Artificial Intelligence for Operation and Maintenance of PV Plants

Deliverable D2.3

Out of normality analysis report

Lead Beneficiary	ISOTROL
Delivery Date	31/03/2023
Dissemination Level	Public
Status	Released
Version	1.0
Keywords	Data Analysis; Normality Analysis; Data Cleansing; Machine Learning; Modelling; Digital Twin

Disclaimer

This work is financed by the ERDF - European Regional Development Fund through -the Operational Programme for Competitiveness and Internationalisation COMPETE 2020 under the Portugal 2020 Partnership Agreement within project AI4PV, with reference POCI-01-0247-FEDER-111936 – and Spain's Multi-regional Operational Programme 2014-2020. International collaborative project EUR 2020058 with the seal of the AI EUREKA CLUSTER.

Document Information

Project Acronym	AI4PV
Work Package	WP 2
Related Task(s)	T2.3
Deliverable	D2.3
Title	Out of normality analysis report
Author(s)	Miguel Angel Delgado (ISOTROL), Sergio Raigón (ISOTROL), Ricardo Morales (ISOTROL), Jose Garcia Franquelo (ISOTROL), Rubén González (ISOTROL), Louelson Costa (INESC TEC), Christian Verrecchia (EDP NEW)

Revision History

Revision	Date	Description	Reviewer
0.1	18 January 2023	Outline of report content	ISOTROL
0.2	22 February 2023	Full draft of full content	ISOTROL
0.3	17 March 2023	Partner inputs	EDP NEW, INESCTEC
0.4	24 March 2023	Second Version	ISOTROL
1.0	31 March 2023	Final version	

EXECUTIVE SUMMARY

This deliverable includes the main results obtained in the task **T2.3 Out of normality analysis at critical components through digital twin analysis** from the project AI4PV. The work carried out in this task has focused on the normality analysis in the different assets in PV plants and quantities related to the operation have been analysed.

The analysis has been based on the simulation of the critical components through the digital twin, enabling the detection of deviations from the expected values and the identification of the problems associated with these deviations. The difference between the expected data and the measured one has helped to quantify the losses corresponding to the failures or inefficiencies that may affect a PV plant.

The use cases studied in this deliverable include inverter stops and underperformances, clipping, soiling, string and stringbox disconnections, panel ageing as well as tracker blocking and inefficiencies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

Acronym	Meaning
AC	Alternating Current
AEP	Annual Energy Production
AI4PV	Artificial Intelligence for Photovoltaic
AOI	Angle of incidence
CAPEX	Capital Expenditure
CP	Change Point
DC	Direct Current
DHI	Diffused Horizontal Irradiance
DNI	Direct Normal Irradiance
dq	synchronous reference frame control
DT	Digital Twin
EVA	Ethylene Vinyl Acetate
KPI	Key Performance Indicator
ML	Machine Learning
MPPT	Maximum Power Point Tracking
O&M	Operations & Maintenance
PLL	Phase Locked Loop
PR	Performance Ratio
PSO	Particle Swarm Optimisation
PV	Photovoltaic
PVPP	Photovoltaic Power Plant
RCA	Root Cause Analysis
SC	Short Circuit
WP	Work Package

GLOSSARY OF KEY TERMS

Artificial Intelligence	Artificial intelligence is a wide-ranging branch of computer science concerned with building smart machines capable of performing tasks that typically require human intelligence.
Machine Learning	Machine learning is a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention.
Deep Learning	Deep learning is a subset of machine learning, which is essentially a neural network with three or more layers. These neural networks attempt to simulate the behaviours of the human brain—albeit far from matching its ability—allowing it to "learn" from large amounts of data.
Fault	A fault is an unpermitted deviation of at least one characteristic property (feature) of the system from the acceptable, usual standard condition.
Failure	Permanent interruption of a system's ability to perform a required function under specified operating conditions.
Malfunction	Intermittent irregularity in fulfilment of a systems desired function.
Fault detection	Determination of faults present in a system and time of detection.
Fault diagnosis	Determination of kind, size, location and time of detection of a fault by evaluating symptoms. Follows fault detection. Includes fault detection, isolation and identification.

1. INTRODUCTION

This document, deliverable **D2.3 Out of normality analysis report**, includes a summary of the results obtained in the process focused on the analysis of technologies, tools, mechanisms and methodologies for the management of data and modelling in the context of the project AI4PV.

1.1 SCOPE OF REPORT

This document focuses on the coverage of the range of use cases that are susceptible to be analised through comparison with the data obtained from the digital twin of the system once the parameters of the normal behaviour of the PV plant have been calculated.

First of all, some considerations about the concept of out of normality in a PV plant are discussed in order to clarify why a certain timestamp is treated as faulty for a given element of the plant at inverter level, which is the reference level that it is employed to classify points as good or faulty in our study. After that, the points classified as faulty are labelled with the specific problem detected by comparison with the previously adjusted parameters of the digital twin. This comparison is then used to estimate the losses linked to each incident. This will allow the development of a recommendation engine that prioritizes those actions which are affecting to a more severe extent the performance of the PV plant.

The results presented here have been achieved in the development of task **T2.3 Out of normality analysis report**, included in the context of the work package **WP2 Descriptive analysis of PV power plant components and operation**. These advancements, in addition to those achieved in the other tasks of the work package (T2.1 and T2.2) allow the generation of a system capable of modelling critical elements in PV plants and the management of all the data, from its acquisition to its preparation and storage.

1.2 OUTLINE OF REPORT

This report is structured as follows,

- ***Chapter 1*** introduces the scope of the document
- ***Chapter 2*** focuses on the considerations regarding out of normality identification
- ***Chapter 3*** provides a complete description of the use cases considered in the project
- ***Chapter 4*** introduces the DT-based out of normality analysis for the inverter
- ***Chapter 5*** presents the DT-based out of normality analysis for the transformer

2. OUT OF NORMALITY OPERATING CONSIDERATIONS

In the normality analysis, criteria is considered for whether assets are operating in their normal zone or whether they are in a failed or degraded situation.

In this regard, two types of "off-normal" operations are considered:

- **Failure:** when there is a fault that can be determined, usually leading to a shutdown of an element, i.e., the disconnection of an inverter.
- **Underperformance:** when the generation element is still operational but has a lower performance than expected, defined as "normal" operation, defined with respect to the reference curve, as indicated in D2.2 [1].

In the case of underperformance, using the reference curve for the consideration of normal operation presents unsatisfactory results, so a lower band is envisaged to consider that at a certain point in time the asset is operating in degraded mode. This band is usually set empirically at 2.5 times the standard deviation of the reference curve. Figure 2-1 shows the distribution of the operating points of an inverter in a solar plant where green points are those classified as "good" whereas the brown ones are those classified as "faulty".

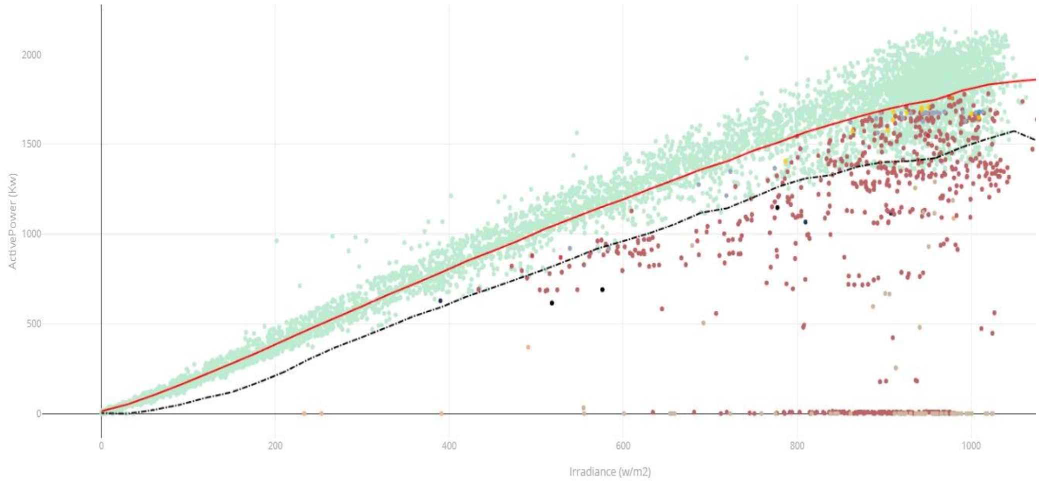


FIGURE 2-1: DISTRIBUTION OF THE OPERATING POINTS OF AN INVERTER IN A PV PLANT

2.1 DIGITAL TWIN BASED NORMALITY

The Digital Twin (DT) is a tool to simulate the response of a device under arbitrary different conditions based on the physical laws that govern processes involved in the functioning of the machine of interest. This is highly advantageous, since the only strictly necessary input is the data provided by the manufacturer and the resource data, instead of an extensive time series of data to train a model which could be subject to overfitting or to learn abnormal behaviours which might be hard to clean in a standard cleansing process, gaining not only in accuracy but also in processing time. In the case of a PV plant, the resource is the energy coming from the sun in the form of photons.

The DT is built as follows. Firstly, the parameters of the solar module and the inverter are determined with the help of the data recorded in a sufficient long period of time. For this task, an approach based on the Particle Swarm Optimisation (PSO) has been employed, which is a technique suitable for the nature of this problem. Once this phase is finished, the meteorological data is treated to obtain the expected energy output taking into account the Angle of Incidence (AOI) of the photons on the solar module, for which it is imperative to distinguish between a fixed tilt ground mount system and a tracking system in which the solar modules follows the movement of the sun up to a certain extent given by the specifications of the manufacturer of the tracker, whose movement is simulated by the DT for each time of the year. This, together with the components of the solar radiation such as the DHI and the DNI, is used to calculate the exact irradiance obtained by the solar modules at each timestamp. Another variable is the temperature of the cell, which can be modelled using the air temperature and solar module characteristics in the cases where a direct measurement is not available.

Secondly, the number of modules per string and per inverter, as well as the parameters fitted in the first step, are given as inputs to calculate the DC power, the DC voltage and the DC current per subsystem. These results enable us to detect string disconnections or module disconnections even when only string combiner data or string data, respectively, is available. For the inverter, the AC power is also calculated, so that an estimation of the efficiency of the inverter is quantified and thus compared with the one stated by the manufacturer in the specifications of the product. Therefore, an out of normality analysis can be performed.

2.2 REFERENCE CURVES DEFINITION AT INVERTER LEVEL

In the management of renewable energy assets, it is necessary to have an exhaustive control of certain performance indicators, which allow the manager to understand the situation of the asset and the actions to be taken to resolve any problems if necessary. In the case of solar PV, the indicators are closely related to the solar resource available or effective in the generation plant.

Perhaps one of the most important KPIs in this sense is the definition of the producible energy of our system. Obtaining the producible energy from the inverters and therefore from the photovoltaic panels associated with them allows us to make production estimates and detect deviations that translate into energy losses associated with the entire transformation process: from the correct absorption of irradiance by the panels, losses in the conductors due to the Joule effect, losses in the transformation from direct to alternating current and including all possible technical incidents that may occur.

The first approach to calculation is to obtain the theoretical energy based on the resource and take this as a reference for producible energy. This is achieved by plotting the historical AC power data coming from the inverter against the irradiance received by the solar modules associated with each inverter. The data used for this curve is previously cleaned with the *mahalanobis distance*, as in other parts of the analysis. Figure 2-2 shows an example of the data cleansing process, where outliers due to bad readings or sensors errors are eliminated for the analysis of the DT.

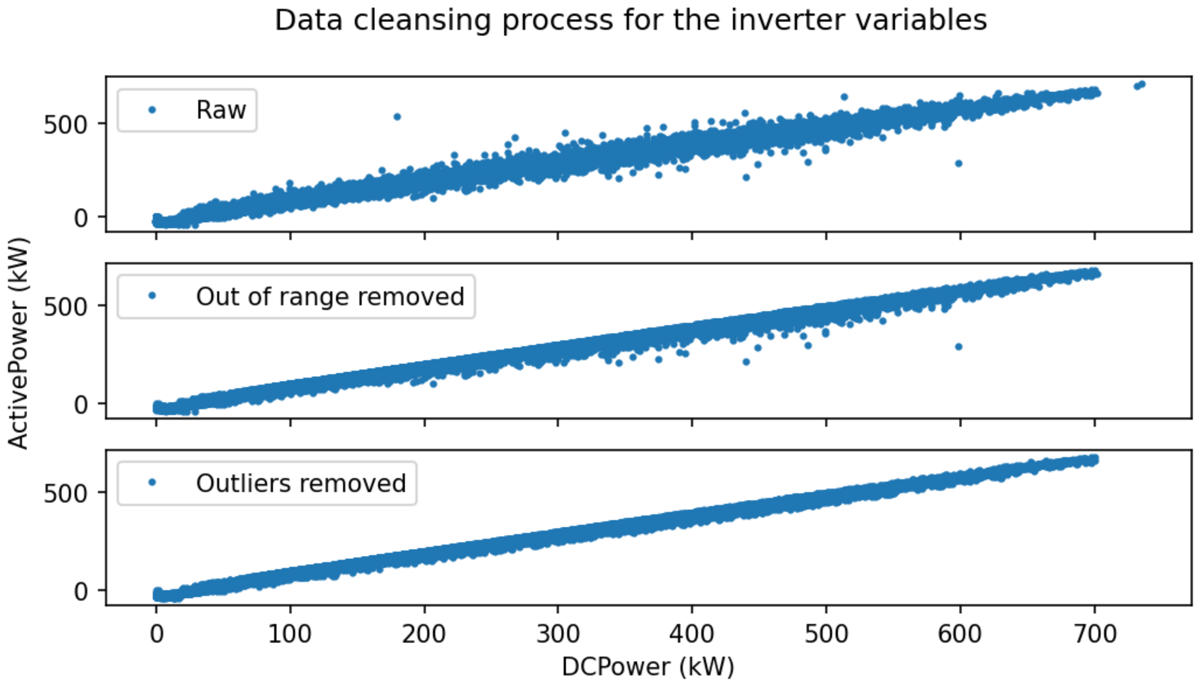


FIGURE 2-2: DATA CLEANSING PROCESS FOR THE INVERTER 1.1 OF THE PV PLANT LOCATED IN MONTE DAS FLORES, ÉVORA (PORTUGAL).

3. SPECIFIC ANALYSES OF OUT OF NORMALITY OPERATION

3.1 OUT OF NORMALITY OPERATION FOR THE PV PANELS

3.2.1 SOILING

A solar panel's performance depends mainly on weather conditions such as irradiance and cell temperature. Nevertheless, there are material and ambient conditions that can affect this performance. A material problem could be a cracked cell, a hot-spot or a panel's browning, in which the EVA film - used as an encapsulant in crystalline silicon modules - experiences a discoloration which leads to a power loss [2]. On the other hand, there are also pernicious ambient conditions such as pollution and dust accumulation, which leads to soiling losses. These losses can amount up to almost a third of the module's output [3].

The dust accumulation is supposed to follow a certain degradation pattern with respect to factors such as the time elapsed since the last rainfall or the dust concentration in the air susceptible to be analysed through a simulation model [3] [4] [5].

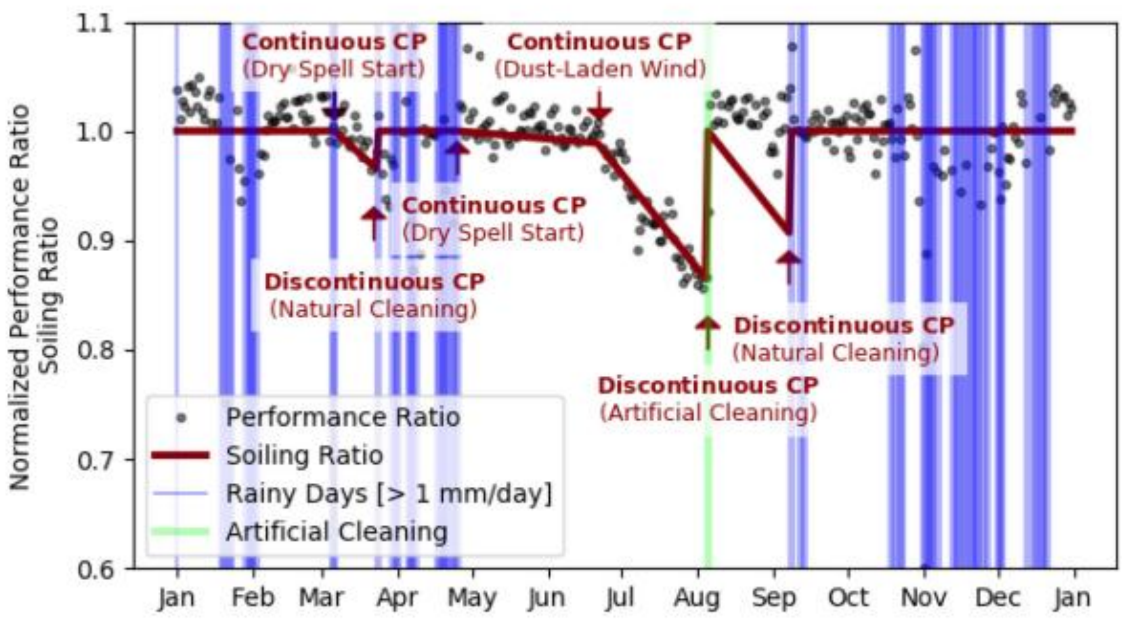


FIGURE 3-1: DAILY PERFORMANCE RATIO AND SOILING RATIO OVER A FULL YEAR

Figure 3-1 shows the trend of the Performance Ratio (PR), the change point (CP) accounts for the possible change in the soiling rate over time [4], such as natural cleaning, artificial cleaning, etc.

3.1.2 STRING DISCONNECTION

A string is a set of solar modules connected in series, so that the generated current is the same across all modules within the same string. There are two possible situations. The first one is when the current or DC power of the string is measured by a sensor. The second one is when only the current or DC power of the string combiner or inverter is measured.

In the first case, string disconnections are easy to detect. However, in the second one an adjustment must be done to determine which number of strings is compatible with the current or DC power detected. For this second one, a DT to simulate the number of working strings is required.

Figure 3-2 shows an example of string disconnection in a PV farm where the DC power of the strings is not directly measured. It can be seen as the DT with 13 strings is close to the real measurement.

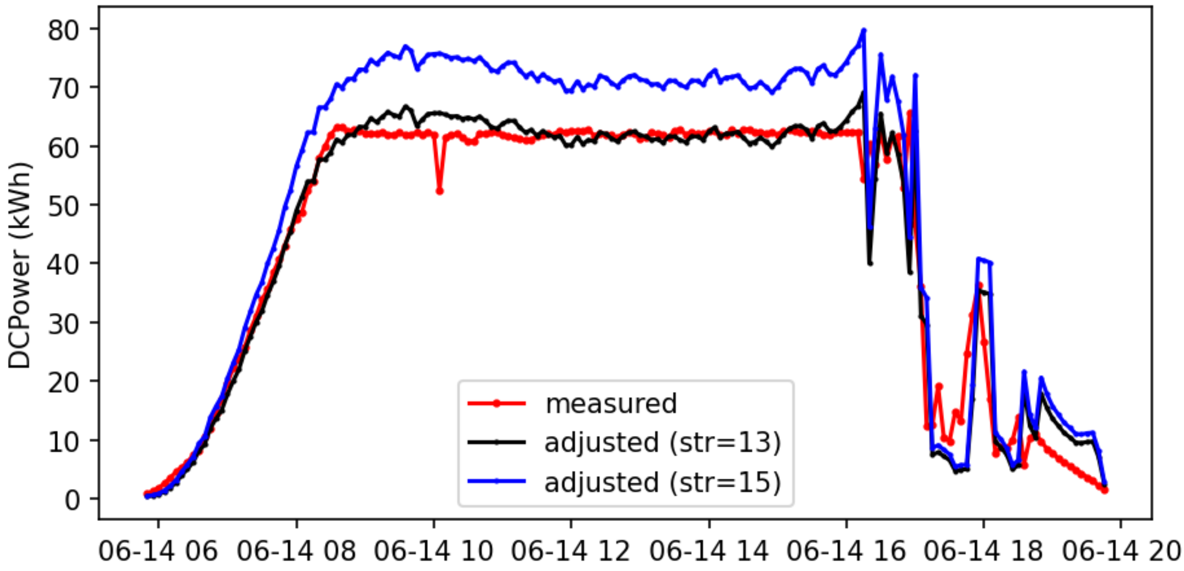


FIGURE 3-2: EXAMPLE OF STRING DISCONNECTION

3.1.3 STRINGBOX DISCONNECTION

A stringbox or string combiner connects a set of strings in parallel, so that the currents generated in each string are added together. This subsystem of a solar farm is usually equipped with sensors to measure magnitudes such as the current and the voltage and/or the DC power. Its detection is analogous to the string case.

3.1.4 TRACKER BLOCKING

A solar tracker is a system that allows solar modules to follow the movement of the sun throughout the day, up to a certain limit given by the manufacturer, but usually around 50 degrees, i.e. solar modules can vary their angle with respect to the sun's mean incident angle up to 50 or -50 degrees. This movement is adjusted for every day of the year, since the movement of the sun is predictable given a latitude and longitude. However, this system is far from being perfect. One of the main problems that affects these devices is known as "tracker blocking", in which the solar modules are stuck in a certain position for a long period of time. This diminishes the performance of solar modules since the direction of the sunlight is not as perpendicular to the modules as it could be.

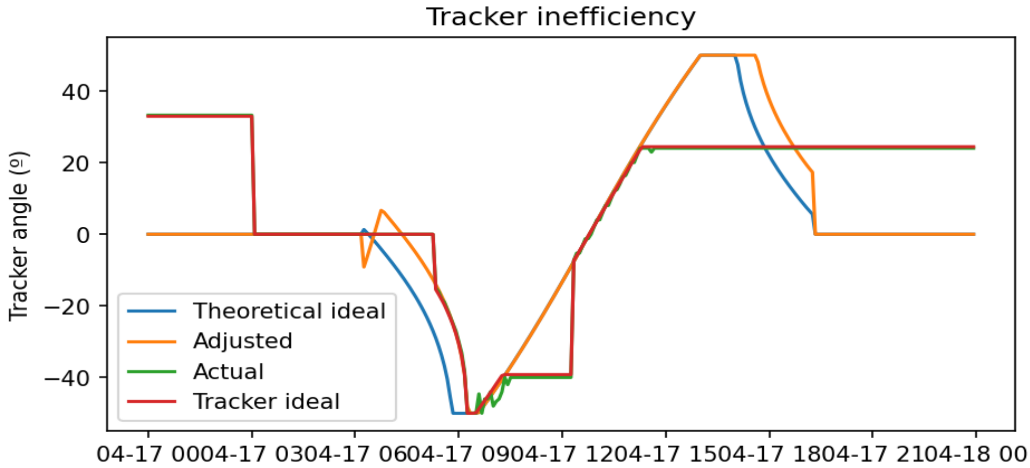


FIGURE 3-3: EXAMPLE OF TRACKER BLOCKING

Figure 3-3 shows an example of tracker malfunction. In blue, tracker angle obtained with the DT based on the orography of the terrain. In orange, tracker angle obtained fitting the tracker angle data. In green, the tracker angle data and in red, the tracker angle's setpoint as given by the device.

With the help of the DT we are able to precisely detect these problems, as shown in Figure 3-4. Even small inefficiencies are detectable since the DT reproduces the theoretical setpoint angle of the tracker given a certain time and a position on Earth.

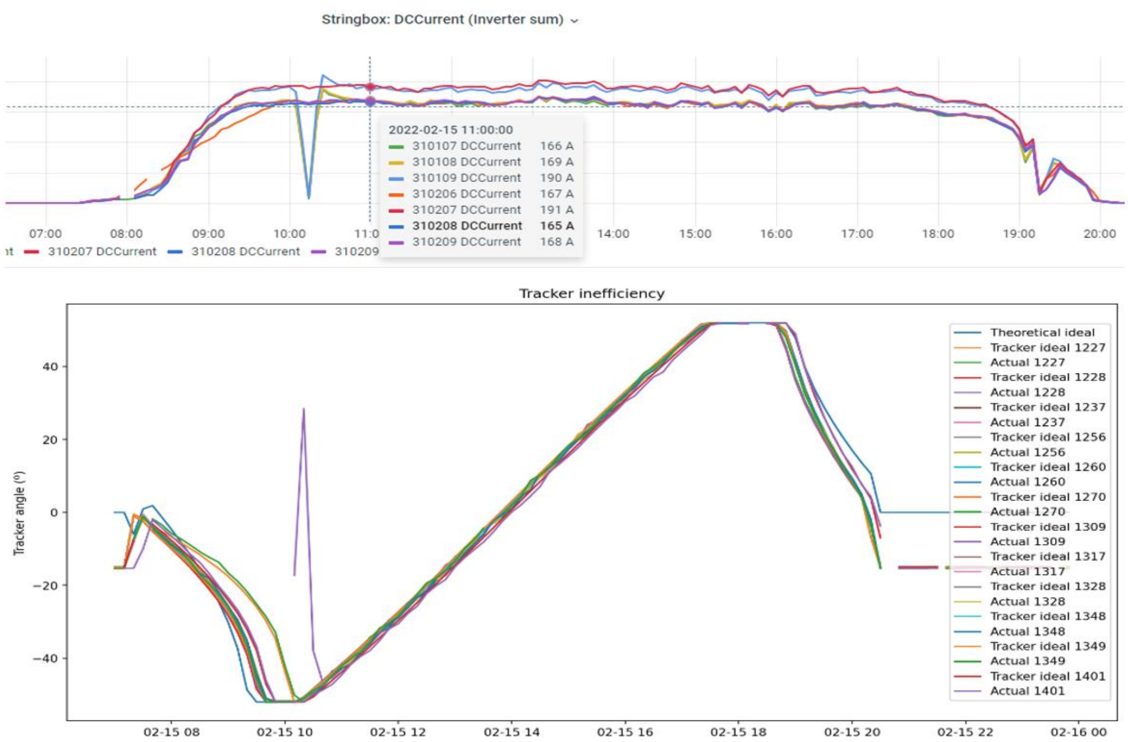


FIGURE 3-4: INEFFICIENCY DETECTED IN THE TRACKER ANGLE AND ITS FOOTPRINT IN THE DC CURRENT OF THE STRING COMBINER.

Apart from that, if the setpoint signal of the tracker is not reliable, the digital twin is able to detect if a tracker blocking is happening. This situation is depicted in Figure 3-5, where a tracker blocking was detected from 4pm onwards.

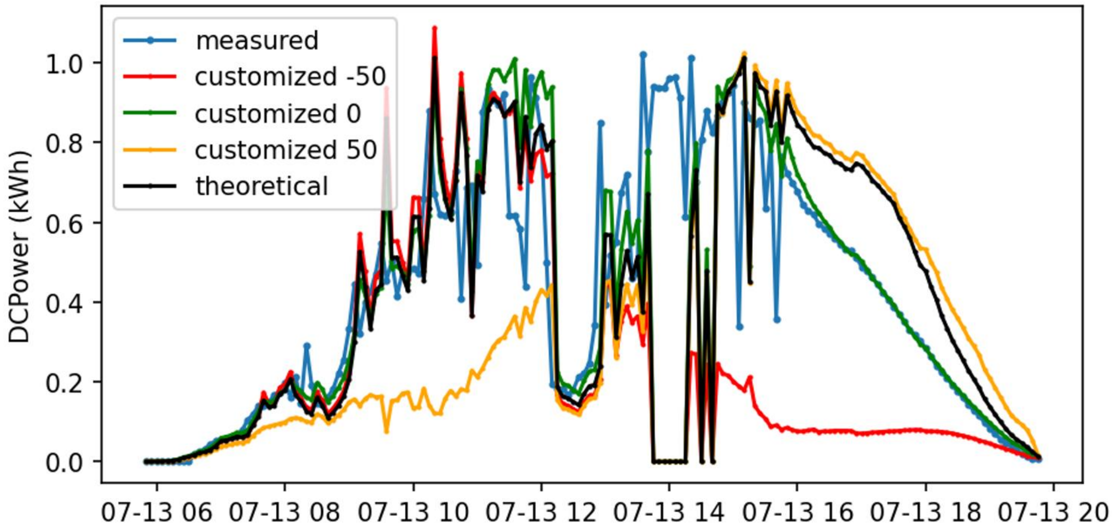


FIGURE 3-5: EXAMPLE OF TRACKER BLOCKING

3.1.5 TRACKER MISALIGNMENT

The tracker is positioned according to the orography of the terrain. We can distinguish between an angle which is parallel to the horizontal of the terrain and the perpendicular one. Based on those two angles, the tracker can be characterized. With the help of the DT, the position of the tracker can be studied to determine whether the tracker is positioned in its optimal position.

This analysis can be done by computing the R^2 of the different choices for the optimal and measured tracker angles throughout a year as shown in Figure 3-6.

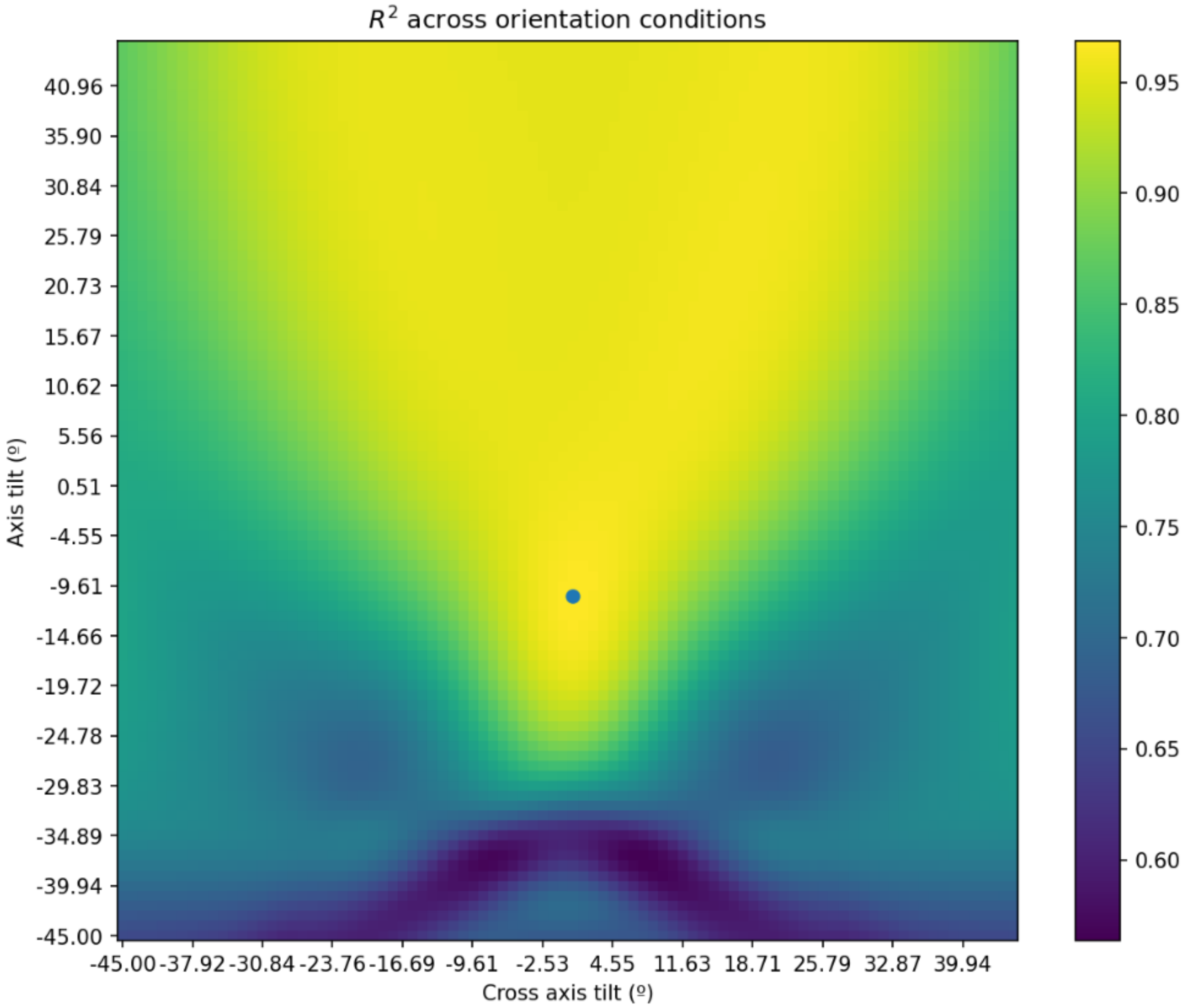


FIGURE 3-6: COLORMAP OF THE R^2 BETWEEN THE OPTIMAL AND REAL TRACKER ANGLE

3.1.6 PANEL AGEING

The solar module experiences a degradation with time specified by the manufacturer with the "linear performance warranty", which describes the expected degradation in the output energy of the module throughout its lifetime. This output can be studied both with the data collected by the plant's SCADA and with the DT, comparing the expected performance with the real one.

An example of this analysis is depicted in Figure 3-7.

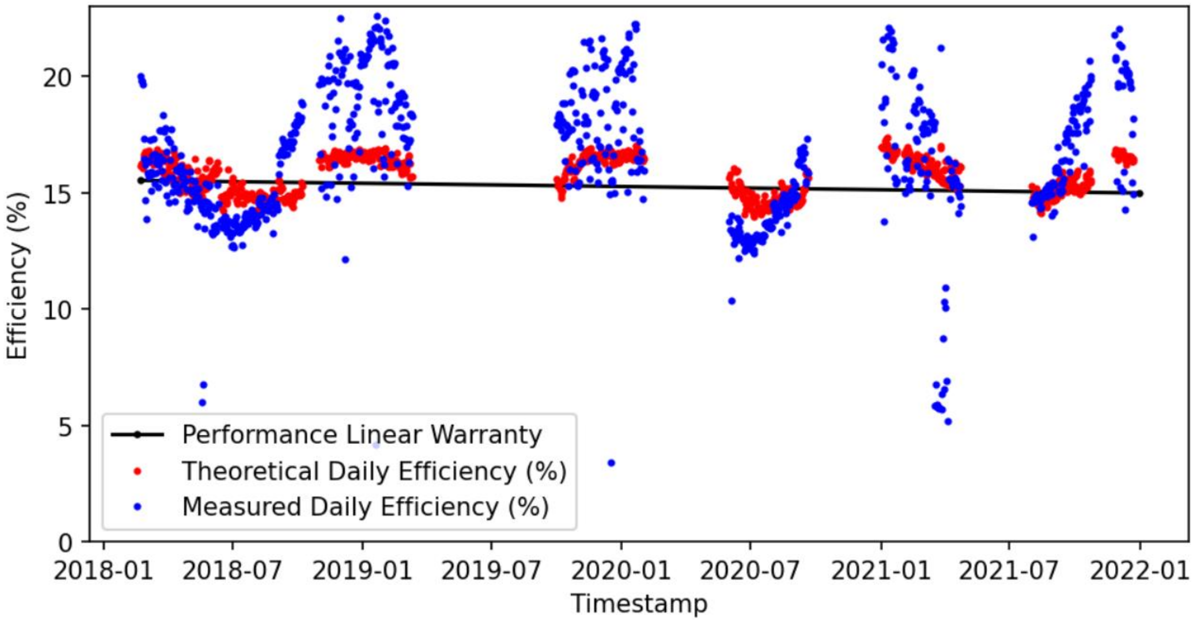


FIGURE 3-7: DECLINE IN THE EFFICIENCY OF ONE STRING

3.2 OUT OF NORMALITY OPERATION FOR THE INVERTER

3.2.1 INVERTER SHUTDOWN

Inverter shutdown occurs in situations such as an excessive internal temperature, too high or low voltage, current or grid frequency, unbalanced AC currents or grounding faults, among others. [6]. These events are usually noticed by an inverter alarm system, which indicates the fault code associated with the specific event that causes the inverter shutdown. Nonetheless, the data available for this project lacks such alarms. Thus, only a basic detection (without root cause in most cases) can be addressed. In order to overcome this problem and have a proper Root Cause Analysis (RCA) a different approach has been proposed in Section 4.

This type of accidents tends to be onerous since inverters usually manage an important part of the plant's resources. It is therefore critical to be able to detect and solve the root cause of these shutdowns as soon as possible.

Figure 3-8 shows an example of the power output for one inverter of the PV plant at stake, in case of normal conditions (left side) and shutdown (right side).

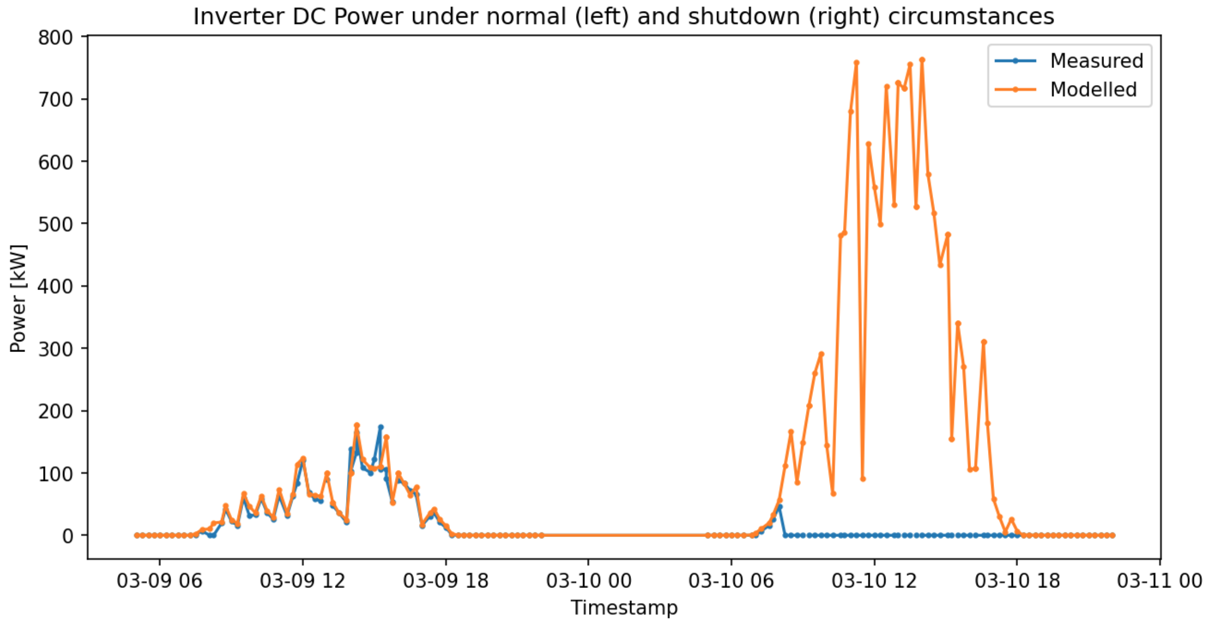


FIGURE 3-8: INVERTER DC POWER UNDER NORMAL CONDITIONS (LEFT) AND SHUTDOWN CONDITIONS (RIGHT)

3.2.2 TEMPERATURE DISCONNECTION

When the internal temperature of the inverter is too high, above a value which is specified by the manufacturer, the inverter shutdowns to protect itself from damage. Factors such as high air temperatures or ventilation problems can lead to a temperature above the threshold given in the datasheet and provoke a temperature disconnection.

In the time frame received from the plant located in Monte das Flores, no temperature disconnection incidents were detected.

3.2.3 MAINTENANCE STOP

Maintenance stops are those stops that are planned by the O&M operators to check the status or to substitute parts from the inverters. Since the decision whether or not to make a maintenance stop is arbitrary from the point of view of the routine of the inverter, only a detection can be addressed instead of a prediction of this unavailability.

In the time frame received from the plant located in Monte das Flores, no maintenance stops incidents were detected.

3.2.4 LATE AWAKENING

Another "shutdown" situation that may affect the inverter is what can be named as "late awakening". During such an event, the inverter does not start to work at all after sunrise, as one could expect.

An example of late awakening is reported in Figure 3-9.

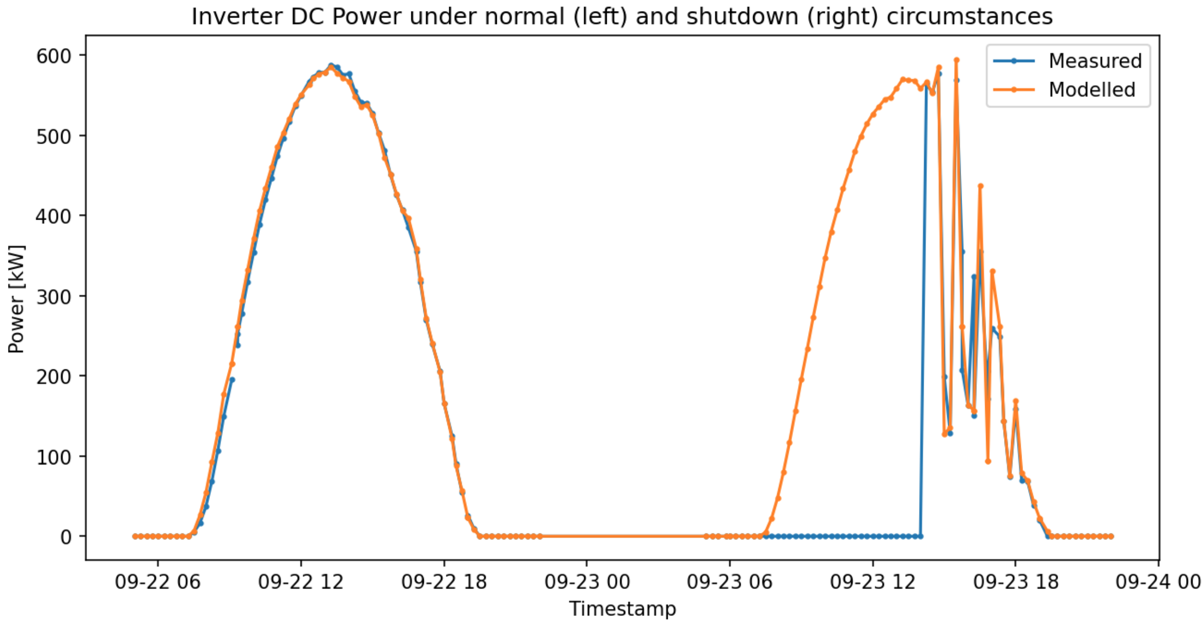


FIGURE 3-9: EXAMPLE OF LATE AWAKENING OF AN INVERTER

3.2.5 CLIPPING

An inverter is a device which is responsible for transforming DC power into AC power. Under low irradiance conditions, DC power generated by solar panels arrays is well below the maximum inverter's input if the installed rated DC capacity coincides with the input power that an inverter is able to manage. This causes an induced underperforming that can be mitigated by oversizing the DC power side, i.e., installing more panels whose total rated DC power (most of the times) is 10-20% higher than the inverter's maximum input. This way, more power will be generated during these low irradiance conditions (for example, in the mornings and evenings).

On the contrary, unused DC power will be lost during peak irradiance conditions, known as inverter clipping, or inverter saturation, since the inverter will be unable to transform the whole DC power received into AC power. Actually, lost power increases with the DC/AC ratio [7]. Apart from that, it does not compromise the normal behaviour of the inverter, but if the oversizing is too high, it may affect the reliability (understood as mean time between failures) and lifetime of the inverter [8]. Thus, the oversizing must be carefully designed in order to maximise the ROI.

In the day-to-day running of a solar power plant in production, the clipping events would be like the one depicted in Figure 3-10.

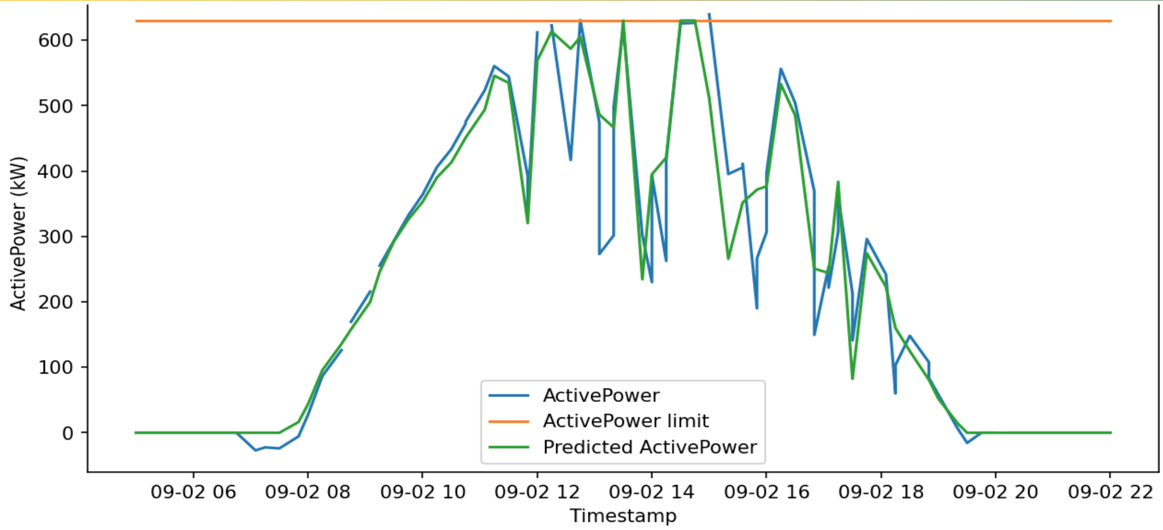


FIGURE 3-10: EXAMPLE OF CLIPPING EVENT

3.2.6 MPPT

The operating principle of an inverter is relatively simple, it transforms DC power into AC power. Nevertheless, depending on weather conditions (irradiance, air temperature, wind speed, etc) the DC current varies. Therefore, since DC power is current times voltage, this voltage must be adapted to the incoming current in order to obtain the maximum possible power from the panels. This can be seen in the following graph corresponding to a whole inverter and obtained with the digital twin of the plant under study in this project (725.4kW rated power at STC[1]).

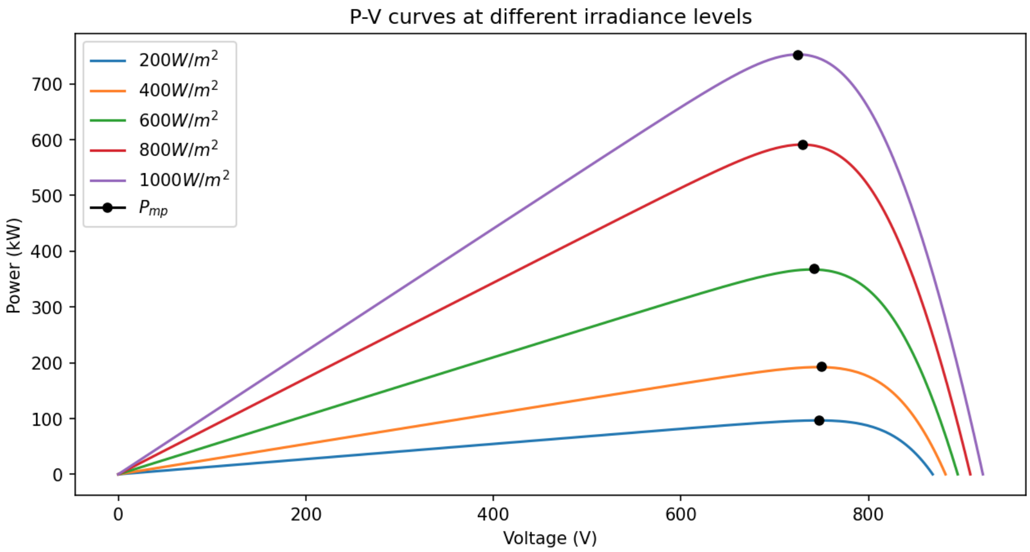


FIGURE 3-11: P-V CURVES AT DIFFERENT IRRADIANCE LEVELS OBTAINED THROUGH SIMULATION WITH THE DIGITAL TWIN OF THE PLANT.

[1] STC: Irradiance 1000W/m², Module temperature 25ºC, AM = 1.5.

When meteorological conditions change, the MPPT algorithm must readapt voltage, decreasing and increasing it until a local (ideally global) power maximum is achieved (the black point in Figure 3-11). It is not always easy to find the global maximum due to the presence of shadows or other problems in the panels [9].

As a consequence of weather's intrinsic variability and stochastic nature, it is fairly difficult for an inverter to work in the proper MPPT point.

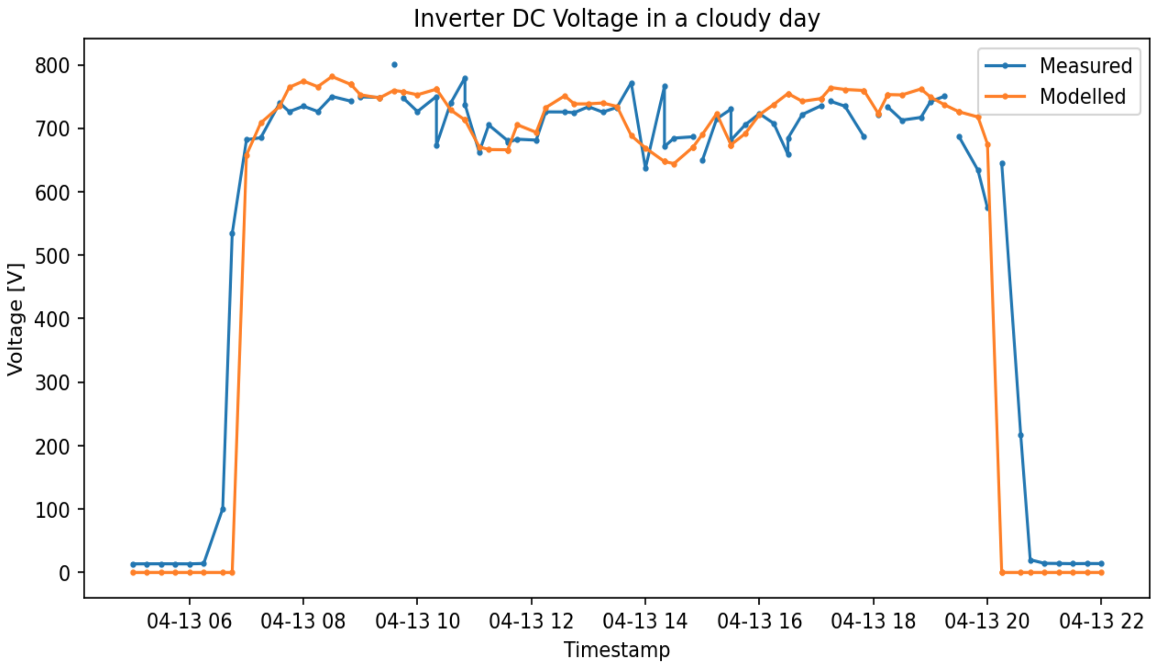


FIGURE 3-12: DC VOLTAGE OF THE INVERTER: A) IN BLUE THE MEASURED VALUE, B) IN ORANGE THE ONE COMPUTED BY THE DT

The out of normality analysis is the state in which the device is performing outside its regular behaviour. Usually, the signals are bounded in regions of normal operation. The further a point is from this normal region, the surer we are that the point is an outlier or a sign of a problem.

This type of analysis is implicitly employed when detecting a failure or underperformance. In those analyses, the signal given by the digital twin is compared with the actual one and then the incident is reported. Nonetheless, the technique used here is more specific for the cases in which the process of incident detection is more subtle, as in the case of the MPPT.

Figure 3-13 shows the out of normality analysis of the MPPT: the lighter the background colour, the greater the guarantee that it is an out-of-normality point.

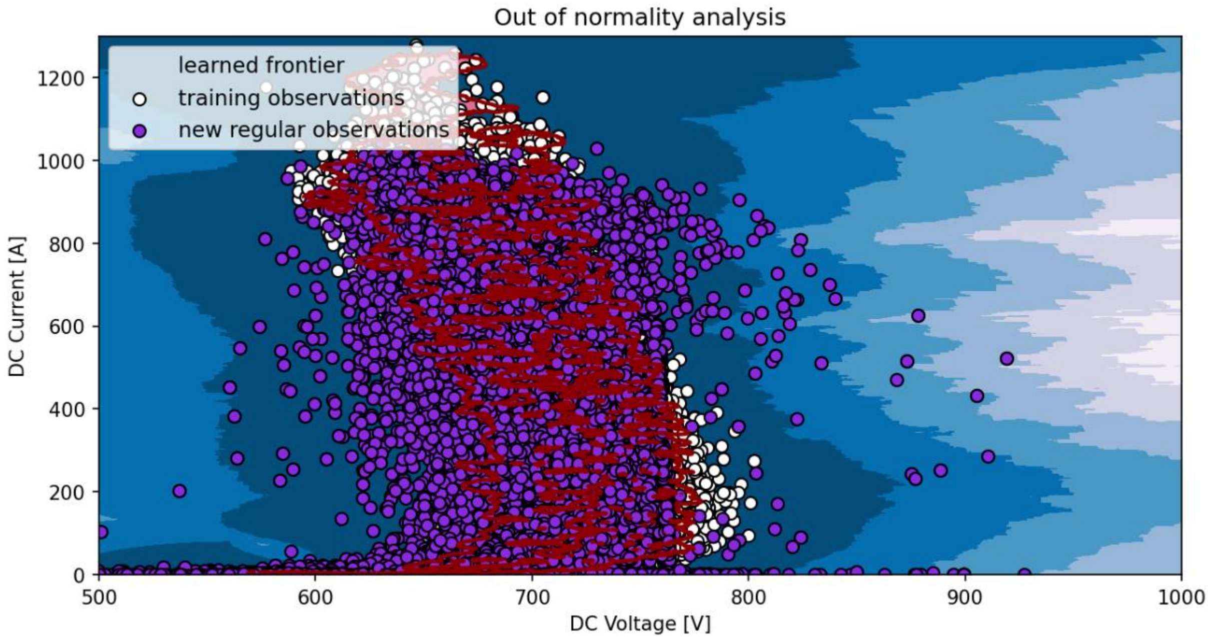


FIGURE 3-13: OUT OF NORMALITY ANALYSIS OF THE MPPT

3.2.7 SENSOR MALFUNCTION DETECTION & MEASUREMENT CORRECTION

If a sensor is not properly calibrated, an offset will appear in the measurements. This sort of sensor malfunction is easily detected, nevertheless, there are other more subtle problems. For instance, it can be considered the irradiance detected by the pyranometers of the plant under study, in which it is evident from the measurements plotted in Figure 3-14 that a sensor is not working as it is supposed to, deviating from its expected operational behaviour.

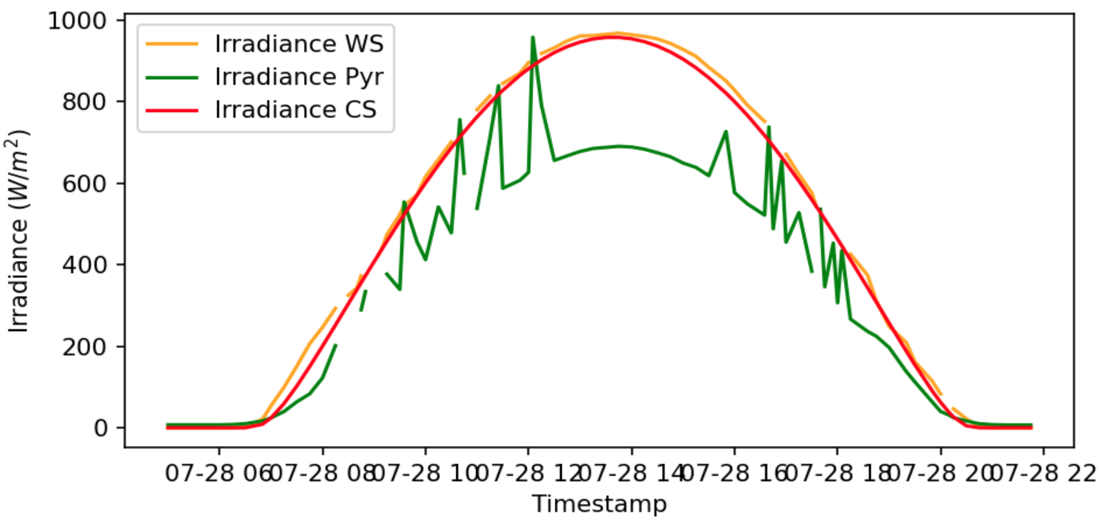


FIGURE 3-14: IRRADIANCE MEASUREMENTS: A) WEATHER STATION (WS), B) PYRANOMETER (PYR) AND THE MODEL OF IRRADIANCE UNDER CLEAR SKY CONDITIONS (CS) GIVEN BY DT

4. INVERTER DIGITAL TWIN-BASED OUT OF NORMALITY ANALYSIS

There are multiple faults and failures that can occur in a PVPP, from the bond-wiring of a transistor to the disconnection of a cable due to external factors. Nevertheless, this project searched for the most common faults and failures and implemented some of them. As the scope of this chapter/report is focused on the PV inverter, only faults directly related to the inverter are considered, thus faults in the PV modules or "before" the connection to the junction boxes, and faults in the transformer or "after" the connection to the inverter are not considered. Of course, both PV modules and transformers are very important elements of the system, so when they are simulated, they are operating under normal conditions. When a fault occurs, it is related to the junction boxes, DC cables, PV inverter and its switches, AC filter, or to the point of common coupling.

Given the lack of data for out-of-normality conditions, a DT based analysis was performed. In particular, the DT model of the inverter described in D2.2 [1], was used to simulate fault and failures of such component and generate synthetic datasets to train the AI algorithms presented in D3.1 [10].

All the faults that were implemented are triggered by a pre-programmed function that inserts the fault in the circuitry of the simulation. Thus, the inverter can have a fault happening all day long, or just under a given time, or be repeatedly applied and removed, etc. This flexibility highlights one of the advantages of developing the DT, as multiple scenarios can be investigated under multiple conditions.

As of the writing of this report, the faults and failures around the PV inverter are targeted at 9, being 4 fully implemented, 3 under evaluation and revision, ad 2 more planned for the final version of the PV inverter DT. Thus, there are 8 operation conditions to be presented in this report which are listed and summarized in Table 4-1. The short circuit to ground faults are still under investigation, and thus not included in this report.

TABLE 4-1: SIMULATED FAULTS SUMMARY

Fault name or condition	Fault acronym	Fault number
Fault-free	*noFault*	00
DC cable degradation	*dcCabDeg*	01
DC cable open-circuit	*DcCabOC*	02
Switch/transistor degradation	*switchDeg*	03
Switch/transistor open-circuit	*switchOC*	04
DC cable short circuit	*dcCabSC*	05

Phase-phase short circuit	*phphSC*	06
MPPT saturation	*MPPT*	07

The noFault condition is the same one that was used for the validation of the simulation described in D2.2 [1]. Thus, in this scenario, no fault is inserted. Besides validation, this simulation was the base for all the others simulations' faults.

The dcCabs faults (Deg, OC and SC) are very common and can happen by electrical stress (repeatedly overvoltage and overcurrent), weather conditions, ageing, etc. The degradation can easily evolve into an open-circuit, as degradation is commonly modelled as series resistance, such resistance will dissipate power through the Joule effect, worsening the degradation that generated it in the first place, and, eventually, leading to the cable disconnection. In a similar way, a short circuit in the DC cable can happen due to bad soldering/assembling in the junction boxes, poor DC cable insulation, etc. The overcurrent generated by such short circuit is very harmful to the system, which may lead to other faults in the cable, including degradation, or even propagating the issue back to the PV modules, causing mismatch faults or overcurrent. Such faults are illustrated in Figure 4-1. The strategies for the dcCabs faults implementation are:

- *Degradation*: series connection of resistance between the connection of a given junction box and the PV inverter.
- *Open circuit*: sudden circuit branch opening between the junction box and the inverter, dropping the current to zero and generating a transient response of the system.
- *Short circuit*: a connection between the positive and negative poles of a junction box by a small resistance.

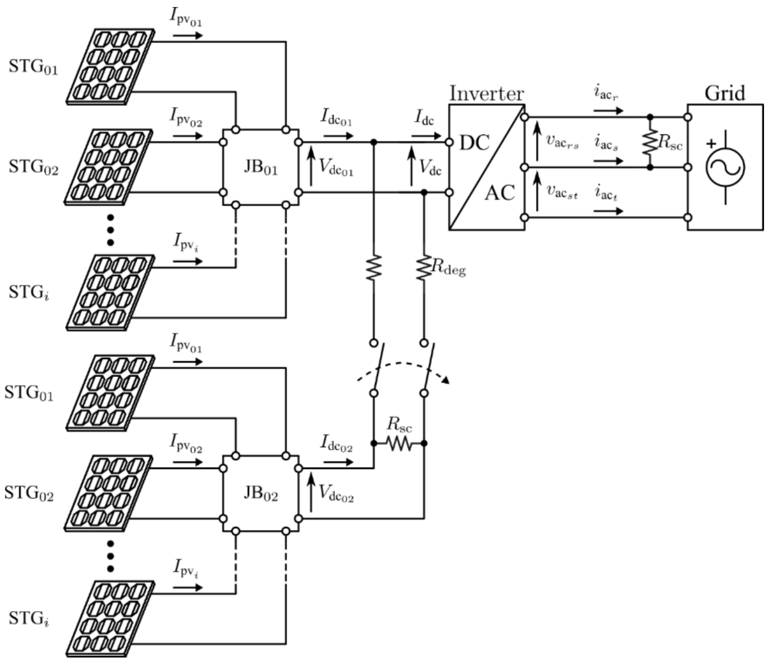


FIGURE 4-1: DC CABLE FAULTS AND PHASE-PHASE SC FAULT ILLUSTRATION

The switch faults are intrinsic to the inverters. There are multiple causes that may lead to such failures, from wire bonding failure to poor switching commands/gating signals. Nevertheless, being a problem with the inverter, those type of faults are perceived across the whole system. A switch failure will cause an unbalanced current generation by the inverter, those inverters are usually three-phase inverters (at least the one under study), thus, such unbalance will cause an asymmetrical stress on the AC-side components (AC filters, transformer, point of common coupling), whilst the DC-side will have a DC-link voltage with a larger oscillation. Under normal condition, the DC-link voltage oscillation should be minimal, as a function of DC-link capacitance and switching frequency, however under an unbalanced condition, the AC-side current harmonics will be reflected in the DC-link voltage, increasing their ripple, which results in performance worsening of the control (MPPT, thus, dq control), stress over the DC-link components (capacitor, cables, junction boxes terminals, etc.). In that sense, it can be noticed that a problem within the inverter have multiple implications across the whole system. As the inverter is one of the components most prone to failure, and the one that causes the higher production losses when underperforming, the switch faults are a delicate area of study with great impact in the condition-based monitoring. Such faults are illustrated in Figure 4-2. The strategies for the switch faults implementation are:

- <u>*Degradation:*</u> increase of the ON resistance of the switches, which are usually in the range of mΩ to the range of hundreds of mΩ.
- <u>*Open circuit:*</u> the gating signal of the switch under failure is clamped at zero, i.e., the switch will be always OFF.

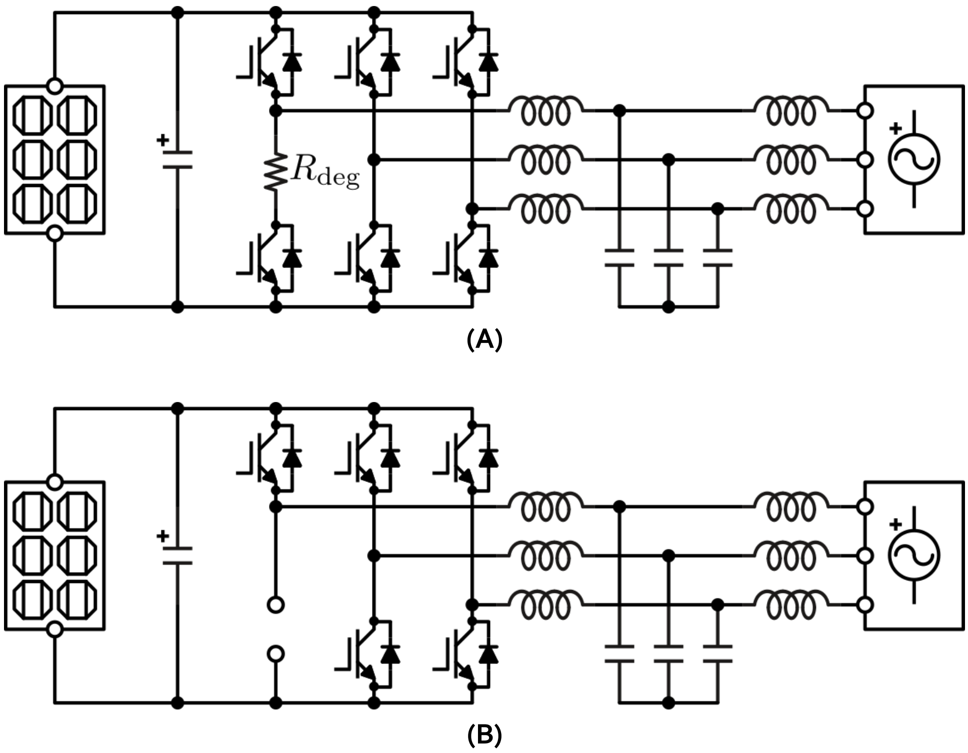


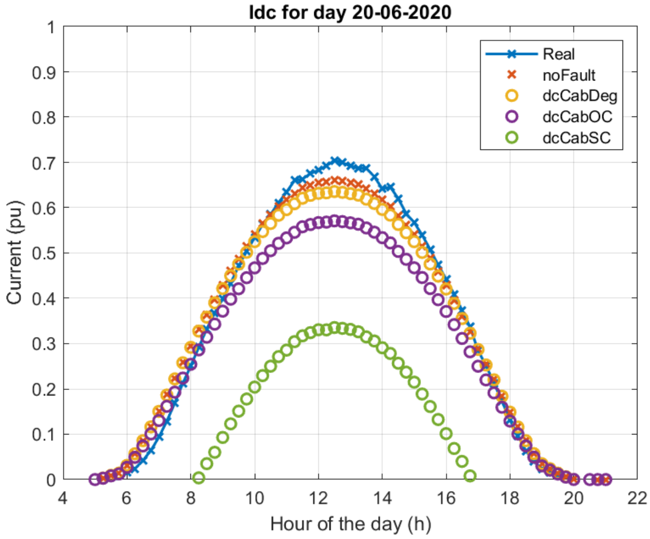
FIGURE 4-2: SWITCH FAULTS REPRESENTATION: (A) DEGRADATION; AND (B) OPEN CIRCUIT.

The MPPT fault represents the subset of faults related to the control, i.e., the firmware of the inverter. Within the scope of the project, it is not trivial to access such faults, however, they are very common and can have multiple causes. A misreading of a sensor may lead to an MPPT operating under the wrong region, a poor Phase Locked Loop (PLL) performance may also affect that, or even a poor tuning of the MPPT controller may cause issues related to it. The consequence of a malfunction of the MPPT algorithm is lower power production, as the whole system will be operating under a non-ideal region. Another challenge is that in the configuration of the PVPP under study in this project, the central inverter "sees" the multiple PV modules as a single large PV module, i.e., there is a single MPPT algorithm for the whole PVPP. As the multiple modules are not identical, the soiling is not identical, the irradiance over them is not the same, etc., it is expected that the MPPT does not have an ideal performance, even though it should be as best as possible. In this project, the MPPT fault selected is the saturation one, i.e., most of the time it will not present the best performance, but at some given combinations of irradiance and temperature, the MPPT may behave as expected under a coincidence circumstance. The strategy for the MPPT fault implementation is to saturate the reference voltage generated by the algorithm at its maximum value.

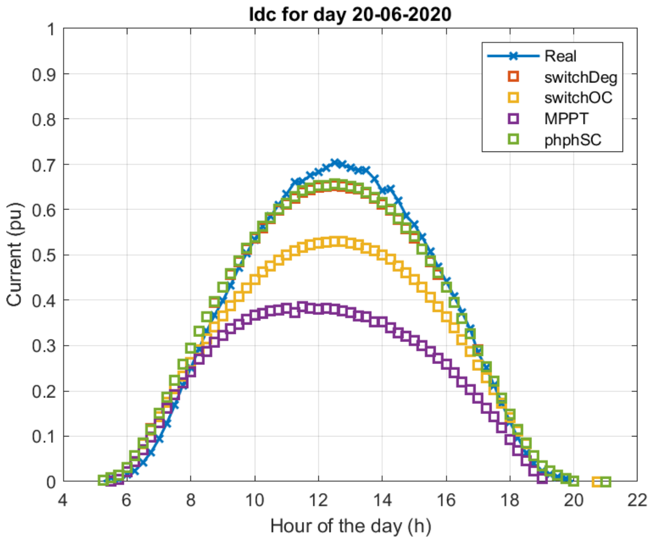
The phphSC is a classical short circuit between two phases of the AC side of the inverter. They are vastly studied in the power system community and are common in PVPP, thus they were included. Such problems are external to the inverter, but have similar consequences as the switch failures, as the unbalancing of the three-phase inverter is a critical condition. They will result in a larger ripple on the DC-side, reducing MPPT efficiency, increasing the stress over the DC-link capacitors, etc., whilst causing overcurrent on the switches, which may lead to the complete failure of the switch, i.e., the switchOC condition. The strategy for the phphSC fault implementation is to suddenly connect two phases of the AC side of the inverter through a low resistance path.

By using those strategies, the dataset containing the real fault-free data and the faulty data can be assembled. The dataset format is of a time-series, depicting the behaviour of the PV inverter during days, weeks, months, years, etc.

The DC current, i.e., the current generated by the PV modules and the input current of the inverter, are shown in Figure 4-3. It can be noticed that for most of the measurements, up to 0.3 pu of the DC-link current there is little difference between them, except for the dcCabSC. However, under higher irradiance levels, i.e., close to the middle of the day, the difference in the DC-link current across multiple scenarios is more discernible. This is an expected behaviour, as the patterns should be easier to recognize the closer the PV inverter is to its nominal value. However, the DC-link current alone may not be enough to distinguish the multiple scenarios, and that's why other measurements are needed to be considered when feeding the ML algorithms.



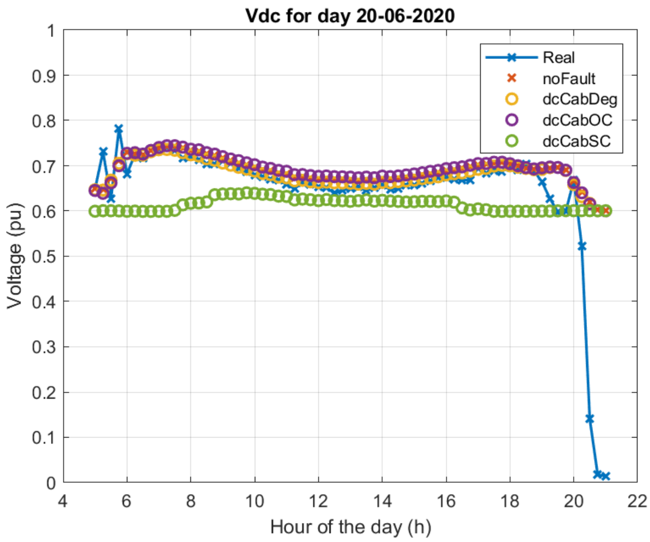
(A)



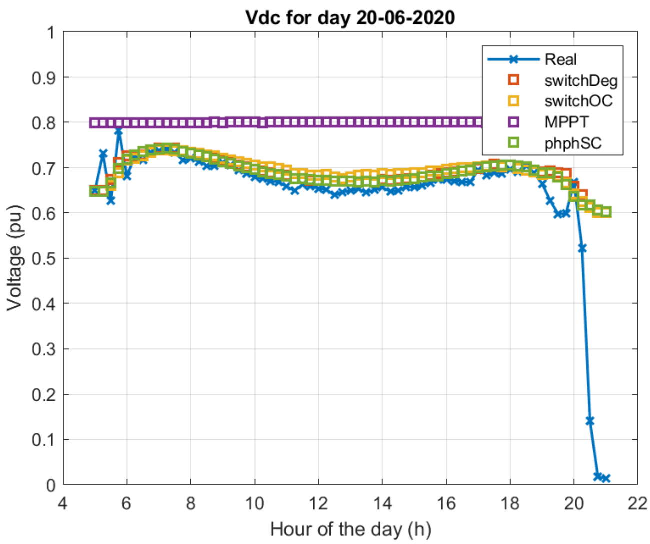
(B)

FIGURE 4-3: TIMESERIES DATA FOR THE PVPP DC-LINK CURRENT FOR THE DAY 20-06-2020

Still, on the DC side, the DC-link voltage is a good indicator of the control algorithm, which can be seen in Figure 4-4. The reference DC-link voltage is directly generated by the MPPT algorithms, thus under an MPPT fault within the scope of this project, it can be noticed that the DC-link voltage is clamped at its maximum allowed value. On the other hand, under a dcCabSC, the voltage is at its lowest, as the Short-Circuit (SC) in one of the junction boxes is literally a load to the system, thus part of the power is not being supplied to the grid, but rather being consumed by the SC. As for the other faults, they fluctuate around the fault-free voltage, since they will be a DC-link voltage with a high ripple, so their mean value may be a little higher or lower, depending on the scenario. The DC-link voltage presents some other interesting patterns, but it still might not be enough for the ML algorithms, so more features (or measurements) must be considered.



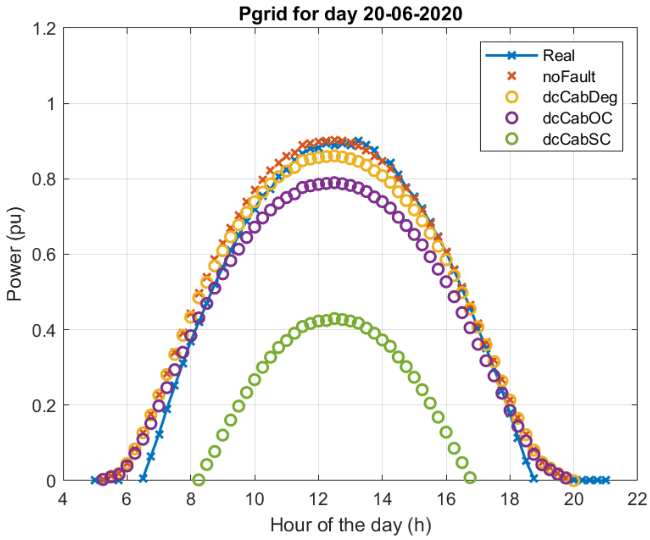
(A)



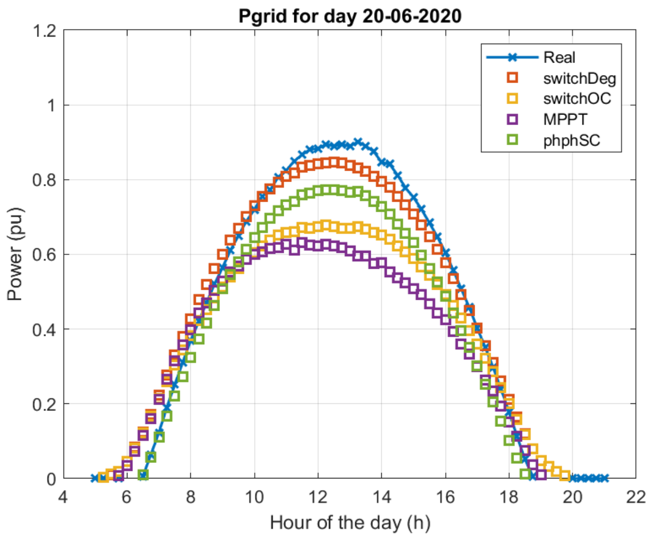
(B)

FIGURE 4-4: TIMESERIES DATA FOR THE PVPP DC-LINK VOLTAGE FOR THE DAY 20-06-2020

Regarding the AC power, or the power supplied to the grid, Figure 4-5 demonstrates that they are slightly more distinguishable from the DC-link current and voltage, as the AC power is the culmination of the whole process happening inside the inverter. The levels of power supplied to the grid for each condition shows that this would be a good lead or a good feature for the ML algorithms, however, they still fall under the problem of the low irradiance range, where the less irradiance, the less distinguishable they are. Even though in the right track, more measurements are needed to be added to the list of features, so the classification algorithms can properly detect the patterns of each operating condition.



(A)



(B)

FIGURE 4-5: TIMESERIES DATA FOR THE PVPP AC POWER FOR THE DAY 20-06-2020

It can be noticed that some of the conditions present a distinct pattern from the others. On the other hand, some conditions have similar behaviours. That is why multiple other features (currents, voltages, power, temperatures, irradiance, etc.) are used for the ML algorithms. The algorithms are trying to identify a pattern within the data of the PVPP, analysing all the measured variables (which are the features of the algorithms) to correctly classify the condition of the PVPP.

5. POWER TRANSFORMER DIGITAL TWIN-BASED OUT OF NORMALITY ANALYSIS

Due to the lack of data on out-of-normality conditions, a synthetic dataset was created based on the DT explained in D2.2 [1]. In particular, the following faults and failures were simulated:

- *Open circuit:* sudden circuit branch opening on one or multiple phases of the transformer windings, dropping the current of that (those) phase(s) to zero.
- *Short circuit:* a connection between one (or more) phase(s) to ground or between two (or three) phases by a neglectable resistance.

All the possible combination among these faults were simulated and the generated data was stored in order to train the AI algorithms for fault detection and classification, as explained in D3.1 [10].

An example is reported in Figure 5-1, where the current of a faulty phase is displayed for the day 20-06-2020. It can be noticed in the case of short circuit that the current of the faulty phase reaches values well above the nominal current of the transformer, whilst in the case of open circuit the current that flows in that phase is really low. It can also be observed that the fault-free value is quite close to the real one, proving the effectiveness of the proposed model.

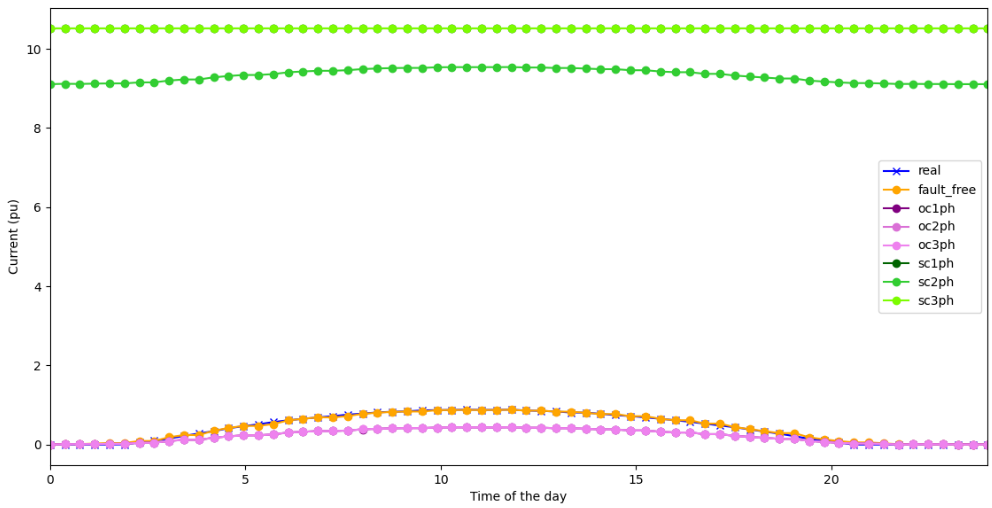


FIGURE 5-1: TIMESERIES DATA FOR THE TRANSFORMER HIGH-VOLTAGE WINDING CURRENT FOR THE DAY 20-06-2020

Similarly, Figure 5-2, shows the voltage on the faulty phase. Whilst in the case of open circuit, the voltage of the faulty phase is around its nominal value, it drops to lower values in the case of short circuit due to the high voltage drop induced on the windings by the high short circuit current.

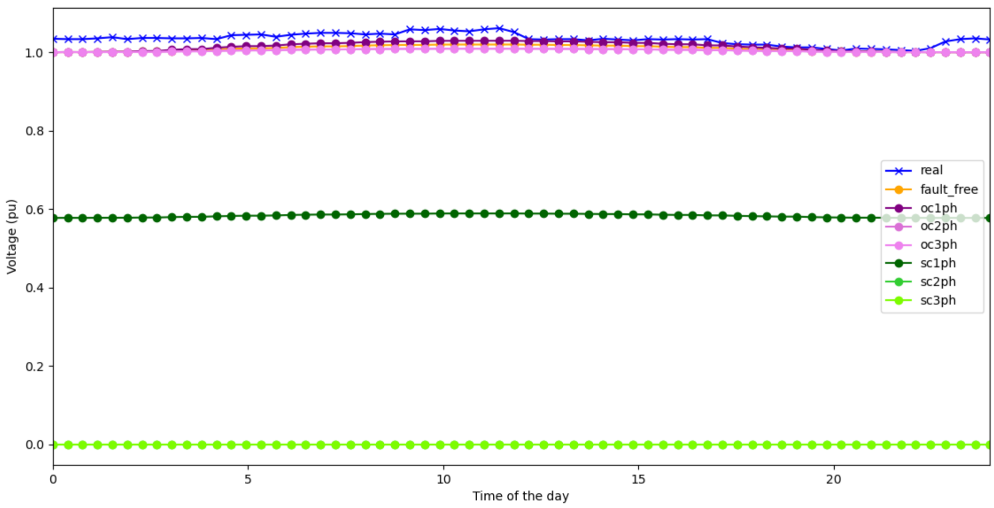


FIGURE 5-2: TIMESERIES DATA OF THE VOLTAGE FOR THE TRANSFORMER HIGH-VOLTAGE WINDING FOR THE DAY 20-06-2020

6. CONCLUSIONS

This report includes the use cases covered in the project AI4PV with the digital twin. It contains a brief description of the different cases and the solution offered by the digital twin in each of them considering the peculiarities and diversity of situations that may apply. Sometimes, it has been necessary to use data from other solar farms to develop the possibilities of the digital twin. One example would be the tracker system, which is not a feature available in the solar farm of the project.

The out of normality analysis begins with the simulation of the elements of the solar farm. After that, the data generated by the digital twin and the data collected in the solar farm is compared to study possible deviations. Once these deviations are detected, the specific failures or underperformances are addressed based on its characteristics. The entire process is structured in functions that sequentially analyse the data.

The analysis completed in this deliverable is the basis for the next stage of the project, in which a recommendation engine is developed to optimize O&M in PV utility-scale plants. The digital twin is employed then as a reference to establish the loss associated with each failure or underperformance, providing a method to precisely calculate the aforementioned loss.

7. REFERENCES

[1] Miguel Angel Delgado (ISOTROL), Sergio Raigon (ISOTROL), Ricardo Morales (ISOTROL), Jose Garcia Franquelo (ISOTROL), Rubén González (ISOTROL), Christian Verrecchia (EDP NEW), Louelson Costa (INESCTEC), D2.2 - Data management and modelling tools, AI4PV project.

[2] Dolia, K., Sinha, A., Tatapudi, S., Oh, J., & TamizhMani, G, "Early Detection of Encapsulant Discoloration by UV Fluorescence Imaging and Yellowness Index Measurements.," in *IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC)*, 2018.

[3] Dehghan, M., Rashidi, S., & Waqas, A., "Modeling of soiling losses in solar energy systems.," *Sustainable Energy Technologies and Assessments,* 2022.

[4] Micheli, L., Theristis, M., Livera, A., Stein, J. S., Georghiou, G. E., Muller, M., Almonacid, F., & Fernández, E. F. , "Improved PV Soiling Extraction Through the Detection of Cleanings and Change Points," *IEEE Journal of Photovoltaics,* vol. 11, no. 2, pp. 519-526, 2021.

[5] Kimber, A., Mitchell, L., Nogradi, S., & Wenger, H., "The Effect of Soiling on Large Grid-Connected Photovoltaic Systems in California and the Southwest Region of the United States," *2006 IEEE 4th World Conference on Photovoltaic Energy Conference,* pp. 2391-2395, 2006.

[6] Pani, B. B., Giri, N. C., Nayak, S. R., & Mishra, S. P, "Fault Detection and Troubleshooting in a PV Grid-Tied Inverter," *Indian Journal of Science and Technology,* vol. 14, no. 22, pp. 1829-1838, 2021.

[7] Good, J., & Johnson, J. X., ". Impact of inverter loading ratio on solar photovoltaic system performance," *Applied Energy,* vol. 177, pp. 475-486, 2016.

[8] Sangwongwanich, A., Yang, Y., Sera, D., Blaabjerg, F., & Zhou, D., "On the Impacts of PV Array Sizing on the Inverter Reliability and Lifetime," *IEEE Transactions on Industry Applications,*, vol. 54, no. 4, pp. 3656 - 3667, 2018.

[9] Wang, M., Zhang, X., Lian, X., Duan, Z., & Yang, J., "Design of MPPT algorithm under partial shadows," *2011 International Conference on Electric Information and Control Engineering,* pp. 309-312, (2011)..

[10] Louelson Costa (INECTEC), Ana Silva (INESCTEC), Christian Verrecchia (EDP NEW), D3.1 - Models for root-cause analysis with data analytics, AI4PV project.