Investigating linear and non-linear degradation rates using Combined degradation and soiling (CODS) in a utility-scale PV system

by Vincent Simensen

Thesis submitted for the degree of
Master in Renewable Energy Systems
60 credits

Department of Technology Systems
Faculty of mathematics and natural sciences

UNIVERSITY OF OSLO

Spring 2023

Investigating linear and non-linear degradation rates using Combined degradation and soiling (CODS) in a utility-scale PV system

by Vincent Simensen

Investigating linear and non-linear degradation rates using Combined degradation and soiling (CODS) in a utility-scale PV system

http://www.duo.uio.no/

Printed: Reprosentralen, University of Oslo

Abstract

Degradation rates in photovoltaic (PV) systems are generally assumed to be linear and can be estimated from calculated performance data. The assumption of linearity has been challenged by methods allowing for non-linear degradation rates. Therefore, a closer look at how the estimated degradation rate changes depending on the performance time series for a utility-scale PV system is of interest. In this work, a recently proposed algorithm called Combined Degradation and Soiling (CODS) was used on performance data from a real-utility-scale PV system. 7 years of meteorological and energy time series data for a 60 MWp utility-scale PV power plant was analyzed with the aim of investigating the assumption of a linear degradation rate for a PV system, by comparing different sizes of calculated performance time series. A selection of thresholds and filters were applied to raw data before calculating the temperature corrected performance ratio (PR'_{STC}) for 4500+ stringsets with a 10-minute resolution. The PR'_{STC} data was aggregated to daily stringset values $PR'_{str,d}$, and the median of all stringsets was used as a representation of the system $PR'_{sys,d}$. CODS was used for an arbitrarily chosen $PR'_{str,d}$ and for $PR'_{sys,d}$. A data shift was identified, possibly from sensor calibration or change, and corrected for. A renewed CODS analysis was performed for the arbitrarily chosen $PR'_{str,d}$ and $PR'_{sys,d}$, which estimated a degradation rate D_d of -0.411 %/year and -0.177 %year, respectively. A rolling CODS analysis was then performed on the same two time series, with a 3 year time window and an increment of 1 month, on the the two different time series by adjusting for the estimated seasonal component SC_d from the CODS results. The estimated D_d varied between -1.306 %/year and 1.300 %/year for the seasonally corrected $PR'_{str,d}$, and between -0.674 %/year and 0.773 %/year for the seasonally corrected $PR'_{sys,d}$. Finally. a selection of 10 stringsets where each was from a different inverter, was analyzed with the rolling CODS method. The estimated D_d for the 10 selected $PR'_{str,d}$ varied between -1.309 %/year and 1.300 %/year. By comparing the results from using CODS with the rolling CODS method, the estimated D_d was shown to vary depending on which years were analyzed. This demonstrated that a linear degradation rate for the entire time series could vary significantly from non-linear estimates Therefore, assuming a linear degradation rate would be less representative in the first few years of an operational PV system, but will likely be more representative after this initial period. The results in this thesis support the need for exploring non-linear changes in the degradation of PV systems and how this can improve O&M in the future.

Contents

Chapter 1

Introduction

1.1 Background and motivation

Since the 1950s, there has been an ever increasing amount of greenhouse gas (GHG) emissions released into the Earth's atmosphere which has lead to the global average surface temperature increasing over the same period [1]. The large increase in GHG emissions has been attributed to anthropogenic sources including power generation, transport, agriculture, construction and retail [2]. Electrical power generation is however the largest contributor to emissions, as oil and gas combined were 59 % of the share of fuels for global electrical power generation in 2021 [3]. The demand for renewable power generation is high, and increasing, as policy makers look to meet their goals of reaching carbon emission neutrality claimed in the Paris agreement or in their own national climate pledges. To reduce GHG emissions, processes that traditionally used fossil fuels as an energy source can transition to using electricity instead. Measures such as these contribute to larger demands for renewable power generation, not to mention the spike in oil and gas spot prices in 2022 motivating further the installment of more renewable power generation [2].

The installed capacity of solar power systems has experienced a rapid growth globally over the past 20 years going from 1.09 GW in 2001 to 843.09 GW in 2021 [3] [4]. A considerable driver for this change has been the substantial decrease in cost of solar panels over the last decades [5]. This has made PV systems increasingly competitive in a market looking for renewable alternatives to fossil fuels as a power source for generation of electricity. For PV systems to further improve its standing as an alternative for electrical power generation, efficient tools for operation and maintenance (O&M) are needed

to minimize losses during production. PV systems can be installed locally on home owners' rooftops, as well as in less populated areas where they may cover large areas of land. For utility-scale PV systems, the requirement of area for installation combined with a significant amount of solar irradiation may often necessitate remote monitoring with periodic inspection by qualified laborers with specialized equipment [6]. As PV systems are comprised of many electrical components, there is a myriad of problems that may occur when operating a PV plant. The remoteness of a plant and the skilled labor required for maintenance presents challenges, from an O&M standpoint, when evaluating the best course of action for minimizing losses of power generation. Over time, the compounded losses can ultimately lead to substantial forfeited income for the owner of a PV plant. The irreversible losses caused by faults in a PV system are collectively considered as degradation and can be difficult to estimate. Diagnostic tools for accurately estimating degradation are therefore necessary, and can be used to determine whether the equipment is performing within the warranted range of degradation. Degradation analysis, along with monitoring equipment, can provide the information to secure well-informed solutions that may optimize the operation of a PV plant, possibly secure a longer lifetime and will likely improve the competitive standpoint for operators in a growing market.

1.2 Scope of work

Tools for degradation analysis are currently in development and methods for calculating degradation are often based on a year-on-year (YoY) method [7]. However, such methods would not take into account the influence of soiling, i.e., the accumulation of dust and other particles on the solar panels. This will call into question the validity of estimated degradation rates without accounting for soiling. The issue is particularly challenging for newly operational PV systems which will require time to establish reliable estimates of soiling in order to optimize maintenance costs against production losses due to soiling. Recently, an algorithm that combines the analysis of degradation and soiling in a PV system has been proposed [8]. The Combined Degradation and Soiling (CODS) algorithm was jointly developed by the Institute for Energy Technology (IFE) and the National Renewable Energy Laboratory (NREL), and is currently available for use in an open-source Python library called Rdtools [9]. CODS shows potential for giving a more accurately estimated degradation rate, but needs to be further tested with real-world data to be able to establish itself as an industry standard [10] [11]. The YoY method assumes a linear degradation rate which may not be accurate for PV

power plants, especially when newly operational, as PV modules are known to experience a higher degree of degradation in this period and towards the end of life [12]. Additionally, climate and temperature would also have an effect on the degradation rate [13]. Module warranties often provide a guaranteed performance level the first year of output, followed by a limit on yearly degradation until the end of the warranty period. However, it may be more accurate to assume a linear degradation over time, as the PV modules and equipment of the power plant transition out of the infancy stage. Nevertheless, standard practice is to consider the degradation rate to be linear, at least until better methods can be established. By way of accurately estimating the degradation rate using data from a PV power plant, the use of CODS can be investigated and the analysis can be used to comment on the change in degradation.

Therefore, the aim of this thesis will be to investigate the use of CODS on real-world time series data gathered from a utility-scale PV power plant. The objectives of the thesis will be to:

- Investigate meteorological and energy generation data from a large utility-scale PV power plant
- Choose and develop a selection of filters to be applied in order to enable CODS analysis
- Investigate the assumption of linear degradation by comparing the size of the time series

This thesis has limited itself to work within a certain scope, in order to focus the discussion and adhere to a time frame for completing this work. The method will not include a section on optimizing each filter for improving data quality before subsequent analysis using CODS. In this work, it is assumed that the default parameters for the CODS algorithm will provide an acceptable degradation analysis based on the recorded data from a PV plant. Furthermore, it is assumed that a demonstration of which data is filtered will suffice to prove the filter's relevance for the method.

Chapter 2

Theory and method

2.1 The solar power source

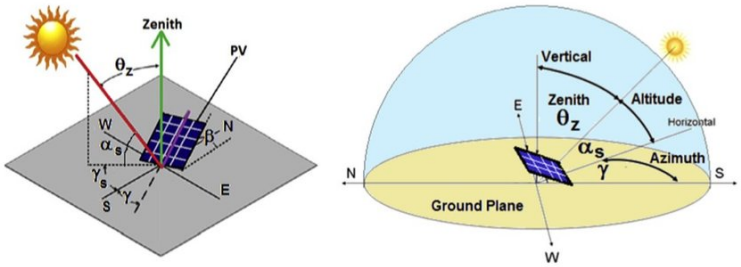


Figure 2.1: Different angles between the sun and a position on a plane with their symbols including; solar zenith θ_z, elevation α_s and azimuth γ angles. The tilt angle β is commonly used when describing the angle of inclination for a solar panel. The figure is adopted from Yilmaz et al. [14].

When solar radiation travels through the atmosphere, a significant portion is attenuated and scattered by different mechanisms. Due to this, solar irradiance is the measure of solar energy received at Earth's surface. The angle between the Sun's current position relative to it's zenith, is called the solar zenith angle θ_z. The solar altitude, or elevation angle, α_s, is the complementary angle to the solar zenith angle and is more commonly used when considering the tilt of a solar panel. The reference point, or panel in question, is also positioned in facing a horizontal direction. The angle formed between the direction normal to a panel and a south-facing direction is called the azimuth angle γ, and the angle formed between the surface and the back of the panel is called the tilt angle β. Another commonly used angle is the angle of incidence (AOI), which is formed between the Sun's position and a line perpendicular to the surface of a solar panel. High AOI angles are

generally not desired as this leads to a higher degree of reflection on the solar panel's surface and lowering the incoming radiation to the solar cell. The path travelled by the sunlight can be described in the terms of the air mass (AM) ratio and is expressed in equation 2.1:

$$AM = \frac{1}{cos(\theta_Z)} \tag{2.1}$$

Here, θ_z is the solar zenith angle. A value of $AM = 0$, simply written AM 0, means that the sunlight does not cross the atmosphere. AM 1 would mean that the sunlight travels directly from over the reference point on the Earth's surface and essentially travels the shortest path. An industry standard of AM 1.5 is used for testing and certification purposes, and is visualized in figure 2.2 to show the spectral distribution of AM 1.5. The figures shows that only certain wavelengths of the solar radiation reaches the Earth's surface.

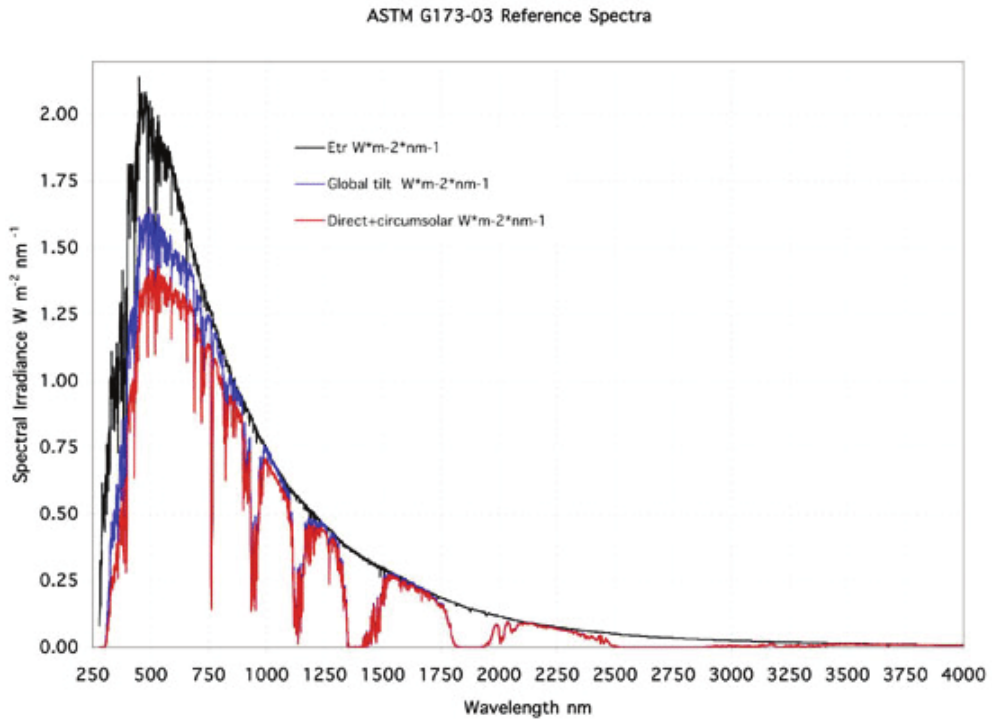


Figure 2.2: The ASTM G-173-03 reference spectra for solar irradiance. "Etr" stands for extraterrestrial irradiance, otherwise known as $AmM0$, while "Global tilt" represents $AM1.5$ in this standard. The figure is adopted from Gueymard et al. [15].

The total irradiance by a horizontal surface is known as the Global Horizontal Irradiance (GHI) and encompasses the two components; Direct Normal

Irradiance (DNI) and Diffuse Horizontal Irradiance (DHI) by the relation:

$$GHI = DHI + DNI \cdot cos(\theta_z) \tag{2.2}$$

The DNI signifies the irradiance arriving on a surface normal to the sun's current position, while the DHI signifies the irradiance scattered by the atmosphere. The ratio between irradiance scattered by a surface and the total irradiance received by the same surface is called the albedo. An albedo of 1 means that all incident irradiance is reflected off the surface. The GHI implies that the surface in question is parallel to the Earth's surface and the contribution of DNI and DHI can vary depending on the solar altitude and zenith angles. However, PV systems are often positioned differently than a horizontal surface and the irradiance is typically then measured in the plane-of-array irradiance (POAI). The POAI consists of 3 components that can be calculated using GHI, DNI and DHI:

$$POAI = E_b + E_g + E_d \tag{2.3}$$

Here, E_b is the beam component, related to DNI by:

$$E_b = DNI \cdot cos(AOI) \tag{2.4}$$

While the ground reflected component E_g can be determined by GHI, albedo A, and the tilt angle β:

$$E_g = GHI \cdot A \cdot \frac{(1 - cos(\beta))}{2} \tag{2.5}$$

Finally, the diffuse sky component E_d is not as simple to calculate. There are different models that estimate this component but generally assume that the component can be split into 3 different contributing factors. The first factor is the uniform irradiance from all directions and the second factor is the circumsolar brightening, as the irradiance close to the DNI will be slightly higher. The third and final factor is the horizontal brightening, as the irradiance will increase slightly as it approaches the horizon. Examples of models that are used to estimate the diffuse sky component E_d include the Perez model [16] and the isotropic model [17].

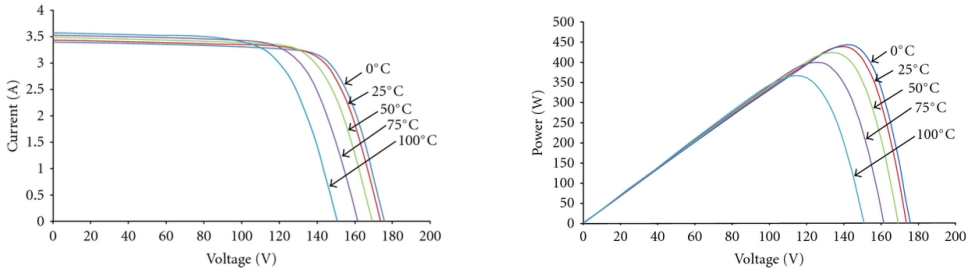
The irradiance measured at a site will be influenced by geographical features like vegetation, water bodies, clouds, snow, rock formations and anthropogenic structures. Different type of actinometers; instruments used to measure the intensity of radiant energy, are typically used in PV plants for measuring irradiance. For instance, pyranometers measure GHI, or DHI, and can be placed on an array in order to measure the POAI. Pyrheliometers are used to measure the DNI and are typically mounted on arrays that track the Sun's movement throughout the day [18].

2.2 PV systems

2.2.1 Solar PV systems

Solar PV systems exist in many different parts of the world. The size and capacity of each system can vary greatly depending on their application. For instance, smaller systems installed on residential rooftops can cater to local demands while larger systems spanning large swathes of the country side can supply the power demand of an entire city. Some systems may be isolated from the power grid as they are designed to supply power to remote areas, while others may be integrated into buildings and structures. However, it is the utility-scale power plants that have the largest capacities and are designed to meet a planned power generation. Every PV system contains different components that together form a system capable of generating and delivering power.

The working principle on how PV systems generate power comes from a property some materials exhibit called the photovoltaic effect. This means that they can absorb photons from sunlight and convert the energy to excite electrons into a state where they can act as charge carriers. A photovoltaic cell, or solar cell, is generally composed of a junction between a p-type and an n-type semiconductor and connecting electrodes with an external circuit. A forward bias applied to the cell will result in current flow across the cell. The commercial standard for the PV material in solar cells today is polycrystalline silicon (c-Si) with a smaller portion of the market constituting of monocrystalline silicon (m-Si) and thin film cells [19]. In regards to temperature, an increasing cell temperature is well-known to cause a reduction in the efficiency of the solar cells [20]. In figure 2.3, the effect can be seen to have a slight increase on the current in the solar cell as temperature increases, but the voltage has a stronger decreasing effect. This results in an overall decrease in power as temperature increases.
The structural design of a solar cell can vary as texturing surfaces to trap light are often applied. Solar cell architectures with silicon based (Si) structures, can include additional layers in the cell in order to improve the cell efficiency. This is done by passivating carrier recombination, which is the process of losing charge carriers before they have reached an electrode. Additionally, choosing different metals as electrodes and how they are placed along the cell surfaces can also lead to passivating carrier recombination. For instance, adding a back surface field (BSF) of Aluminum will reduce carrier recombination at the back of the cell [23]. The passivated emitter and rear contact (PERC) structure introduces localized BSF contacts surrounded by



(a) (b)

Figure 2.3: Effect of temperature on the a) Current-Voltage and b) Power-Voltage curve of a solar cell under different temperatures. The figures are adopted from Wang and Chen [21].

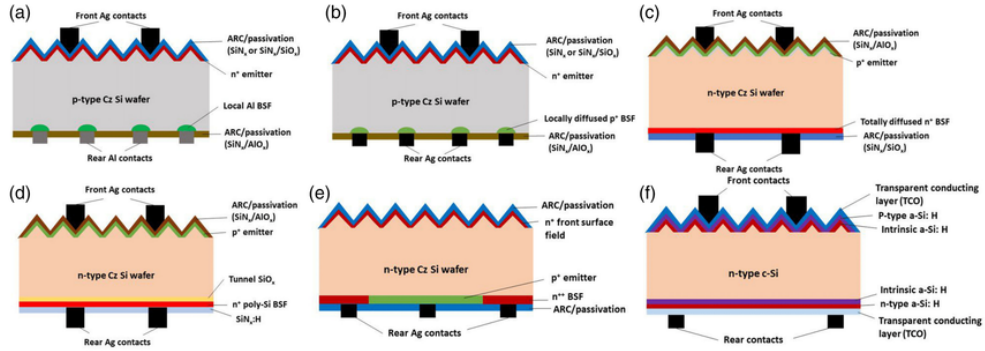


Figure 2.4: Examples of different silicon-based bifacial solar cell designs. a) PERC b) PERL c) PERT d) TOPCon e) Interdigitated back contact (IBC) and f) silicon heterojunction (SHJ). The figure is adopted from Molto et al. [22].

a passivation layer that reduces recombination and increases the reflection of light at the back surface [24]. Different to the PERC is the passivated emitter rear locally diffused (PERL) where the local back contacts are doped material, while the passivated emitter rear totally diffused (PERT) structure has a complete doped back layer [25]. By adding a thin passivating oxide layer to the diffused back layer, the cell is known as a tunnel oxide passivated contact (TOPCon) cell and further reduces the surface recombination between the metal electrode and semiconductor[26]. Rear contact cells, like the interdigitated back contact (IBC) cell, remove the electrode contacts at the front of the cell and thereby eliminate shading caused by the front contacts [27]. The heterojunction cells are different in the sense that they are formed by two different semiconductors [23]. For example, the silicon heterojunction (SHJ) cells are formed with crystalline silicon wrapped with amorphous silicon on each side [28]. The solar cell designs even permit both

surfaces to convert sunlight into electricity, which are called bifacial solar cells [29]. Even though PERC designs have the largest market share today, TOPCon and SHJ designs are expected to increase over the next decade [30].

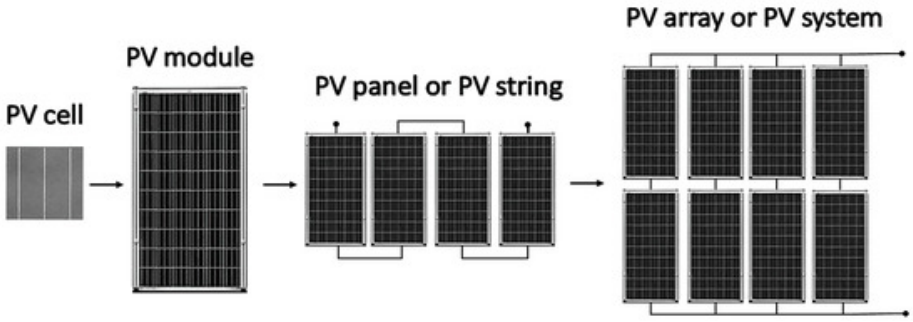


Figure 2.5: Configuration and difference between the parts of a PV array from cells to modules and strings. The figure is adopted from Zsiborács et al. [31].

By connecting multiple cells one after the other, the series can now be considered a solar module as seen in figure 2.5. If a module is placed within a frame which provides protection and stability, it is now considered a solar panel. Multiple strings, which are panels connected in a series, may be connected in parallel to each other with an inverter. At the inverter, the type of electrical current is changed from the produced direct current (DC) to an alternating current (AC). Multiple inverters and strings can be present in what is now considered a PV system, and further require a form of controller module that adjusts the target power, or power set point, of the entire system. Inverters can also include power tracking functionalities in order to optimize the power production [32]. The entire fleet of panels or strings in a PV system is called the array and can be mounted and faced in different ways. The tilt β and azimuth γ angles determine the direction of which the array is facing, as the PV modules can be fixed while others are built with tracking systems. Such tracking systems, being either 1 axis or 2 axis systems, change the tilt β and azimuth γ angles of an array in an attempt to maximize the POAI, and thereby, the power produced. A few examples of tracking systems can be seen in figure 2.6.

2.3 PV system performance metrics

As the PV system generates electrical power, the performance can be monitored and may vary greatly throughout the day. The performance of a PV

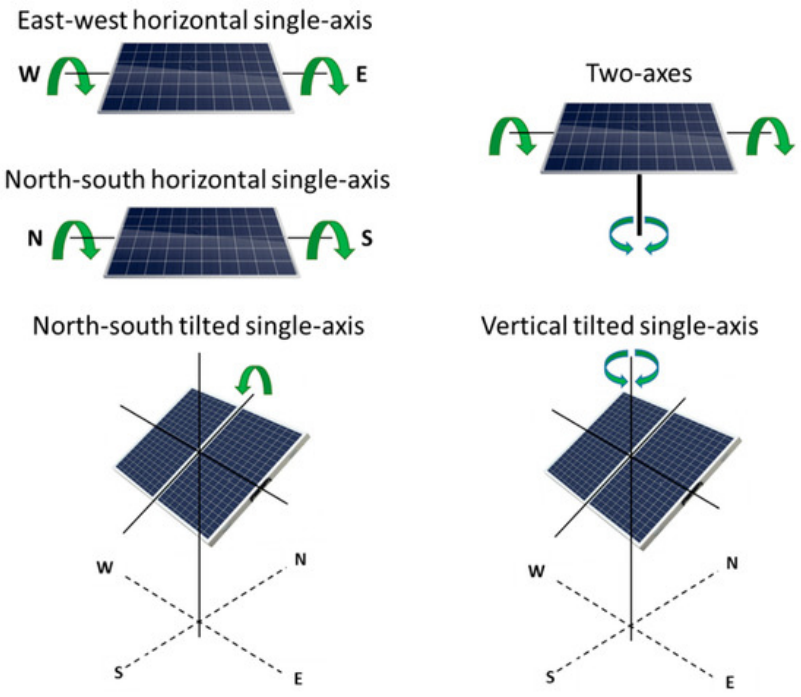


Figure 2.6: Different tracking configurations for 1 or 2 axis systems. The figure is adopted from Gutierrez et al. [33].

system is expected to degrade over time as the natural wearing of the modules and electrical components in the PV system causes a loss in performance. The influence of faults and available irradiation also influences measurements which necessitates a selection of methods for calculating performance indicators that can accurately describe the system performance.

Energy yield Y_A and reference yield Y_{ref}

A PV system or array that generates energy will have a rated energy that is normally given under STC conditions and gives an indication of the expected quantity of energy to be generated. The actual energy generated compared to the rated energy is called the yield Y and can be expressed in various ways. An overview of yield and associated terms is provided in the IEC standard for Photovoltaic System Performance [34]. If considering the measured energy produced by the array E_k over a defined time period k compared to the rated energy of the array P_0, then it is called the array energy yield Y_A where:

$$Y_A = \frac{E_k}{P_0} \tag{2.6}$$

A different way of considering the yield is comparing the total in-plane irradiance H_i to the array's rated reference in-plane irradiance $G_{i,ref}$. The latter is set by the irradiance at which P_0 is determined; 1000 W / m^2 for standard test conditions (STC). This is called the reference yield Y_{ref} where:

$$Y_{ref} = \frac{H_i}{G_{i,ref}} \tag{2.7}$$

The performance ratio PR and temperature corrected performance ratio PR'_{STC}

Power produced by a stringset, inverter or entire system can vary greatly throughout the year, hour, or minute depending on the meteorological conditions and available irradiation. Therefore, power production solely estimated by a model is not always an accurate estimate of performance and could be compared to normalized generated energy for the expected production of the array in question. A measure of performance can be represented by a performance index PI_k for a defined time period k:

$$PI_k = \frac{E_k}{E'_k} \tag{2.8}$$

Where E'_k is the expected energy generated by the array, and is calculated by using a chosen performance model. The performance ratio PR is a normalized performance measurement and is the quotient of the array energy yield Y_A to the reference yield Y_{ref}:

$$PR = \frac{Y_A}{Y_{ref}} \tag{2.9}$$

Here, the performance of the system is displayed and includes the faults or inefficiencies due to components and the array temperature of the system. Further correction for the module temperature is therefore pertinent by calculating the temperature corrected performance ratio PR'_{STC}:

$$PR'_{STC} = \frac{PR}{(1 + (\gamma * [T_{mod} - T_{ref}]))} \tag{2.10}$$

Here, γ is the relative maximum power temperature coefficient of the module, T_{mod} is the measured module temperature, and T_{ref} is the reference temperature under STC conditions which is 25 °C. An overview of performance ratios and associated terms is also provided in the IEC standard for Photovoltaic

System Performance [34]. Since the PR'_{STC} takes into account the effect of module temperature, it can be used as a model of expected energy and will thereby be considered a performance index PI.

For sun-daily measurements, the chosen performance index can vary greatly and may be difficult to use further for analysis purposes. An aggregation of performance indices into daily values is a more robust alternative and less sensitive to outliers of intra-daily variation. Therefore, the intra-daily stringset PR'_{STC} values can be used to calculate the median for all data points within a day to represent the daily performance value for a stringset $PR'_{str,d}$:

$$PR'_{str,d} = Median[PR'_{STC}{}_{k=1}^{K}] \tag{2.11}$$

Here, K stands for all the data points in a day that belong to the stringset. By the same principle, the median for all $PR'_{str,d}$ within a PV system can be calculated to represent the daily system PR'_{STC}, written as $PR'_{sys,d}$. For N number of stringsets in a PV system:

$$PR'_{sys,d} = Median[PR'_{str,d}{}_{n=1}^{N}] \tag{2.12}$$

2.4 Time series analysis

Time series analysis is a discipline within data analysis that considers the temporal development of variables, in order to run statistical analysis which can be used to establish trends in the historical data and to forecast future trends. A time series, as the name suggests, includes data points that are time stamped and logged by sensors measuring values such as temperature, voltage, or rainfall. A typical time series can be decomposed into different components, all of which together make up a model that fits the time series. For instance a classical model of a time series can be expressed by the following:

$$y_t = T_t * S_t * I_t \tag{2.13}$$

In this equation, the time series y_t for a time period of t is made up of a trend component T_t, a seasonal component S_t and an irregular component I_t. T_t displays the overall movement of the time series while S_t reflects the seasonal variation throughout a time series. Finally, I_t is the remaining signal that represents small irregular variations in the data set, otherwise called the noise, residual or random variation trend. The practice in the PV

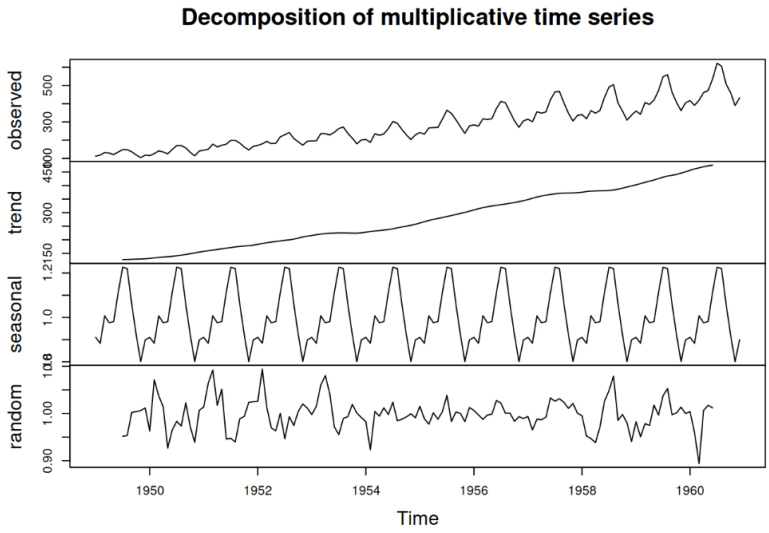


Figure 2.7: An example of a typical time series decomposition into a trend, seasonal and random component. The figure is adopted from Svetunkov [35].

industry is to determine the degradation in a PV system by decomposition of the time series as in equation 2.13. For the case of PV systems, there is an expected seasonal variation of irradiance throughout the year while the trend component T_t represents the decline of irreversible performance of the system, otherwise known as the degradation of the PV system. It is important for operators and investors to determine the degradation of a PV system in order to estimate the expected output of a PV system over its lifetime. The challenge arises when attempting to factor out the seasonality S_t of the time series and the irregular signal I_t. In order to establish the performance loss of a PV system, time series data is viewed and analyzed to model the PV system performance over time. External influences to the PV system like curtailment and power outages can be corrected for, in order to examine the PV system performance influenced only by the system and module level faults. This information would provide a more accurate estimate of the PV system performance and can be used to determine a course of action to improve the power production by the PV system. In addition, filters can also be used to remove data that appear as noise and are not representative of the system performance.

2.4.1 Data screening and filters

There is a large variety of sensors on the market to select from when looking to monitor factors like temperature and irradiation. A sensor would regardless require maintenance and calibration to ensure its function and durability over time. Therefore, sensors are likely to have some issues over their life times and for this reason it becomes necessary to screen the data before further investigation. Unreasonable large or small data points can be logged by the sensor due to a various of reasons like internal malfunctions, poorly connected wiring or a loss of power triggering a shut down. Data screening can be initially performed by applying maximum and minimum thresholds to remove such points and improve how representative the collected data are in terms of the operational conditions of the PV system. Data filtration is used prior to calculating system performance as there are many sources of influence on the energy produced from each array. The balance between reducing noise in the time series and keeping a significant amount of data points is an important concern when applying filters to data sets. The balance stems from the fact that data points representative of the performance will likely be masked by effects occurring to real PV systems thus warranting the use of data filtration. On the other hand, uncontrolled or overt use of filtration will remove significant amounts of data that are representative of the PV system's performance. The consequence of data filtration may have significant effects on the outcome of calculations based on filtered data sets [36]. Therefore, an examination of filters and their effect after applying said filters on performance data, is necessary to understand the final results.

AOI filter

An angle of incidence filter can be used to filter out time stamps where the incident direct solar radiation has an AOI that is too high. The spectral response of the module at varying AOI is also dependent on the degree of soiling [37]. For sub-daily measurements, a typical AOI filter will remove data points from late in the evening, at night, and during the early morning. Data points during these hours are associated with low power generation and are not indicative of the current performance.

Solar elevation filter

Similar to the AOI filter, a solar elevation filter will remove data points until the sun has reached a given solar elevation α_s. Therefore, for sub-daily measurements, a typical solar elevation filter will remove data points during late in the evening, at night, and during the early morning.

Clear sky filter

A clear sky filter calculates first times associated with clear sky conditions, where there is little to no shading, e.g. from cloud coverage, and uses this to filter out time stamps that do not experience clear sky conditions. For sub-daily measurements, a clear sky filter will contribute to reduce unrepresentative performance data that appear as noise compared to performance data during clear sky conditions. This can be seen in figure 2.8 where the start of the day is experiencing clear sky conditions visible by the smooth increase in GHI. However, in the afternoon, the GHI measurements are jagged and vary often from one hour to the next, thereby indicated as non-clear sky conditions.

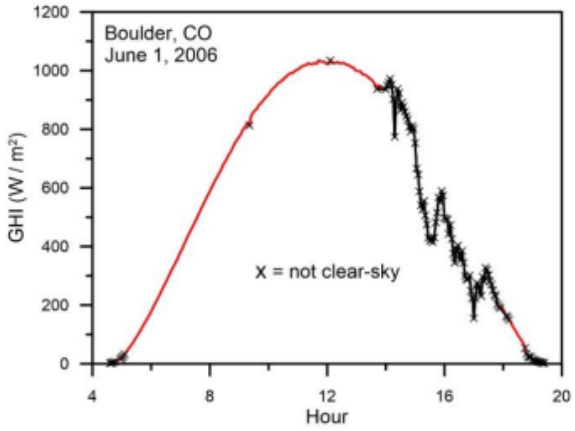
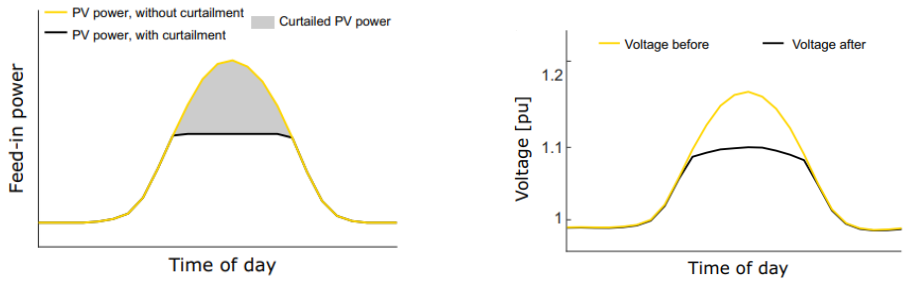


Figure 2.8: Example of a clear sky filter detecting non-clear sky conditions, in Boulder, Colorado on June 1st, 2006. The black crosses represent time stamps that were determined to be experiencing non-clear sky conditions. The figure is adopted from Marion [38].

Curtailment filter

Curtailment occurs when the TSO imposes a reduction in power production and the reason is primarily due to a lack of available capacity on the transmission grid at certain times, but can also be due to issues with voltage or interconnections [39]. A curtailment filter would take into account when the system purposely reduces its' output and filters out these data points, as the power produced during these time stamps are not representative of the actual system performance. This can be seen in practice in figure 2.9 where the power output of the system is leveled due to curtailment, and can also be seen by a drop in voltage.



(a) (b)

Figure 2.9: Effect of curtailment a) on the power generated and b) on the voltage of a PV system. The figures are adapted from Luthander et al. [40].

Clipping filter

Clipping occurs when the DC to AC capacity of an inverter is met or is about to be exceeded. To prevent overload and faults to the electrical components, the inverter reduces the voltage and thereby containing the AC power output of the inverter. An example of this can be seen in figure 2.10 where a system with a high enough DC to AC ratio will experience clipping if the power generation reaches the striped line. Like curtailment, during clipping the power generation at inverter level will remain leveled until the production reaches the determined DC to AC ratio.

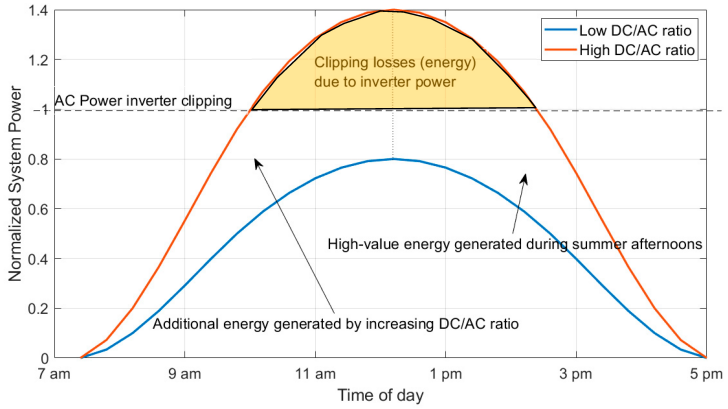


Figure 2.10: Effect of inverter clipping on the system power output. The dashed line shows the capacity of the inverter and the clipping losses during peak production hours, highlighted in yellow during a high DC to AC ratio. The figure is adopted from Hazim et al. [41].

2.4.2 Time series modelling

Working with large data sets can be done in a variety of ways including database programs and libraries developed specifically for data analysis. There are many different types of libraries created for this in mind, but Python is a popular alternative for time series analysis in the PV industry. From data visualization to data manipulation, Python has become a well supported and popularly used open-source programming language for data analysis. The following are available resources for modelling and analyzing time series data.

NumPy NumPy is a package that contains different mathematical functions and routines that are commonly used in other libraries developed for data analysis. This means that functionalities in such libraries are often reliant on using NumPy and its functions. Although NumPy may not be the literal library being called in the code, it certainly often works in the background for Python to work as the intended tool [42].

Pandas The Pandas library contains tools to read, analyze and manipulate large data structures into tabular objects called data frames or series. Developing filters and applying them on the data frames is a use-case in working with times series data from PV systems. In particular to working with time series data, date-time indexation streamlines the use of builtin functions applied to the data. Other python packages often take use of Pandas in their modules in conjunction with NumPy [43].

Pvlib The Pvlib package developed and maintained by Sandia National Laboratories is an open source library for simulating the performance of PV systems. Tools provided by the package include functionality for time series analysis, modelling and filters used for this purpose. Additional functionalities include forecasting and diode modelling [44].

Rdtools Rdtools is an open source library also for time series analysis in PV systems developed by researchers at NREL to produce better estimates of degradation for a PV system. The library itself relies on using Pandas and Pvlib, as well as the contributions of different researchers that continue to improve upon the library [9].

2.5 Soiling, faults and degradation

2.5.1 Soiling

The importance of soiling in PV systems

According to a recent report by the International Energy Agency (IEA), soiling is the most influential factor on the loss of energy yield from PV systems after irradiance [45]. Especially in dry climates or areas that experience snowfall, a buildup of soiling necessitates measures to counter the impact it has on the energy yield. However, measures like applying anti-soiling coatings or increasing the frequency of cleaning by maintenance crews or automated systems will result in higher maintenance costs for the PV system. In a study by Ilse et al. [46], it was estimated that in 2018, soiling caused a loss of annual PV energy generation of at least 3-4 %, even when optimized by weighing cleaning costs and revenue losses from lowered production. This number was also expected to grow in 2023 as it was reported that installation of PV modules is increasing in high-insolation areas which are generally more exposed to soiling. Furthermore, the reduction of electricity prices from governmental policies like feed-in tariffs lowers the incentive for cleaning, as the revenues from energy yield recovery would be lower. Also, modules with higher efficiencies are consequently prone to higher losses due to soiling compared to modules with lower efficiencies.

A metric of describing the loses due to soiling is the soiling ratio SR. The soiling ratio SR is a dimensionless value generally defined as the ratio between the performance of a system under soiling conditions $Performance_{dirty}$ to the performance output of a system without soiling conditions $Performance_{clean}$:

$$SR = \frac{Performance_{dirty}}{Performance_{clean}} \in [0, 1] \tag{2.14}$$

Soiling mechanisms

Airborne dust particles are the primary source of soiling and relative humidity, wind speed, wind direction, precipitation have all been shown to influence the degree of soiling [45]. Other sources of soiling include bird droppings, algae, mosses, pollen and snow [47][48]. The accumulated soiling of dust is the net result of deposition, rebound and resuspension of particles on the panel surface over time [49]. Deposited particles are those that come in contact with the panel surface, and the particles that rebound are those that do not adhere to the surface. The resuspension of particles occurs when wind gusts

cause the particles to become airborne again. Whether particles are washed or resuspended, the effects of soiling are reversible and the transmission losses of light due to soiling can be consequently be recovered. A separate natural cleaning mechanism is rainfall, but dew formation by increased humidity is reported to increase the degree of soiling as it causes increased particle adhesion to the panel surfaces [46]. The degree of soiling within the same PV system can be non-uniform [50], typically concentrated on the first rows facing the wind direction transporting dust particles and generally towards the lower end of the panels as seen in figure 2.11. Non-uniform distribution of soiling presents a challenge in accurately measuring the degree of soiling for the PV system. It is therefore beneficial to have multiple measurement points throughout the PV system to account for non-uniform distribution of soiling.

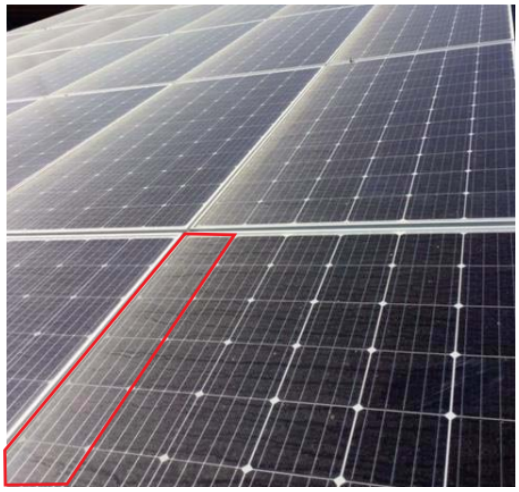


Figure 2.11: Soiling concentrated on lower end of solar panels. The figure is adopted from Kagan et al. [51]

The cleaning of panel surfaces can be done manually or automatically, and generally requires water for washing but efforts to establish methods without the need for water have been reported [52]. This is of interest as the supply and use of water for cleaning soiled modules in a PV array can be significantly more difficult and costly to supply in remote areas. Methods are also generally separated into manual cleaning, performed by maintenance crews, semi-manual, by using machinery to clean multiple panels at the same time, or fully automated systems. The latter of the three generally requires higher capital costs and is therefore less frequently seen, but might be a viable option in highly remote areas [47].

Soiling measurements

Soiling can be measured through soiling stations, and by optical methods like optical soiling measurements (OSM) or soiling image analysis (SIA). Soiling stations are described in standard IEC 61724-1 [34] and involve using a pair of reference PV devices, like cells or modules, where one is cleaned regularly while the other is permitted to soil naturally. OSM uses photosensors that measure the degree of soiling on small glass surfaces which are allowed to soil naturally. These sensors have no moving parts and generally do not require cleaning to function [53]. However, OSM sensors are relatively new and are still being field tested, but show promise in being used for PV systems [54]. The SIA sensors leverage imaging and post-processing of aerial photographs. The images can be taken using drones [55] or satellite imagery [56] and are the youngest type of soiling measurement sensors to have been tested.

Figure 2.12: A soiling station showing an automated periodical cleaning functionality for the lower cell. The figure was adapted from Barnes [57].

2.5.2 Faults in a PV system

From when modules leave the assembly line until a PV system is built, there can be be faults occurring on a system level and others occurring at the module level. Over the lifetime of a PV system, there will also likely be faults occurring that affect the power production of the system. Faults can be categorized by system or module level faults, as the causes and measures for handling faults will vary accordingly. Faults on a system level include faulty connection and junction box wiring, inverter faults, and sensor faults like drift, bias or complete failure [32]. Faults on a module level include soiling, corrosion, cracking, delamination, discoloration, hot spots, light induced

degradation (LID) and power induced degradation (PID) [58].

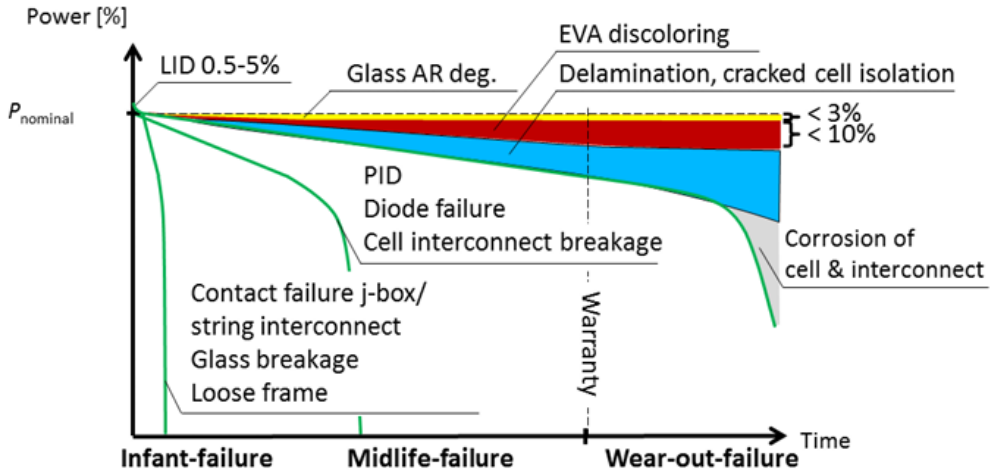


Figure 2.13: Typical fault scenarios for PV modules and systems over the lifetime of the system. The figure is adopted from Köntges et al. [58].

The individual parts of a PV system are tested and rated under standard testing conditions (STC) to establish expected performance when used in operation. The system nominal power is thereby a known quantity along with the inverter nominal power and the stringset nominal power. However, in operation the PV system will rarely produce at the STC rated power as these conditions are rarely seen in the field. Primarily, the availability of direct solar radiation is limited to a certain amount of hours in a day. This implies that the number of peak power production hours, where the PV system produces power close to or at the nominal system power, is limited. The ambient temperature of the site will also have an impact on the PV system, as Light- and elevated Temperature Induced lifetime Degradation (LeTID) is known to occur over the lifetime of a silicon-based PV modules [59][60]. Furthermore, clouds as well as shading can occur during daylight hours which will further reduce the power produced by the PV system. Depending on the local climate, a site will also experience a degree of soiling, where the PV modules become gradually covered by dust particles and thereby reduce the portion of incoming solar radiation on the PV array.

Considering performance, irreversible faults will compound the production losses over time which gives rise to the term degradation rate. Performance losses can be a product of any number of faults, but some may be reversible, like soiling. The degradation of PV systems is known to depend on climate,

as hotter and more humid climates are reported to have high degradation of PV modules [61] [62]. In terms of the climate and irradiation classifying Kõppen Geiger PV (KGPV) zones [63], PV systems in AH (A - Tropical climate, H - High irradiance) zones are modelled to have the highest degradation rates, even higher than AK (K - Very high irradiance), as the increased humidity contributes to increased degradation [64]. Therefore, considering climate and its' impact on the faults that may occur in a PV system is key to better understand the degradation of a PV system over time.

2.5.3 Degradation rates

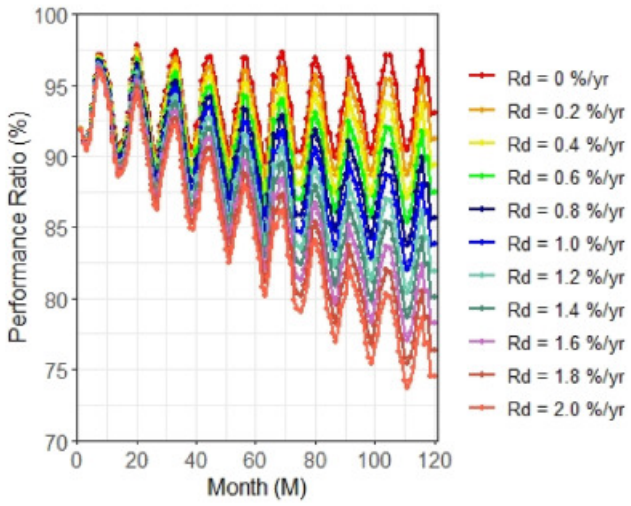


Figure 2.14: An illustration of the effect of simulated degradation rates *Rd* on the performance ratio over time. The figure is adopted from Romero-Fiances et al. [65].

Degradation is commonly assumed to be linear, where time series data is used to determine a single degradation rate for the entire time series. This assumption might prove to be problematic as PV systems experience both reversible performance loss, like soiling, and irreversible performance loss, like module degradation. The impact of degradation can be severe over time and is demonstrated by figure 2.14. Degradation rates have been estimated by linear regression methods like ordinary least squares (OLS) which includes multiple predictors [66], and robust linear regression (R-LR) which which is less affected by outliers while being more demanding computationally [7]. Classical signal decomposition (CSD) and auto-regressive integrated moving averages (ARIMA) [66] [67] [68] have also seen its use and are both based

on moving averages to estimate a degradation trend. A sort of continuation of CSD is the seasonal trend decomposition with LOESS (STL) [69]. Here, LOESS stands for locally estimated scatterplot smoothening and handles seasonality in a time series better as it is more robust [70]. Attempts at reducing the underlying data structure in order to retain variables that show the most substantial effect on degradation have also been made through robust principal component analysis [65] [71]. Year-on-year (YoY) methods have shown to outperform regressions based methods in the presence of outliers, strong soiling and pronounced seasonality [72][73]. However, it is difficult for the YoY-method to detect non-linearities in the degradation rates, as the method relies on at least two years worth of data, and compares the change for a data points between consecutive years. Adler et al. [74] reported having used a YoY-method on daily PR'_{STC} data in an AH KGPV zone, estimating a median degradation rate of -1.45 %/year for the nominal system degradation measured on the DC power of the inverters. The authors referenced the first year of the performance time series experiencing a more rapid degradation than the subsequent years, and that the chosen YoY-method would likely not be representative of this section. They also compared the degradation rates to a PV system in a BK/CK (B - Desert climate, C - Steppe climate) KGPV zone, where the PV system in the AH KGPV zone appeared to degradation faster. However, the article reported to assume negligible soiling, as the neither park was reported to be located in an area with much soiling and that both practiced regular cleaning, which was determined using soiling sensors.

There has also been published work investigating methods for estimating non-linear degradation. The principal motivation for investigating non-linear behavior is the financial impact it may have in regards to deviations from expected production and increased costs for operations and maintenance. Non-linearities have been reported for modules, especially in the beginning of life and during the wear-out phase, and will have an effect on the finances of the PV system [12]. Theristis et al. [75] applied the Facebook Prophet (FBP) algorithm to detect change points in the time series. Change points can be found automatically or for a specific amount given to the the FBP algorithm and it assumes a trend, seasonal, holiday (not used) and error component to decompose the time series. The trend component is then divided into sections using the identified change points and OLS is used to estimate the degradation rate for each section. The change point method was then used, restricted to 1 and 2 change-points, on monthly PR values for different PV systems over an 8-year period. The results showed that the PV systems with c-Si modules displayed a linear degradation, while the m-Si and thin film modules displayed non-linearities. It is worth mentioning that all of the

tested PV systems were of around 1000 kWp and situated in a CH KGPV zone. It can be assumed that for larger PV systems situated in areas with more humidity will experience a higher degree of degradation, thus increasing the chance of non-linearities in the modules. More recently, different change-points methods were compared using synthetic data sets and for different climates [76]. The methods included were FBP, piece-wise linear regression (PW-L), the Bayesian estimation of abrupt change, seasonality and trend (RBeast) and the Sequential and Batch Change Detection Using Parametric and Nonparametric methods (CPM). Here, the FBP and PW-L methods performed the best in terms of mean absolute errors (MAE) of change-point location and degradation rate in each segment. However, it was mentioned in the study that the synthetic data sets do not include any influence from soiling, temperature, sensor drift and other factors that have an impact on real PV systems. FBP was also pointed at as a computationally demanding method for estimating non-linear degradation rates. This may not be a favorable approach for degradation analysis using many years of data, not to mention if soiling is to be accounted for as well.

Testing non-linear methods for estimating degradation on real data has also been investigated by by Lindig et al. [77]. Here, a multi step performance loss (MS-PL) algorithm was used, which estimates the optimal amount of change points for dividing the time series into separate linear degradation segments. The MS-PL algorithm was used on PR and PR'_{STC} data for PV systems with various solar cell technologies in a DM (D-Temperate climate, M-Medium irradiation), and without accounting for soiling. For the c-Si PV systems, the estimated linear degradation rates by using STL varied between -1.56 %/year and -0.76%/year, while the non-linear degradation rates using MS-PL varied between -4.55 %/year and 1.06 %/year. It was reported that all c-Si PV systems exhibited an increase in performance during the first 3 years, before exhibiting a rapid degradation for about 4 years, before settling at a smaller degradation rate. The study also pointed out that change points were sensitive to maintenance events, which would be beneficial for servicing concerns, but do not increase the understanding of degradation and system characteristics. It was therefore concluded that the MS-PL would be a good choice for analysts who are not familiar with the PV system to get a quick status analysis.

Livera et al. [78] reported the use of methods for finding non-linear degradation rates on 7 years of performance ratio PR data, for 11 PV systems in a CH KGPV zone. For the c-Si PV system displaying a non-linear degradation trend, results from using the FBP method showed that the degradation

rates varied between -1.21 %/year and -0.90 %year, without accounting for soiling. Here, an interesting point from the results was that the number and location of identified change-points, and thus the estimated degradation rates, depended on the choice of algorithm. The study showed how identified change-points could be detecting faults, soiling or maintenance events. This would warrant accounting for soiling and maintenance in performance data if representative degradation rates are to be estimated.

2.5.4 Combined degradation and soiling (CODS)

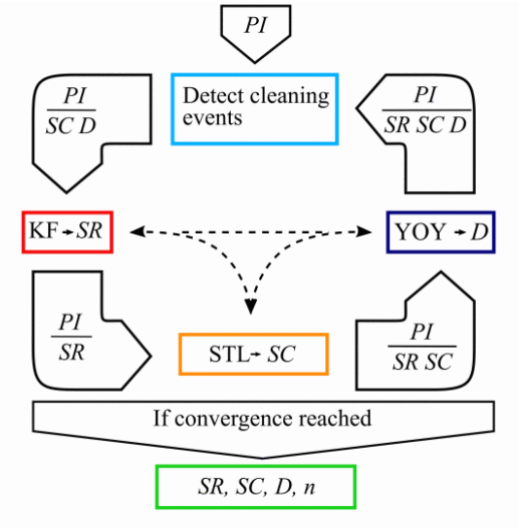


Figure 2.15: A flowchart explained the steps in the iterative signal decomposition in the CODS algorithm. The figure is adopted from Skomedal and Deceglie [8].

In an effort to combine the effects of soiling and degradation on performance, CODS is a recently developed an algorithm that estimates a degradation rate of a PV system [8] and has been tested on data from utility-scale PV systems [10][11][79]. The algorithm assumes that the performance index can be divided into separate components, that each have a multiplicative effect on the performance index. The time series components for an amount of days d are as follows:

$$PI_d = SR_d * SC_d * D_d * n_d \tag{2.15}$$

Here SR_d is the daily soiling ratio, SC_d is the daily seasonal component, D_d is the daily degradation and n_d is the daily noise that is left after accounting for the other components. The work flow of the algorithm is summarized

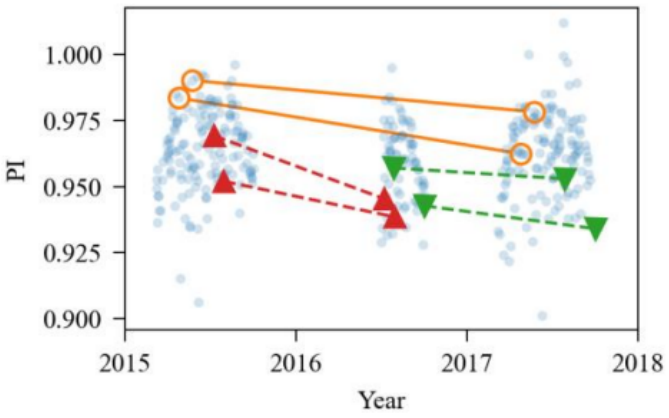


Figure 2.16: Illustrating the working principle of the YoY-method. Performance indices separated by subsequent years are compared to calculate the rate of change. If missing data is present, the YoY-method will use the data from the remaining years. The figure is adopted from Sveen et al. [81].

in figure 2.15 In order to divide the performance index into these separate components, the algorithm applies a series of steps to establish initial estimates for each time series component. First, it attempts to detect cleaning events by considering them as outliers based on a rolling 9-day rolling median. Then, the algorithm estimates missing data points that need to be filled in. Next, a Kalman filter is used for estimating SR_d. In essence, the Kalman filter takes a series of noisy data and estimates the variables that give rise to the data, and in the case of the CODS algorithm this would be the estimated soiling ratio SR_d and its rate of change [8]. STL is then used to find the seasonal variation in a times series. Finally, for finding D_d, a YoY method is applied. In YoY, data points for the same time stamp separated by subsequent years are compared to see if there is a significant change from one year to another, as seen in figure 2.16. Therefore, The YoY-method requires at least 2 years worth of data and creates a minimum amount of data required for applying CODS [80].

The algorithm in CODS will iteratively use previously estimated components in an attempt to estimate more accurate values for each component. The algorithm will then converge on a fitted model that does not change above a threshold root mean square error (RMSE of 0.5 % compared to the previous model. The remaining noise n_d not accounted for in the SR_d, SC_d and D_d is also tested for stationarity, meaning if the probability distribution of the noise changes over time. After the model fitting, the algorithm proceeds to calculate uncertainty of each model by bootstrapping, where randomly

selected samples with replacement from a data set are taken to create new variations of the original data set. This is done to calculate 95 % confidence intervals for the fitted models. The results include estimated model fit, degradation rate, soiling rate, seasonal component and remaining noise for each time stamp, in addition to 95 % confidence intervals, along with other estimates like soiling rates and cleaning events. A final degradation rate D_d is calculated for the best fit model along with a 95 % confidence interval.

CODS has been reported to estimate the soiling well compared to measurements from soiling stations when applied to performance indices from inverters in a 150+ MWp PV system situated in a hot and humid climate [10]. Here, the median degradation rate among all inverters was -0.12 %/year, and ranged between -0.66 %/year and 0.61 %/year. Further more, CODS has been used on inverter level PR'_{STC} in a high-soiling environment in the Middle East. The degradation rates were reported to be between -0.86 %/year and -0.15 %/year with a mean degradation rate of -0.48 %/year. It was noted that CODS estimated a more consistent degradation rate and with a smaller confidence interval compared to only using a YoY-method. CODS has also been applied to performance data in a high-soiling environment in California, and found that CODS estimated less negative degradation rates compared to the YoY-method [11].

2.6 Operations and maintenance (O&M)

The operations and maintenance (O&M) of a PV system consists of continuous as well as retrospective monitoring of performance. Performance can be monitored on the level of a string, array, inverter or for the entire system. It is then evaluated by comparing measured energy generation to a model that takes into account various data inputs from sensors that are placed in and around the site. Power output on system level can be monitored on a continuous basis for the interest of power production, but also for grid operation. A grid operator must maintain balance and control of feeding enough power to a grid in order to meet demand, but not too much as to overload the capacity of the local transmission grid. Curtailment may occur if the transmission grid system operator (TSO) imposes a reduction in power production by the plant operator. Therefore, any deviance from expected power input from producers to the transmission grid must be reported to the grid operator. The power producers will likely have a Power Purchasing Agreement (PPA) in place, where a supplier commits to producing a certain

amount of electricity, and a buyer commits to buying the electricity at an agreed upon price. Optimizing operation of a PV system to meet the expected production values in a PPA will likely be the primary interest of the plant operator.

In addition to power production, parameters that further influence performance of a PV system are monitored to establish relative system performance, meaning how well the system performs given the current conditions. Irradiance can be measured by satellite measurements or locally at the site typically in terms of GHI or POAI. Temperature can be measured as the ambient temperature around the PV system, or on the back of modules in the PV system. Cleaning events can be logged to establish when and how often the modules are cleaned by maintenance crews. The power set point tells the targeted production of the plant and is set by a central controller. Additionally, there are many other signals that are measured and monitored including precipitation and humidity, wind speeds, and power outages. These signals and events can be logged for monitoring purposes and may influence design choices for future PV systems to be built in similar conditions.

2.7 Method

2.7.1 Data collection

Data sets were collected from a PV power plant situated in a *AH* KGPV zone, meaning *A* - Tropical and *H* - high irradiation area [63]. The area was situated in a lower grass land or steppe, and experienced dry winters as well as a rain season during the summer. The yearly average temperature was around 29 °C and the area experienced around 2900 sunshine hours each year. The power plant consisted of a total of ∼50 inverters and 4500+ stringsets in addition to pyranometers, soiling stations and thermocouples. The PV plant had been in operation since 2015 and had an overall installed capacity of 60 MWp. The solar cell technology and supplier was not known for the modules in the PV system, but could be assumed to be Al-BSF considering the time of installation [82]. The data sets were made available by the plant operator for research purposes and the time range of the data sets was from August 2015 up until October 2022.

The data sets were comprised of 10 minute averages for the following; DC power for stringsets, DC voltages of inverters, GHI measurements by 7 pyranometers, POAI measurements by 7 pyranometers, module temperature by

7 thermocouples, 7 soiling stations measurements, active power and power set point measured by the plant controller.

2.7.2 Performance metrics

The PR'_{STC} was chosen as the performance index since it takes into account the influence of irradiance and temperature. After thresholds and filters were applied, the 10-minute PR'_{STC} time series from equation 2.10 was calculated for each stringset in the entire PV system. This was done by first aggregating module temperature data and POAI data. The median across all sensors was calculated to represent the data point for module temperature and POAI. Next, the 10-minute PR'_{STC} time series was used to calculate $PR'_{str,d}$ for all stringsets by using equation 2.11. All $PR'_{str,d}$ time series were then subject to removal of outliers. This was done by considering $Z_n = PI_n - \tilde{p}_{7,n}$ for n days, where $\tilde{p_{7,n}}$ was the rolling 7-day median of the PI_n time series. Data points were removed if $z_n > Q_3 * 3IQR$ or if $z_n < Q_1 * 3IQR$ where Q_1, Q_3 were the first and third quartile, and IQR was the interquartile range.

After the $PR'_{str,d}$ was calculated for each stringset, and then the $PR'_{sys,d}$ was calculated by equation 2.12. An arbitrarily selected $PR'_{str,d}$ and the $PR'_{sys,d}$ was then used as input for CODS analysis. The soiling ratios SR for soiling stations were calculated to determine the degree of soiling present in the plant. For each soiling station, the median for all data points constituting a day was calculated to represent the daily value. Then, for each day the median for all soiling stations was calculated and used to represent the soiling ratio measured by soiling stations SR'.

2.7.3 CODS analysis

CODS was not a part of the official latest release of Rdtools (version 2.2.0-beta.1), but could be accessed from a development branch of the git repository [9]. The CODS signal decomposition and subsequent bootstrapping for estimating confidence intervals was performed using default parameters through the run_bootstrap() method. For all CODS results, the default 512 repetitions of bootstrapping was used. CODS analysis for the entire time series was run for the arbitrarily selected stringset's $PR'_{str,d}$ and the $PR'_{sys,d}$.

2.7.4 Rolling CODS analysis

Followed by this, the seasonal component SC_d estimated by the CODS algorithm was used to find the seasonally adjusted $PR'_{str,d}$, or $PR'_{sys,d}$, by employing equation 2.15 and calculating $\frac{PI_d}{SC_d}$. This was done principally because the seasonality was assumed to remain the same for every year and would improve the computation time when running the rolling CODS analysis on performance data. Therefore, CODS analysis for only soiling rate SR_d and degradation rate D_d was then run for the seasonally adjusted $PR'_{str,d}$, the seasonally adjusted $PR'_{sys,d}$, and an arbitrary selection of 10 seasonally adjusted $PR'_{str,d}$, each from a different inverter, with a 3 year window that moves one month forward for each run. Estimated degradation rates D_d for each run were then compared to the estimated degradation rate D_d for the CODS analysis using the entire time series.

2.7.5 Scale factor correction

In order to correct for sudden changes in POAI sensor data due to a calibration or replaced sensor, a correction factor y_α was found to account for such a sudden changes. This was done based on a method proposed by Øgaard et al. [83] where the measured POAI data was compared to modelled POAI during times that were clear sky conditions. To demonstrate the relative difference ΔI in irradiance the following was calculated:

$$\Delta I = \frac{I_{meas} - I_{CS}}{I_{meas}} \tag{2.16}$$

Here, I_{meas} was the measured POAI, and I_{CS} was the modelled POAI, both during clear sky times. Additionally, for each year, a scale factor α for fitting the modelled POAI to the measured POAI was calculated by minimizing the RMSE between the two during clear sky times by:

$$RMSE(\alpha) = \frac{\sqrt{\sum_{i=1}^{n}(\alpha * I_{meas} - I_{CS})^2}}{n} \tag{2.17}$$

A significant shift in sensor data was corrected for by first calculating the scale factors α for each year using equaiton 2.17. The time period with a significant change in scale factor α_t was identified by looking at the change in scale factors $\Delta\alpha$. In order to correct the data from the point with a significant shift, it was necessary to find a correction factor y_α that could be

applied to the data set. The correction factor y_α was found through:

$$y_\alpha = \frac{\alpha_y}{\alpha_t} \tag{2.18}$$

Here, α_y represented the estimated corrected value for α_t. The former was found by first calculating the average rate of change $\overline{\Delta\dot{\alpha}_i}$ for the years prior to the significant data shift. For a set of years $i = 1, 2, ..., n$ before the shift in data t:

$$\overline{\Delta\dot{\alpha}_i} = \frac{\Delta\alpha_{t-1} - \Delta\alpha_{t-n}}{n} \tag{2.19}$$

This was then added to the change in scale factor right before the shift in data $\Delta\alpha_{t-1}$, in order to find the new and corrected scale factor α_y .

$$\alpha_y = \alpha_{t-1} + (\Delta\alpha_{t-1} + \overline{\Delta\dot{\alpha}_i}) \tag{2.20}$$

Chapter 3

Results and discussion

3.1 Data preparation

3.1.1 Data quality inspection

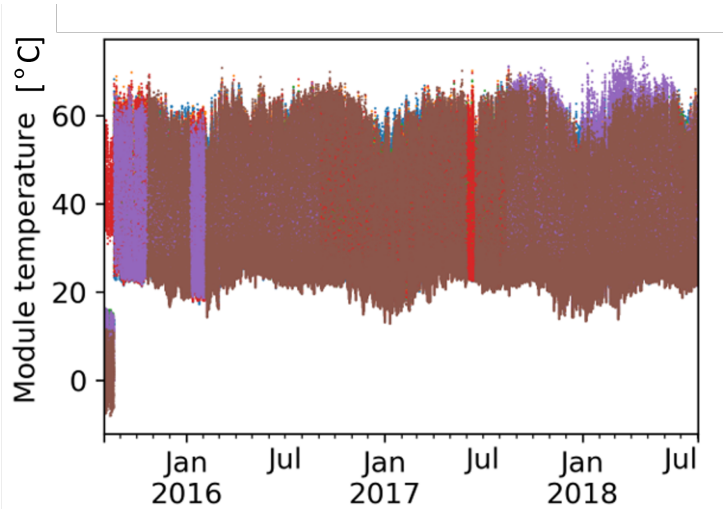


Figure 3.1: 10-minute averages of module temperature measured by 7 different thermocouples. Each color represents a separate thermocouple.

Data sets collected from the site were visually investigated to identify periods with irregularly recorded data. The first 3 years of data for module temperature and POAI can be seen in figure 3.1 and 3.2, respectively. The first month of raw sensor data, August 2015, appeared to be anomalous in most cases by either missing recorded data or seemingly erroneous measurements. Therefore, the first month of raw data was discarded from all data sets.

Measured active power and power set point raw data sets were also checked

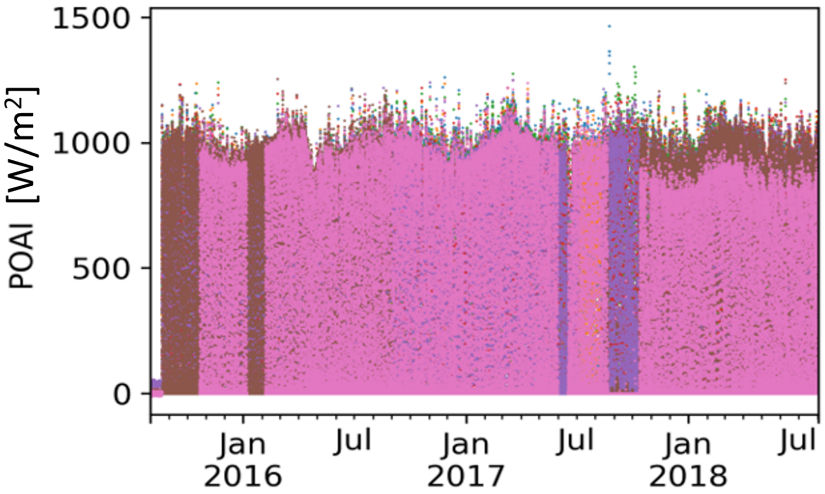
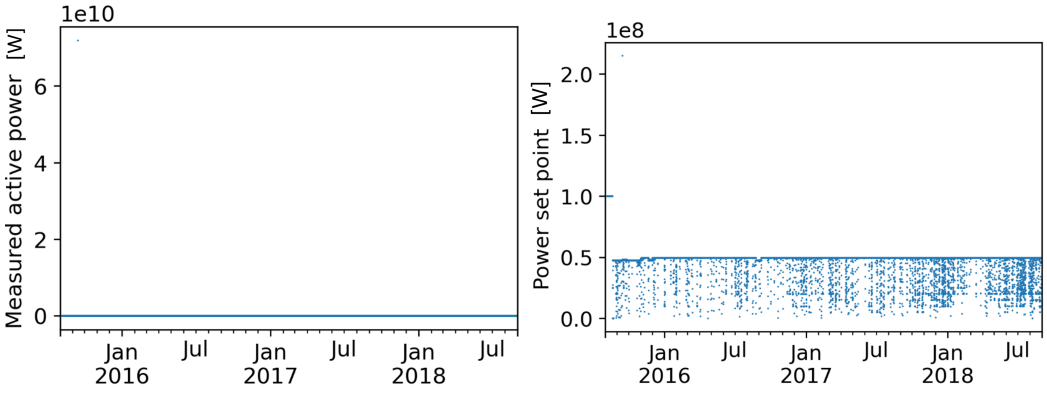


Figure 3.2: 10-minutes averages of POAI measured by 7 different pyranometers, Each color represents a separate pyranometer.Each color represents a separate pyranometer.

and displayed a need for thresholds as well. By looking at the first 3 years, there were some outliers in the raw data sets in 2015 for instance, as seen in figure 3.3.



(a) Measured active power

(b) Power set point

Figure 3.3: Initial quality check of the a) measured active power of the PV plant and b) power set point of the PV plant.

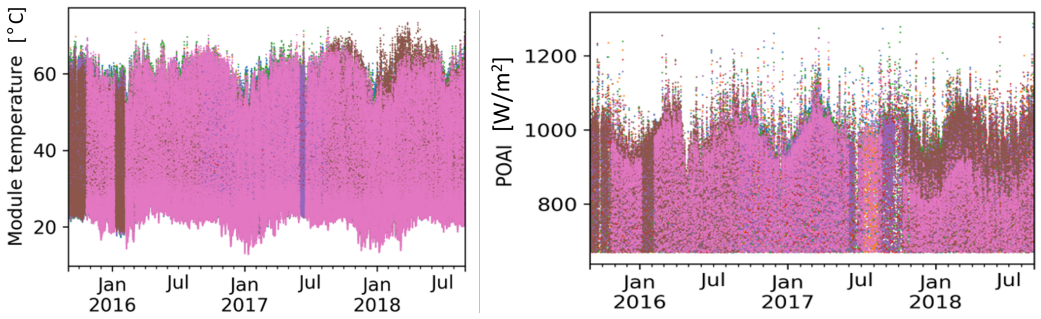
Initial data quality was improved by using a maximum and minimum limit based on recommended thresholds for initial data quality checks [84]. The data thresholds were summarized in table 3.1. Thresholds removed data that were outside the range of -20 and 100 °C for measurements of module temperature and outside the range of 670 to 1300 W/m^2 for measurements

of POAI. For the stringset DC power, the limits were based on the nominal capacity of the stringset $C_{str,STC}$ and varied depending on the size of the stringset. The lower limit was set to 0.04*$C_{str,STC}$ and the upper limit was set to 1.2*$C_{S,STC}$. Thresholds were also used on the measured active power and power set points data sets by using the system nominal capacity $C_{sys,STC}$ as a maximum and 0 as a minimum. 4 out of 7 soiling stations were discarded as the data was too scattered for acceptable quality. The median of the remaining 3 soiling stations was calculated to represent the estimated soiling measured by soiling stations SR'. To be able to be compared to daily soiling estimates, the median of measurements throughout a day was taken as a representation of the daily SR'.

Sensor data type	Minimum	Maximum
Module temperature (°C)	-25	100
POAI (W/m^2)	10	1300
DC stringset power (W)	0.04*$C_{str,STC}$	1.2*$C_{str,STC}$
GHI (W/m^2)	0	1300
Measured active power (W)	0	$C_{sys,STC}$
Power set point (W)	0	$C_{sys,STC}$

Table 3.1: Overview of data thresholds applied to time series data

After the thresholds were applied, the data sets were reexamined. Sensor data for module temperature and POAI after applying thresholds from 3.1 can be seen in figure 3.4a and 3.4b, respectively. Unreasonably high or low data points were not present in the raw data sets any more.



(a) Module temperature

(b) POAI

Figure 3.4: Sensor data after applied thresholds in table 3.1 for a) 7 different thermocouples. Each color represented a separate thermocouple, and b) POAI measured by 7 different pyranometers. Each color represented a separate pyranometer.

3.1.2 Data filtration

Different filters were used to reduce noise in the data set and to improve the quality of the time series before the use of CODS. A suitable suite of data filters was established by examining the effect of each filter on the data sets. The following data filters were examined for use; AOI filter, solar elevation filter, clear sky filter, curtailment filter, and clipping filter.

By using the irradiance.aoi() method from Pvlib (version 0.9.5), the solar path throughout the year was calculated for the site. For the AOI filter, a maximum AOI of 55° was used, and any data points associated with a higher AOI would be filtered out. For the solar elevation filter, a minimum solar elevation angle α_s of 22° was used, and any data points with a lower solar elevation would be filtered out. Addtionally, a recommended threshold in the IEC standard [84] for POAI data of 670 W/m^2 was investigated.

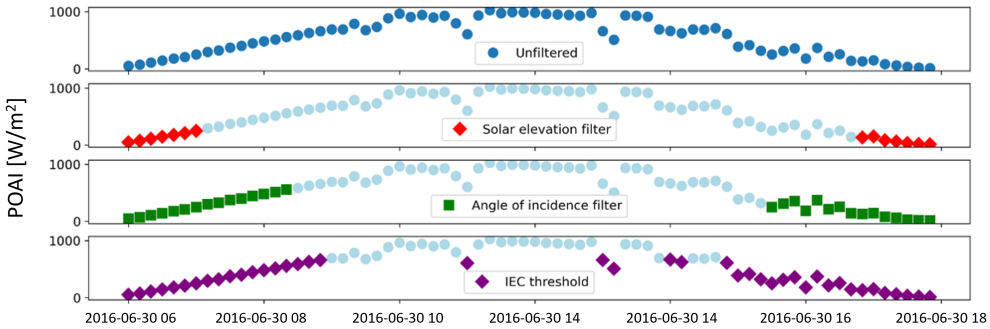


Figure 3.5: Demonstrating the effect of the solar elevation filter, AOI filter, and IEC threshold on the POAI throughout a day in June 2016.

In figure 3.5, the effect of different filters can be seen on the POAI data on a chosen day. The colored data points show which data each filter removes from the time series. Firstly, the solar elevation filter appeared to remove the least amount of data points, while the IEC screening threshold of 670 W/m^2 removed the most data points. Not only did the IEC threshold remove the most data points towards the beginning and end of the day, but also data points where the POAI dropped during in the middle of the day. This made it difficult to gauge the effect of applying filters including the IEC screening threshold. It was determined that the IEC threshold was set too high for the effects of the other filters to be determined. Therefore, the IEC threshold was substituted by a cutoff threshold of 10 W/m^2 as a measure to remove background noise from the data sets. Furthermore, the AOI filter appeared

to filter out more data points throughout the day compared to the solar elevation filter. It seemed therefore unnecessary to apply both the AOI and solar elevation filter, when the AOI filter would suffice.

Clear sky times were found by using location.get_clearsky() method from Pvlib (version 0.9.5), as described by Reno and Hansen [85], using a clear sky period of 30 minutes. The median across all GHI sensors was calculated and used to represent the GHI in the PV system. The method then used the measured GHI data and compared it to modelled GHI data for the location to calculate clear sky times and could then be used as a filter. In figure 3.6, the effect of the clear sky filter on the measured POAI was displayed. The clear sky filter appeared to remove data points that were not a part of a smooth GHI curve, which is expected throughout a day. However, it also appeared to remove some data points that did were not certain to be non-clear sky times. This was indicative of the notion that the clear sky filter would remove many data points compared to the other filters.

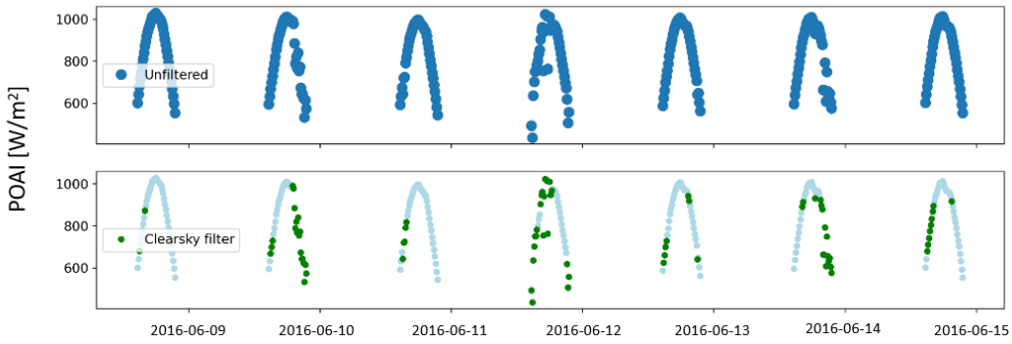


Figure 3.6: Demonstrating the effect of the clear sky filter. The top window shows the measured POAI data for a week in February 2016 and the bottom window shows which data points, colored in green, are removed by the clear sky filter.

Curtailment was detected using the method proposed by Nygård et al. [10] using two conditions to identify curtailment. The conditions were firstly if the set point was below 5 % of the AC system nominal capacity $C_{sys,STC}$, and secondly if the measured active power was below the $C_{sys,STC}$, but higher than 95 % of the set point. If either condition was fulfilled, then the data was associated with curtailment and subsequently filtered out. In figure 3.7, the effect of the curtailment filter was shown on the stringset DC power generation, by using the aforementioned conditions on the PV system's measured active power and power set point. The filter appeared to identify data points

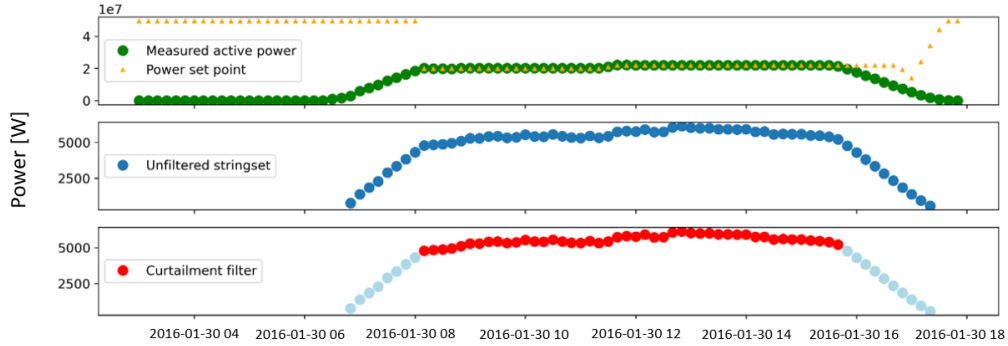


Figure 3.7: Demonstrating the effect of the curtailment filter. The top window shows the power set point and the measured active power of the plant. The middle window shows the stringset DC power without the curtailment filter and the lower window shows which data points, colored in red, were removed by the curtailment filter.

related to curtailment as seen by the flattened section in the measure active power, and compared to the DC power generation of a stringset. Even though the power set point was kept at the lowered level for a while after 16:00 in the afternoon, the filter managed to avoid removing these data points as the measured active power was dropping below the power set point and and no longer showed signs of curtailment.

Inverter voltages were investigate for signs of clipping, but were not found to display signs of clipping. Therefore, it was determined that the use of a clipping filter was not necessary as it would not result in a significant reduction in noise and was no longer considered for the final selection of filters. It was also a concern at this point that a further reduction in data would reduce the quality of the time series before the use of CODS.

The final selection of methods for reducing noise before calculating the performance metrics were; the thresholds found in table 3.1, an AOI filter, a clear sky filter and a curtailment filter. In the final selection of filters, the clear sky filter removed the most data points indicating that the PV system experienced a significant amount of cloud coverage, or other non-clear sky conditions like shadowing. Following the clear sky filter, was the AOI filter in terms of amount of raw data points removed, and expectedly removed more data points than the solar elevation filter considering figure 3.5. To clarify, the amount of data points listed next to each method for data removal in table 3.2 was after only applying the individual method. That meant that in the final selection of methods, the same data point could for instance be filtered by both the AOI filter and the clear sky filter. The curtailment filter

Table 3.2: The effect of filters and thresholds on the amount of removed raw data summarized for the different methods investigated and the final selection of methods selected for raw data removal.

Method for raw data removal	Amount of data points remaining
Without data removal	1,658,666,704
Data thresholds from table 3.1	765,882,421
AOI filter	425,702,624
Solar elevation filter	548,181,032
Clear sky filter	198,152,804
Curtailment filter	638,869,144
Final selection of methods	71,981,112

showed that the PV system also experienced some curtailment which corroborated the need for using a curtailment filter. As seen in table 3.2, the final selection of filters removed close to 90.6 % of data points after having applied the thresholds from table 3.1. Therefore, there were likely numerous days that would not contain any data points.

Prior to the calculation of the 10-minute stringset PR'_{STC}, the use of filters and data thresholds were used in an effort to remove unreasonable and noisy data points from sensor data. However, a numeric indication of whether the individual filter or threshold improved the quality of the data set was not specifically calculated. Rather, the visual effect on the 10-minute stringset PR'_{STC} for the whole PV system was the deciding factor by considering the change in apparent noisy data points. This would affect the 10-minute stringset PR'_{STC} before aggregation to $PR'_{str,d}$, but the effect was considered minimal. This was due to the fact that taking the median of each 10-minute stringset PR'_{STC} throughout a day to represent the $PR'_{str,d}$, was a robust practice that was less sensitive to outliers. Expecting the complete removal of outliers in data sets when using real data was not considered reasonable as data measured from real PV systems, and with real sensors, would likely experience artefacts that could not always be expected to be accounted or corrected for. Efforts to establish a standard methodology for filtration exist, but have yet to be finalized [7]. Therefore, the effect of data filtration comes down to the individual analyst who must determine the critical values by which each filter relies on and whether or not this improves the quality of the data set before analysis. Alternatively, an optimization process could be performed where a calculation of a signal-to-noise ratio is done and iteratively finding the optimized critical values for each filter.

Through filtering and applying thresholds, the noise in the data sets was reduced and consequentially the performance was estimated based on somewhat ideal conditions, like a clear sky and without curtailment. The fact remains that real PV systems will experience variable conditions, which can lead to a discussion on whether an estimated degradation rate based on ideal conditions can be used to estimate power generation from a real PV system. Such a degradation rate would be better used to described the condition of the PV system rather than a method for predicting long-term production values. The variability of factors like soiling or downtime due to maintenance can change drastically from one year to the next, which would complicate estimating future production values significantly.

A study by Jordan and Kurtz [36] investigated the effect of data thresholds and performance metrics on the estimated degradation rate for a PV system. The study showed the dependence on the different critical values chosen in each threshold can have a significant effect on the estimated degradation rate. They also demonstrated how the uncertainty rises when reducing too many or too few data points from the raw data set. This indicated that the final selection of filters and thresholds could be on the lower side of a preferred amount of data points for degradation analysis. However, the 10-minute stringset PR'_{STC} was deemed sufficient for analysis as the effect of each filter had been investigated and the CODS algorithm was capable on running the different filtered time series used in the following chapters.

3.2 Soiling and performance

The soiling station estimates SR' calculated from soiling stations were shown in figure 3.8a. There was a significant increase in the SR' at around March every year followed by a gradual further increase until peaking at around June, The first rapid increase likely due to a regularly scheduled cleaning campaign after the dry winter season, while the gradual increase could be explained by the gradual reduction of dust particles in the air as time approached the rainy summer season. Nevertheless, it was warranted to account for soiling in the PV system when investigating the performance and ultimately the degradation of the PV system. An arbitrarily chosen $PR'_{str,d}$ was also calculated and displayed in 3.8b. The $PR'_{str,d}$ appeared to have a seasonal yearly behavior with increases in performance coinciding with the SR' during the first years until mid 2020. In 2021 and 2022, the $PR'_{str,d}$ appeared more scattered and the seasonality was not as evident. Looking at the 1-

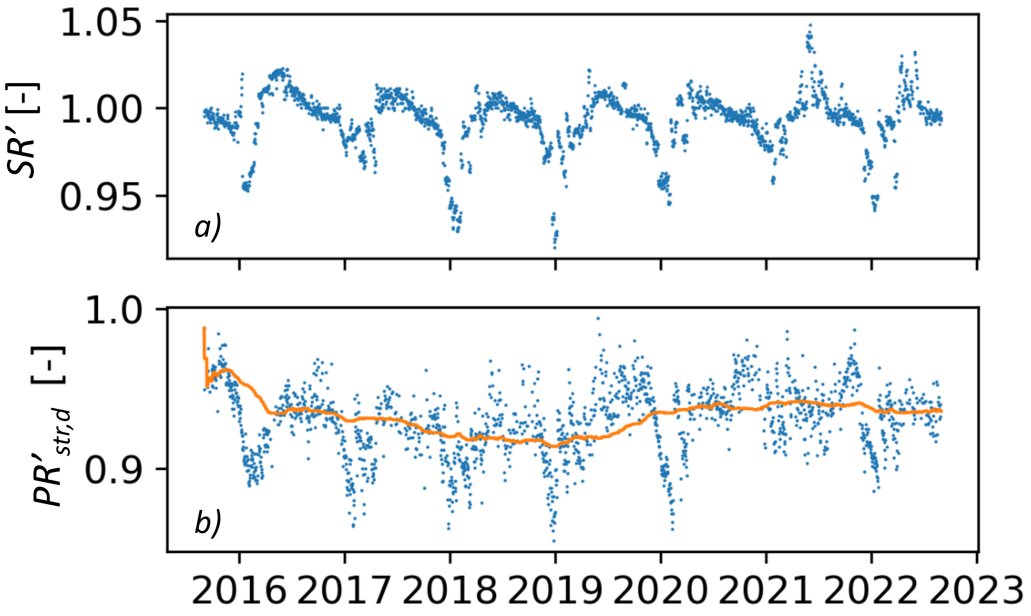


Figure 3.8: Soiling station estimates SR' and an arbitrarily chosen $PR'_{str,d}$ with a 1-year rolling median shown by the orange line.

year moving median, there was also an apparent increase in the $PR'_{str,d}$ from around 2019 through 2020. Concerning increased performance, Theristis et al. [86] reported that underrated modules are present in the market and exist among multiple producers. Underrating modules would allow the modules to degrade for a longer period of time before a possible warranty claim can be made by the plant operator or installer. They observed that the degradation rate would stabilize after 3-4 years. This means that positive degradation rates were possible, but would likely only be seen in the first couple of years before degrading normally. A change in performance could also occur due to significant changes in the PV system, like a module replacement campaign or a change in the sensor data used to calculate the performance metric, like replacement or a calibration. It was also unlikely that the data shift was a result of the aggregating method in equation 2.11, as the robust practice made it less sensitive to outliers and a such a shift would have to affect the majority of data before aggregation. Considering equation 2.10, there were various sources that may have caused an increase in the $PR'_{str,d}$ from the sensor data used in the equation. Either one or all of module temperature, POAI or DC power measurements could have contributed to such an apparent change in the $PR'_{str,d}$.

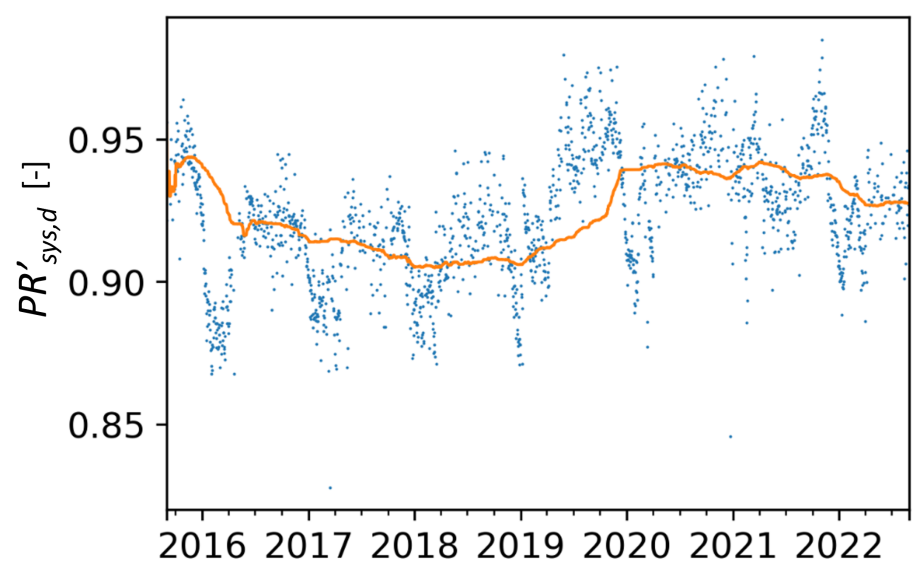


Figure 3.9: The $PR'_{sys,d}$ with a 1-year rolling median in orange.

In figure 3.9, the $PR'_{sys,d}$ was calculated and displayed a stronger increase in the 1-year rolling median during 2019 compared to the $PR'_{str,d}$ in figure 3.8. This indicated that there were likely other stringsets with an increased $PR'_{str,d}$ during 2019 than the arbitrarily selected stringset in figure 3.8. The increased $PR'_{sys,d}$ in 2019 could also have been a result of reduced soiling, but this notion was not supported by the SR' in figure 3.8, which peaked mid-2019 before falling until the end of the year. It was difficult to discern the role in which soiling played merely by analyzing the performance trends in figures 3.8 and 3.9. Even though a significant shift in the performance appeared to be present in the data sets, both $PR'_{str,d}$ and $PR'_{sys,d}$ were investigated using CODS to see if the role of soiling could be better understood in the performance data.

3.3 CODS analysis

In figure 3.10, CODS analysis results for the arbitrarily chosen $PR'_{str,d}$ in figure 3.8 were shown, resulting in an estimated degradation rate D_d of -0.059 %/year. The run time for CODS in the $PR'_{str,d}$ was 4.5 min. The seasonal component SC_d appeared to follow the $PR'_{str,d}$ more closely during 2019 through 2022 than compared to the beginning of the time series until 2019. The fitted model PI_d followed the SC_d closely during the peak performance periods in 2015, 2019, 2020 and 2021. This behavior appeared to suggest

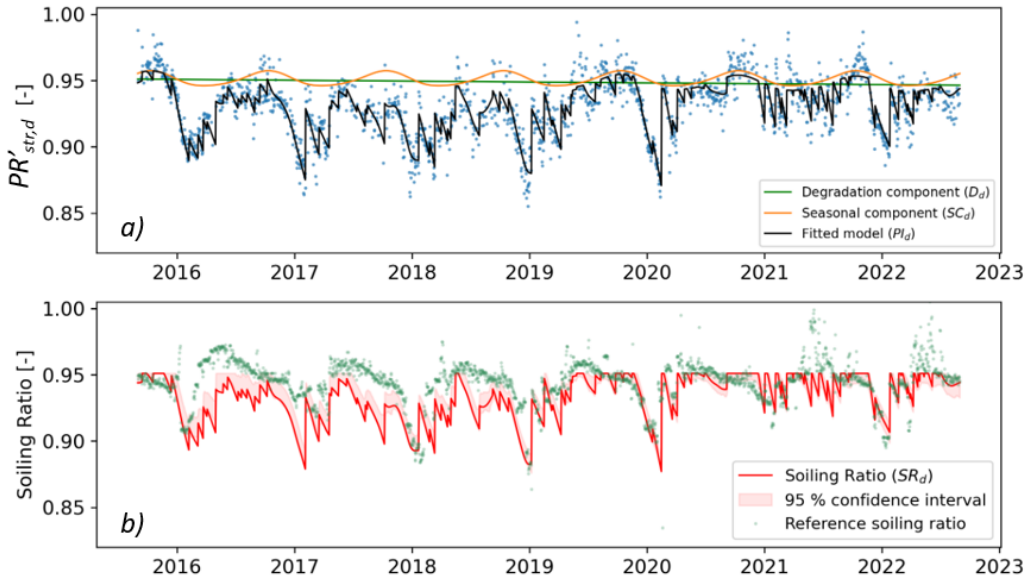


Figure 3.10: CODS results for a $PR'_{str,d}$ displayed for the a) daily seasonal component SC_d, degradation component D_d and total fitted model PI_d in orange, green and black lines, respectively. b) The daily soiling component SR_d was shown by the red line with a 95 % confidence interval indicated by the red shaded area, along with daily soiling station estimates SR' asa reference. The estimated degradation rate D_d was calculated to be -0.059 %/year, with a 95 % confidence interval between -0.170 %/year and 0.053%/year.

that the SC_d restricted the PI_d during the peak performance periods, as seen my the smoothed PI_d in the shape of the SC_d. The PI_d also appeared to follow the $PR'_{str,d}$ well until 2019, 2020 and 2021 where it estimated a lower the performance, especially seen during the peak performance periods in the middle of each year. The estimated soiling component SR_d appeared to have a periodic drop during or close to the end of each year and start of the next, which correlated with the periodic drops in the soiling station estimates SR'. Otherwise, the SR_d appeared to follow the general trend of the SR', but seemed to estimate a higher degree of soiling during peak $PR'_{str,d}$ periods. A different point of interest is that the SR' had more scattered data points during 2021 and 2022, and coincided with the PI_d and the SR_d which displayed a sawtooth-trend without any significant drops during these years. An exception to this behavior was seen in the start of 2022 when a significant drop in SR' and SR_d also correlated with a drop in the $PR'_{str,d}$ and PI_d. In fact, it appeared that the CODS algorithm managed to pick up on such a drop in SR' at the beginning of each year, again, likely due to a regularly scheduled cleaning campaign of the PV system after the dry winter season.

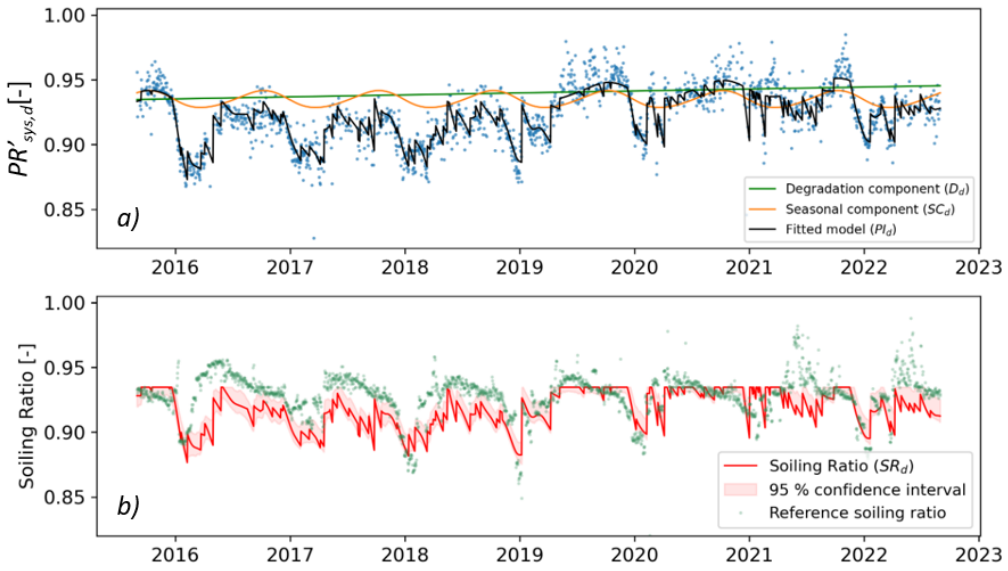


Figure 3.11: CODS results for $PR'_{sys,d}$ displayed for the a) daily seasonal component SC_d, degradation component D_d and total fitted model PI_d in orange, green and black lines, respectively. b) The daily soiling component SR_d was shown by the red line with a 95 % confidence interval indicated by the red shaded area, along with daily soiling station estimates SR' as a reference. The estimated degradation rate D_d was estimated to be 0.181 %/year, with a 95 % confidence interval between 0.064 %/year and 0.298 %/year.

In figure 3.11, CODS analysis results for the $PR'_{sys,d}$ were shown, resulting in an estimated degradation rate D_d of 0.181 %/year. Much like the CODS analysis for the $PR'_{str,d}$ in figure 3.10, the estimated seasonal component SC_d appeared to favor the $PR'_{sys,d}$ during the later years of the time series. However, this time it was more clearly shown that the fitted model PI_d did not follow the $PR'_{sys,d}$ during the peak performance periods in 2019, 2020 and 2021 compared to the fitted model PI_d in 3.10. This did not agree with the possibility of reduced soiling as the trend in the soiling station estimates SR' appeared more or less unchanged until 2021. The estimated soiling component SR_d also appeared to estimate a higher degree of soiling during peak performance periods of 2016, 2017 and 2018. An important point to emphasize was that the SR' was calculated from soiling stations which were point estimates from the entire PV system. This meant that they were not likely to represent the soiling experienced by every single stringset, but functioned as a point of reference to which the SR_d could be compared to. A positive D_d for the $PR'_{sys,d}$ seemed unlikely as the performance was expected to fall gradually over time, and considering that the PV system lied in a hot and humid climate. A significant shift in data points during 2019 seen in

figure 3.9 could have caused the different estimated time series components in CODS to skew the results towards a section before or after the shift, thus not being able to fit the entire time series. Further investigation of what may have been the cause of the shift in $PR'_{sys,d}$ was necessary to better explain the D_d of the system.

3.4 Calculation of correction factor y_α

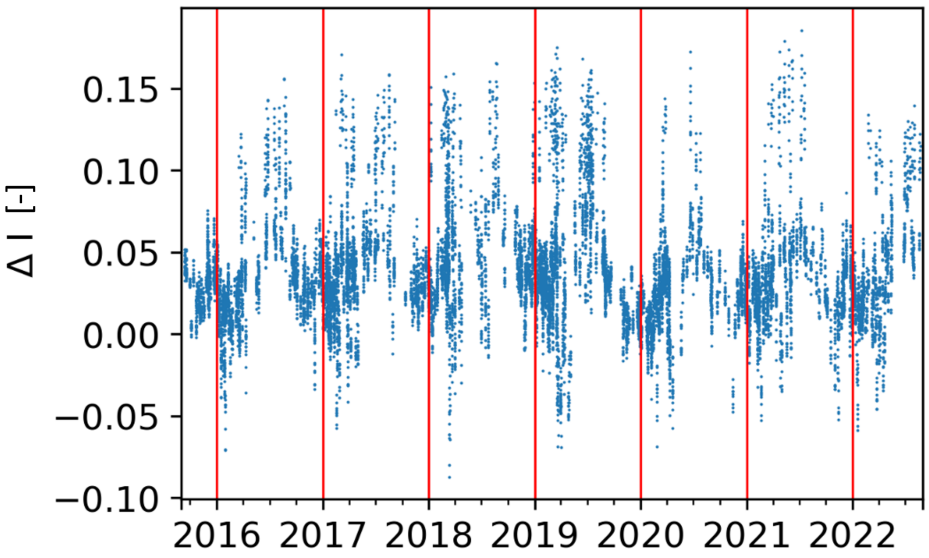


Figure 3.12: The difference ΔI between measured POAI I_{meas} and modelled POAI I_{CS} during clear sky times. Vertical red lines are included to distinctly show where each year ends and the next begins.

Module temperature, POAI and DC power data sets used to calculate the 10-minute PR'_{STC} were investigated for any noticeable shifts. It was determined that the variation in module temperature did not change significantly, while the DC power appeared to have a negative trend over times as expected, possibly due to degradation of the modules. The data set for POAI did not show any clear evidence of a trend. Therefore, POAI data sets were further investigated for determining any data shifts based on the method proposed by Øgaard et al. [83], described in section 2.7.5.

In figure 3.12, ΔI was calculated for the measured POAI data I_{meas} and the modelled POAI data I_{CS} during clear sky times. The red lines show where the data was segmented by each year's end. There were no evident changes to the ΔI, which was expected as the POAI data was not found to display

any change in trend. Each section was used to find the scale factor α for each year using equation 2.17. The α found for each year and change in scale factor $\Delta\alpha$ for each year were shown in table 3.3, with omission of 2015 and 2022 as the POAI time series did not include the entirety of these years.

Table 3.3: Yearly scale factors α and change in scale factors $\Delta\alpha$ compared to the previous year, all rounded to closest third decimal.

Year	Scale factor α	Change in scale factor $\Delta\alpha$
2016	1.029	-
2017	1.034	0.005
2018	1.044	0.010
2019	1.035	-0.009
2020	1.024	-0.011
2021	1.029	0.005

In table 3.3, the calculation of each scale factor showed that there was a significant shift in the scale factor α during 2019. Prior to that, the α was increasing for each year, which could indicate a sensor drift. In 2019 it dropped instead of increasing and continued to do so in 2020 as well. According to equation 2.9 and 2.10, a lower POAI would explain the increased performance during this period. However, it was difficult to say what the source of the change in POAI came from without verification from the plant operator, but a correction was warranted for giving an accurate estimation of the estimated degradation rates D_d using CODS.

By using equations 2.18, 2.19 and 2.20, a α_y of 1.058 was estimated and a correction factor y_α of 1.022 was applied to all 10-minute stringset PR'_{STC} from January 1st 2019 and throughout the time series. To reiterate; all of the results in the following sections included a correction factor y_α.

In figure 3.13, the $PR'_{str,d}$ with applied correction factor y_α was shown. The change in 2019 was no longer as clear, but the irregularity in 2020 and 2021 was stilled present. The former was expected as the y_α would simply change all data points by the same magnitude. The 1-year rolling median now showed a steady decrease until around March 2019, where it started to increase slightly until around June 2020. From here on, the 1-year rolling median appeared to continue stably. The y_α appeared to have caused a form of correction for the increasing trend in 2019 through 2020 for the $PR'_{str,d}$, and results from CODS analysis were expected to be more representative with the applied y_α.

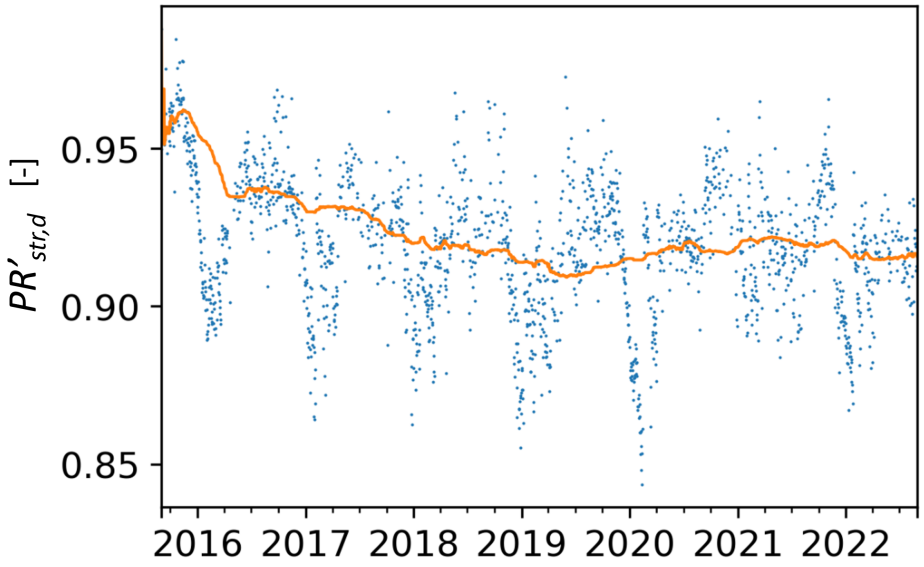


Figure 3.13: $PR'_{str,d}$ after having applied a y_α from 2019 and onward, in addition to a 1-year rolling median shown by the orange line.

Sensor drift, calibration or replacement could all occur during the lifetime of a PV system and could affect performance metrics unless corrected for [72]. This information was not available, and a form of correction was deemed necessary for further analysis. Correcting the PR'_{STC} time series further could have been done, as a decreased change in scale factor $\Delta\alpha$ was seen for 2020 as well. The procedure using equations 2.16 and 2.17, could have been performed after the first correction and subsequently used to find a second correction factor to be applied from 2020 and on. However, it became necessary to consider the fact that the scaling factors α were calculated by fitting the modelled POAI I_{CS} to the measured POAI I_{meas} during clear sky times. If the significant shift in data actually occurred in the middle of the year, then an applied correction from the start of year would change a large portion of data that did not need to be corrected. It was therefore questionable at this point how many corrections were necessary to account for in the data sets. From an O&M standpoint, sensor calibrations and changes would need to be kept track of and used to correct performance data. An alternative method to correct for the POAI data, would have been to utilize satellite data and estimate the GHI for the site, and then transpose the irradiance to the plane-of-array. This way, the POAI data would not be affected by local sensor drift and calibrations, However, one would be reliant on the choice of model and accuracy of the data, as satellite data estimates the irradiance

for an area [87] [88]. The method used for finding and applying a correction factor y_α did not specify how large of a shift in α would warrant a correction. An alternative way of implmenting this method could have been to apply an individual y_α for each year in an effort to correct for any data shift, including sensor drift or calibrations. The question would then remain whether all years were to be corrected according to the first year's α or some other way.

The fact remains that a real PV system will see calibrations to sensor data throughout the system's lifetime, and should be considered important for any degradation analysis thereof. In the IEC standard for PV system performance [34], it is specified that the highest standard of irradiance sensor maintenance requires a calibration each year, and every other year for the second highest standard. Consequently, any degradation analysis performed over multiple years would likely need to include corrections for data shifts due to irradiance sensor maintenance. A clearer guideline or methods for performing sensor corrections is of importance to ensure higher quality estimates of degradation rates. Due to consideration and the time constraint of the thesis, any further corrections were not performed on the 10-minute PR'_{STC} time series. The $PR'_{str,d}$ and the $PR'_{sys,d}$ with applied y_α were subject to CODS analysis in order to investigate whether the correction had caused a better fit in the different estimated time series components.

3.5 CODS analysis with applied correction factor y_α

In figure 3.14, CODS results for the $PR'_{str,d}$ with the applied correction factor y_α were shown with an estimated degradation rate D_d of -0.411%/year. Compared to the CODS results for the $PR'_{str,d}$ prior to applying the y_α in figure 3.10, the D_d was now expectedly lower. Similar to figure 3.10, the fitted model PI_d followed the SC_d closely during the peak performance periods in 2015, 2020 and 2021, but not in 2019 any more. Therefore, the SC_d appeared to govern the trend of the PI_d during peak performance periods in 2020, and 2021, reminiscent of peak shaving even after the applied y_α. The SR_d still seemed to estimate a higher degree of soiling compared to the daily soiling station estimates SR'. Keeping in mind that the SR' was based on soiling stations placed throughout the PV system, it was possible that the chosen stringset in figures 3.10 and 3.14 was experiencing a higher degree of soiling than the soiling stations. If the stringset was among the first rows facing a particular wind direction, it was reasonable to expect that a higher

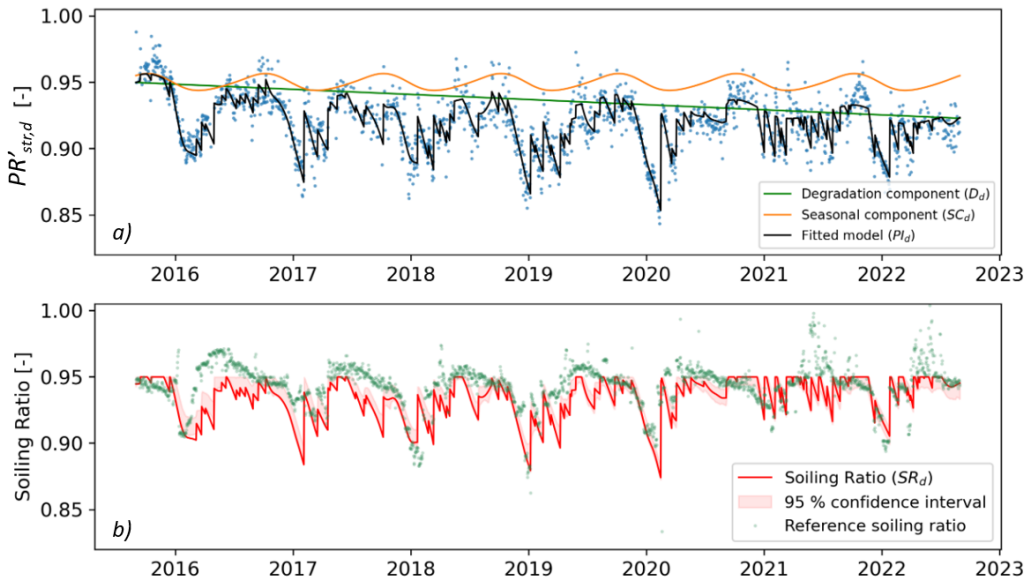


Figure 3.14: CODS results for an arbitrarily chosen $PR'_{str,d}$ after applied correction factor y_α. a) The daily estimated seasonal component SC_d, degradation component D_d and total fitted model PI_d in orange, green and black lines, respectively. b) The daily estimated soiling component SR_d was shown by the red line with a 95 % confidence interval indicated by the red shaded area, along with daily soiling station estimates SR' as a reference. The estimated degradation rate was estimated to be -0.411 %/year, with a 95 % confidence interval between -0.524 %/year and -0.302%/year.

degree of soiling would occur for modules in this area. Both the PI_d and the SR_d showed a noisy sawtooth-trend in 2021 and 2022 with exception of the drop during the beginning of 2022. This was likely due to the $PR'_{str,d}$ data being too scattered for the CODS algorithm to discern a soiling trend in the data set, which impacted the SR_d and thus the PI_d.

CODS attempts to account for soiling when estimating the degradation rate D_d in a time series of performance data. However, the influence of maintenance on the data that was used to calculate the PR'_{STC} was a significant factor and was demonstrated by the results. Sensor drift, replacement and calibration are all known to happen in real world PV systems, and have been discussed by Jordan et al.[72]. Here, a data shift in performance data was demonstrated and different methods for correction were evaluated. The study found, among other conclusions, that the YoY-method applied to different types of performance ratios PRs was less sensitive to bias from data shifts than OLS. On the other hand, figures 3.10 and especially 3.11 would indicate that the YoY-method can still be significantly biased by data shifts

even when applying robust data aggregation practices. The study by Jordan et al. also suggested that a performance ratio based on modelled POAI and modelled cell temperature would be independent of sensor drift and data shifts. This metric could however only be used during clear sky times, thus being an inadequate choice for a PV system experiencing many hours of non-clear sky conditions.

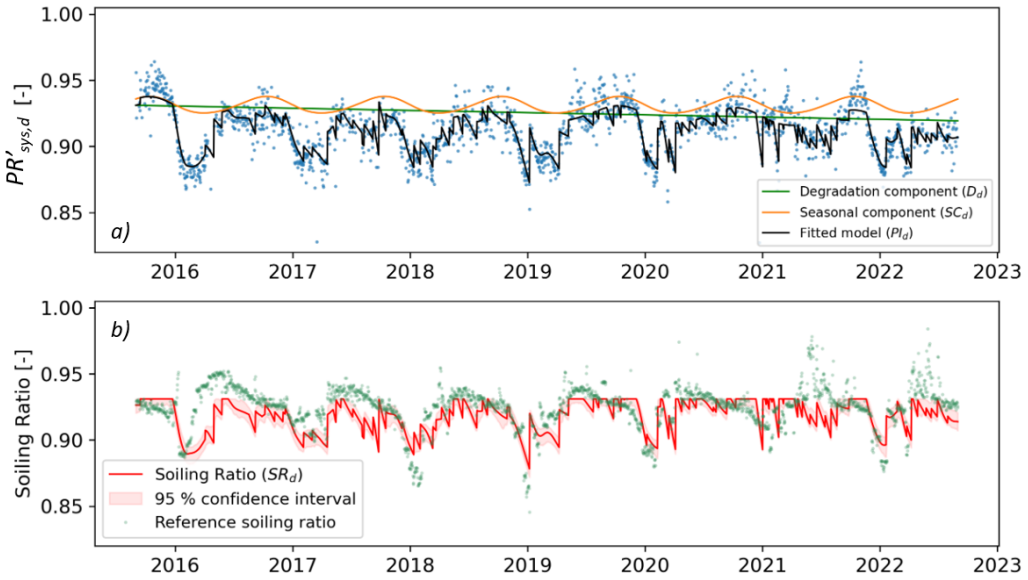


Figure 3.15: CODS results for $PR'_{sys,d}$ after applied correction factor y_α. a) The daily seasonal component SC_d, degradation component D_d and total fitted model PI_d in orange, green and black lines, respectively. b) The daily soiling component SR_d was shown by the red line with a 95 % confidence interval indicated by the red shaded area, along with daily soiling station estimates SR' as a reference. The estimate degradation rate was calculated to be -0.177 %/year, with a 95 % confidence interval between -0.278 %/year and -0.071%/year

In figure 3.15, CODS results for the $PR'_{sys,d}$ with the applied correction factor y_α were shown with an estimated degradation rate D_d of -0.177 %/year. Although the D_d was no longer positive like in figure 3.11, it was still not as negative as in figure 3.14. This further supported the notion that there was a variation in daily stringset D_d for the stringsets in the PV system. The estimated soiling component SR_d appeared to follow the general trend of the soiling stations estimates SR', but estimated slightly higher soiling during peak performance periods shown by the PR'_{STC}. The fitted model PI_d in 3.14 appears to follow the performance better than the PI_d in figure 3.15, likely due to the SC_d influencing the PI_d as previously seen.

3.5.1 Rolling CODS analysis

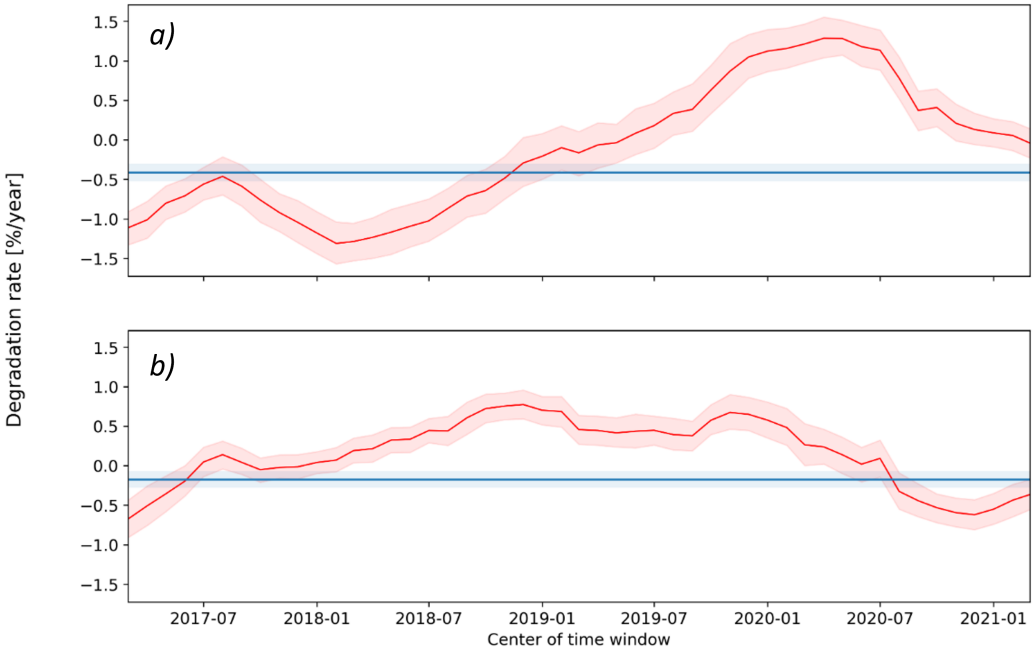


Figure 3.16: Degradation rates D_d estimated by the rolling CODS method for a) seasonally adjusted $PR'_{str,d}$ and b) seasonally adjusted $PR'_{sys,d}$, shown by the red lines, compared to the linear estimated degradation rate D_d based on the entire time series, shown by the blue line. Both data sets had an applied correction factor y_α and included a 95 % confidence interval shown by the shaded areas. The degradation rates D_d were plotted using the center of the time window.

In figure 3.16a, estimated degradation rates D_d by the rolling CODS method for the seasonally adjusted $PR'_{str,d}$ was shown. and was the same stringset seen in figure 3.14. The D_d based on the entire $PR'_{str,d}$ time series from figure 3.14 was also shown and clearly showing the differences over time. The run time for rolling CODS on the seasonally adjusted $PR'_{str,d}$ was around 2.5 h and the D_d varied between -1.311 %/year and 1.285 %/year. The largest difference in D_d appeared to be when the center of the time window was around mid 2020. The figure showed that a steady increase in D_d was estimated from when the center of the window was around early 2018 and peaking around May 2020. Afterwards, the D_d decreased steadily back towards a degradation rate of 0 %/year. Comparing to the $PR'_{str,d}$ in figure 3.13, the start of 2020 appeared to be about halfway of the increasing trend starting in 2019 and going into 2020. After this period, the 1-year rolling median appeared to level off somewhat, but the D_d in 3.16a would indicate that the performance was declining again. The difference in these results could be explained by

soiling of the system, but remained uncertain. This was due to both the soiling estimates from soiling stations SR' and estimated soiling components SR_d from the CODS analysis in figure 3.14 appeared to be scattered in 2021 and 2022. The correction factor y_α applied to the time series did not account for sensor drift of the POAI, which was indicated in table 3.3. The combined effect of sensor drift could explain the increasing estimated degradation rate in 3.16a, but could also be an artefact of other, more subtle, calibrations or changes to the sensor equipment.

In figure 3.16b, estimated degradation rates D_d by the rolling CODS method for the seasonally adjusted $PR'_{sys,d}$ was shown as well as the D_d based on the entire $PR'_{sys,d}$ time series from 3.15. The D_d varied between -0.674 %/year and 0.773 %/year for the time series and the general trend for the D_d was reminiscent to the trend shown in figure 3.16a, but peaked at two different times instead. The peaks observed in 3.16b appeared when the center of the time window was around the end of 2018 and end of 2019. The latter was observed as a shoulder to the global peak in figure 3.16a, but the former only barely appeared as a local peak in a otherwise increasing trend. This further indicated that there was a distribution of degradation rates for the stringsets in the PV system.

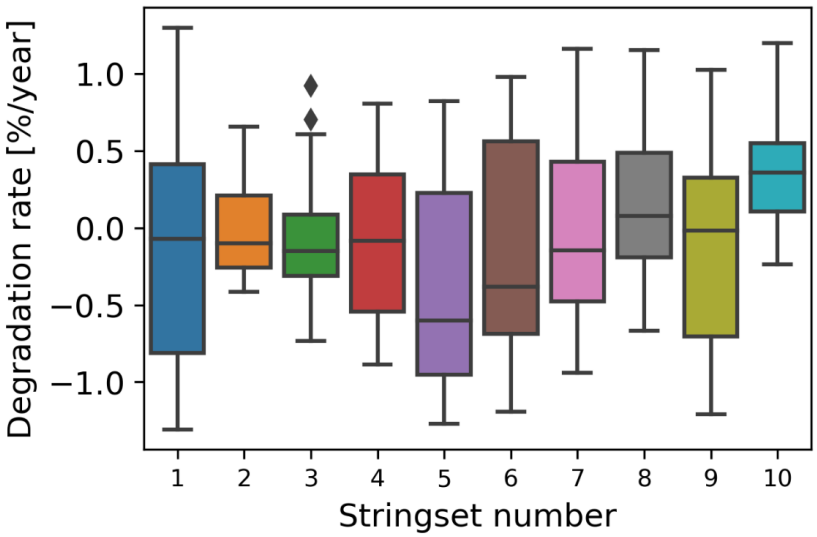


Figure 3.17: Boxplots of 10 arbitrarily selected $PR'_{str,d}$ degradation rates D_d estimated by the rolling CODS method, and each from a different inverter in the PV system.

A selection of 10 stringsets' seasonally adjusted $PR'_{str,d}$, each selected arbitrarily from different inverters, was further investigated to better understand

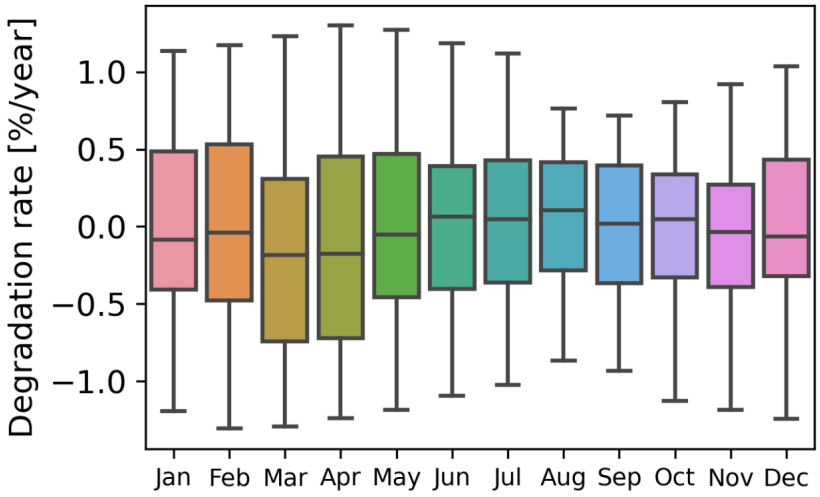


Figure 3.18: Boxplots of the monthly distribution of degradation rates D_d from 10 arbitrarily selected $PR'_{str,d}$ estimated by the rolling CODS method, and each from a different inverter in the PV system.

the distribution of degradation rates D_d for the stringsets and how this related to the D_d in figure 3.16b. Figure 3.17 showed boxplots demonstrating the distribution of D_d using the rolling CODS method for each selected stringset. Here, stringset nr. 1 was the same stringset shown in figures 3.8, 3.10, 3.13 and 3.14. The D_d varied from -1.311 %/year and 1.285 %/year. The distribution in figure 3.17 showed that D_d varied depending on the stringset in question. This was expected considering the differences between CODS results for the $PR'_{str,d}$ and for the $PR'_{sys,d}$ in figures 3.14 and 3.15, respectively.

In figure 3.18, the same stringsets' estimated degradation rates D_d in figure 3.17 were shown distributed by month. Here, the distribution throughout the year did not appear to be significantly different for any month. The D_d IQR for each month was within the range of -0.741 %/year and 0.532 %/year, but none appeared to be significantly different to each other. Rather, compared to figure 3.17, the temporal variation in in the D_d was smaller than the total variation on D_d for each stringset.

Literature values of degradation rates for crystalline silicon technologies in hot and humid areas have been reported to range between -2.10 to 0.27 %/year [73]. Notably, this range was reported for a variety of crystalline silicon technologies installed from 1980s until the 2010s. The study also showed that longer field exposure for modules resulted in decreased estimated degradation. The estimated degradation rates D_d for the corrected $PR'_{str,d}$ and

$PR'_{sys,d}$ were found to be -0.411 %/year and -0.177 %/year, respectively, and are both within literature values, but on the less negative end. This could possibly be due to longer field exposure in the range of mentioned literature values compared to the analyzed PV system in this thesis and the influence of the filters used on the raw data sets, particularly the clear sky filter which was the filtered that removed the most data points by itself. The two D_d were compared to a study using CODS for estimating degradation rates in a PV system situated in a hot and humid climate [10]. In the study, the median D_d for all inverters was -0.12 %/year, which was less negative than the D_d for the $PR'_{sys,d}$ in this thesis, being the median for all stringsets in the analyzed PV system. For the rolling CODS method, the estimated D_d varied between -1.311 %/year and 1.285 %/year for the seasonally adjusted $PR'_{str,d}$, and between -0.674 %/year and 0.773 %/year for the seasonally adjusted $PR'_{sys,d}$, respectively. Therefore, the linear estimations were within literature values, but this was not entirely the case for the range of degradation rates from the rolling CODS. It was expected that the PV system exhibited non-linear behavior in terms of degradation partly due to the influence of soiling of climatic conditions. The results were generally small in terms of degradation, and many were positive. This can possibly be due to the influence of sensor measurements, like drift, calibrations or replacements, or underrated modules.

Highlighted in figure 3.16, the rolling CODS method shows that if assuming the degradation to be linear, then the first few years will likely exhibit deviating results from this estimate. For both cases in figure 3.16, the estimated degradation rates D_d end close to the linear degradation rate. Therefore, the results indicate that assuming a linear degradation rate would be less representative in the first few years of an operational PV system, but will likely be more representative after this initial period. This supports the need for understanding non-linear changes in a PV system's degradation using methods that do not simply force a linear estimate, but also elucidate the possible change in degradation over time.

Chapter 4

Conclusion

7 years of meteorological and energy time series data for a 60 MWp utility-scale PV power plant has been analyzed with the aim of investigating the assumption of a linear degradation rate for a PV system by comparing different sizes of analyzed performance time series. A selection of thresholds were applied to remove unreasonable or noisy raw data and a selection of filters has been investigated. The final selection included an AOI filter, a curtailment filter and a clear sky filter. The clear sky filter removed the most data points suggesting a lot of non-clear sky conditions, and a significant portion of raw data points was removed by the final selection of filters and data thresholds. The 10-minute stringset PR'_{STC} was then calculated for all 4500+ stringsets in the PV system and aggregated to $PR'_{str,d}$, The median of all $PR'_{str,d}$ was used to represent the $PR'_{sys,d}$. An arbitrarily chosen $PR'_{str,d}$ and the $PR'_{sys,d}$ time series were used to run CODS, a combined degradation and soiling algorithm. From this CODS analysis, the estimated degradation rates D_d for the arbitrarily selected $PR'_{str,d}$ and $PR'_{sys,d}$ were found to be higher than expected, with the latter estimated to be positive. A shift in the POAI data was found based on a method by Øgaard et al. [83], and subsequently, a correction factor y_α was calculated and applied to the 10-minute PR'_{STC} time series. The method for finding a correction factor also suggested that a secondary correction could be warranted. However, as maintenance logs were not available for analysis, any further corrections would risk questioning the validity of the results thereof. Considering this, and due to time constraints, no further corrections were performed on the data sets.

A renewed CODS analysis was performed for the arbitrarily chosen $PR'_{str,d}$ and $PR'_{sys,d}$ which estimated a D_d of -0.411 %/year and -0.177 %year. The estimated degradation rates were found to be within literature values, but were less negative possibly due to the relative short exposure time compared

to the literature values, as well as the influence of sensor maintenance. A rolling CODS analysis was then performed on the seasonally adjusted $PR'_{str,d}$ and $PR'_{sys,d}$, with a 3 year time window and an increment of 1 month. The estimated D_d varied between -1.311 %/year and 1.285 %/year for the seasonally adjusted $PR'_{str,d}$, and between -0.674 %/year and 0.773 %/year for the seasonally adjusted $PR'_{sys,d}$. The difference in estimated D_d was further investigated by taking an arbitrary selection of 10 stringsets, each from a different inverter, and used in the rolling CODS method. The temporal variation of D_d for the stringsets was not found to be significant. By use of the rolling CODS method, the estimated D_d was shown to vary depending on the size of the time series. This demonstrated that a linear degradation rate for the entire time series could vary significantly from non-linear estimates. Consequently, the risk of deviating from budgeted O&M costs could be significant in the initial, and arguably most critical, phase of the a PV system's lifetime. The results in this thesis support the need for exploring non-linear changes in the degradation of PV systems and how this can improve O&M in the future. PV system planners and operators would benefit from considering these results when determining O&M procedures for future projects.

4.1 Future work

An interesting approach to the rolling CODS method would be to investigate increasing the analyzed time series window, instead of moving it, by an increment at a time. This could further elucidate how the estimated degradation rate can vary depending on the size of the analyzed time series, taking into account a weighting of measurements to arrive at a more representative linear degradation rate.
For the PV system investigated, a closer inspection on inverter performance and how this compares to aggregated stringset performance within the same inverter would be of particular interest. Provided information on data shifts, corrections or changes, the influence of maintenance could be corrected for and more representative estimates for degradation could be found as a result. Additionally, an attempt at using satellite data to model POAI would be an alternative worth exploring in a effort to work around the issue of data shifts. A comprehensive direct comparison of non-linear methods for estimating degradation rates included a soiling estimation would also be of interest. This could include change point analysis using FBP or the MS-PL method, and can be compared to the rolling CODS method.

Bibliography

[1] IPCC, *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland: IPCC, 2014, pp. 1–151, [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)] [Online]. Available: https://www.ipcc.ch/report/ar5/syr/.

[2] IEA, *World Energy Outlook 2022*. IEA, Paris, 2022, License: CC BY 4.0 (report); CC BY NC SA 4.0 (Annex A). [Online]. Available: https://www.iea.org/reports/world-energy-outlook-2022.

[3] bp, *bp Statistical Review of World Energy*, 71st ed. BP, London, 2022. [Online]. Available: https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html.

[4] IRENA, *Renewable Capacity Statistics 2022*. The International Renewable Energy Agency, Abu Dhabi, 2022, ISBN: 978-92-9260-428-8. [Online]. Available: https://www.irena.org/publications/2022/Apr/Renewable-Capacity-Statistics-2022.

[5] G. Masson, I. Kaizuka, E. Bosch, *et al.*, *TRENDS IN PHOTOVOLTAIC APPLICATIONS 2022*. IEA PVPS TCP, 2022, ISBN: 978-3-907281-35-2. [Online]. Available: https://iea-pvps.org/trends_reports/trends-2022/.

[6] U. Jahn, M. Herz, M. Köntges, *et al.*, *Review on Infrared and Electroluminescence Imaging - for PV Field Applications*. IEA-PVPS, 2018, ISBN: 978-3-906042-53-4. [Online]. Available: https://iea-pvps.org/key-topics/review-on-ir-and-el-imaging-for-pv-field-applications/.

[7] R. H. French, L. S. Bruckmann, D. Moser, *et al.*, *Assessment of Performance Loss Rateof PV Power Systems*. IEA-PVPS, 2021, ISBN: 978-3-907281-10-9. [Online]. Available: https://iea-pvps.org/key-topics/assessment-of-performance-loss-rate-of-pv-power-systems/.

[8] Å. Skomedal and M. G. Deceglie, "Combined estimation of degradation and soiling losses in photovoltaic systems," *IEEE Journal of Photovoltaics*, vol. 10, no. 6, pp. 1788–1796, 2020. DOI: 10.1109/JPHOTOV.2020.3018219.

[9] M. Deceglie, K. Anderson, A. B. Shinn, *et al.*, *Rdtools:version 2.2.0-beta.1/development branch*, Dec. 2022. [Online]. Available: https://github.com/NREL/rdtools/tree/development.

[10] M. M. Nygård, Å. F. Skomedal, M. S. Wiig, and E. S. Marstein, "Combined degradation and soiling with validation against independent soiling station measurements," *IEEE Journal of Photovoltaics*, vol. 13, no. 2, pp. 296–304, 2023. DOI: 10.1109/JPHOTOV.2023.3239752.

[11] D. C. Jordan, K. Anderson, K. Perry, *et al.*, "Photovoltaic fleet degradation insights," *Progress in Photovoltaics: Research and Applications*, vol. 30, no. 10, pp. 1166–1175, 2022. DOI: https://doi.org/10.1002/pip.3566. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pip.3566. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/pip.3566.

[12] 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, 2017. DOI: https://doi.org/10.1002/pip.2835. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pip.2835. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/pip.2835.

[13] D. C. Jordan, J. H. Wohlgemuth, and S. R. Kurtz, "Technology and climate trends in pv module degradation: Preprint," Oct. 2012. [Online]. Available: https://www.osti.gov/biblio/1053315.

[14] S. Yilmaz, H. R. O., O. Dogmus, F. Dincer, O. Akgol, and M. Karaaslan, "Design of two axes sun tracking controller with analytically solar radiation calculations," *Renewable and Sustainable Energy Reviews*, vol. 43, pp. 997–1005, 2015, ISSN: 1364-0321. DOI: https://doi.org/10.1016/j.rser.2014.11.090. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032114010375.

[15] C. Gueymard, D. Myers, and K. Emery, "Proposed reference irradiance spectra for solar energy systems testing," *Solar Energy*, vol. 73, no. 6, pp. 443–467, 2002, ISSN: 0038-092X. DOI: https://doi.org/10.1016/S0038-092X(03)00005-7. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0038092X03000057.

[16] R. Perez, P. Ineichen, R. Seals, J. Michalsky, and R. Stewart, "Modeling daylight availability and irradiance components from direct and global irradiance," *Solar Energy*, vol. 44, no. 5, pp. 271–289, 1990, ISSN: 0038-092X. DOI: https://doi.org/10.1016/0038-092X(90)90055-H. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0038092X9090055H.

[17] H. Hottel and B. Woertz, "Performance of flat-plate solar-heat collectors," *Trans. ASME (Am. Soc. Mech. Eng.); (United States)*, vol. 64, Jan. 1942. [Online]. Available: https://www.osti.gov/biblio/5052689.

[18] M. A. Korevaar, "Measuring solar irradiance for photovoltaics," in *Solar Radiation*, M. Aghaei, Ed., Rijeka: IntechOpen, 2022, ch. 2. DOI: 10.5772/intechopen.105580. [Online]. Available: https://doi.org/10.5772/intechopen.105580.

[19] S. Phillips, "Photovoltaics report," Frauenhofer ISE and Werner Warmuth, PSE Projects GmbH, Tech. Rep., 2023.

[20] F. Shan, F. Tang, L. Cao, and G. Fang, "Comparative simulation analyses on dynamic performances of photovoltaic–thermal solar collectors with different configurations," *Energy Conversion and Management*, vol. 87, pp. 778–786, 2014, ISSN: 0196-8904. DOI: https://doi.org/10.1016/j.enconman.2014.07.077. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S019689041400716X.

[21] M.-H. Wang and M.-J. Chen, "Two-stage fault diagnosis method based on the extension theory for PV power systems," *International Journal of Photoenergy*, vol. 2012, pp. 1–10, 2012. DOI: 10.1155/2012/892690. [Online]. Available: https://doi.org/10.1155/2012/892690.

[22] C. Molto, J. Oh, F. I. Mahmood, *et al.*, "Review of potential-induced degradation in bifacial photovoltaic modules," *Energy Technology*, vol. 11, no. 4, p. 2200943, 2023. DOI: https://doi.org/10.1002/ente.202200943. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ente.202200943. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/ente.202200943.

[23] J. Nelson, *The Physics of Solar Cells*, eng. London, England: Imperial College Press, May 2003, ISBN: 978-1-86094-340-9.

[24] A. Blakers, "Development of the perc solar cell," *IEEE Journal of Photovoltaics*, vol. 9, no. 3, pp. 629–635, 2019. DOI: 10.1109/JPHOTOV.2019.2899460.

[25] J. Benick, B. Steinhauser, R. Müller, *et al.*, "High efficiency n-type pert and perl solar cells," in *2014 IEEE 40th Photovoltaic Specialist Conference (PVSC)*, 2014, pp. 3637–3640. DOI: 10.1109/PVSC.2014.6924895.

[26] F. Feldmann, M. Bivour, C. Reichel, M. Hermle, and S. W. Glunz, "Passivated rear contacts for high-efficiency n-type si solar cells providing high interface passivation quality and excellent transport characteristics," *Solar Energy Materials and Solar Cells*, vol. 120, pp. 270–274, 2014, ISSN: 0927-0248. DOI: https://doi.org/10.1016/j.solmat.2013.09.017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0927024813004868.

[27] E. Franklin, K. Fong, K. McIntosh, *et al.*, "Design, fabrication and characterisation of a 24.4% efficient interdigitated back contact solar cell," *Progress in Photovoltaics: Research and Applications*, vol. 24, no. 4, pp. 411–427, 2016. DOI: https://doi.org/10.1002/pip.2556. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pip.2556. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/pip.2556.

[28] W. Fuhs, K. Niemann, and J. Stuke, "Heterojunctions of amorphous silicon and silicon single crystals," *AIP Conference Proceedings*, vol. 20, no. 1, pp. 345–350, 1974. DOI: 10.1063/1.2945985. eprint: https://aip.scitation.org/doi/pdf/10.1063/1.2945985. [Online]. Available: https://aip.scitation.org/doi/abs/10.1063/1.2945985.

[29] R. Guerrero-Lemus, R. Vega, T. Kim, A. Kimm, and L. Shephard, "Bifacial solar photovoltaics – a technology review," *Renewable and Sustainable Energy Reviews*, vol. 60, pp. 1533–1549, 2016, ISSN: 1364-0321. DOI: https://doi.org/10.1016/j.rser.2016.03.041. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032116002768.

[30] I. International, "International technology roadmap for photovoltaic (itrpv) - 2023 results," VDMA, Tech. Rep., 2023. [Online]. Available: https://www.vdma.org/international-technology-roadmap-photovoltaic.

[31] H. Zsiborács, N. Hegedűsné Baranyai, S. Csányi, A. Vincze, and G. Pintér, "Economic analysis of grid-connected pv system regulations: A hungarian case study," *Electronics*, vol. 8, no. 2, 2019, ISSN: 2079-9292. DOI: 10.3390/electronics8020149. [Online]. Available: https://www.mdpi.com/2079-9292/8/2/149.

[32] N. Pearsall, *The Performance of Photovoltaic (PV) Systems: Modelling, Measurement and Assessment*, eng. Cambridge: Elsevier Science & Technology, 2016, ISBN: 9781782423362.

[33] S. Gutierrez, P. M. Rodrigo, J. Alvarez, A. Acero, and A. Montoya, "Development and testing of a single-axis photovoltaic sun tracker through the internet of things," *Energies*, vol. 13, no. 10, 2020, ISSN: 1996-1073. DOI: 10.3390/en13102547. [Online]. Available: https://www.mdpi.com/1996-1073/13/10/2547.

[34] "Photovoltaic system performance - part 1: Monitoring," International Electrochemical Comission, Standard, 2021.

[35] I. Svetunkov, *Forecasting and analytics with adam*, Monograph. OpenForecast, (version: [current date]), 2022. [Online]. Available: https://openforecast.org/adam/.

[36] D. C. Jordan and S. R. Kurtz, "The dark horse of evaluating long-term field performance—data filtering," *IEEE Journal of Photovoltaics*, vol. 4, no. 1, pp. 317–323, 2014. DOI: 10.1109/JPHOTOV.2013.2282741.

[37] J. John, V. Rajasekar, S. Boppana, S. Tatapudi, and G. Tamizhmani, "Angle of incidence effects on soiled PV modules," in *Reliability of Photovoltaic Cells, Modules, Components, and Systems VII*, N. G. Dhere, Ed., International Society for Optics and Photonics, vol. 9179, SPIE, 2014, p. 91790D. DOI: 10.1117/12.2063351. [Online]. Available: https://doi.org/10.1117/12.2063351.

[38] B. Marion, "Evaluation of clear-sky and satellite-derived irradiance data for determining the degradation of photovoltaic system performance," *Solar Energy*, vol. 223, Jun. 2021. DOI: 10.1016/j.solener.2021.05.071.

[39] L. Bird, J. Cochran, and X. Wang, "Wind and solar energy curtailment: Experience and practices in the united states," Mar. 2014. DOI: 10.2172/1126842. [Online]. Available: https://www.osti.gov/biblio/1126842.

[40] R. Luthander, D. Lingfors, and J. Widén, "Large-scale integration of photovoltaic power in a distribution grid using power curtailment and energy storage," *Solar Energy*, vol. 155, pp. 1319–1325, 2017, ISSN: 0038-092X. DOI: https://doi.org/10.1016/j.solener.2017.07.083. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0038092X17306680.

[41] H. I. Hazim, K. A. Baharin, C. K. Gan, A. H. Sabry, and A. J. Humaidi, "Review on optimization techniques of pv/inverter ratio for grid-tie pv systems," *Applied Sciences*, vol. 13, no. 3155, 2023, ISSN: 2076-3417. [Online]. Available: https://www.mdpi.com/2076-3417/13/5/3155.

[42] C. R. Harris, K. J. Millman, S. J. van der Walt, *et al.*, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020. DOI: 10.1038/s41586-020-2649-2. [Online]. Available: https://doi.org/10.1038/s41586-020-2649-2.

[43] T. pandas development team, *Pandas*, version 2.0.1, Feb. 2020. [Online]. Available: https://github.com/pandas-dev/pandas/.

[44] W. Holmgren, K. Anderson, Calama-Consulting, *et al.*, *Pvlib python*, version 0.9.4, Dec. 2022. [Online]. Available: https://github.com/pvlib/pvlib-python.

[45] C. Schill, A. Anderson, C. Baldus-Jeursen, *et al.*, *Soiling Losses – Impact on the Performance of Photovoltaic Power Plants*. IEA-PVPS, 2022, ISBN: 978-3-907281-09-3. [Online]. Available: https://iea-pvps.org/key-topics/soiling-losses-impact-on-the-performance-of-photovoltaic-power-plants/.

[46] K. Ilse, L. Micheli, B. W. Figgis, *et al.*, "Techno-economic assessment of soiling losses and mitigation strategies for solar power generation," *Joule*, vol. 3, no. 10, pp. 2303–2321, 2019, ISSN: 2542-4351. DOI: https://doi.org/10.1016/j.joule.2019.08.019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542435119304222.

[47] K. K. Ilse, B. W. Figgis, V. Naumann, C. Hagendorf, and J. Bagdahn, "Fundamentals of soiling processes on photovoltaic modules," *Renewable and Sustainable Energy Reviews*, vol. 98, pp. 239–254, 2018, ISSN: 1364-0321. DOI: https://doi.org/10.1016/j.rser.2018.09.015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032118306634.

[48] B. Stridh, "Evaluation of economical benefit of cleaning of soiling and snow in pv plants at three european locations," in *2012 38th IEEE Photovoltaic Specialists Conference*, 2012, pp. 001 448–001 451. DOI: 10.1109/PVSC.2012.6317869.

[49] B. Figgis, A. Ennaoui, S. Ahzi, and Y. Rémond, "Review of pv soiling particle mechanics in desert environments," *Renewable and Sustainable Energy Reviews*, vol. 76, pp. 872–881, 2017, ISSN: 1364-0321. DOI:

https://doi.org/10.1016/j.rser.2017.03.100. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032117304355.

[50] M. Gostein, J. R. Caron, and B. Littmann, "Measuring soiling losses at utility-scale pv power plants," in *2014 IEEE 40th Photovoltaic Specialist Conference (PVSC)*, 2014, pp. 0885–0890. DOI: 10.1109/PVSC.2014.6925056.

[51] S. Kagan, E. Giosa, R. Flottemesch, *et al.*, "Impact of non-uniform soiling on pv system performance and soiling measurement," in *2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC & 34th EU PVSEC)*, 2018, pp. 3432–3435. DOI: 10.1109/PVSC.2018.8547728.

[52] S. Toth, M. Hannigan, M. Vance, and M. Deceglie, "Predicting photovoltaic soiling from air quality measurements," *IEEE Journal of Photovoltaics*, vol. 10, no. 4, pp. 1142–1147, 2020. DOI: 10.1109/JPHOTOV.2020.2983990.

[53] M. Gostein, S. Faullin, K. Miller, J. Schneider, and B. Stueve, "Mars soiling sensor™," in *2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC & 34th EU PVSEC)*, 2018, pp. 3417–3420. DOI: 10.1109/PVSC.2018.8547767.

[54] M. Gostein, B. Bourne, F. Farina, and B. Stueve, "Field testing of mars™ soiling sensor," in *2020 47th IEEE Photovoltaic Specialists Conference (PVSC)*, 2020, pp. 0524–0527. DOI: 10.1109/PVSC45281.2020.9300975.

[55] S. Mehta, A. P. Azad, S. A. Chemmengath, V. Raykar, and S. Kalyanaraman, "Deepsolareye: Power loss prediction and weakly supervised soiling localization via fully convolutional networks for solar panels," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2018, pp. 333–342. DOI: 10.1109/WACV.2018.00043.

[56] H. Supe, R. Avtar, D. Singh, *et al.*, "Google earth engine for the detection of soiling on photovoltaic solar panels in arid environments," *Remote Sensing*, vol. 12, no. 9, 2020, ISSN: 2072-4292. DOI: 10.3390/rs12091466. [Online]. Available: https://www.mdpi.com/2072-4292/12/9/1466.

[57] R. Barnes, *Implementing soiling monitoring systems in solar pv plants*, Apr. 2023. [Online]. Available: https://constructionreviewonline.com/installations-materials/implementing-soiling-monitoring-systems-in-solar-pv-plants/.

[58] M. Köntges, S. Kurtz, C. Packard, *et al.*, *Review of Failures of Photovoltaic Modules*. IEA-PVPS, 2014, ISBN: 978-3-906042-16-9. [Online]. Available: https://iea-pvps.org/key-topics/review-of-failures-of-photovoltaic-modules-final/.

[59] D. Chen, M. Vaqueiro Contreras, A. Ciesla, *et al.*, "Progress in the understanding of light- and elevated temperature-induced degradation in silicon solar cells: A review," *Progress in Photovoltaics: Research and Applications*, vol. 29, no. 11, pp. 1180–1201, 2021. DOI: https://doi.org/10.1002/pip.3362. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pip.3362. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/pip.3362.

[60] M. Winter, D. C. Walter, B. Min, R. Peibst, R. Brendel, and J. Schmidt, "Light and elevated temperature induced degradation and recovery of gallium-doped czochralski-silicon solar cells," *Scientific Reports*, vol. 12, no. 1, May 2022. DOI: 10.1038/s41598-022-11831-3. [Online]. Available: https://doi.org/10.1038/s41598-022-11831-3.

[61] A. Limmanee, N. Udomdachanut, S. Songtrai, *et al.*, "Field performance and degradation rates of different types of photovoltaic modules: A case study in thailand," *Renewable Energy*, vol. 89, pp. 12–17, 2016, ISSN: 0960-1481. DOI: https://doi.org/10.1016/j.renene.2015.11.088. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960148115305048.

[62] 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.

[63] J. Ascencio-Vásquez, K. Brecl, and M. Topič, "Methodology of köppen-geiger-photovoltaic climate classification and implications to worldwide mapping of pv system performance," *Solar Energy*, vol. 191, pp. 672–685, 2019, ISSN: 0038-092X. DOI: https://doi.org/10.1016/j.solener.2019.08.072. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0038092X19308527.

[64] J. Ascencio-Vásquez, I. Kaaya, K. Brecl, K.-A. Weiss, and M. Topič, "Global climate data processing and mapping of degradation mechanisms and degradation rates of pv modules," *Energies*, vol. 12, no. 24, 2019, ISSN: 1996-1073. DOI: 10.3390/en12244749. [Online]. Available: https://www.mdpi.com/1996-1073/12/24/4749.

[65] I. Romero-Fiances, A. Livera, M. Theristis, *et al.*, "Impact of duration and missing data on the long-term photovoltaic degradation rate estimation," *Renewable Energy*, vol. 181, pp. 738–748, 2022, ISSN: 0960-1481. DOI: https://doi.org/10.1016/j.renene.2021.09.078. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S096014812101404X.

[66] D. Jordan and S. Kurtz, "Analytical improvements in pv degradation rate determination," in *2010 35th IEEE Photovoltaic Specialists Conference*, 2010, pp. 002 688–002 693. DOI: 10.1109/PVSC.2010.5617074.

[67] E. Pieri, A. Kyprianou, A. Phinikarides, G. Makrides, and G. E. Georghiou, "Forecasting degradation rates of different photovoltaic systems using robust principal component analysis and arima," *IET Renewable Power Generation*, vol. 11, no. 10, pp. 1245–1252, 2017. DOI: https://doi.org/10.1049/iet-rpg.2017.0090. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-rpg.2017.0090.

[68] A. Frick, G. Makrides, M. Schubert, M. Schlecht, and G. E. Georghiou, "Degradation rate location dependency of photovoltaic systems," *Energies*, vol. 13, no. 24, 2020, ISSN: 1996-1073. DOI: 10.3390/en13246751. [Online]. Available: https://www.mdpi.com/1996-1073/13/24/6751.

[69] R. B. Cleveland, W. S. Cleveland, J. E. McRae, and I. Terpenning, "Stl: A seasonal-trend decomposition procedure based on loess (with discussion)," *Journal of Official Statistics*, vol. 6, pp. 3–73, 1990.

[70] S. Lindig, I. Kaaya, K.-A. Weiss, D. Moser, and M. Topic, "Review of statistical and analytical degradation models for photovoltaic modules and systems as well as related improvements," *IEEE Journal of Photovoltaics*, vol. 8, no. 6, pp. 1773–1786, 2018. DOI: 10.1109/JPHOTOV.2018.2870532.

[71] A. Kyprianou, A. Phinikarides, G. Makrides, and G. E. Georghiou, "Definition and computation of the degradation rates of photovoltaic systems of different technologies with robust principal component analysis," *IEEE Journal of Photovoltaics*, vol. 5, no. 6, pp. 1698–1705, 2015. DOI: 10.1109/JPHOTOV.2015.2478065.

[72] D. C. Jordan, C. Deline, S. R. Kurtz, G. M. Kimball, and M. Anderson, "Robust pv degradation methodology and application," *IEEE Journal of Photovoltaics*, vol. 8, no. 2, pp. 525–531, 2018. DOI: 10.1109/JPHOTOV.2017.2779779.

[73] 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, pp. 978–989, 2016. DOI: https://doi.org/10.1002/pip.2744. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pip.2744. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/pip.2744.

[74] S. W. Adler, M. S. Wiig, Å. Skomedal, H. Haug, and E. S. Marstein, "Degradation analysis of utility-scale pv plants in different climate zones," *IEEE Journal of Photovoltaics*, vol. 11, no. 2, pp. 513–518, 2021. DOI: 10.1109/JPHOTOV.2020.3043120.

[75] M. Theristis, A. Livera, C. B. Jones, G. Makrides, G. E. Georghiou, and J. S. Stein, "Nonlinear photovoltaic degradation rates: Modeling and comparison against conventional methods," *IEEE Journal of Photovoltaics*, vol. 10, no. 4, pp. 1112–1118, 2020. DOI: 10.1109/JPHOTOV.2020.2992432.

[76] M. Theristis, A. Livera, L. Micheli, *et al.*, "Comparative analysis of change-point techniques for nonlinear photovoltaic performance degradation rate estimations," *IEEE Journal of Photovoltaics*, vol. 11, no. 6, pp. 1511–1518, 2021. DOI: 10.1109/JPHOTOV.2021.3112037.

[77] S. Lindig, A. Louwen, D. Moser, and M. Topic, "New pv performance loss methodology applying a self-regulated multistep algorithm," *IEEE Journal of Photovoltaics*, vol. 11, no. 4, pp. 1087–1096, 2021. DOI: 10.1109/JPHOTOV.2021.3078075.

[78] A. Livera, G. Tziolis, M. Theristis, J. S. Stein, and G. E. Georghiou, "Estimating the performance loss rate of photovoltaic systems using time series change point analysis," *Energies*, vol. 16, no. 9, 2023, ISSN: 1996-1073. DOI: 10.3390/en16093724. [Online]. Available: https://www.mdpi.com/1996-1073/16/9/3724.

[79] A. Skomedal, M. Deceglie, H. Haug, and E. Marstein, "Iterative and self-consistent estimation of degradation and soiling loss in pv systems - a case study," in *37th European Photovoltaic Solar Energy Conference and Exhibition*, (Online), P. Helm, P. Kenny, and P.Verlinden, Eds., WIP, Sep. 2020, pp. 1339–1344, ISBN: 3-936338-50-7. DOI: 10.4229/EUPVSEC20202020-5DO.1.3. [Online]. Available: https://www.eupvsec-proceedings.com/proceedings?char=I&paper=49034.

[80] E. Hasselbrink, M. Anderson, Z. Defreitas, *et al.*, "Validation of the pvlife model using 3 million module-years of live site data," in *2013 IEEE 39th Photovoltaic Specialists Conference (PVSC)*, 2013, pp. 0007–0012. DOI: 10.1109/PVSC.2013.6744087.

[81] E. Sveen, M. Øgaard, J. Selj, and G. Otnes, "Pv system degradation rates in the nordics," in *37th European Photovoltaic Solar Energy Conference and Exhibition*, (Online), P. Helm, P. Kenny, and P.Verlinden, Eds., WIP, Sep. 2020, pp. 1563–1566, ISBN: 3-936338-50-7. DOI: 10.4229/EUPVSEC20202020-5CV.3.36. [Online]. Available: https://www.eupvsec-proceedings.com/proceedings?paper=49494.

[82] I. International, "International technology roadmap for photovoltaic (itrpv) - 2018 results," VDMA, Tech. Rep., 2018. [Online]. Available: https://www.vdma.org/international-technology-roadmap-photovoltaic.

[83] M. Øgaard, H. Haug, and J. Selj, "Methods for quality control of monitoring data from commercial pv systems," in *"35th European Photovoltaic Solar Energy Conference and Exhibition"*, (Brussels, Belgium), P. Helm, P. Kenny, and P.Verlinden, Eds., EU PVSEC, Sep. 2018, pp. 2083–2088, ISBN: 3-936338-50-7. DOI: 10.4229/35thEUPVSEC20182018-6DV.1.53.

[84] "Photovoltaic system performance - part 4: Degradation rate evaluation method," International Electrochemical Comission, Standard, 2017.

[85] M. J. Reno and C. W. Hansen, "Identification of periods of clear sky irradiance in time series of ghi measurements," *Renewable Energy*, vol. 90, pp. 520–531, 2016, ISSN: 0960-1481. DOI: https://doi.org/10.1016/j.renene.2015.12.031. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960148115305358.

[86] M. Theristis, J. S. Stein, C. Deline, *et al.*, "Onymous early-life performance degradation analysis of recent photovoltaic module technologies," *Progress in Photovoltaics: Research and Applications*, vol. 31, no. 2, pp. 149–160, 2023. DOI: https://doi.org/10.1002/pip.3615. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/pip.3615.

[87] D. Palmer, E. Koubli, I. Cole, T. Betts, and R. Gottschalg, "Satellite or ground-based measurements for production of site specific hourly irradiance data: Which is most accurate and where?" *Solar Energy*, vol. 165, pp. 240–255, 2018, ISSN: 0038-092X. DOI: https://doi.org/10.1016/j.solener.2018.03.029. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0038092X18302408.

[88] A. Ayaz, F. Ahmad, M. A. A. Irfan, Z. Rehman, K. Rajski, and J. Danielewicz, "Comparison of ground-based global horizontal irradiance and direct normal irradiance with satellite-based suny model," *Energies*, vol. 15, no. 7, 2022, ISSN: 1996-1073. DOI: 10.3390/en15072528. [Online]. Available: https://www.mdpi.com/1996-1073/15/7/2528.

[89] P. Helm, P. Kenny, and P.Verlinden, Eds., *37th European Photovoltaic Solar Energy Conference and Exhibition*, (Online), WIP, Sep. 2020, ISBN: 3-936338-50-7.