable Energy*, vol. 133. 2019, doi: 10.1016/j.renene.2018.09.101.

[270] T. Mukai, S. Kawamoto, Y. Ueda, M. Saijo, and N. Abe, “Residential PV system users’ perception of profitability, reliability, and failure risk: An empirical survey in a local Japanese municipality,” *Energy Policy*, vol. 39, no. 9, 2011, doi: 10.1016/j.enpol.2011.05.019.

[271] K. Kiefer, N. H. Reich, D. Dirnberger, and C. Reise, “Quality assurance of large scale PV power plants,” 2011, doi: 10.1109/PVSC.2011.6186343.

[272] M. Fuentes, M. Vivar, J. M. Burgos, J. Aguilera, and J. A. Vacas, “Design of an accurate, low-cost autonomous data logger for PV system monitoring using Arduino™ that complies with IEC standards,” *Solar Energy Materials and Solar Cells*, vol. 130, 2014, doi: 10.1016/j.solmat.2014.08.008.

[273] S. Vergura, “Hypothesis tests-based analysis for anomaly detection in photovoltaic systems in the absence of environmental parameters,” *Energies*, vol. 11, no. 3, 2018, doi: 10.3390/en11030485.

[274] E. Koubli, D. Palmer, P. Rowley, and R. Gottschalg, “Inference of missing data in photovoltaic monitoring datasets,” in *IET Renewable Power Generation*, 2016, vol. 10, no. 4, doi: 10.1049/iet-rpg.2015.0355.

[275] S. Killinger *et al.*, “On the search for representative characteristics of PV systems: Data collection and analysis of PV system azimuth, tilt, capacity, yield and shading,” *Solar Energy*, vol. 173, 2018, doi: 10.1016/j.solener.2018.08.051.

[276] Y. M. Saint-Drenan, G. H. Good, M. Braun, and T. Freisinger, "Analysis of the uncertainty in the estimates of regional PV power generation evaluated with the upscaling method," *Solar Energy*, vol. 135, 2016, doi: 10.1016/j.solener.2016.05.052.

[277] A. Capozzoli, F. Lauro, and I. Khan, "Fault detection analysis using data mining techniques for a cluster of smart office buildings," *Expert Systems with Applications*, vol. 42, no. 9, 2015, doi: 10.1016/j.eswa.2015.01.010.

[278] A. Golnas, J. Bryan, R. Wimbrow, C. Hansen, and S. Voss, "Performance assessment without pyranometers: Predicting energy output based on historical correlation," 2011, doi: 10.1109/PVSC.2011.6186347.

[279] V. P. A. Lonij *et al.*, "Forecasts of PV power output using power measurements of 80 residential PV installs," 2012, doi: 10.1109/PVSC.2012.6318280.

[280] B. Elsinga and W. G. J. H. M. van Sark, "Inter-system time lag due to clouds in an urban PV ensemble," 2014, doi: 10.1109/PVSC.2014.6925029.

[281] N. A. Engerer and F. P. Mills, "KPV: A clear-sky index for photovoltaics," *Solar Energy*, vol. 105, 2014, doi: 10.1016/j.solener.2014.04.019.

[282] J. Leloux, L. Narvarte Fernández, A. Luna, and A. Desportes, "Automatic fault detection on BIPV systems without solar irradiation data," *29th European Photovoltaic Solar Energy Conference and Exhibition (EU PVSEC 2014) | 29th European Photovoltaic Solar Energy Conference and Exhibition | 22/09/2014 - 26/09/2014 | Amsterdam, Netherlands*, 2014, doi: 10.13140/2.1.2280.3200.

[283] I. Popovic and I. Radovanovic, "Methodology for detection of photovoltaic systems underperformance operation based on the correlation of irradiance estimates of neighboring systems," *Journal of Renewable and Sustainable Energy*, vol. 10, no. 5, 2018, doi: 10.1063/1.5042579.

[284] F. Harrou, A. Dairi, B. Taghezouit, and Y. Sun, "An unsupervised monitoring procedure for detecting anomalies in photovoltaic systems using a one-class Support Vector Machine," *Solar Energy*, vol. 179, 2019, doi: 10.1016/j.solener.2018.12.045.

[285] G. J. Ball, R. M. Hudson, and M. R. Behnke, "Recent application and performance of large, grid-connected commercial PV systems," 2002, doi: 10.1109/pvsc.2002.1190950.

[286] J. Leloux, L. Narvarte, A. Desportes, and D. Trebosc, "Performance to Peers (P2P): A benchmark approach to fault detections applied to photovoltaic system fleets," *Solar Energy*, vol. 202, 2020, doi: 10.1016/j.solener.2020.03.015.

[287] N. Haghdadi, J. Copper, A. Bruce, and I. MacGill, "A method to estimate the location and orientation of distributed photovoltaic systems from their generation output data," *Renewable Energy*, vol. 108, 2017, doi: 10.1016/j.renene.2017.02.080.

[288] S. Killinger, F. Braam, B. Müller, B. Wille-Haussmann, and R. McKenna, "Projection of power generation between differently-oriented PV systems," *Solar Energy*, vol. 136, 2016, doi: 10.1016/j.solener.2016.06.075.

[289] M. Journée and C. Bertrand, "Improving the spatio-temporal distribution of surface solar radiation data by merging ground and satellite measurements," *Remote Sensing of Environment*, vol. 114, no. 11, 2010, doi: 10.1016/j.rse.2010.06.010.

[290] M. Journée, R. Müller, and C. Bertrand, "Solar resource assessment in the Benelux by merging Meteosat-derived climate data and ground measurements," *Solar Energy*, vol. 86, no. 12, 2012, doi: 10.1016/j.solener.2012.06.023.

[291] IEC, "IEC 62446-3: Photovoltaic (PV) systems - Requirements for testing, documentation and maintenance - Part 3: Photovoltaic modules and plants - Outdoor infrared thermography." 2017.