DEGRADATION & PARTIAL SHADING STUDY OF PHOTOVOLTAIC MODULES IN THE FIELD: ENABLED BY TIME-SERIES CURRENT-VOLTAGE & POWER ANALYSIS

by

JIQI LIU

Submitted in partial fulfillment of the requirements

For the degree of Master of Science

Department of Materials Science and Engineering

CASE WESTERN RESERVE UNIVERSITY

August, 2020

Degradation & Partial Shading Study of Photovoltaic Modules in the Field: Enabled by Time-series Current-Voltage & Power Analysis

Case Western Reserve University

We hereby approve the thesis[1] of

JIQI LIU

for the degree of

Master of Science

Dr. Roger H. French

Committee Chair, Advisor	May 28, 2020
Department of Materials Science and Engineering	

Dr. Alp Sehirlioglu

Committee Member	May 28, 2020
Department of Materials Science and Engineering	

Dr. Daniela Calvetti

Committee Member	May 28, 2020
Department of Mathematics, Applied Mathematics and Statistics	

Dr. Jennifer L. Braid

Committee Member	May 28, 2020
Department of Materials Science and Engineering	

---

[1]We certify that written approval has been obtained for any proprietary material contained therein.

Table of Contents

List of Tables

List of Figures

0.1 Acknowledgements

I would like to thank my adviser, Dr. Roger French for his patience in communicating with me in scientific problems and my committee members Prof. Alp Sehirlioglu, Prof. Daniela Calvetti and Prof. Jennifer Braid. I would like to thank some SDLE members especially, Menghong Wang and Alan Curran for their cooperation in this study, Ahmad Maroof Karimi and Wei-Heng Huang for their help in using HBase and setting Slurm jobs in HPC, and Yu Wang especially for her help in life the first year I was in the U.S.. I am very grateful for the many things I have learned in the Solar Durability and Lifetime Extension Center over the past three years. I would like to thank Michael Köhl and Erdmut Schnabel in Fraunhofer-ISE for providing the dataset and their collaboration.

I would like to thank my parents, who provide me support and understanding whenever I need it. And my best friend Yue Ning, who is always willing to listen my endless worries and complaints even when we have a twelve hour time difference.

Degradation & Partial Shading Study of Photovoltaic Modules in the Field: Enabled by Time-series Current-Voltage & Power Analysis

Abstract

by

JIQI LIU

0.2 Abstract

The time-series power (P) and current-voltage (I-V) curve datastreams of photovoltaic modules exposed in the field for long times are very useful for studying degradation behavior and fault detection of photovoltaic modules. In this study, over two million I-V curves were processed using the data-driven I-V feature extraction algorithm. A statistical analysis method was developed and applied to detect partial shading by returning shading profiles, the relative orientation of the shading object, detailed information on shading scenarios, and the duration of partial shading based on multistep I-V curves. Three out of eight PV modules studied exhibit partial shading which is consistent across a number of years in the shading profile. The extracted I-V features are processed by the outdoor I_{sc}-V_{oc} and power loss calculation method to obtain the power losses due to four degradation modes including uniform current loss, recombination loss, series resistance loss and current mismatch loss. Then the month-by-month regression method is applied to the four types of time-series power loss to obtain the rate of change of each degradation mode of each module. The

degradation results from both I-V and P_{mp} analysis are compared across module brands, architectures and their climate zone locations. In addition, the performance loss rates (PLR) was determined from the power time-series and show that the BWh and BSh Köppen-Geiger climate zones cause similar performance loss rates, which are larger than those in the ET climate for brand F (glass-backsheet architecture) modules. For Brand G (double glass architecture) modules, BSh causes more aggressive degradation than BWh and both are more aggressive than ET. Brand G (double glass) has significantly better performance than brand F (glass-backsheet) ones especially in BWh. From the outdoor I_{sc}-V_{oc} analysis, the dominant degradation mode is obtained for each brand (architecture) PV modules under exposure in each climate zone. These I-V,P_{mp} time-series datastream analyses represent a new approach to remotely identify degradation rates and mechanisms, and partial shading problems of real-world PV power plants.

1 Introduction

Solar energy is a kind of clean and renewable energy, and its market share has continued to increase for many years. In order to reduce the levelized cost of electricity produced and make solar energy more competitive in the energy market, modules with higher efficiency were developed and research was conducted to study the degradation of photovoltaic modules to reduce the power loss rate in long term degradation by improving in manufacturing. For studies of degradation, various indoor accelerated exposures and characterization techniques are most commonly used because exposure condition variables are more easily controlled in the lab. However, the real working environment of a PV module is much more complex and in most cases, there are a lot of environmental factors acting together. These factors also change over time, and they can form different combinations due to the local weather conditions. In order to overcome this gap, sequential indoor exposures have been developed to understand the degradation behavior of PV modules or related materials under multiple stressors, and thanks to the rapid growing application of photovoltaic modules, more and more time-series data recording the performance of solar panels in the field have become available for research and many of these are long enough for degradation studies[1].

This provides new opportunities for studying the degradation of solar panels in practical applications, which will bring insights into the degradation of solar panels under exposure to real world conditions. There are two types of time-series data measuring the electrical performance of solar panels: maximum power (P_{mp}) and current-voltage (I-V) curves. P_{mp} time-series datasets are more common and accessible compared with time-series current-voltage datastreams since they can be obtained through the AC/DC inverter, and can be used to evaluate the degradation of the general performance of solar panels. A lot of studies have been proposed on methods to obtain the performance loss rate (PLR) of PV power plants using time-series P_{mp} datastreams.

However, P_{mp} time-series data have limited ability to reveal the causes of performance loss, on the another hand, time-series I-V curves have advantages for this, the change in I-V curves can be closely related to the specific physical or chemical changes of the modules. However, compared with studies using P_{mp}, degradation studies using long term time-series I-V are very limited. Usually, some features are extracted from the I-V curves, and processed in a similar way to time-series P_{mp} to obtain the percentage of change or a rate of change for each feature[2]. This is helpful to understand the changes that happens in the module, but these percentages won't be able to be directly compared with each other considering their contribution to power loss. The outdoor I_{sc}-V_{oc} and power loss method can obtain the power loss due to different changes in some I-V features which make these changes comparable regarding their contribution in power loss[3]. Moreover, outdoor field studies usually

involve different modules located at a single location, and there is a lack of studies using the same method to analyze modules under completely different weather conditions at multiple locations.

In addition, considering the safety and influence on degradation of commercial and research PV systems, it is also crucial to ensure the normal operation of solar panels. Partial shading, as one kind of power loss or fault condition, not only reduces the power generation but can also lead to hot spots and reverse bias conditions which can cause permanent damage to the PV cells and aging of the PV module. It could be caused by lots of common things around the module such as trees, snow, grass, or bird dung. In order to avoid and detect these abnormal operation situations, there are lots of studies about fault detection taking advantages of the P_{mp} time-series data from the PV module under study and another time-series data from a reference object, and the methods developed can already be applied to real systems[4].

However, the partial shading studies by *I-V* curves still mainly link the shape of *I-V* curves to specific shading examples, and there is a lack of research on methods using time-series *I-V* curves to detect shading, even the *I-V* data itself when applied to actual PV systems, have advantages as a signal in partial shading detection[5].

In this research, the structure of which is shown in Fig. 1.1, millions of time-series *I-V* curves are processed, using a data-driven approach for detecting steps and extracting *I-V* features. A partial shading detection method is developed by taking multistep *I-V* curves and returning the shading profile, obstacle orientation, and more detailed information regarding duration and shading scenarios. Three of eight studied modules are identified as experiencing partial shading, which impacts power. Extracted time-series *I-V* features are processed using the outdoor I_{sc}-V_{oc} method
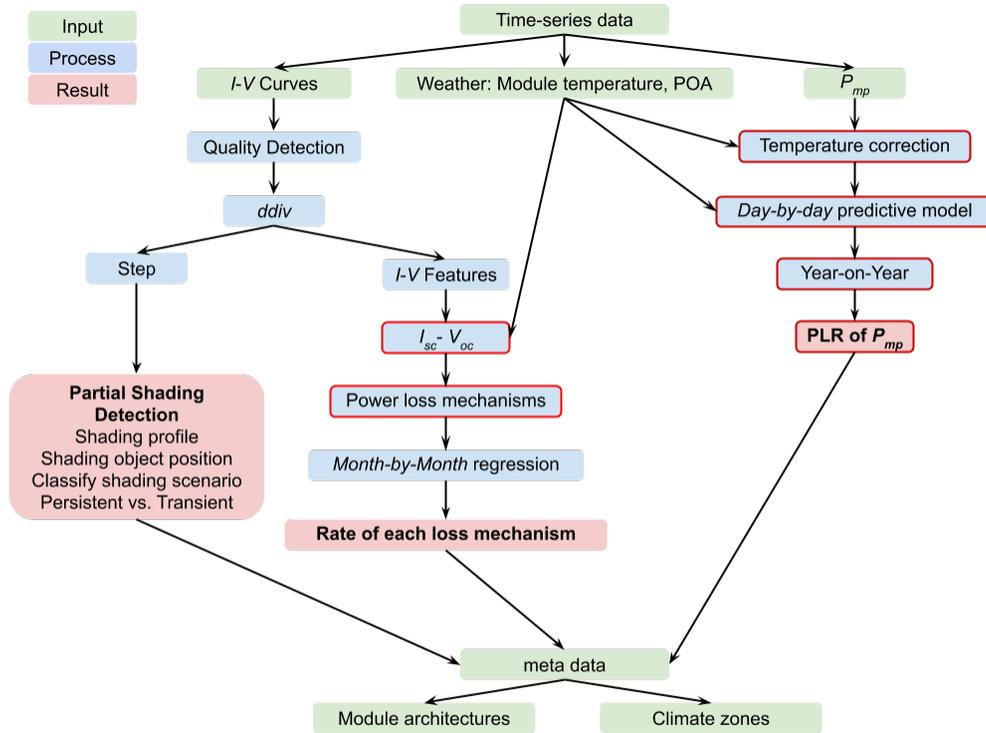
Figure 1.1. Flow chart of study process enabled by time-series *I-V* curves and P_{mp}.

to obtain the power loss due to four types of degradation modes, which are further processed by a regression method to obtain the dominant degradation mode for each module. The time-series P_{mp} is then processed to obtain the performance loss rate. The comparison of degradation results reveals the climate-zone dependent and brand dependent degradation behaviors of these eight modules including both their general performance and the dominant degradation mode.

2 Literature Review

In this chapter, published research on degradation of photovoltaic (PV) modules using time-series electrical data obtained from modules installed in the field is reviewed. The degradation behavior of both double glass (DG) and glass backsheet (GB) PV modules and PV module degradation across Köppen-Geiger climate zones is summarized. And power fault detection, and partial shading of PV modules as studied using time-series data are reviewed.

2.1 Time-series Datastreams & Related PV Degradation Studies

Due to the rapid growth of solar energy[6,7] and development of inverters and low-cost I-V tracing instruments in the past few years, more and more time-series data have became accessible and many studies are of sufficient duration for PV degradation studies. Time-series electrical data for PV modules installed in the field are mainly divided in two categories, maximum power tracking (P_{mp})[8] and current-voltage (I-V) curves[2]. The former can be AC or DC power, usually measured at the inverter, and sometimes with voltage and current at maximum power (V_{mp} and I_{mp}) and the later studies require additional I-V tracing equipment. These time-series datastreams can be used to study operational status of the system, for fault detection and for

degradation studies of solar panels, in the following we mainly focus on degradation studies.

Time-series P_{mp} data are usually analyzed to calculate the performance loss rate (PLR)[9]. First the data is filtered by an irradiance cutoff to remove low irradiance data, sometimes combined with a clear sky filter, then processed to produce a standardized performance metric so that different systems located in different climate zones can be compared to each other[9,10]. There are several common performance metrics used, which include correction to standard test conditions[11], empirical metrics such as PVUSA[12], 6k-values performance model[13] and Sandia models[14], and performance ratio[15]. After this conversion, a statistical modeling method such as regression is used to obtain the rate of change. However, due to seasonality and the inherent noise present in the time-series data, different methods are also used instead of simple linear regression for PLR determination, such as classical seasonal decomposition[16], auto-regressive integrated moving average (ARIMA)[15], Year-on-year regression[17,18] and so on. PLR values obtained by the same method can be compared across different types of modules for degradation study[1,19].

Contrasting with P_{mp}, which indicates the general performance, *I-V* curves are able to provide more insight on PV module degradation behavior[1]. The time-series *I-V* curves are usually not obtained from the AC/DC inverter but instead recorded by special *I-V* tracing systems[20]. In total, there are eight *I-V* features defined, which can be extracted from the curve, even though these are not all independent features. They are maximum power (P_{mp}), voltage at maximum power (V_{mp}), current at maximum power (I_{mp}), short-circuit current (I_{sc}), open-circuit voltage (V_{oc}), series resistance (R_s), shunting resistance (R_{sh}) and fill factor (FF). FF is equal to P_{mp}

divided by the product of I_{sc} and V_{oc} [21]. These *I-V* features are important parameters for evaluating the specific changes in the PV modules [2].

There are mainly two types of approaches that are developed for extracting these *I-V* features, and these methods use either physical models or data-driven models. The former one includes the single diode [22–24] and the two diode [25] models which are based on the equivalent electrical circuit of the PV module, the later one focuses on using the definition of each feature [26] and extracting the features as fitted parameters using partial data of the curve. These features need to be corrected [23] or predicted [27] to be at some reference conditions for module degradation studies. The change of each feature over time is related to different degradation modes and mechanisms of the photovoltaic modules under the specific exposure conditions. For example, yellowing of the encapsulant layer can cause I_{sc} to decrease [28], and corrosion may cause R_s to increase [29]. In additional, there are studies using the extracted I_{sc} and V_{oc} to construct *Suns*-V_{oc} [30] or I_{sc}-V_{oc} curve [31] for degradation studies, and M. Wang proposes a power loss calculation method [3] that can convert the change in some *I-V* features to be the change in power for a more direct comparison using the outdoor I_{sc}-V_{oc} curves.

It is worth mentioning here that there is a very recent study [32] by Xingshu Sun that focuses on reconstructing *I-V* curves using time-series P_{mp}, V_{mp}, and weather variables and then extracting features from them for other later study. These data are potentially more abundant and easier to obtain than time-series *I-V* curves.

2.2 Degradation Behavior Dependencies of PV Modules

The degradation of PV modules is influenced by solar cells' material, such as monocrystalline silicon (mono-Si) or multicrystalline silicon (multi-Si) or CdTe thin films[33], the cell type such as Passivated Emitter and Rear Contact (PERC) or Aluminum Back Surface Field (Al-BSF)[29]. In addition the packaging materials and strategies used play a critical role in degradation. These include the encapsulants[34], the module architecture such as double glass or glass-backsheet[1] and the exposure conditions including both indoor accelerated[35] and outdoor real-world[1] conditions.

According to Jordan[33], who studied reported degradation rates over a 40 year period, the median degradation rate of modules and systems made of silicon type cells and thin-film technology is the same. In addition, module-level data show that the degradation rate of mono-Si type is lower than multi-Si, and is further lower than amorphous-Si.

The standard solar cell structure that has been in use for decades and also occupies the majority of the solar cell market is Al-BSF which stands for Aluminum Back Surface Field cell. In recent years, a new cell structure, first proposed in the 1989[36,37] and now becoming the dominant cell type in the market, is PERC which stands for Passivated Emitter and Rear Contact cell[38]. This new cell technology[39] increases the efficiency of solar cell by adding a rear side dielectric passivization layer. However, studies have shown that p-type PERC cells are more sensitive to the LeTID degradation mode[40,41] (LeTID is light and elevated temperature induced degradation). This may cause p-type PERC cells have a higher degradation rate when the operating temperature is relatively high.

The choice of packaging materials for the PV module, such as encapsulants[34] and backside materials such as the backsheet, and the type of backsheet[42] or glass also affects the modules' degradation behavior. Study of Yingbin, Zhang[43] shows that double glass modules have higher durability over four kinds of indoor accelerated exposures including 600 temperature cycles, 4000 hours of damp-heat exposure, 600 hours potential induced degradation and 50 cycles humidity freeze. These exposure conditions are defined in the IEC 61215 standard[35]. But the chemical composition of the multilayer polymer laminate backsheet in this study was not defined in the paper. Their results showed that this backsheet cracked after 3000 hours damp heat, indicating that this backsheet has quite low durability since the cracks in the backsheet usually should be activated by sequential exposure[44].

Analogous to how various indoor exposure conditions could cause different degradation modes to be activated in the PV modules[45], the outdoor environment, depending on the location, can be vary widely, which can activate different degradation modes and mechanisms even for PV modules of the same brand[46].

2.3 Partial Shading Detection

A power fault, if not be detected and corrected will not only cause a loss in performance of the PV array but also threaten the security, safety and reliability of the whole PV power plant[47]. Shading, which is a kind of fault operation scenario, should be detected and eliminated. Depending on the scenario including partial (or non-uniform) shading and uniform shading of a PV module as Fig. 2.1, can also affect the power plant differently.
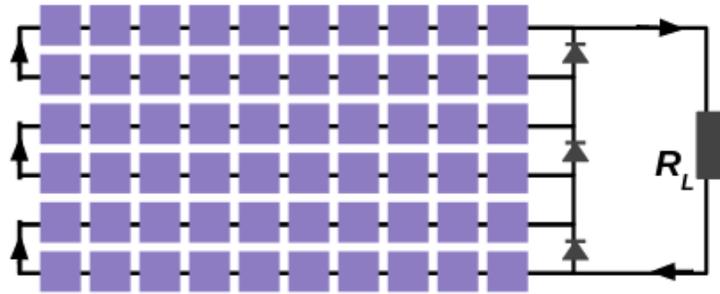
(a) Uniform shading
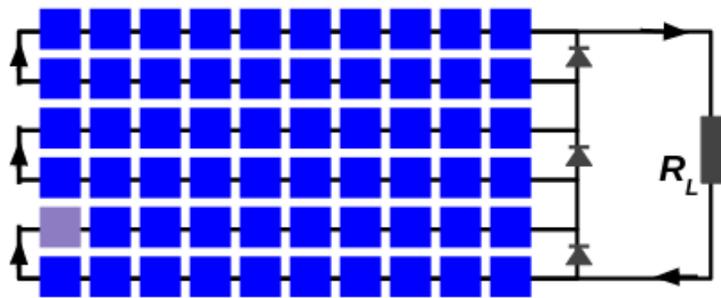
(b) Partial shading

Figure 2.1. Schematic diagram of uniform shading and partial shading (the blue square represents not shaded cell and the purple square represents shaded cell).

Partial shading is the condition of nearby objects casting shade onto part of a photovoltaic module, causing the PV module to receive spatially non-uniform irradiance[48]. Uniform shading, on another hand, means the PV modules has uniform irradiance under shadow. Especially, in the case of several PV modules connected in series to form a string of a PV array, they can be treated as a large PV module as a whole and partial shading is a similar issue for a string of modules[48]. First, both of them cause a reduction in power generation[49], but partial shading generates more concerns including local hot spots, which can lead to lower module efficiency[5], overheating[50], non-uniform severe degradation of packaging material especially for

polymer, safety issue, and reverse bias which could bring the risk of permanent destruction to the cell and cause the whole string to be open-circuit in later operation[51].

For commercial PV modules and most modules studied in research, one solar cell is connected in series with others. If one, or some, cells but not all cells connected in series are shaded, the shaded cells will be forced into reverse bias, consuming power and generating heat[51]. In the worst case, the voltage applied to the shaded cell reaches the breakdown voltage, and the cell can be permanently damaged. Thanks to the use of by-pass diodes in commercial PV modules, which are connected in anti-parallel, each with typically 20 cells in a string (for a 60 cell PV module) to guarantee the maximum negative voltage would not exceed the breakdown voltage[52]. The presence of by-pass diodes also can reduce the dissipation power[53]. However, even in this kind of module configuration, studies with both experiment[54] and simulation[50] shows that, under 1 sun illumination level, the temperature increase in the completed shaded cell can have an increase of 183 $°C$ in one thousand seconds.

Many studies use time-series data generated by power plants to monitor operations and detect shadows, rather than relying on on-site inspections by maintenance personnel. Studies for shading detection, based on the data used, can be divided in two categories, using time-series P_{mp} or I-V curves, the former one sometimes take the voltage at maximum power (V_{mp}) and current at maximum power (I_{mp}) into also[47,55].

Since shading can cause a reduction in P_{mp}[56,57], comparing the P_{mp} of the studied PV module or the module string (composed of modules connected in series) with a reference PV module[4] or pyranometer[58,59] that is guaranteed not be shaded, or

with simulation data[60], when shading occurs can be detected to the studied module. Studies using time-series P_{mp} to detect shading are already quite mature and can be applied to time-series data obtained from the real operating systems[53,61] and obtain the shading profile which has information of when shading happens to the module[4]. Because time-series P_{mp} is abundant and easy to obtain now due to the fast development of PV in the electricity market in recent years[6], the application for this branch of study could be very wide. But it has two main disadvantages, first, it does not distinguish partial shading and uniform shading cases, second, it requires a guaranteed reference object.

However, partial shading in most cases will cause distortion in I-V curves[47,62], which looks like a "step" in the I-V curves, while uniform shading won't cause "step". This phenomena provides an opportunity to distinguish the partial shading cases from uniform ones and detect the partial shading problem without a reference system. A lot of studies using I-V curve for shading detection focus on using simulation and experiments to link the different shading cases with different changes in the I-V curve[5,47,51,55,63,64] but there is a lack of studies that have analysis of real time-series I-V datastreams. The main shortcoming of these methods is the I-V data itself. Time-series I-V curves of PV modules in the field are much less common than P_{mp} and they are usually owned by research facilities, who buy the commercial PV modules and install them in the field with I-V monitoring for degradation study. In addition, an I-V curve compared to a single value is more difficult to process and the significant increase in data in time-series I-V also increases the requirement of computation.

Last but not least, the time-series I-V curves contain more operational diagnostic information than P_{mp}, V_{mp} , and I_{mp} , which are usually already monitored in the

commercial PV inverter. Studies show that, from the hardware perspective, the PV inverter has the necessary ability to measure the *I-V* curve but it's not been introduced by most PV inverter manufacturers at this time[65].

3   Dataset Description

This chapter provides a description of the dataset used in this study, which was provided by our collaborators Michael Köhl and Erdmut Schnabel at Fraunhofer-ISE in Freiburg Germany.

3.1   Time-series Datastreams

The Fraunhoer-ISE dataset comes from a long-term study of photovoltaic modules by Fraunhofer-ISE. It consists of time-series data of eight PV modules installed outdoors, including current-voltage (I-V) curves data, I-V features reported (by the I-V tracer equipment's own fitting of the I-V curves using an undisclosed and unknown algorithm), P_{mp} data and weather data of air mass, angle of incident (aoi), zenith angle, delta irradiance, which is irradiance fluctuation in percentage, effective irradiance measured by a reference cell, global irradiance in the plane of array (POA) and ambient temperature. In addition, the module temperature and the temperature of the reference cell are also recorded.

The I-V curves are measured at a time interval of approximately 5 minutes by an I-V curve tracer instrument, the Electronic Load ESL-Solar 500[66]. Each I-V curve usually contains 40 to 70 data points. Air mass, aoi, delta irradiance, POA, effective

irradiance and temperature of studied module and reference cell, ambient temperature all are measured by local sensors at the same time interval. The maximum power is measured by the same instrument with a time interval about 1 to 2 minutes time interval together with ambient temperature, module temperature, POA, zenith angle, and aoi.

In this thesis, the datastreams of time-series I-V curves, maximum power P_{mp}, irradiance POA ,and module temperature are used.

3.2 Meta Data for Observed Modules

The acquired time-series data is from eight commercial PV modules located at three different sites that belong to three different Köppen-Geiger climatic zones[67]: two modules are on Mount Zugspitze in the Bavarian Alps (ET climate zone), three are in Gran Canaria (the Canary Islands of Spain) (BWh climate zone), and the other three are in the Negev desert (BSh climate zone), the climate zone is identified by the *kgc* package[68] using the longitude and latitude provided. The latitude for the three sites in BWh, BSh, and ET are 27.82 °, 30.86 °, 47.42 ° respectively. The longitude for them are -15.42 °, 34.78 °, 10.89 ° respectively.

The Köppen climate classification divides climates into 5 main groups indicates by the first letter: A (tropical), B (dry), C (temperate), D (continental), and E (polar). The second letter in all groups except for group E is for the seasonal precipitation. The second letter in group E and the third letter in all other groups is for a temperature subgroup.

BWh, BSh, and ET climate stand for hot desert climate, hot semi-arid climate, and tundra climate respectively. Group B climates are defined by little precipitation, but there is a difference between the BWh and BSh climates, the BSh climate has more precipitation. In comparison, ET is quite different from BWh and BSh, because of the very low temperature with the average temperature of every month in a year is lower than 10 $^{\circ}C$[69].

The eight modules belong to two commercial brands, the brand F module has a glass backsheet (GB) module architecture, while the brand G module has a double glass (DG) module architecture. There are four brand F: GB modules which are glass-backsheet (GB) ones with 60 cells in total and three by-pass diodes and four brand G modules, which are double glass (DG) modules, three of these four have 80 cells and four by-pass diodes and another one has 72 cells and three by-pass diodes. Depending on each specific module, the system ages vary from three to nine years of outdoor exposure. Table 3.1 has more detailed description of these eight PV modules.

Table 3.1. START AND END DATES, BRAND (ARCHITECTURE), NUMBER OF CELLS AND BY-PASS DIODE INFORMATION OF EACH MODULE

ID	Start	End	System Age (Year)	Climate Zone	Brand	Cells Number	By-pass Diode Number
1	2010-10-19	2018-10-31	8.03	BWh	G:DG	80	4
2	2010-02-05	2018-10-31	8.74	BWh	G:DG	72	3
3	2010-09-28	2016-11-24	6.16	BWh	F:GB	60	3
4	2012-06-11	2018-10-31	6.39	BSh	G:DG	80	4
5	2012-06-11	2015-05-17	2.93	BSh	F:GB	60	3
6	2012-06-11	2018-10-31	6.39	BSh	F:GB	60	3
7	2010-06-16	2013-01-31	2.63	ET	G:DG	80	4
8	2010-06-16	2015-02-18	4.69	ET	F:GB	60	3

4   Analytical Methods

This chapter introduces several analytical methods used in the study represented in Fig. 1.1. These include performance loss rate (PLR) determination, I-V curve quality detection, data-driven I-V features extraction ($ddiv$), partial shading detection based on time-series multistep I-V curves, and outdoor I_{sc}-V_{oc} power loss factor calculation. The methods of PLR calculation[9], $ddiv$[70], outdoor I_{sc}-V_{oc} loss factor[3] are published work developed by Solar Durability and Lifetime Extension (SDLE) Research Center. While the I-V curve quality detection method and the partial shading detection method are developed in this thesis research.

4.1   Performance Loss Rate (PLR) Calculation

The maximum power (P_{mp}) of PV modules in the field varies with the environmental conditions such as temperature and irradiance. In order to study the power loss due to degradation, P_{mp} under certain environmental conditions is the basis for comparison.

First, the power and irradiance data is filtered such that the POA irradiance should be higher than 200 W/m^2 so as to remove night time and low irradiance observations. In order to maximize the utilization of the data, instead of simply filtering, we obtain the P_{mp} at certain conditions in two steps.

The first step is temperature correction of P_{mp}, then the *day-by-day* predictive method is applied to get the daily predicted P_{mp} at a certain irradiance value (900 W/m^2) for each day from the linear model fitted by time-series P_{mp} and POA of each day with P_{mp} as response variable and POA as predictor variable as Eq. 4.2. The temperature correction coefficient is obtained from a simple linear model of P_{mp} versus module temperature in the subset with POA in the range of 890 W/m^2 to 910 W/m^2. Then, the P_{mp} is corrected to 40 °C module temperature by the Eq. 4.1, where γ_T is the module temperature correction coefficient.

$$P_{mp,corrected} = \frac{P_{mp}}{1 + \gamma_T(T_{obs} - T_{ref}) \times (\frac{POA}{900})} \qquad (4.1)$$

Next, using the data of each day, we fit the model as Eq. 4.2 and input POA as 900 W/m^2, we get the predicted P_{mp} under the reference condition, which are 900 W/m^2 of POA and 40 °C of module temperature for each day.

$$P_{mp,corrected} = \beta_0 + \beta_1 \times POA \qquad (4.2)$$

Then, *year-on-year* method[17,18,71] is applied to the daily predicted P_{mp} and returns the distribution of PLR. Each PLR is the slope of the linear model fitted by the corrected P_{mp} of one day and another day that is one year after and then divided by the initial predicted P_{mp}, which is obtained by taking the intercept of the y-axis from the linear model regression on all daily predicted P_{mp} versus time. Because this method gets each PLR value from points separated by exactly one year, the final result can form a distribution of PLR, from which we could get a median PLR and standard error for each module to describe the module performance loss. For

more detailed information about the *PLR* calculation method and comparison with other similar methods using the same outdoor data set, please see the paper of *PLR* Consistency and Uncertainty in Methods and Filtering Standards[9] and the *PVplr* R package[72].

4.2 Current-Voltage Curve Quality Detection

The single diode model, also called the five parameter model as shown in Eq. 4.3, is the simplest physical model for describing the PV module's operation under working conditions. I and V are current and voltage of the module, I_{ph} is the photo-current which is strongly correlated with irradiance with a coefficient influenced by cell temperature, I_0 is the inverse saturation current depending on temperature also, R_s and R_{sh} are the series resistance and shunt resistance of the cell, q is the charge of the electron, k is the Boltzmann's constant, γ is the diode quality factor, normally between 1 and 2, N_{cs} is the number of cells in series, and T_c is the effective temperature of the cells in units of Kelvin[73]. The five parameters for the model are I_{ph}, I_0, R_s, R_{sh}, and γ.

The equivalent electrical circuit is shown in Fig. 4.1, where the R_L is the load resistance connected to the module. During an *I-V* tracer voltage sweep, the measurement can be understood as the process of adjusting the load resistance from zero to an extremely large value compared to the R_s of the module itself. Since the duration of sweep is in seconds, it is assumed that I_{ph}, R_s, R_{sh}, T_c, and all other variables, except I and V, remain unchanged. Then from Eq. 4.3, we arrive at the conclusion that $\partial I/\partial V$ would be negative as Eq. 4.4, which indicates that with the increasing

voltage, the current of every point in an *I-V* curve should decrease. And this constraint on an *I-V* curve, that is the first derivative must be negative to be physically meaningful, is the first consideration for the method derived in this thesis to detect the "quality" of measured *I-V* curves.

$$I = I_{ph} - I_0[exp(q \cdot (V + I \cdot R_s)N_{cs} \cdot \gamma \cdot k \cdot T_c) - 1] - (V + I \cdot R_s)/R_{sh} \qquad (4.3)$$

$$\frac{\partial I}{\partial V} = -\frac{a_1 + \frac{a_4}{a_3 - a_4 \cdot V - a_5 \cdot I}}{a_2 + \frac{a_5}{a_3 - a_4 \cdot V - a_5 \cdot I}}$$

$$a_1 = \frac{q}{N_{cs} \cdot \gamma \cdot k \cdot T_c} > 0$$

$$a_2 = a_1 \cdot R_s > 0$$

$$a_3 = \frac{I_{ph}}{I_{sc}} - 1 > 0 \qquad (4.4)$$

$$a_4 = \frac{I_0}{R_{sh}} > 0$$

$$a_5 = \frac{R_{sh} + R_s}{I_0 \cdot R_{sh}} > 0$$

$$a_3 - a_4 \cdot V - a_5 \cdot I = exp(\frac{q \cdot V + q \cdot R_s \cdot I}{N_{cs}\gamma kT_c}) > 0$$

However, the measurement of current has an finite uncertainty determined by the *I-V* tracer instrument, so comparing one point on an *I-V* curve, with it's lower voltage neighbor datapoint, the difference in the current should be negative, but many not be, due to the inherent measurement uncertainty. The curve has data points that can not be explained even by taking the current measurement accuracy into consideration
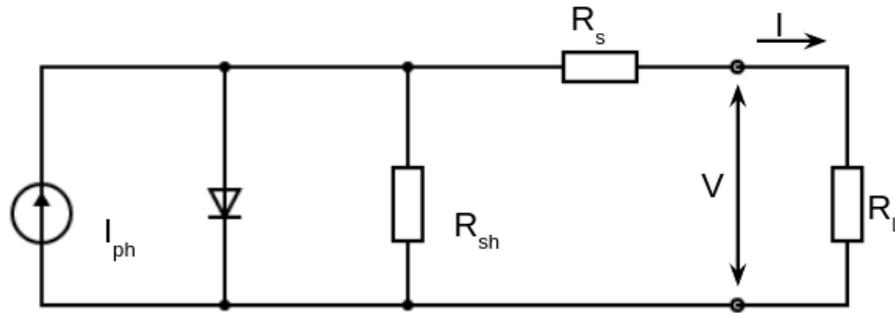
Figure 4.1. Single diode model equivalent electrical circuit for PV module.

is a mistake occurrence during a sweep. This is the second consideration we have for quality detection.

A third criteria is dependent on subsequent analysis steps, and is the number of abnormal points or outliers we can tolerant in an *I-V* curve. This parameter should vary from zero to a small positive integer considering the case that too many curves are filtered out and create problems for further study due to too fewer, or an inadequate number of observations. This is the third consideration for our testing of the quality of *I-V* curve.

The overall process of *I-V* curve quality detection is in Fig. 4.2. $P1$ is the limitation of measured current accuracy, $P2$ is the number of tolerant abnormal points.

As an example of the P1 criteria, consider the specific case of the electronic load ESL-Solar 500 *I-V* tracer used to measure *I-V* curves of PV modules in this study. Its current accuracy is 0.2% of its full scale, and full scale for this instrument is 10 A, so the maximum positive offset we can reliably read is 0.02 A, which is the value for $P1$. And for the $P2$ criteria, we decide that we have no tolerance for outliers in an *I-V* curve analyzed in this study, so $P2 = 0$.
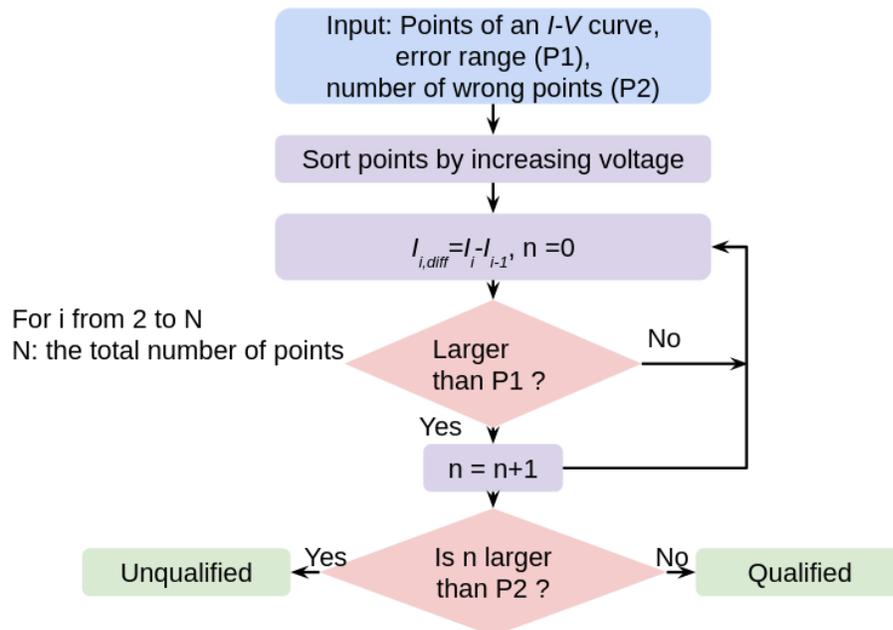
Figure 4.2. Flow chart for I-V curve quality detection ($P1$ is the limitation of measured current accuracy, $P2$ is the number of tolerant abnormal points).

4.3 Data-driven I-V Features Extraction Algorithm

The data-driven I-V features extraction algorithm ($ddiv$) is an R package developed at the Solar Durability and Lifetime Extension (SDLE) Research Center and published in the comprehensive R package archive (CRAN)[70]. The ddiv algorithm has two hyperparameters that can be adjusted for detecting steps in the I-V curves: the maximum number of change points (k) and the critical value of slope difference (m_Δ^a). The ddiv algorithm first fits the points using a 500 point smoothing spline model describing the curve continuously by a model with voltage as the predictor variable and current as the response variable. Five hundred points are evenly generated from the fitted model, then a change-point model, using the segmented R package[74,75], is applied to the generated points with the maximum change points number controlled

by the input parameter k. Next, for each change point, they are two linear models before and after the change point. If the slope of linear model to the left of the change point is negative and the absolute value is larger than that of the linear model to the right of the change point, where the difference is larger than the input m_Δ^a 4.5, then the change point will be counted as a "step" in the curve and the voltage value of its position is recorded as the voltage of the "step". m_1^a is the left slope of step a and m_2^a is the right slope of step a. Using the red curve in Fig. 4.3 as an example, at $V = 21.1$, m_1 is the slope to the left of 21.1 V and m_2 is the slope to the right of 21.1 V.

$$m_\Delta^a = |m_1^a - m_2^a| \tag{4.5}$$

Finally, for each step, a moving window is applied for detecting when the region close to I_{sc} is flat enough to extract I_{sc} and R_{sh} then the linear model is built for some region in the curve partially and I-V features are extracted for each "step" based on the definition which include maximum power (P_{mp}), current at maximum power (I_{mp}), voltage at maximum power (V_{mp}), short-circuit current (I_{sc}), open circuit voltage (V_{oc}), series resistance (R_s), shunt resistance (R_{sh}) and fill factor (FF). More detailed information about how the algorithm works can be found in the document for the package and corresponding paper[26].

Table 4.1 shows the result of two curves by the default setting in the package, one is a single step I-V curve, another is a multiple steps' one, and points of the curves are present in Fig. 4.3 and connected by straight lines.

From how the algorithm works, we know that the I-V step detection results will be dependent on how we choose the values of k and m_Δ^a.

Table 4.1. EXTRACTED *I-V* FEATURES RESULT FOR SINGLE STEP AND MULTISTEP *I-V* CURVES

Step	I_{sc}	R_{sh}	V_{oc}	R_s	P_{mp}	I_{mp}	V_{mp}	FF	$Cutoff$
3	2.915	477.963	43.919	10.389	50.184	2.674	18.767	39.2	21.111
	3.636	14.661	81.756	27.06	55.823	1.979	28.21	18.78	28.27
	2.869	31.698	46.87	1.408	82.129	1.96	41.896	61.08	NA
1	3.109	852.741	47.076	1.24	111.422	2.853	39.059	76.13	NA
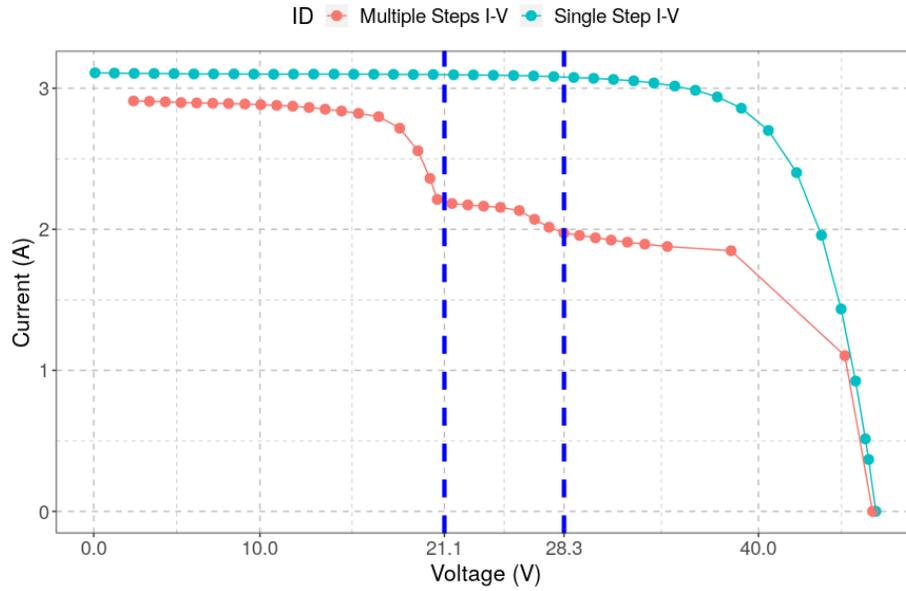
Figure 4.3. multistep and single step *I-V* curves.

If most of the observations in a dataset are single step *I-V* curves, a smaller k value and larger m_Δ^a value gives higher detection accuracy. And when there is an increasing percentage of multistep curves in the dataset, a larger k value and proper m_Δ^a value will increase the accuracy. The way we set the values of these hyperparameters is using a supervised training and testing method, where one manually labels some *I-V* curves as single step curves or multistep curves and then partition the labeled data set into a training and testing dataset. By balancing the accuracy of detecting single step and multistep *I-V* curves in the training data set, and checking the performance

in testing, we determine the optimal values for the two hyperparameters. In addition, in order to avoid the situation where the determined parameter values would limit the physical possible cases, the minimum choice of k should be equal to two times the by-pass diode number minus one. Since in our dataset, the maximum number of by-pass diodes (in Brand G: DG PV module) is four, we choose the minimum k to be seven.

4.4 Partial Shading Detection with Time-series Multistep I-V Curves

Partial shading on the PV module would cause the I-V curves to have multiple steps because of the activation of the by-pass diodes[62] in most cases. Therefore, using the time-series multistep I-V curves, we can detect the occurrence of partial shading during the system lifetime. Studying protocol of partial shading detection includes several aspects: the time dependence, solar angle dependence of occurrence of multistep I-V curves, classification of shading scenarios, and duration of multistep I-V curves.

First, from the *ddiv* result, the percentage of multistep I-V curves (MS) is calculated for each module, and because of the limitation of accuracy in the *ddiv* algorithm, only a module with MS larger than 20% can be reasonably considered to have a partial shading problem during operation. The values of the criteria depend on specific dataset. Next, we study the time dependence of multistep I-V curves occurrence. For each year, the MS of each time point in a day is calculated based on Eq. 4.6, in which n is the number of I-V curves. The point with a local maximum in a spline

model of MS versus time is extracted by a peak finding function "findpeaks" in the *pracma* R package[76] to derive the shading profile which indicates the most probable time to see partial shading happening on a specific PV module.

$$MS_t = \frac{N_{ms,t}}{N_{total,t}} \tag{4.6}$$

Next we study the solar angle dependence of multistep *I-V* curve to obtain information on the relative orientation of the shading obstacles, i.e. the object causing shadows on the PV module[22]. The solar angle is obtained by using "getSunlight-Position" function in *suncalc* R package[77] with the date and time, longitude, and latitude of the module's location. Both the PV module and the obstruction are stationary, but the solar angle, which includes both elevation angle and azimuth angles are changing over time. The angle with the highest MS corresponds to the situation where the sun, the obstacle ,and the PV module most often fall in the same line, so the sun angle with high MS will indicate the relative position of the obstacle and the observed module. In this study, the solar azimuth angle that has local maximum MS, calculated in each azimuth degree interval, is used to represent the relative orientation of the obstacle.

Then, using the $Cutoff$ column from the *ddiv* result, we build the voltage cluster of steps' location expect V_{oc} for all multistep *I-V* curves. Previous research shows that the voltage of steps in *I-V* curves corresponds to different shading scenarios, so based on how the steps voltages vary, we can classify the multistep *I-V* curves into three cases for a three by-pass diode PV module[51]. If there are two steps in the *I-V* curves, labeled as a and b at increasing voltage, the step a falls into the higher voltage cluster (close to V_{oc}), it is classified into the case where only one of three

by-pass diode is activated, which is case 1. If the step a falls into the lower voltage cluster (close to 0 V), it is classified as a case where two by-pass diodes are activated with the shading scenario cause similar change in current in the corresponding two strings, which is case 2. If there are three steps in the I-V curve, labeled as a to c at increasing voltage, the step a falls into the lower voltage cluster, the step b falls into the higher voltage cluster, then the I-V curve is classified into the case where two by-pass diodes are activated with quite different shading scenario that causes different current change in the corresponding two strings, which is case 3. To find the dominant shading scenario for the observed module, the percentage of each case is calculated.

In addition multistep I-V curves, are further classified as being persistent or transient, based on whether at least one time-series neighbor is also multistep. The transient multistep curve case can be caused by a very short-lived obstacle such as a person walking by. While the persistent multistep curve case is more likely to be caused by some stable objects and the time duration of it is meaningful. The results of this classification can indicate whether the measurement interval is long enough to capture the duration of partial shading.

4.5 Outdoor I_{sc}-V_{oc} Loss Factor

The I_{sc} (short circuit current) and V_{oc} (open circuit voltage) curve is determined from I_{sc} and V_{oc} measured at each illumination level. Note that series resistance (R_s) has no effect on V_{oc} since there is no current drawn from the solar cell, and also has no influence on I_{sc} when R_s is less than 10 Ω/cm^2 since the I-V curve is

flat around I_{sc}. The *Pseudo I-V* curve composed from I_{sc}-V_{oc} curve shows the *I-V* curve of the measured cell without the series resistance (R_s) influence[78]. Another measurement that is very similar is the *Suns-V_{oc}* curve, which uses a separate solar cell to monitor the illumination intensity instead of the I_{sc} of the measured cell to simplify the measurement. The I_{sc}-V_{oc} and *Suns-V_{oc}* curves are usually measured indoors to track contact formation, shunting, materials quality, surface passivation, and minority carrier lifetime.

The outdoor I_{sc}-V_{oc} method developed by M. Wang[3], and released as an R package on CRAN[79], uses extracted I_{sc} and V_{oc} from time-series *I-V* curves measured outdoors over a time period to construct the I_{sc}-V_{oc} curve[80]. The time interval is usually one or two weeks depending on the time interval between the measured *I-V* curves. Temperature correction is then applied to the I_{sc}- V_{oc} curve based on Eq. 4.7, T_m is the module temperature in Celsius, the curve is corrected to module temperature at $40\,°C$.

$$V_{oc}(I_{sc}, T_m) = \alpha_0 + \alpha_1 \cdot (T_m + 273.15) \cdot ln(I_{sc}) + \alpha_2 \cdot (T_m + 273.15) + \varepsilon \qquad (4.7)$$

Then *Pseudo I-V* curve is obtained using Eq. 4.8. I_{sc}^0 is the I_{sc} predicted at 1 sun irradiance and $40\,°C$ module temperature. The *Pseudo I-V* curve obtained from the *I-V* curves measured in the first time period (first one or two weeks) is referred to as the initial *Pseudo I-V* curve, and the *Pseudo I-V* curves obtained from the *I-V* curves measured in any other time periods are referred to as degraded *Pseudo I-V* curves.

$$I_{psd}(V_{psd}) = I_{sc}^0 - I_{sc}(V_{oc}) \tag{4.8}$$

By comparing the initial *Pseudo I-V* curve and degraded *Pseudo I-V* curve, the difference between I_{sc}, V_{oc} can be calculated based on the one diode model equation, Eq. 4.3. Then the I_{sc}, I_{sc} and V_{oc} corrected *Pseudo I-V* curve can be calculated. Compared with the initial *Pseudo I-V* curve, the differences in P_{mp} due to the change in I_{sc} and V_{oc} when module is degrading are obtained, they are the power loss due to uniform current loss and recombination respectively. Next, R_s is extracted from *I-V* curves measured in each time period, can then be predicted at the reference condition, which is 40 $°C$ of module temperature and one sun irradiance based on Eq. 4.9.

$$R_s(I_{sc}, T_m) = \zeta_0 + \zeta_1 \frac{T_m + 273.15}{I_{sc}} + \varepsilon \tag{4.9}$$

Compared with the one extracted from the degraded *Pseudo I-V* curve, there is a R_s difference for the predicted R_s. The P_{mp} difference caused by this R_s difference correction to the degraded *Pseudo I-V* curve is the power loss due to series resistance (R_s). Finally, the P_{mp} difference between the P_{mp} calculated from I_{mp} and V_{mp} extracted and predicted from degraded real *I-V* curve and the one extracted from the degraded *Pseudo I-V* curve after correction for R_s difference is the power loss caused by current mismatch, which is the short circuit current difference in the cells connected in series. Fig. 4.4 shows the process for obtaining the power loss of each loss mode for each degraded period.

In short, the outdoor I_{sc} - V_{oc} curve contains information from a large number of *I-V* curves measured at varying irradiance, so it reduces the dimension of the

Figure 4.4. Flow chart of power loss calculation for each power loss mode from outdoor I_{sc}-V_{oc} analysis of time-series I-V curves.

data from three to two, and the result is the time-series power loss of four types. In addition, the calculation of power loss modes has the advantages that changes caused by I-V features can be directly comparable on power basis. In the outdoor I_{sc}-V_{oc} study, the power loss due to the difference in I_{sc} is called uniform current loss (ΔP_{Isc}), the power loss due to the difference in V_{oc} is called recombination loss (ΔP_{rec}) and the power loss due to difference in R_s is called R_s loss (ΔP_{Rs}), and other P_{mp} difference is called current mismatch loss (ΔP_{Imis}). Other previous research also discussed in detail determination of composing the outdoor I_{sc}-V_{oc} curve and various power loss calculation methods[31]. To compare systems with different total power outputs, the percentage change in power is calculated by using the power loss divided by the initial

predicted power under the same reference conditions. And this is obtained from a linear model fit to the predicted P_{mp} versus time, the predicted P_{mp} is also provided in the process of the I_{sc}-V_{oc} analysis demonstrated in Fig. 4.4.

5   Results

This chapter is divided into two parts. The first part (5.1) focuses on the application of these methods to obtain the results of analyzing time-series data, using a small dataset of a single PV module as an example. We illustrate the results of I-V curve quality detection, showing several unqualified I-V curves, and also the way to determine the optimal hyperparameters' values k and m_Δ^a for $ddiv$ using the labeled training and testing dataset, partial shading detection using data of module 3 (GB: DG), and the rate of change in power loss of each degradation mode from power loss obtained from outdoor I_{sc}-V_{oc} analysis using module 2 (BWh: DG). The second section (5.2) presents the complete results of analyzing time-series data of all eight modules in the Fraunhofer-ISE dataset, which include PLR, I-V curves quality detection, rates of change of degradation modes, and partial shading detection.

5.1   Illustration of Analytical Methods

5.1.1 I-V Curve Quality Detection

Fig. 5.1 shows some randomly selected I-V curves identified as unqualified ones in the setting of the error range $P1 = 0.02\ A$ and the number of tolerable abnormal

points $P2 = 0$. In the figure the red points are the ones identified as abnormal points by the quality detection algorithm, which means they have a current difference larger than 0.02 A compared with the adjacent higher voltage datapoint. The I-V curves that are identified as unqualified have a region of increasing current with increasing voltage and the I-V curves have regions of positive radius of curvature.

(a) a

(b) b

(c) c

(d) d

(e) e

(f) f
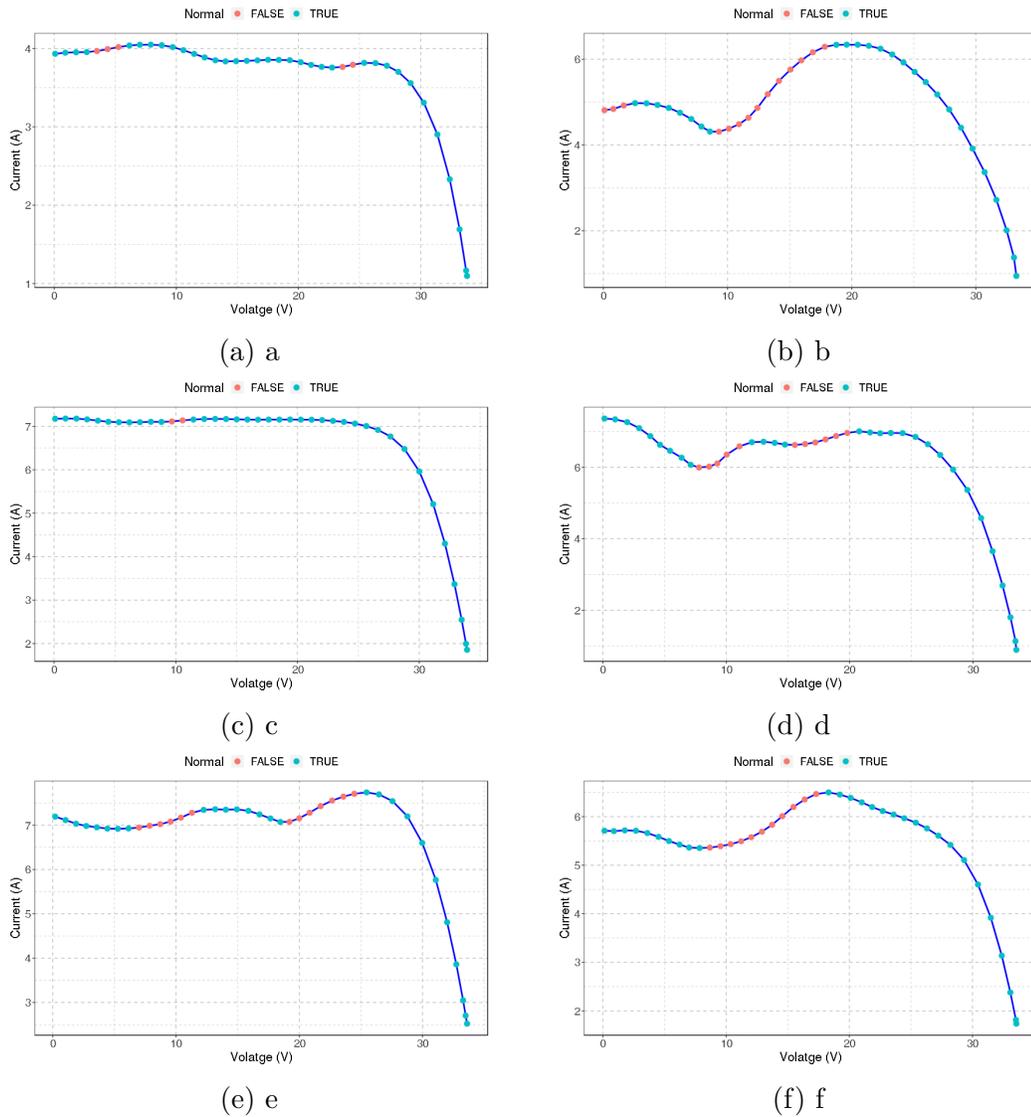
Figure 5.1. Examples of I-V curves identified as unqualified with P1 = $0.02A$ and P2 = 0, red points are abnormal points demonstrating the anomolous areas of the curve, and blue are normal datapoints.

If setting the number of tolerable abnormal points as 0 in the quality detection algorithm removes too many observations in a particular dataset cause a problem for further studies then it is reasonable to consider increasing the $P2$ value. From Fig. 5.1, if we set the value of $P2$ to 5 instead of 0, then Fig. 5.1a and Fig. 5.1c will be identified as qualified *I-V* curves and they do look more normal than the others.

5.1.2 Determination of Input Hyperparameters k and m_Δ^a for $ddiv$

k and m_Δ^a are two important input hyperparameters for the $ddiv$ algorithm to detect steps in the *I-V* curves, and determining their values is important to arrive at meaningful results. Therefore to determine these hyperparameters, we take ten thousand *I-V* curves randomly selected from all two million *I-V* curves of the eight modules in the Fraunhofer-ISE dataset (about 0.5%) and manually label them as single step *I-V* curves or multistep *I-V* curves. One thousand single steps and another one thousand multistep *I-V* curves are randomly selected from these ten thousand labeled observations. Next, we divided each type of the *I-V* curves into a training and testing datasets with 80% and 20% partitioning, so in each dataset, it has the same number of single step and multistep observations. Then, we perform a grid search to determine the optimal hyperparameter values of k and m_Δ^a with the range of k from five to 14 by one and the range of m_Δ^a from 0.005 to 0.030 by 0.001.

For the training data set, we calculate the accuracy of predicting the single step *I-V* curves and the multistep *I-V* curves as shown in Fig. 5.2 and the overall accuracy. The accuracy of detecting single steps is defined as the number of *I-V* curves predicted to be single step that are labeled as single step divided by the total number of single step *I-V* curves. The accuracy of detecting multistep is defined as the number of *I-V*

curves predicted as multistep, also labeled as multistep, divided by the total number of multistep I-V curves. The accuracy of the dataset is defined as the number of curves predicted the same as the label divided by the total number of curves.
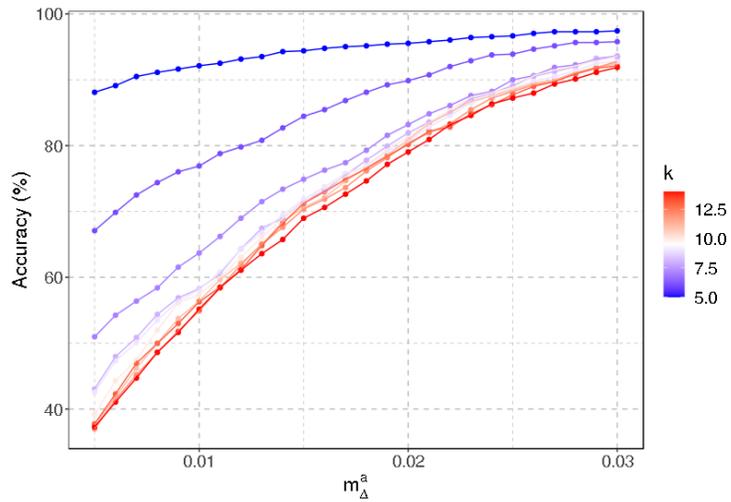
(a) Single step accuracy
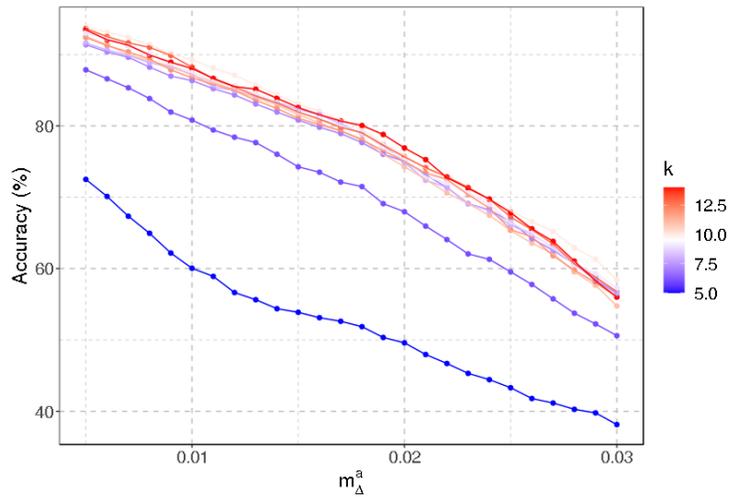
(b) Multiple steps accuracy

Figure 5.2. Accuracy of detecting single step and multistep I-V curves in the training dataset as a function of the two *ddiv* hyperparameters k and m_Δ^a.

From Fig. 5.2a, we can see that the accuracy of the detection of single step *I-V* curves increases as k decreases and m_Δ^a increases, which is consistent with the design of *ddiv* algorithm. We also find that, with increasing m_Δ^a, k has less influence on the accuracy of detecting single-step observations. From Fig. 5.2b, the trend for detecting multistep *I-V* curve is the converse, and it is in competition with the accuracy of detecting single step *I-V* curves. However the m_Δ^a seems to have less influence on the importance of k to the accuracy of detecting multistep observations.

Since detection of single step and multistep curves are in competition, the k and m_Δ^a hyperparameters should be choosen to maximize the overall accuracy, with the secondary condition that the two types of accuracy be balanced. In addition, due to the number of by-pass diodes in the eight modules we studied, in order to include all possible physical cases, the minimum value of k for processing our dataset is determined to be seven. Based on these three considerations, there are two sets of hyperparameters, shown in Table 5.1 that performed similarly and very well in training and testing.

Table 5.1. PERFORMANCE IN TRAINING AND TESTING DATASET FOR THE BEST TWO SET OF HYPERPARAMETERS

k	m_Δ^a	Training AS (%)	Training AM (%)	Training Accuracy (%)	Testing AS (%)	Testing AM (%)	Testing Accuracy (%)
7	0.018	79.3	77.7	78.5	94.2	71.1	82.7
8	0.018	77.8	78.9	78.4	91.2	74.7	83.0

The performance difference between these two sets of hyperparameters is small, and seven is the limitation we set for k. In testing, k is equal to eight has better balance between the accuracy of detecting single-step *I-V* curve (AS) and the accuracy of detecting multistep *I-V* curves (AM) while also achieving higher overall accuracy.

Finally, $k = 8$, and $m_\Delta^a = 0.018$ are choosen as the optimal hyperparameter values for processing the complete time-series *I-V* curves dataset for all eight modules.

5.1.3 Partial Shading Study Process

Due to the accuracy limitations in detecting single step and multistep *I-V* curves, we set a criteria for selecting modules which have the partial shading problem that is likely to be caused by a stationary object. Assuming the accuracy of detecting single step *I-V* curves is AS_i and that of multistep ones is AM_i, the percentage of multistep *I-V* curves for module i is $MS_{Real,i}$, and the percentage of multistep *I-V* curves calculated from *ddiv* result(MS) would be as Eq. 5.1.

From Eq. 5.1, we can see that if AS_i and AM_i are both 100%, then the MS will be equal to MS_{Real}. However, according to the result of the training data set in Table 5.1, AS and AM are competitive. In the training data set, they are 77.8% and 78.9%. From Eq. 5.1, we can see that when there is no multistep *I-V* curves in the dataset, we would still get a MS equal to 22.2% because of the accuracy limitation of the algorithm. This 22.2% is considered as an average estimation across all the studied modules. For our dataset, only in the case of MS of one module being higher than 20%, we are confident that the module has an amount of true multistep *I-V* curves that are not detecting error. The MS for each module is possibly to be lower than 22.2% because the AS_i and AM_i for each module could be different from the estimate obtained through the training dataset which contains *I-V* curves randomly selected out from the whole dataset.

$$MS_i = (1 - AS_i) \times (1 - MS_{Real,i}) + AM_i \times MS_{Real,i} \tag{5.1}$$

The *MS* of each module is shown in Table 5.12. Using 20% as a criterion, module 3 (BWh: GB), module 5 (BSh: GB) and module 6 (BSh: GB) are very likely to have severe partial shading. In this section, we use module 3 (BWh: GB) as an example to illustrate the partial shading detection result including the shading profile, obstacle orientation, shading scenarios classification, and duration of partial shading.

A partial shading diagram, with time of day as the x-axis, and date in a year as the y-axis, Fig. 5.3 shows the occurrence of multistep *I-V* curves in each year for module 3 (BWh: GB). The red points are for single step observations and the green points are for multistep ones, with the datapoints having transparency, so that the color intensity is also shows the density of observations.
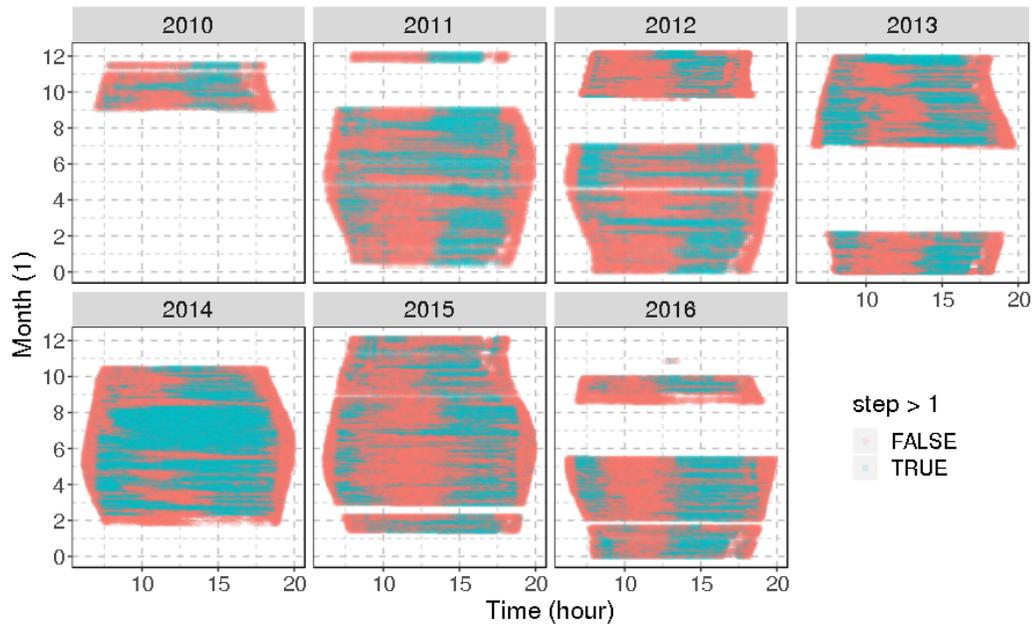
Figure 5.3. A partial shading diagram visualizes the occurrence of multistep *I-V* curves for module 3(BWh: GB).

The gaps along the y-axis of the partial shading diagram indicate the *I-V* tracer was offline. In addition the the pear shaped border of the partial shading diagram

arises from the seasonal time change of sunrise and sunset, and arises due to the the irradiance cutoff of 5 W/m^2, which we used for removing nighttime. The occurrence of multistep *I-V* curves is strongly dependent on the time of day, for module 3 (BWh: GB), most of them showed up in the morning and afternoon, which indicates there are shading obstructions around the module and the afternoon one usually persists for longer times.

Next, we want to quantify when the partial shading happens most often. For each year, we will determine the shading profile, using a natural spline model fitted using *MS* as the response variable and time of day as the predictor[81]. Fig. 5.4 is an example for 2011. The red dashed line is the shading profile spline model. There are two peaks in the shading profile for *MS* in the year of 2011, one peak is located at about 8:00 am and another is located at about 16:00 pm. Using the "findpeaks" function in *pracma* packages[76], we can extract the exact peak position and create a shading profile of each year, the result of module 3 (BWh: GB) is in Table 5.2. If one year has observations of less than 100 days, we think the number of observations is too low to evaluate the module for that year, and the shading profile will not be created for it. The peak location is the time most likely to correspond to a partial shading condition if someone comes to do on-site inspection. The value of 100 days can be modified, but the number of observation will directly influence the resolution in *MS*, when it's too small, the curve in Fig. 5.4 will show high fluctuation instead of a smooth change.
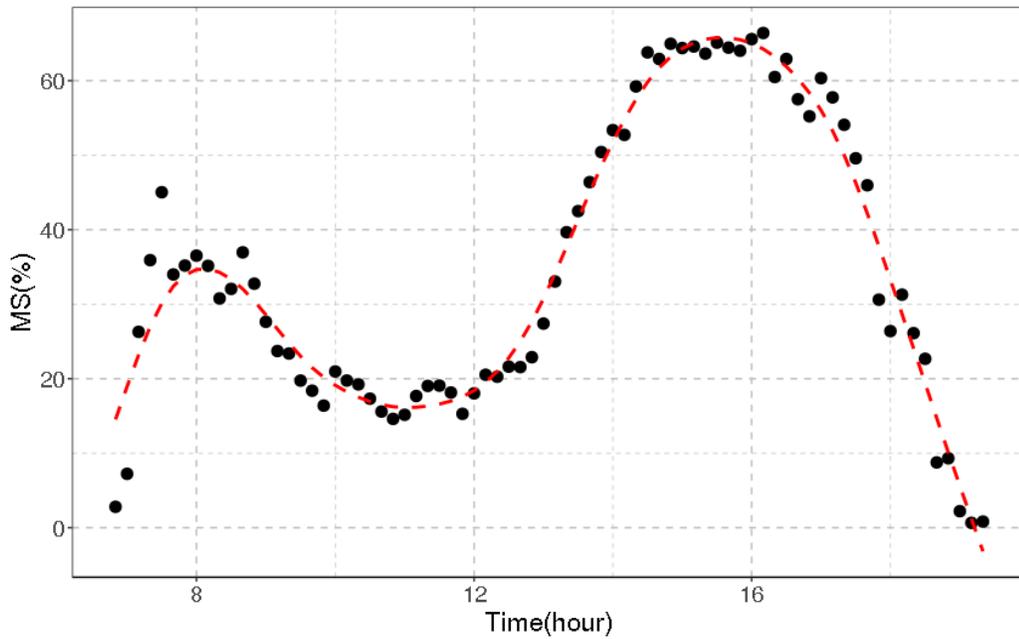
Figure 5.4. Partial shading profile of *MS* vs. time in a day for module 3 (BWh: GB) in the year of 2011.

Table 5.2. SHADING PROFILE FOR MODULE 3 (BWH: GB)

Time in a Day	$MS(\%)$	Year
08:10:00	35.1	2011
15:30:00	65.1	2011
07:10:00	46.9	2012
14:46:00	57.1	2012
08:10:00	55.7	2013
15:20:00	61.9	2013
08:40:00	58.9	2014
15:45:00	78.0	2014
07:55:00	32.1	2015
16:10:00	52.1	2015
07:50:00	40.0	2016
15:15:00	64.0	2016

We can also visualize the relative orientation of the Sun, the PV module and the stationary shading obstacle, using what we will call a Shading Poynting Vector diagram in honor of John Henry Poynting[82]. After converting date and time into solar
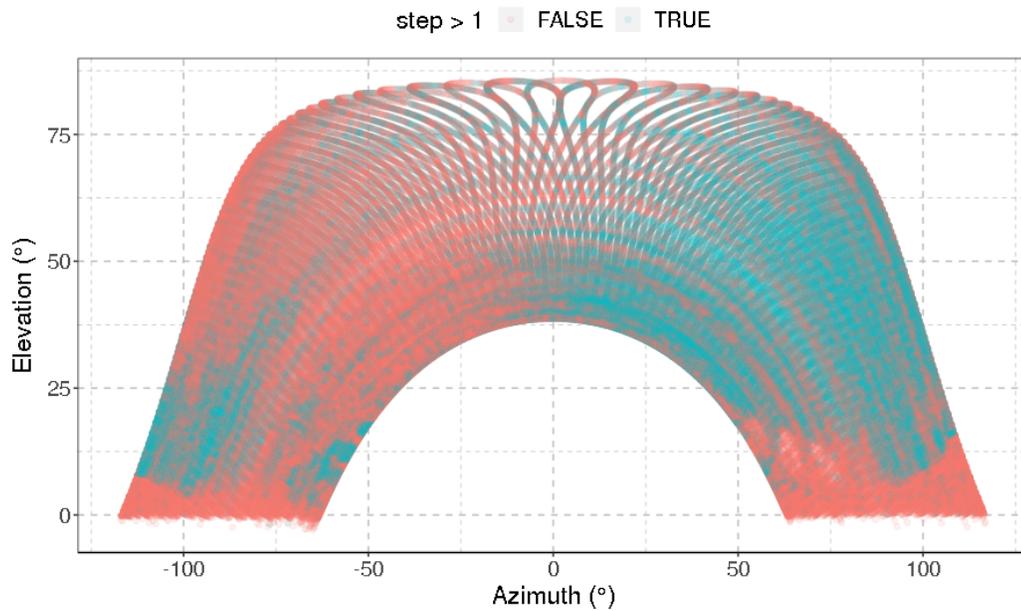
Figure 5.5. A shading Poynting vector diagram shoing the occurrence of multistep *I-V* curves for module 3(BWh: GB) with solar angles as coordination.

elevation and azimuth angles using the information of the longitude and latitude of the module, we plot the occurrence of multistep *I-V* curves in Fig. 5.5. Fig. 5.5 uses all observations of module 3 (BWh: GB) as an example. Since both the solar panel and the shading obstacle are stationary, the two angles corresponding to the green points' cluster indicate the relative position of the obstacle which can cause shadow on the module.

From Fig. 5.5, the obstacle on the east of module 3 (BWh: GB) usually shades the module when the solar elevation angle is in the range of 0 ° to 30 °, which is quite low, but the obstacle on west of the module shades the module in a much wider range of higher solar elevation angle. In contrast to the solar elevation angle, the solar azimuth angle is more helpful to locate where the obstacle is located. We calculate
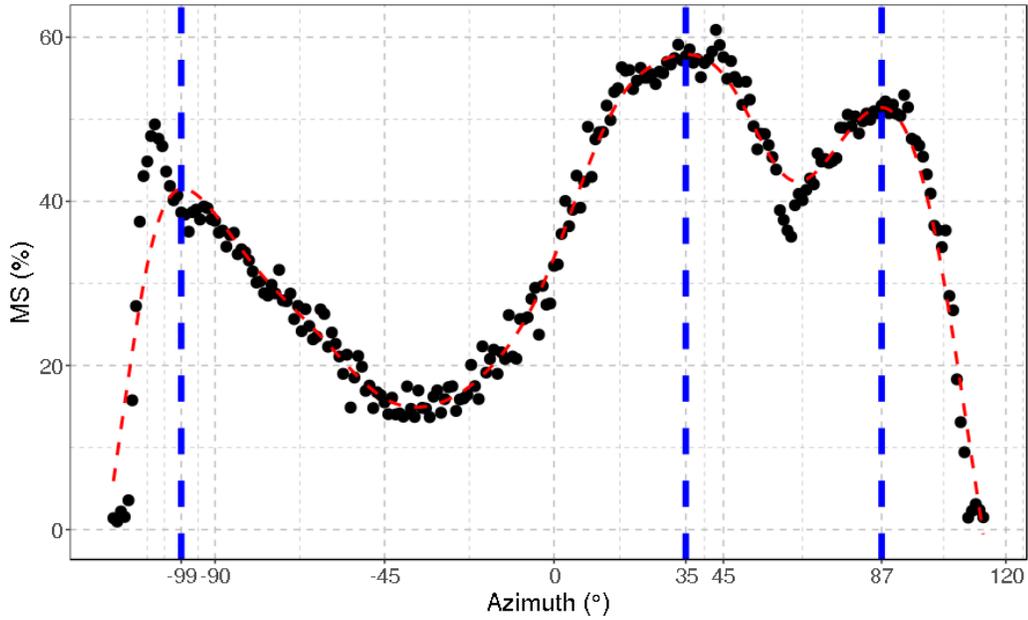
Figure 5.6. *MS* vs. solar azimuth angle.

the *MS* of each degree of the solar azimuth angle by Eq. 5.2, in which N is the number of *I-V* curves.

$$MS_a = \frac{N_{ms,[a,a+1)}}{N_{total,[a,a+1)}} \tag{5.2}$$

Fig. 5.6 shows how the *MS* changes with solar azimuth angle, using the similar method to get the shading profile, we can extract the solar azimuth localized peak location, but this time we directly do this to all observations at once instead of group by each year. They are -99 °, 35 °, 87 ° for module 3 (BWh: GB).

A previous study[51] showed that, for a module with three by-pass diodes, the *I-V* curve has two steps and the first step's location is closer to V_{oc} when one by-pass diode is activated. And that it will be move closer to 0 volts when two by-pass diode are activated if the corresponding two strings have similar shading scenarios that cause similar change in current. If there are a total of three steps in the *I-V* curves,
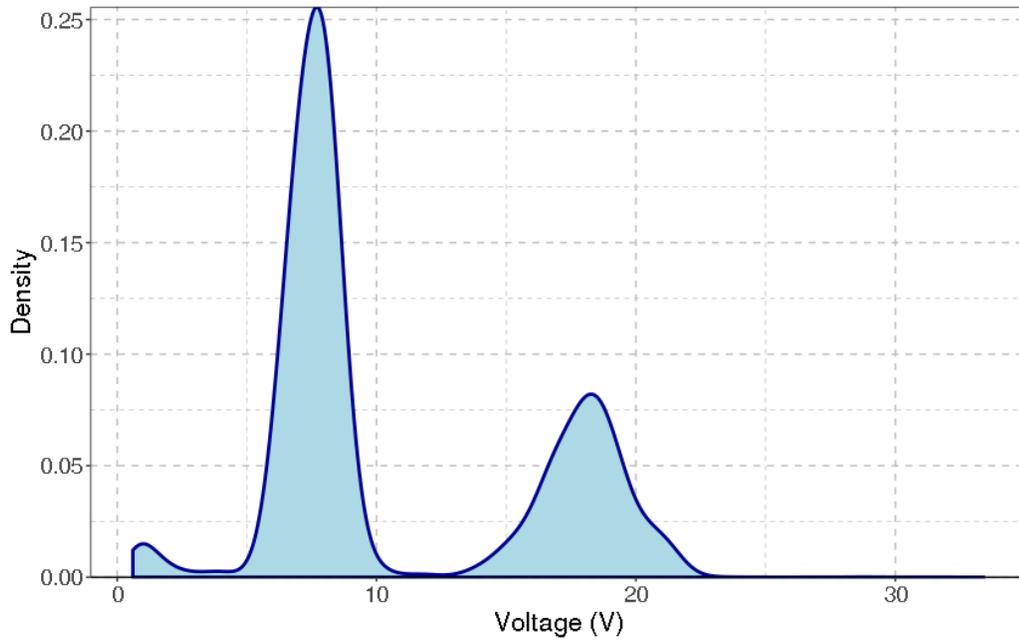
Figure 5.7. *MS* versus solar azimuth angle.

it indicates that two by-pass diodes are activated and the shading scenarios on the corresponding two strings cause quite different current changes. Fig. 5.7 shows the density of steps' location in voltage, which is saved in the "*cutoff*" column of *ddiv* result for two and three steps *I-V* curves of module 3 (BWh: GB). There are only 0.1973% multistep *I-V* curves which have more than three steps.

Similar to the way we created the shading profile, we can calculate the percentage of observations with steps located at each 0.1 voltage interval and then extract the voltages corresponding to the start, end ,and maximum density value for each peak. For the module 3 (BWh: GB), the density distribution shows three clusters, the first one is generated by the 500 points generated from *ddiv* algorithm because the shortage in stability of the region that is very close to I_{sc} due to tracking systems, and the other two peaks are the clusters of step location that correspond to the by-pass diodes status. The result of voltage clusters is in Table 5.3 for module 3 (BWh: GB).

Table 5.3. VOLTAGE CLUSTERS FOR MODULE 3 (BWH: GB)

Cluster No.	Start (V)	End (V)	Peak (V)
1	3.9	11.1	7.6
2	11.1	23.9	18.1

Using the start and end voltages of each cluster, we sort the multistep *I-V* curves into three different shading scenarios, which are one by-pass diode activation (case 1), two by-pass diodes activated with the same (case 2) or different shading scenarios (case 3). The classification result of module 3 (BWh: GB) is listed in Table 5.17. For module 3 (BWh: GB), case 1 counts for 22.86%, case 2 counts for 61.19% and case 3 counts for 15.95%. So it's more common to have two by-pass diodes activated roughly equal in module 3 (BWh: GB).
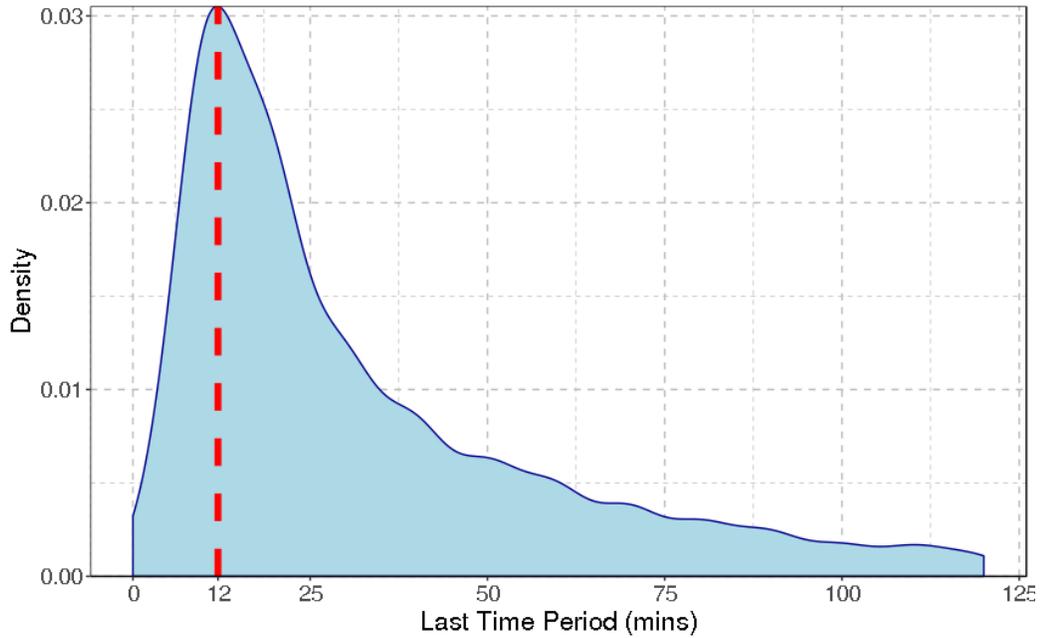
Figure 5.8. Duration of persistent multistep *I-V* curves for module 3 (BWh: GB).

Next, according to whether there is at least one adjacent observation that is also multi-steps', the multistep *I-V* curves are classified into persistent or transient curves.

For all multistep *I-V* curves of module 3 (BWh: GB), persistent ones are 89.91% and transient ones are 10.09%. This finding indicates that the by-pass diode activation duration usually is long enough to be captured by our time resolution. Fig. 5.8 shows density distribution of duration of persistent multistep *I-V* curves. The peak is at 12 minutes.

5.1.4 Outdoor I_{sc}-V_{oc} & Power Loss Factors Calculation

Fig. 5.9 shows the percentage change in power caused by four degradation modes including uniform current loss (ΔP_{Isc}), recombination (ΔP_{rec}), power loss due to R_s increase (ΔP_{Rs})and non-uniform current loss (ΔP_{Imis}) for module 2 (BWh: DG) from outdoor I_{sc}-V_{oc} analysis. A negative value indicates the module has loss in power. The lines in Fig. 5.9 are simple linear models fitted for each degradation mode.
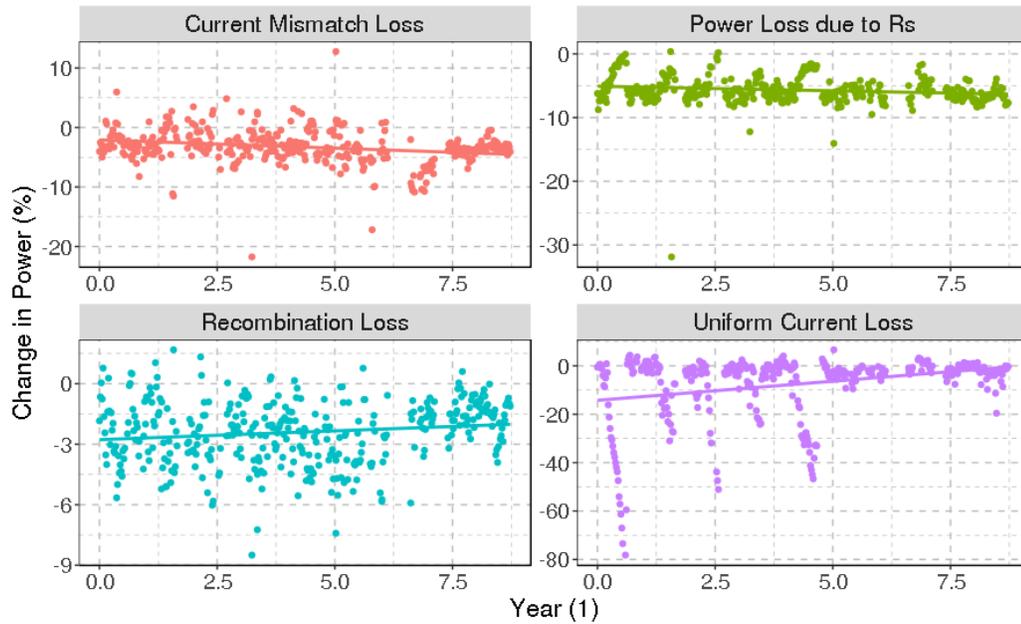
Figure 5.9. Time-series power change from outdoor I_{sc}-V_{oc} analysis for module 2 (BWh: DG).

In the next step, we use Eq.5.3 to remove outliers with β is equal to 1.5, Q_1 and Q_3 are the lower and upper quartiles respectively[83].

$$[Q_1 - \beta(Q_3 - Q_1), Q_3 + \beta(Q_3 - Q_1)] \tag{5.3}$$

In order to handle the seasonality that may exist in different degrees for different modes and modules, we apply a month-by-month linear regression. The specific implementation is as follows: first, fit a simple linear model using data in each month from all years, for example January across all years, and then get 12 slopes of each power loss mode of each module, which are the slopes from models of each month. Next, we remove the slopes which are estimated from the data of the month that miss observations of the beginning and the ending year or more than two years. Using the rest of the slopes, we calculate the average and the standard deviation to get the 95% confidence interval of the rate of change in unit of $\%/a$.

Fig. 5.10 shows the *month-by-month* simple linear regression on the uniform current loss in module 2 (BWh: DG). The slopes of each month is recorded in Table 5.4. Age is the difference between the ending year and the beginning year and count is the number of years have observations. The system age of module 2 (BWh: DG) is 8.73 years, so it had observations in 8 complete years. The linear slopes of July and August are removed because they are missing more than 2 years and the age for August is 3 years shorter, which indicates that it missed both beginning and ending years. Using the remaining 10 slopes, we calculate the average and standard deviation. They are 0.1049 $\%/a$ and 0.3439 $\%/a$ respectively for uniform current loss of module 2 (BWh: DG).
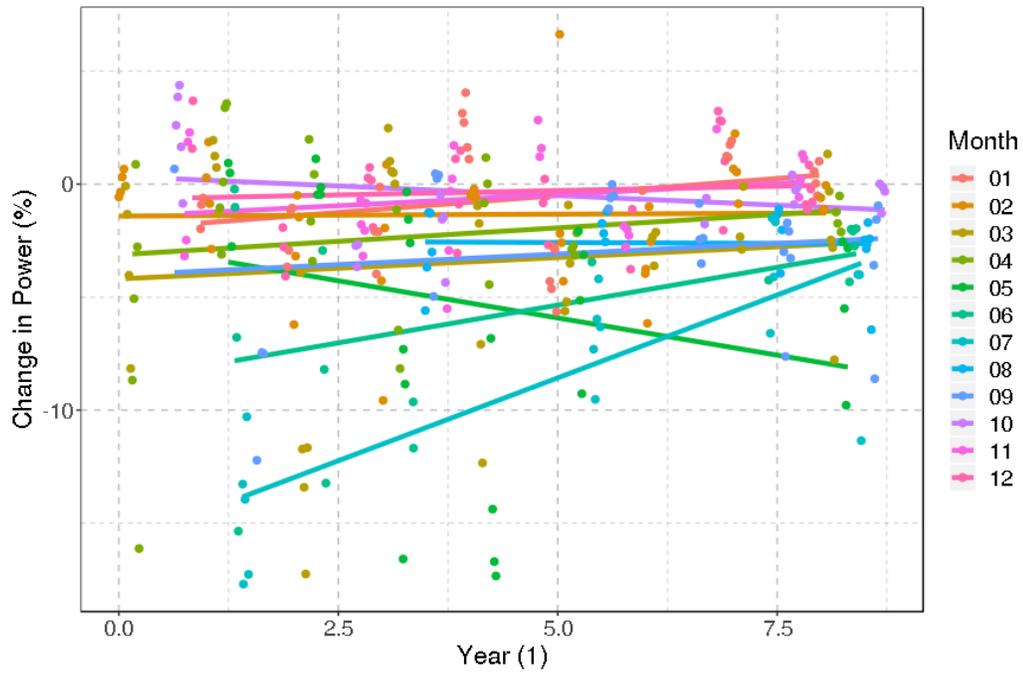
Figure 5.10. *month-by-month* simple linear regression on the uniform current loss (ΔP_{Isc}) in module 2 (BWh: DG).

Table 5.4. RATE OF CHANGE FROM *month-by-month* SIMPLE LINEAR MODEL OF UNIFORM CURRENT LOSS OF MODULE 2 (BWH: DG)

Month	Rate of Change ($\%/a$)	Age(year)	Count
01	0.3009	7	8
02	0.0164	8	9
03	0.1897	8	9
04	0.2383	8	7
05	-0.6586	7	6
06	0.6697	7	6
07	1.4678	7	4
08	-0.0157	5	4
09	0.1870	8	7
10	-0.1697	8	6
11	0.1979	7	7
12	0.0771	7	8

5.2 Time-series Result of All Eight Modules

This sections contains the results of all eight modules including *PLR*, pre-processing *I-V* curves by quality detection and *ddiv*, outdoor I_{sc}-V_{oc} analysis with month-by-month linear regression, and partial shading detection.

5.2.1 Performance Loss Rate (*PLR*) Result

For each module, the *PLR* calculated using *year-on-year* method could form a distribution, and the median *PLR* are listed in Table 5.5.

Table 5.5. MEDIAN *PLR* FOR ALL EIGHT MODULE

Module ID	Climate Zone	Brand	Median *PLR* (%/a)
1	BWh	G:DG	-0.039
2	BWh	G:DG	-0.085
3	BWh	F:GB	-0.473
4	BSh	G:DG	-0.465
5	BSh	F:GB	-0.286
6	BSh	F:GB	-0.600
7	ET	G:DG	0.292
8	ET	F:GB	0.321

Fig. 5.11 shows the median *PLR* and the 95% confidence interval colored by located climate zones and in different shapes by the module architectures.
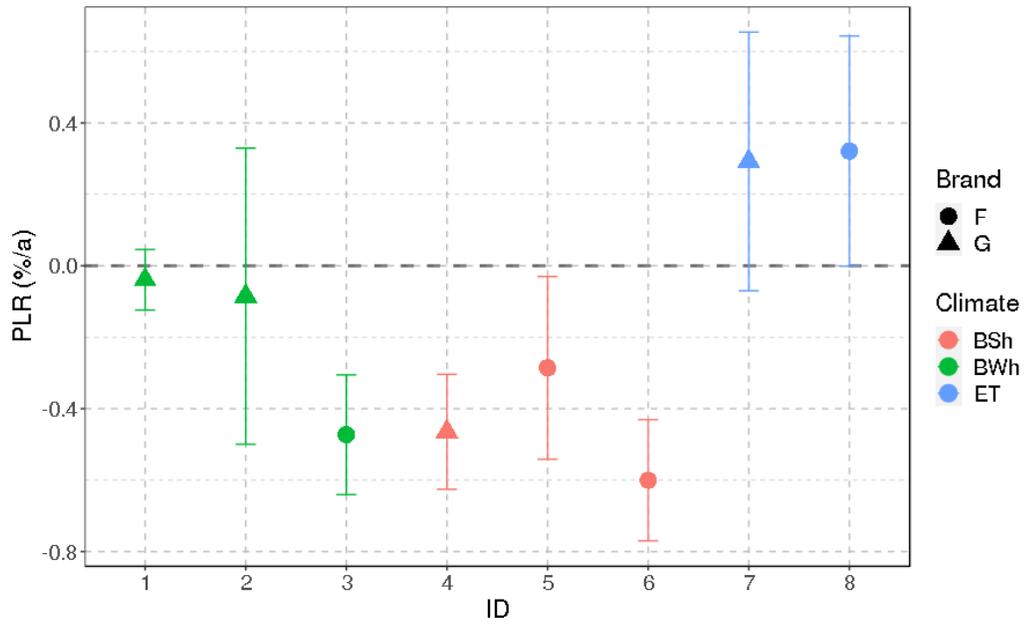
Figure 5.11. Median *PLR* and 95% confidence interval for all eight modules.

5.2.2 Preprocessing of Time-series I-V Curves

The time-series I-V curves are first preprocessed using the I-V curves quality detection and then using *ddiv* algorithm[70] for the I-V features extraction followed by the outdoor I_{sc}-V_{oc} analysis and partial shading studies. The time-series I-V features are used in the outdoor I_{sc}-V_{oc} and power loss factor calculation to study the degradation mechanisms leading to the power loss using both qualified and unqualified I-V curves. Using the I-V curves quality detection method, with the error range ($P1$) set as 0.02 A and the number of tolerable wrong points ($P2$) in an I-V curve set as 0, the percentage of unqualified I-V curves among all I-V curves for each module are listed in Table 5.6. The value varies from 1.2% to 11.7%, and the average is 4.65%. The I-V curves labeled as unqualified are removed for partial shading study.

Table 5.6. Percentage of Unqualified *I-V* curves for each module

ID	System Age (Year)	Climate Zone	Brand	Unqualified (%)
1	8.03	BWh	G:DG	3.8
2	8.74	BWh	G:DG	3.1
3	6.16	BWh	F:GB	11.4
4	6.39	BSh	G:DG	1.2
5	2.93	BSh	F:GB	3.2
6	6.39	BSh	F:GB	1.3
7	2.63	ET	G:DG	11.7
8	4.80	ET	F:GB	1.5

Using the *ddiv* hyperparameters k as 8, and m_{Δ}^{a} as 0.018, the time-series *I-V* curves of all eight modules were processed. And only a few of *I-V* curves can not be analyzed by *ddiv*, and the percentage of them for each module is listed in Table 5.7.

Table 5.7. Percentage of *I-V* curves that can not be analyzed by *ddiv*

ID	System Age (Year)	Climate Zone	Brand	NA in *ddiv* (%)
1	8.03	BWh	G:DG	0.401
2	8.74	BWh	G:DG	0.771
3	6.16	BWh	F:GB	1.826
4	6.39	BSh	G:DG	0.063
5	2.93	BSh	F:GB	0.739
6	6.39	BSh	F:GB	0.079
7	2.63	ET	G:DG	0.015
8	4.80	ET	F:GB	0.168

We randomly selected out four *I-V* curves from module 1 (BWh: DG) that could not be analyzed using *ddiv*. They are shown in Fig. 5.12. The message returned from *ddiv* is an error in smooth.spline(V, I): " smoothing parameter value too small", so the algorithm stopped when it fit smooth.spline model to the points of the curve. Because there are very few observations that are unable to get *ddiv* results, only

about 1.8% to 0.02% for different modules, the number of observations that could not be used in future study is in an acceptable range.
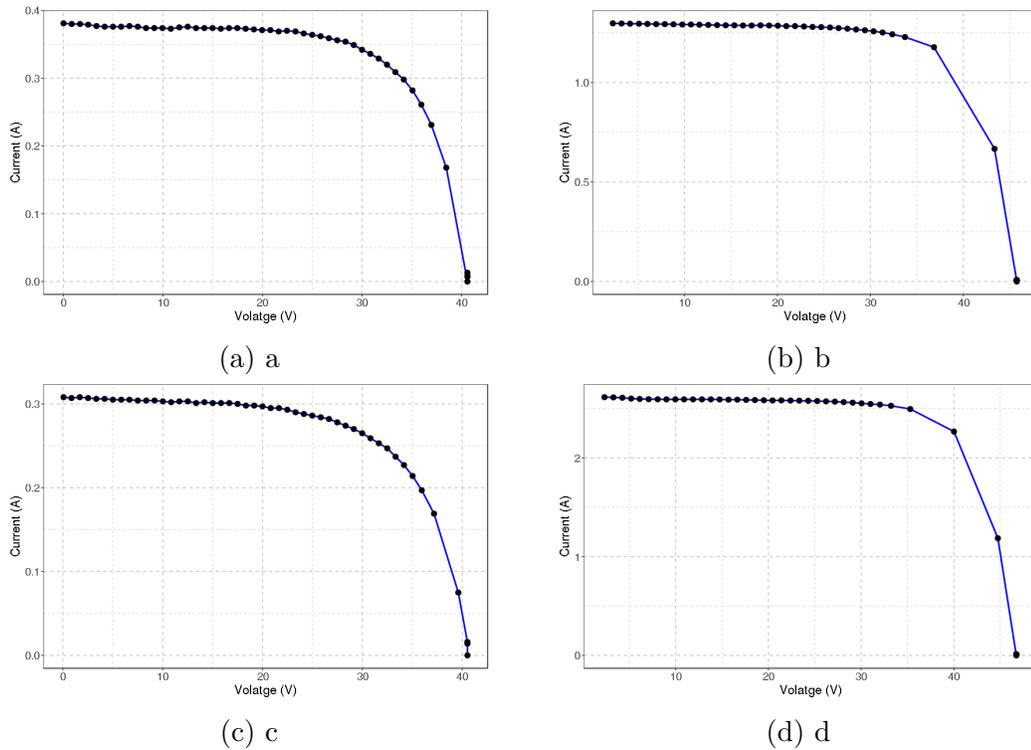
Figure 5.12. Examples of *I-V* curves failed in *ddiv*.

5.2.3 Outdoor I_{sc}-V_{oc} & Loss Factor Result for All Eight Modules

By performing linear regression on the predicted P_{mp} under a reference condition, which is irradiance at 1 sun and module temperature at 40 $°C$, we obtained the intercept of the y-axis as the initial predicted P_{mp} for each module from the fitted linear model. This value is used to normalize the change in power from the unit of W to % in the next step. The initial predicted P_{mp} of each module is listed in Table 5.8. The p_value for all of them are smaller than 1×10^-24.

Table 5.8. INITIAL PREDICTED P_{mp} FOR ALL EIGHT MODULES

Module ID	Climate Zone	Brand	Initial predicted P_{mp}	std_error
1	BWh	G:DG	264.39	1.716
2	BWh	G:DG	138.73	2.062
3	BWh	F:GB	239.90	7.015
4	BSh	G:DG	224.20	0.552
5	BSh	F:GB	275.96	19.311
6	BSh	F:GB	234.91	1.064
7	ET	G:DG	250.26	10.835
8	ET	F:GB	237.69	18.395

Then apply our *Month-by-Month* linear regression to calculate the average and standard deviation for each mode of each module, and further based on Eq. 5.4 to calculate the 95% confidence interval (CI) with z is equal to 1.96 and n is the number of observations. The results of the average rate of change are in Table 5.9 and the standard deviation results are in Table 5.10. The results of the rate of change and their confidence interval for each power loss mode in each module are present in Fig. 5.13.

$$\left(\bar{x} - z \times \frac{\sigma}{\sqrt{n}}, \bar{x} + z \times \frac{\sigma}{\sqrt{n}}\right) \tag{5.4}$$

Table 5.9. AVERAGE RATE OF CHANGE OF EACH POWER LOSS MECH-ANISM FOR EACH MODULE

ID	ΔP_{Isc} (%/a)	ΔP_{rec} (%/a)	ΔP_{Rs} (%/a)	ΔP_{Imis} (%/a)
1	-0.2039	0.0082	-0.0039	-0.1732
2	0.1049	0.0835	-0.1280	-0.1851
3	-0.7469	-0.0532	0.1224	-0.2497
4	-0.5648	0.0278	-0.0308	0.0560
5	-1.6676	0.1275	0.7762	-0.0411
6	-0.7502	-0.0325	-0.0042	-0.0825
7	0.1426	-0.0107	0.2752	0.1266
8	0.1261	0.0792	-0.1272	0.4289

Table 5.10. STANDARD DEVIATION OF RATE OF CHANGE OF EACH
POWER LOSS MECHANISM FOR EACH MODULE

ID	ΔP_{Isc} (%/a)	ΔP_{rec} (%/a)	ΔP_{Rs} (%/a)	ΔP_{Imis} (%/a)
1	0.1223	0.1217	0.1314	0.1829
2	0.3439	0.0953	0.1854	0.1195
3	0.2864	0.1984	0.0927	0.3658
4	0.2288	0.0329	0.0698	0.0974
5	3.0167	0.1939	1.0475	1.0536
6	0.2835	0.0625	0.0939	0.2895
7	3.4289	1.2730	0.9448	2.1629
8	1.0015	0.1036	0.1349	0.3369
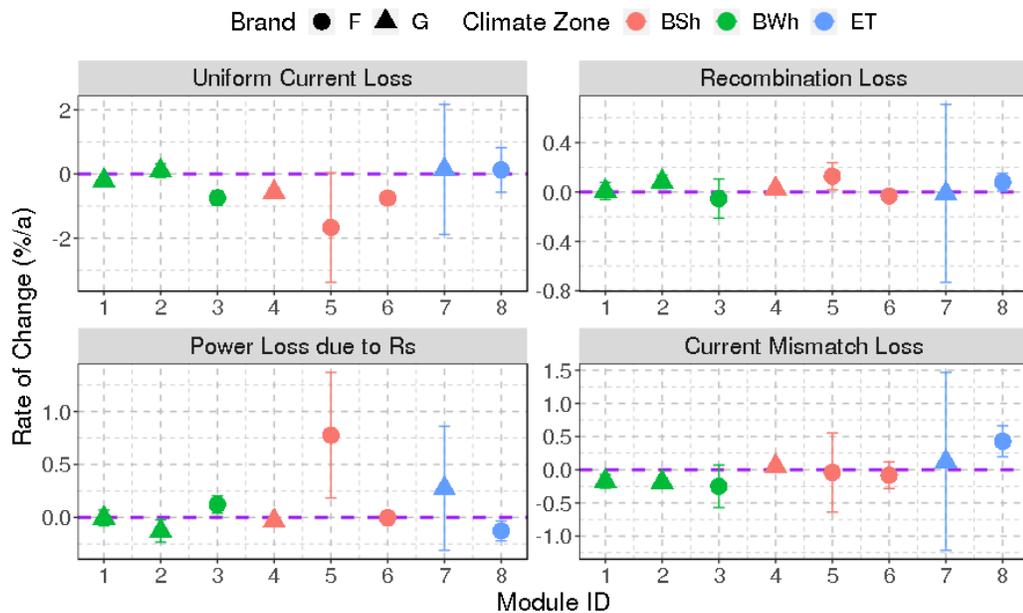
Figure 5.13. Rate of change result of each power loss mechanisms from outdoor I_{sc}-V_{oc} analysis.

Module 5 (BSh: GB) and module 7 (ET: DG) have relatively large standard deviations compared with that of others due to the very short system age that lasts less than 3 years. Removing module 5 (BSh: GB) and module 7 (ET: DG) from further comparison, Fig. 5.14 shows the rate of change result of each power loss

mechanisms for the rest modules, and the result is more likely to relate to long term degradation behavior. Then, we take an average of the rate of change for module 1 (BWh: DG) and 2 (BWh: DG) and compare the values across modules' architecture and climate zones, the result is in Table 5.11. The one has most negative value across four types of degradation mode is in bold.
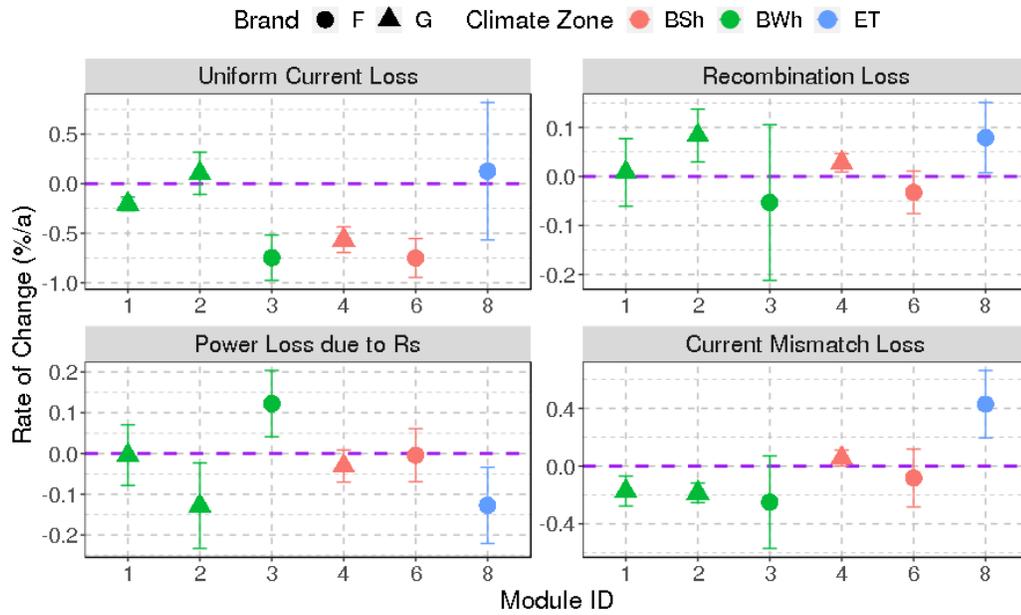
Figure 5.14. Rate of change result of each power loss mechanisms from outdoor I_{sc}-V_{oc} analysis after removing module 5 and 7.

Table 5.11. AVERAGE RATE OF CHANGE OF EACH POWER LOSS MECHANISM ACROSS CLIMATE ZONES AND MODULE ARCHITECTURES

Climate Zone	Brand	ΔP_{Isc} (%/a)	ΔP_{rec} (%/a)	ΔP_{Rs} (%/a)	ΔP_{Imis} (%/a)	Sum of four rates (%/a)
BWh	G:DG	-0.0495	0.0458	-0.0660	**-0.179**	-0.2487
BWh	F:GB	**-0.747**	-0.0532	0.1220	-0.250	-0.9282
BSh	G:DG	**-0.565**	0.0278	-0.0308	0.0560	-0.5120
BSh	F:GB	**-0.750**	-0.0325	0.0042	-0.0825	-0.8608
ET	F:GB	0.1261	0.0792	**-0.1272**	0.4289	0.5210

5.2.4 Partial Shading Detection

After getting the result from *ddiv*, the *MS* of each module is calculated and listed in Table 5.12. Module 3 (BWh: GB), module 5 (BSh: GB) and module 6 (BSh: GB) have *MS* higher than 20% and are identified as the ones that have the partial shading problem. We have done the partial shading study process illustration using module 3 (BWh: GB) as an example in previous section 5.1.3.

Table 5.12. *MS* OF ALL EIGHT MODULES

ID	System Age (Year)	Climate Zone	Brand	*MS* (%)
1	8.03	BWh	G:DG	2.97
2	8.74	BWh	G:DG	4.44
3	6.16	BWh	F:GB	36.38
4	6.39	BSh	G:DG	5.42
5	2.93	BSh	F:GB	50.31
6	6.39	BSh	F:GB	50.65
7	2.63	ET	G:DG	4.55
8	4.80	ET	F:GB	14.55

The shading profile of module 5 (BSh: GB) is in Table 5.13 and module 6 (BSh: GB) is in Table 5.14.

Table 5.13. SHADING PROFILE FOR MODULE 5 (BSH: GB)

Time in a Day	*MS* (%)	Year
09:00:00	76.4	2012
12:35:00	65.3	2012
09:10:00	84.1	2013
12:00:00	65.9	2013
07:05:00	37.8	2014
08:45:00	54.8	2014
12:25:00	66.0	2014
12:35:00	86.6	2015

The solar azimuth angles have peak *MS* for all three modules covering all their system age are listed in Table 5.15.

Table 5.14. Shading profile for module 6 (BSh: GB)

Time in a Day	*MS* (%)	Year
06:10:00	39.7	2012
08:55:00	82.0	2012
14:00:00	60.3	2012
09:10:00	75.0	2013
13:50:00	53.7	2013
06:25:00	45.9	2014
12:25:00	81.7	2014
07:55:00	64.9	2015
10:40:00	69.3	2015
06:30:00	42.7	2016
08:10:00	67.2	2016
17:15:00	52.3	2016
06:35:00	34.0	2017
09:55:00	87.8	2017
17:25:00	44.7	2017
10:35:00	93.4	2018
17:35:00	26.4	2018

Table 5.15. solar azimuth angle with peak *MS* for all three module identified to have partial shading problem

Module ID	solar azimuth angle (°)	*MS* (%)
	-99	34.9
3	35	53.5
	87	52.7
5	-84	60.1
	-11	72.7
	-88	67.4
6	-34	71.6
	81	42.3

The begining, ending and peak voltages for clusters of steps' location in voltage are listed in Table 5.16. All three modules have two voltage clusters, which is consistent with the physical diode model.

Based on the voltage clusters of the steps' voltage locations, the multistep *I-V* curves are classified into three shading scenarios, the result is in Table 5.17. The

Table 5.16. Voltage clusters for module 3 (BWh: GB)

Module ID	Cluster No.	Begin (V)	End (V)	Peak (V)
3	1	3.9	11.1	7.6
	2	11.1	23.9	18.1
5	1	4.5	10.8	7.9
	2	10.8	21.7	16.5
6	1	4.2	10.1	7.1
	2	10.1	26.8	15.7

number of multistep *I-V* curve with the number of steps as two or three but can not be classified into the three cases accounts 3.43%, 3.12% and 0.25% for module 3 (BWh: GB), module 5 (BSh: GB) and module 6 (BSh: GB) respectively.

Table 5.17. Classification of shading scenarios based on voltage clusters of steps' location

Module ID	Case 1(%)	Case 2(%)	Case 3(%)
3	22.86	61.19	15.95
5	26.54	59.93	13.52
6	25.19	36.67	38.13

The classification result of persistent and transient multistep *I-V* curves for all three modules is in Table 5.18.

Table 5.18. Classification of persistent and transient multistep *I-V* curves

Module ID	Persistent (%)	Transient(%)
3	89.9	10.1
5	92.0	8.0
6	95.2	4.8

Fig. 5.15 shows the density distribution of the duration of persistent multistep *I-V* curves for module 5 (BSh: GB) and module 6 (BSh: GB). From Fig. 5.8 and Fig. 5.15, we can see that both module 3 (BWh: GB) and module 6 (BSh: GB) have peak at 12 minutes and module 5 (BSh: GB) has peak at 11 minutes.
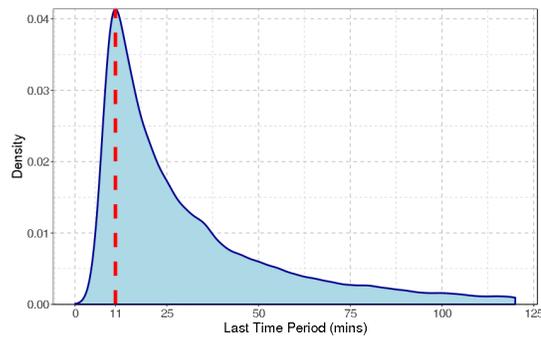
(a) Module 5 (BSh: GB)
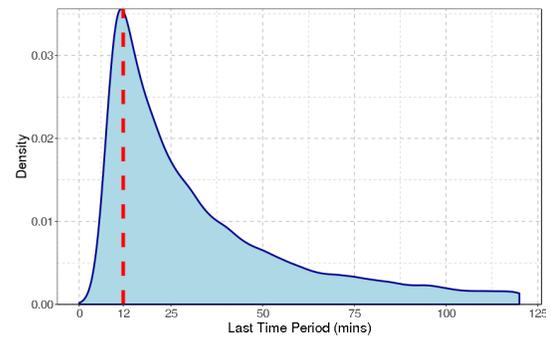
(b) Module 6 (BSh: GB)

Figure 5.15. Density of duration of persistent multistep *I-V* curves.

6   Discussion

The study protocol and analysis is summarized in Fig. 1.1 of the introduction chapter. The green boxes are the input data, which is composed of time-series data and metadata. The time-series data we used in the analysis are composed of time-series datastreams of I-V curves, P_{mp}, module temperature, and POA irradiance values. The blue boxes indicate the analysis process and the red boxes are results from analyzing the time-series data. The I-V curves first are processed by I-V curve quality detection, which is based on equipment accuracy limitations and physical models. Next, $ddiv$ is applied to extract the information of steps and I-V features. The information of *steps* is processed for the partial shading study and result is the percentage of multistep I-V curves (MS) to identify modules that can confidently be assessed as have partial shading conditions and this is summarized with the partial shading diagram, the shading profile, and the shading Poynting vector diagram, which summarize the obstacle orientation, shading scenarios and duration of partial shading. I-V features are the used for the outdoor I_{sc}-V_{oc} analysis and are converted to power loss modes or factors for each degradation mode. The rate of power loss in each power loss mode is calculated using a *month-by-month* linear regression, and the final result is the changing rate of each loss mechanism for each module. Time-series

P_{mp} are then processed to obtain the performance loss rate(PLR) and the process contains temperature correction, *day-by-day* predictive method and *Year-on-Year* regression[9].

In summary, from the input time-series data of P_{mp}, *I-V* curves, *POA*, and module temperature, we obtain the results of partial shading information, *PLR*, and the changing rate of power loss due to each degradation mechanism. Then we compare the time-series results across module architectures and Köppen-Geiger climate zones where the modules are deployed, which are included as metadata of the PV modules studied. What's more, the boxes with border in red are completed by other colleagues, the results are directly taken for further analysis of my part. So the focuses of my study is the pre-processing of *I-V* curves for quality detection, partial shading analysis, further processing of outdoor I_{sc}-V_{oc} results and comparisons across module brands (architectures) and climate zones.

This chapter focuses on the time-series results for each module and compares the partial shading and degradation behavior of modules with different module architectures and deployed in different climate zones.

6.1 Partial Shading and the Local Environment

The PV modules that are considered to have the partial shading problem are modules 3 (BWh: GB), 5 (BSh: GB) and 6 (BSh: GB), as summarized in Table 5.12. All three have *MS* significantly higher than the criteria of the accuracy of the *ddiv* algorithm and also significantly higher than the *MS* of other modules. All three modules selected are brand F modules, and they all have glass backsheet (GB) module architectures.

However, based on the physical causes of multiple steps in *I-V* curves, the huge difference in *MS* between these three modules and the rest is unlikely to arise due to the module architecture and the by-pass diode components in the junction boxes of the modules. This is because the by-pass diodes in commercial PV modules are supposed to work when modules are under partial shading and partial shading is most relevant to the specific local surroundings of the module or system[84,85].

Therefore, the most likely reason for differences in the percentage of *MS* is the local surrounding of the specific installation location of each module. Even for modules located in the same climate zone and installed in the same solar farm, it's very possible that one module experiences partial shading and another one doesn't, such as a module installed on the first row won't be shaded but the ones installed on the frame behind it could experience row-to-row shading[86].

The time dependence of the occurrence of multistep *I-V* curves for module 3 (BWh: GB) is shown in Fig. 5.3. The points in red are for single steps *I-V* curves, and the points in green are for multistep *I-V* curves. The data points are plotted with transparency, so that the color intensity in each sub-graph also indicates increasing density of observations. Each sub-figure is for one year, and the x-axis is time of day, which varies from 0 to 24, and the y-axis is the date in unit of months, which varies from January to December. The pear shape is caused by the *POA* cutoff, so we can see in the summer observations covers a longer time in a day, while in the winter it is the opposite, which is consistent with human experience. The same phenomenon can be found in another study[61], which used sunrise and sunset times as cutoffs. From Fig. 5.3, we can see the green points can be clustered into two parts based on the time, morning or afternoon, and the afternoon time lasts longer. There is variation

seen from year to year, with the most severe occurrences happened in 2014. There is more area in deep green and it shows some degree of continuation across time and month. This finding is also reflected in the shading profile of module 3 (BWh: GB) as Table 5.2. The *MS* corresponding to the 2014 morning and afternoon peaks is higher than that in other years.

The partial shading diagram shown in Fig. 5.3 is very useful for showing the partial shading occurrence of the studied module. However, for system diagnosis, quantitative and specific information is preferred. For this purpose, we generate the shading profile as Table 5.2 for each module to show the time that has a local peak in percentage of multistep *I-V* curves. If someone chooses to visit the site, the shading profile provides information on when they would be most likely to see the partial shading phenomena on the PV modules.

From the shading profile of module 3 (BWh: GB) as Table 5.2, module 5(BSh: GB) as Table 5.13, and module 6 (BSh: GB) as Table 5.14, we find consistency in time across different years for each module. For module 3 (BWh: GB), for every year, there is one peak in the morning located around 8 am and another peak in the afternoon located around 3:30 pm. For module 5 (BWh: GB), all four years have one morning peak around 9:00 am and another peak at noon. Module 6 (BWh: GB) is more complex, but we are still able to see some consistency, in 2012 and 2013 there were two peaks that happened at very similar times around 9:00 am and 14:00 pm. The first and third peaks in 2016 are similar to the two in 2017 and the second and third peaks in 2017 are similar to the ones in 2018.

If the maintenance staff checks on-site according to the time in the shadow outline, it is easy to distinguish which object caused the shadow on the module under study.

But, if no staff is available for site inspection, it would be nice to obtain the relative position of the shading object so as locate the obstacle on a site map. For this reason, we converted the date and time to the solar elevation angle and the solar azimuth angle. The shading Poynting vector diagram, Fig. 5.5, shows the multistep *I-V* curve occurrence for module 3 (BWh: GB). The space at high solar elevation angles is caused by the five minute constant measurement time interval, which will have a low resolution in angle at noon because the solar elevation angle changes faster when it's higher.

Similar as generating a shading profile, we report the solar azimuth angle has a local peak *MS* for each module, which indicates the relative orientation of the shading object to the study module. In this study, we apply this to the PV module using the complete time-series instead for each year, even though it can be evaluated for shorter time periods. The decision depends on the timeframe we want to focus on, which relate to purpose for the analysis, whether it's for monitoring or studying of modules' history, general overview, or a more detailed study. Like for the shading profile, the method of detecting the obstacle's orientation can be applied to different time periods. Since we analyze the shading profile for each year, we compute the azimuth angle using the complete time-series dataset, with the result summarized in Table 5.15.

An obstacle with the absolute value of the azimuth angle close to 90°, corresponds to times very early in the morning or very late in the afternoon, when the solar elevation is quite small and the irradiance is weak, the object's shadow won't cause much current mismatch in the PV module. However, an object with an azimuth angle close to 0 corresponds to an object that needs to be removed, because it shades the

module most of the time and corresponding to time around noon which has strong irradiance. According to Table 5.15, this happens to module 5 (BSh: GB) with a peak *MS* at -11° azimuth.

The multistep *I-V* curves provide more detailed information of shading scenarios by the number of steps they have and the locations of these steps, with these results summarized in Table 5.17. Here we see that module 3 (BSh: GB) and module 5 (BWh: GB) have similar percentages for different shading cases, so compared with module 6 (BWh: GB), their shadowing are more similar to each other.

In order to run this classification, we first obtain the voltage range of the cluster of steps' positions in the multistep *I-V* curves, Fig. 5.7 shows an example for module 3 (GB: BWh). We see three clustering peaks, the left-most one we believe is not a real step but instead is induced due to the fluctuation of measurements close to the I_{sc} region and has later been generated in the spline model used in *ddiv* algorithm. Since we only have around 40 to 70 points in each *I-V* curve with this ESL tracer, we find that there are only 3 or 4 datapoints lower than 1 *V*, and with this step detection is difficult. The voltage range of the steps' location cluster for module 3 (GB: BWh), module 5 (GB: BSh), and module 6 (GB: BSh) is listed in Table 5.16, let us compare the location of peaks, the lower voltage clusters are very similar for all three modules. However, for the one with a higher voltage cluster, module 5 (BWh: GB) and module 6 (BWh: GB) are more similar to each other compared with module 3 (BSh: GB). This is likely caused by similar weather conditions, such as similar temperature distributions, causing the voltage of a curve to have similar amount of temperature influenced shift.

We further ran classification on multistep I-V curves into a persistent steps class and a transient steps class based on whether they have at least one time-series neighbor that is also a multistep curve, the result is shown in Table 5.18. The percentage of persistent step curves is in the range of 89.9% to 95.3%, which indicates our measurement time interval is long enough to capture the duration of partial shading. The distribution on the time period for persistent multistep I-V curves have also been studied, the peaks are at 12 minutes as Fig. 5.15, 11 minutes as Fig. 5.15a and 12 minutes Fig. 5.15b and the distributions are skewed to the right significantly and there are plenty of observations could last over half an hour. Most multistep I-V curves last only 2 to 3 measurement intervals, so 10 to 15 minute duration, and this observation is relatively shorter than an impression of a stable object continuously shading the module, which could be caused by unstable weather condition and error in steps detection to make long period be cut into short pieces of a time period.

The partial shading detection algorithm developed initially for PV module studies, but it has the potential to also being applied in large-scale power plants, if timeseries I-V curves relatively complete and good enough for detecting steps in it. In large power plants, the I-V curve will miss data points I_{sc} due to effect of wire resistance and high current[65]. In our study, we relate the steps in I-V curves with shading scenarios on PV module by the study of M.Bressan[5], in which these two are linked by experiments. The study of M.Bennet[48] can simulate I-V curves from the shading scenarios on a large PV array with series and parallel connection, which means the time-series statistical analysis on multistep I-V curves can be linked with shading scenarios happening in large power plants also.

Compared with studies only using time-series power output with a reference system for obtaining shading profiles[4] and obstacle's orientation[61]. Our partial shading detection algorithm is able to realize this information using time-series *I-V* curves only without needing comparison to a reference systems such as a reference cell, pyranometer, or simulated reference result. Since it's based on multiple steps arising in *I-V* curves, our approach excludes uniform shading cases from being reported. As mentioned before, uniform shading means a PV module under a shadow with uniform irradiance, will reduce the power output the same as partial shading, which has non-uniform irradiance, but uniform shading doesn't cause degradation of the module nor serious safety issues such as local hot spots.

The realization of partial shading detection using the time-series *I-V* curves also expands the research significance of studies linking the shape of *I-V* curves with shading scenarios. This is the research focus of partial shading detection with *I-V* curves in many studies[5,47,51,55,63,64], and now these cases dependent studies can find a way to be applied in the field.

However, partial shading is not necessary to activate the by-pass diode, and the by-pass diode activation is the basis of multiple steps in the *I-V* curve. If there were no by-pass diodes installed in a PV module, partial shading won't cause multiple steps in the *I-V* curve[5]. The by-pass diode is connected in parallel, but with opposite polarity, to a solar cell as shown in Fig. 6.1. The diode outside the blue box is the by-pass diode, the diode and current source within the blue box is the equivalent circuit for a piece solar cell. When under normal operation, as in Fig. 6.1a the cells are in forward bias, and the by-pass diodes are in reverse biased and have no effect. When one of the cells is shaded as Fig. 6.1b, there is a current mismatch between

the two cells, which means $I_{sc,1}$ and $I_{sc,2}$ have a difference, then the by-pass diode of the good (unshaded) cell is reverse biased and has no effect and the diode of the good PV cell is forward biased, but the by-pass diode of the shaded cell is forward biased and conduct current, the shaded cell diode is reverse biased but only to a single diode drop (about 0.6 to 0.7 V).

(a) Normal operation   (b) Partial shaded operation
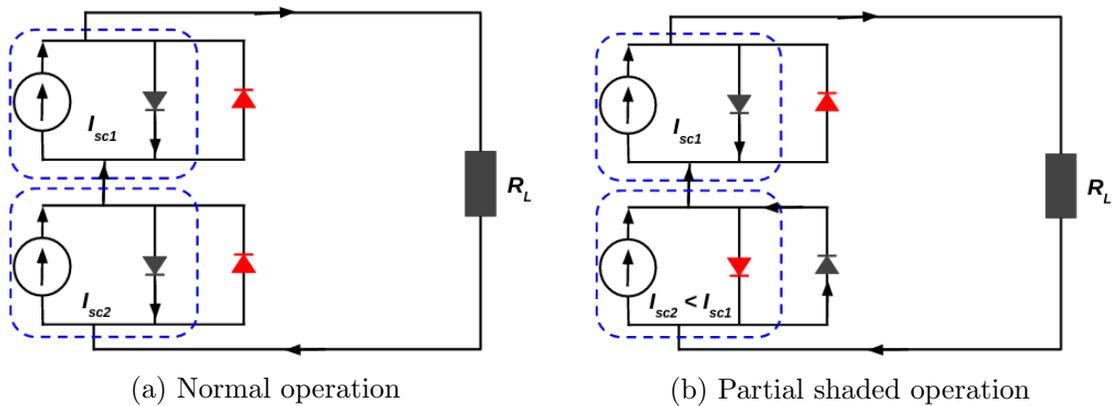
Figure 6.1. Illustration of by-pass diodes status of normal operation and partial shaded operation, the part inside the blue dash box is for solar cells and the diodes outside the blue dash box are by-pass diodes, the diodes that are reversed biased are highlight in red.

In commercial PV modules, it is too expensive to install a by-pass diode for each PV cell, so one installs a by-pass diode in parallel with about 20 cells connected in series string. So from the diode activation discussed above, if the by-pass diode is broken for a module[5], or partial shading happens without current mismatch across different strings, such as the same shading scenario happens on each string at the same time as Fig. 6.2, then the *I-V* curve won't have multiple steps and this partial shading occurrence won't be detected by our method.
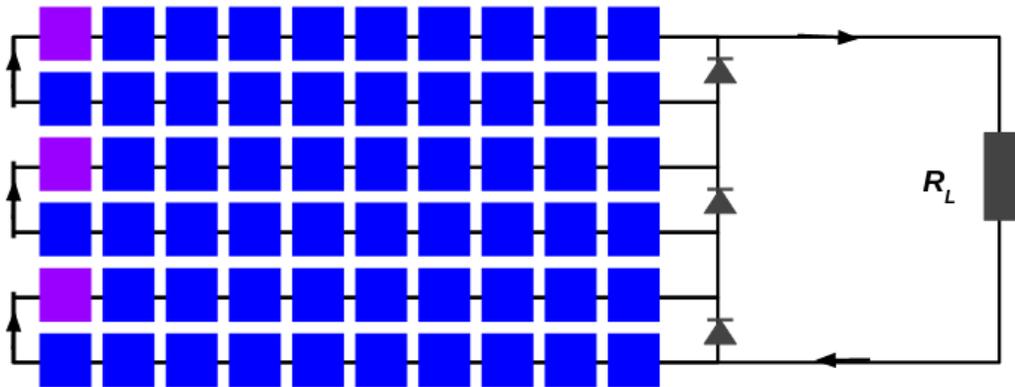
Figure 6.2. Example of shading scenario happens the exactly the same on each string, the cells in purple are shaded cell.

6.2 Long Term Degradation Study of PV Modules

In this section, we compare the *PLR* and rate of change of each degradation mode from outdoor I_{sc}-V_{oc} analysis and power loss calculation across modules of different brands (module architecture) and located climate zones.

6.2.1 Comparison of Performance Loss Rate

The BWh Köppen-Geiger climate zone is a hot desert climate, while BSh is a hot semi-arid climate, with the latter climate zone having more precipitation than the former one, and the precipitation for both is less than any tropical climates. ET corresponds to a tundra climate, which has an average temperature below $10°C$ of every month, which is much lower than that of BWh or BSh climate zones. From the *PLR* result as Fig. 5.11, we can see these different environmental conditions cause a difference in the degradation of the general performance of modules in the two brands. BWh and BSh climates cause more significant performance loss than the ET climate

zone for brand F: GB modules because the *PLR* values for the first two are more negative than that observed in the ET climate zone. For brand G: DG modules, BSh cause more aggressive degradation than BWh, and both are more aggressive than ET. In summary, the degree of aggressiveness of the three climate zones studied here show that for brand F: GB is BSh $\approx$ BWh > ET, and for brand G: DG is BSh > BSh > ET.

The brand G: DG modules have better performance in BWh climate compared with brand F: GB modules, while for the other two climate zones, the performance of these two brands are similar. This suggests that brand G: DG module exhibits better weather resistance considering its better performance in BWh.

What's more, from the *PLR* results, compared with a larger population of commercial PV modules, the modules we studied have better quality[33]. The average *PLR* of all eight modules is -0.103 %/a, and the smallest value is -0.473%/a. From the NREL report published in 2012[33], using 1,920 reported cases of silicon-based photovoltaic systems, the median *PLR* is about -0.5%/a and the average is about -0.8%/a. Comparing the *PLR* values in a more recent study[87], the *PLR* value for 14 different commercial crystalline silicon-based module is in the range from -0.3%/a to -2.9%/a. From another outdoor study of degradation in PV modules, several brands of modules are installed in Nicosia, Cyprus, which belongs to BSh Köppen climates[69], for five years[88]. We are able to find the same brands as the PV modules studied but different specific models. The *PLR* is obtained by using the performance ratio metric, and then two regression methods are applied to get two versions of results. The first is a simple linear model regression, and the second is do a classical seasonal decomposition then fit the linear model. Both results showed brand G (DG) modules has better

performance than brand F (GB) modules. The brand F modules also experienced shading problems during outdoor exposure, which is the same situation as our brand F module in BSh climate[88]. A climate dependent PV degradation study with predicted result shows that polar climates have a slower PLR (lower degradation rate) than the desert climate. For a specific model of a mono-Si PV module, the polar climate exhibited a degradation rate of -0.18 %/a which is much lower than that in a desert climate, which varied from -0.41 %/a to -0.61 %/a based on the irradiance intensity[89].

6.2.2 Dominant Degradation Power Loss Modes

The outdoor I_{sc}-V_{oc} analysis provides us more information on what happens to the module, based on Table 5.11. First, for each brand and architecture (GB or DG) of modules, using the rate of change for each degradation mode results, we can define the one that has the fastest degradation rate, i.e. the most negative value as the degradation mode, which means this degradation mode has largest contribution to the overall power loss of the module. For the BWh climate, the dominant degradation mode is uniform current loss and current mismatch loss for brand F: GB modules and G: DG modules respectively. For the BSh climate, the dominant degradation mode for both is uniform current loss, and for the ET climate zone, the dominant degradation mode is series resistance loss for brand F: GB. However, we also found that for brand F: GB modules, BWh and BSh emphasize the same dominant degradation mode, which is uniform current loss, but it is different for brand G: DG. From the qualitative field observation, which is mainly built on glass backsheet modules, desert

and steppe climates are more easy to trigger the encapsulant discoloration[89], which will cause a decrease in I_{sc}[90].

The four degradation or power loss modes determined from the outdoor time-series I_{sc}-V_{oc} analysis are related to specific degradation or power loss mechanisms activated in the module arising from changes in the PV module packaging materials (encapsulants and polymeric backsheets), changes in the PV cells or causes that are external to the module such as soiling. The I_{sc} of a solar cell is directly dependent on the light intensity so the uniform current loss (ΔP_{Isc}) which is calculated based on the I_{sc} difference is most likely caused by a decline in the irradiance on the PV cell, which could be caused by decreased transmittance of the packaging material such as encapsulant yellowing, or soiling of the modules frontsheet glass. The current mismatch mode (ΔP_{Imis}) happens on the module when the cell connected in series have different I_{sc}, and I_{sc} is equal to short circuit current density times the solar cell area, so the current mismatch usually happens to the module when the module is partly shaded. In addition, the current mismatch power loss (ΔP_{Imis}) obtained from outdoor I_{sc}-V_{oc} analysis includes the shunting resistance(R_{sh}) contribution also.

The negative value in the rate of change of series resistance (ΔP_{Rs}) usually is related to an increase of series resistance due to corrosion of the metal interconnect and metalization. It's a common phenomenon when the module is under the indoor accelerated damp-heat exposure[29], that the series resistance is seen to increase strongly, and this has been quantitatively tracked using machine learning on electroluminescent imaging and I-V analysis[91,92] . However, for the brand F: GB module in the ET climate zone, the reason they are experiencing series resistance power loss is unlikely to be an increase in series resistance considering that the climate zone is so
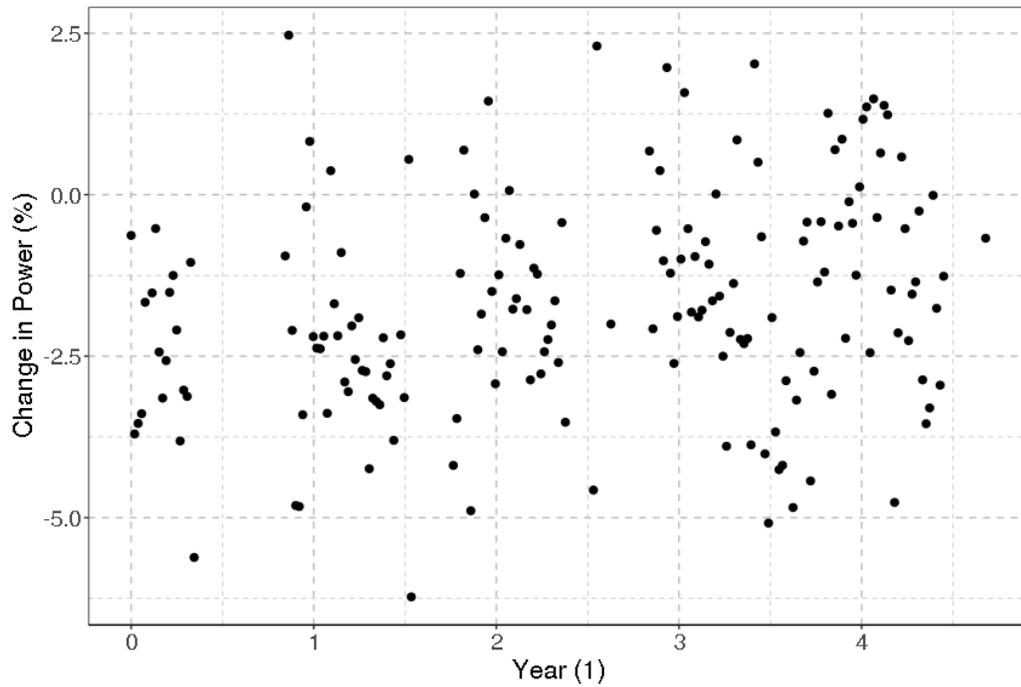
Figure 6.3. Current mismatch for Brand G:DG module in ET climate

cold. Instead, it may be due to an increase in current since power consumed by series resistance is equal to the square of current times resistance. The increasing current can be verified by the positive value in the uniform current loss (ΔP_{Isc}). In fact, the uniform current loss rate plus that of the series resistance loss is very close to zero for brand F: GB module in ET climate. The positive rate of the current mismatch is because of the positive trend happens in the summertime in Fig. 6.3.

If we do the same average across modules for the *PLR* results in Table 5.5: -0.062 %/a (BWh: DG), -0.473 %/a (BWh: GB), -0.465 %/a (BSh: DG), -0.60 %/a (BSh: GB) and 0.321 %/a (ET: GB), and compare these with the sum of the rate of change in the four power loss mechanisms listed in Table 5.11, the correlation between them is 0.96 and the root mean squared error (RMSE) is 0.2654%/a. This demonstrates that

the four power loss modes determined from time-series I-V analysis, are components of the total PLR.

However, the power loss modes have relatively small rates of change over time, and they exhibit noise compared with the trend, such that their rates of change are relatively more uncertain. So we do the similar comparison between PLR result and the rate of change in the dominant degradation or power loss mode. Brand F: GB module has a more negative rate of change in uniform current loss than that of brand G: DG in the current mismatch. And brand F: GB module has a more negative rate of change than that of G: DG module in BSh climate zone, and the module of brand F: GB in ET climate are more positive than the values of dominant degradation mode for all others. This finding from outdoor I_{sc}-V_{oc} analysis agrees with the PLR result of the same modules in Fig. 5.11. In addition, PLR results shows that brand F: GB module shows similar value for BWh and BSh climates and their dominant degradation modes from outdoor I_{sc}-V_{oc} analysis are the same, which are ΔP_{Isc} uniform current loss, and brand G: DG module has a quite different PLR value for BWh and BSh, and their returned dominant degradation modes are also different. Quantitative results shows the correlation is 0.934 and the RMSE is 0.2538%/a. The high correlation and relatively large RMSE from these two kinds of comparison all suggest that outdoor I_{sc}-V_{oc} and power loss calculation method is acceptable for relative comparisons of degradation across PV modules but not yet accurate enough for obtaining a degradation rate. This has not been the focus of our research since, we also can determine the PLR from the power timeseries.

PV degradation science studies using outdoor I_{sc}-V_{oc} analysis and loss factor analysis, compared with other published research using time-series I-V features for studies

of module degradation[2], realize direct comparison of changing in I-V features with respect to their contribution in the change of P_{mp}. Further analysis identifies the dominant degradation mode, which is very helpful for future research on reducing the module degradation rates.

7 Conclusion

In this research, the data-driven I-V feature extraction algorith, $ddiv$, was successfully applied to millions of time series I-V curves of eight modules with system ages vary from 3 to 9 years. And this analysis not only permits extraction of the single and mulitiple I-V "step" information that enables new and more detained temporal partial shading detection, but also identifies the I-V features and four distinct power loss modes of the PV modules. The I-V curve multistep analysis method developed in this research enables partial shading detection while returning the partial shading diagram, the shading profile, the shading Poynting vector diagram, obstacle orientation, and detailed information on shading scenarios and duration of partial shading throughout the year, and over many years. This approach incorporates a new method for quality detection of current-voltage curves, and ways to establish the optimal hyperparameter values for data-driven I-V features extraction ($ddiv$) to obtain the number of "step" in I-V curves.

This approach was successfully applied to multiyear time-series datasets obtained from outdoor PV modules of two brands and module architectures (glass-backhseet (GB) and double glass (DG)) in three distinct Köppen-Geiger climate zones. Three out of eight modules included in our study were detected to have the partial shading

problems, they are, module 3 (BWh: GB), module 5 (BSh: GB), and module 6 (BSh: GB). These module's shading profile is consistent across multiple years, there were often times where more than one by-pass diode had been activated and the peak of, and the duration of, partial shading is longer than the 5 minute time interval between *I-V* curve measurement.

Analysis of the time-series *I-V* curves using the I_{sc}-V_{oc} analysis method enables the determination of the *Suns*-V_{oc} curve of these fielded modules and also the identification of four degradation modes or power loss factors. These include ΔP_{Imis}, ΔP_{Isc}, ΔP_{Rec} and ΔP_{Rs} corresponding to power losses arising from Current Mismatch, Short Circuit Current, Recombination and Series Resistance respectively. These power loss modes are compared to the Performance Loss Rate (*PLR*) determined from analysis of the P_{mp} datastreams using the *month-by-month* linear regression method. The *PLR* is obtained for each module using the XbX with UTC predictive method, where X = 1 day, and the universal temperature correction (UTC) is used, followed by *year-on-year* regression. Comparing the degradation results of various brands and module architectures and different located climates, our brand G, double glass modules showed higher weather resistance and durability. In addition the BWh and BSh climate zone caused more severe degradation than the ET climate zone. The dominant degradation mode for brand F: GB module are the uniform current loss for both BWh and BSh climate zone, and series resistance loss due to increasing current for EL climates, and for brand G: DG module, the dominant degradation modes are current mismatch loss, uniform current loss for the BWh and BSh climate zone respectively. The relative value of the rate of change in dominant degradation modes agree with the *PLR* result.

Appendix A

Preparation of this document

This document was prepared using pdfLaTeX and other open source tools. The (free) programs implemented are as follows:

- LaTeX implementation:

  MiKTeX http://www.miktex.org/

  TeXLive https://www.tug.org/texlive/

- TeX-oriented editing environments:

  TexStudio https://www.texstudio.org/

- Bibliographical:

  BibTeX http://www.bibtex.org/

  Biber http://biblatex-biber.sourceforge.net/

  Zotero https://www.zotero.org/

  Better BibTeX For Zotero https://retorque.re/zotero-better-bibtex/

References

[1] Jiqi Liu, Menghong Wang, Alan J. Curran, Ahmad Maroof Karimi, Wei-heng Huang, Erdmut Schnabel, Michael Khl, Jennifer L. Braid, and Roger H. French. Real-world PV Module Degradation across Climate Zones Determined from Suns-Voc, Loss Factors and I-V Steps Analysis of Eight Years of I-V, Pmp Time-series Datastreams. In 2019 IEEE 46th Photovoltaic Specialists Conference (PVSC), pages 0680–0686, June 2019. ISSN: 0160-8371.

[2] Ryan M. Smith, Dirk C. Jordan, and Sarah R. Kurtz. Outdoor PV Module Degradation of Current-Voltage Parameters: Preprint. Technical Report NREL/CP-5200-53713, National Renewable Energy Lab. (NREL), Golden, CO (United States), April 2012.

[3] Menghong Wang, Xuan Ma, Wei-Heng Huang, Jiqi Liu, Alan J. Curran, Erdmut Schnabel, Michael Khl, Kristopher O. Davis, Jenn Brynjarsdttir, Jennifer L. Braid, and Roger H. French. Evaluation of Photovoltaic Module Performance Using Novel Data-driven I-V Feature Extraction and Suns-VOC Determined from Outdoor Time-Series I-V Curves. In 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC), pages 0778–0783, June 2018. ISSN: 0160-8371.

[4] Odysseas Tsafarakis, Kostas Sinapis, and Wilfried G. J. H. M. van Sark. A Time-Series Data Analysis Methodology for Effective Monitoring of Partially Shaded Photovoltaic Systems. Energies, 12(9):1722, January 2019.

[5] Michael Bressan, Y.el Basri, Alonso Gutirrez Galeano, and Corinne Alonso. A shadow fault detection method based on the standard error analysisof I-V curves. Renewable Energy, December 2016.

[6] Solar Industry Research Data. Library Catalog: www.seia.org.

[7] Austin Perea, Colin Smith, Michelle Davis, Allison Mond, Benjamin Gallagher, Shawn Rumery, Aaron Holm, Rachel Goldstein, and Justin Baca. U. S. Solar Market Insight Report 2018 Year In Review. Technical report, Solar Energy Industries Association, March 2019.

[8] D.C. Jordan, R.M. Smith, C.R. Osterwald, E. Gelak, and S.R. Kurtz. Outdoor PV degradation comparison. In 2010 35th IEEE Photovoltaic Specialists Conference, pages 002694–002697, Honolulu, HI, USA, June 2010. IEEE.

[9] Alan J. Curran, C. Birk Jones, Sascha Lindig, Joshua Stein, David Moser, and Roger H. French. Performance Loss Rate Consistency and Uncertainty Across Multiple Methods and Filtering Criteria. In 2019 IEEE 46th Photovoltaic Specialists Conference (PVSC), pages 1328–1334, June 2019. ISSN: 0160-8371.

[10] Sascha Lindig, Ismail Kaaya, Karl-Anders Weiss, David Moser, and Marko Topic. Review of Statistical and Analytical Degradation Models for Photovoltaic Modules and Systems as Well as Related Improvements. IEEE Journal of Photovoltaics, 8(6):1773–1786, November 2018. Conference Name: IEEE Journal of Photovoltaics.

[11] IEC. IEC60904-1 ed2.0 - Photovoltaic devices - Part 1: Measurement of photovoltaic current-voltage characteristics. Technical report, International Electrotechnical Commission, 2006.

[12] C.M. Whitaker, T.U. Townsend, J.D. Newmiller, D.L. King, W.E. Boyson, J.A. Kratochvil, D.E. Collier, and D.E. Osborn. Application and validation of a new PV performance characterization method. In Conference Record of the Twenty Sixth IEEE Photovoltaic Specialists Conference - 1997, pages 1253–1256, September 1997.

[13] Georgi Hristov Yordanov. Relative efficiency revealed: Equations for k1k6 of the PVGIS model. In 2014 IEEE 40th Photovoltaic Specialist Conference (PVSC), pages 1393–1398, June 2014. ISSN: 0160-8371.

[14] Jay A. Kratochvil, William Earl Boyson, and David L. King. Photovoltaic array performance model. Technical Report SAND2004-3535, Sandia National Laboratories, August 2004.

[15] Alexander Phinikarides, Nitsa Kindyni, George Makrides, and George E. Georghiou. Review of photovoltaic degradation rate methodologies. Renewable and Sustainable Energy Reviews, 40:143–152, December 2014.

[16] D.C. Jordan and S.R. Kurtz. Analytical improvements in PV degradation rate determination. In 2010 35th IEEE Photovoltaic Specialists Conference, pages 002688–002693, June 2010. ISSN: 0160-8371.

[17] Dirk C. Jordan, Chris Deline, Sarah R. Kurtz, Gregory M. Kimball, and Mike Anderson. Robust PV Degradation Methodology and Application. IEEE Journal of Photovoltaics, 8(2):525–531, March 2018. Conference Name: IEEE Journal of Photovoltaics.

[18] Dirk C. Jordan, Michael G. Deceglie, and Sarah R. Kurtz. PV degradation methodology comparison A basis for a standard. In 2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC), pages 0273–0278, June 2016.

[19] Santiago Silvestre, Sofiane Kichou, Letizia Guglielminotti, Gustavo Nofuentes, and Miguel Alonso-Abella. Degradation analysis of thin film photovoltaic modules under outdoor long term exposure in Spanish continental climate conditions. Solar Energy, 139:599–607, December 2016.

[20] E. Duran, M. Piliougine, M. Sidrach-de Cardona, J. Galan, and J.M. Andujar. Different methods to obtain the IV curve of PV modules: A review. In 2008 33rd IEEE Photovoltaic Specialists Conference, pages 1–6, May 2008. ISSN: 0160-8371.

[21] Roland Szabo and Aurel Gontean. Photovoltaic Cell and Module I-V Characteristic Approximation Using Bzier Curves. Applied Sciences, 8(5):655, May 2018. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.

[22] Ahmed A. El Tayyan. A simple method to extract the parameters of the single-diode model of a PV system. Turkish Journal of Physics, (2013) 37:121 – 131, 2013.

[23] Clifford W. Hansen. Estimation of parameters for single diode models using measured IV curves. In 2013 IEEE 39th Photovoltaic Specialists Conference (PVSC), pages 0223–0228, June 2013.

[24] JS Stein, Juergen Sutterlueti, Steve Ransome, CW Hansen, and BH King. Outdoor PV performance evaluation of three different models: Single-diode, SAPM and loss factor model. In EU PVSEC Proccedings, pages 2865–2871, 2013.

[25] T. O. Saetre, O.-M. Midtgård, and G. H. Yordanov. Two-Diode Model Revisited: Parameters Extraction from Semi-Log Plots of I-V Data. 25th European Photovoltaic Solar Energy Conference and Exhibition / 5th World Conference on Photovoltaic Energy Conversion, 6-10 September 2010, Valencia, Spain, pages 4156–4163, October 2010.

[26] Xuan Ma, Wei-Heng Huang, Erdmut Schnabel, Michael Khl, Jenn Brynjarsdttir, Jennifer L. Braid, and Roger H. French. Data-Driven $I$$V$ Feature Extraction for Photovoltaic Modules. IEEE Journal of Photovoltaics, 9(5):1405–1412, September 2019. Conference Name: IEEE Journal of Photovoltaics.

[27] Jiqi Liu, Alan J. Curran, Justin S. Fada, Xuan Ma, Wei-Heng Huang, C. Birk Jones, Erdmut Schnabel, Michael Kohl, Jennifer L. Braid, and Roger H. French. Cross-correlation Analysis of the Indoor Accelerated and Real World Exposed

Photovoltaic Systems Across Multiple Climate Zones. In 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC), pages 3949–3954, June 2018. ISSN: 0160-8371.

[28] F.J. Pern, A.W. Czanderna, K.A. Emery, and R.G. Dhere. Weathering degradation of EVA encapsulant and the effect of its yellowing on solar cell efficiency. In The Conference Record of the Twenty-Second IEEE Photovoltaic Specialists Conference - 1991, pages 557–561 vol.1, October 1991.

[29] Jennifer L. Braid, Alan J. Curran, Jing Sun, Eric J. Schneller, Justin S. Fada, Jiqi Liu, Menghong Wang, Alexandra J. Longacre, Jeff Dai, Bryan D. Huey, Kristopher O. Davis, Jean-Nicolas Jaubert, Laura S. Bruckman, and Roger H. French. EL and I-V Correlation for Degradation of PERC vs. Al-BSF Commercial Modules in Accelerated Exposures. In 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC), pages 1261–1266, June 2018. ISSN: 0160-8371.

[30] M. Keith Forsyth, Matthew Mahaffey, Adrienne L. Blum, Weston A. Dobson, and Ronald A. Sinton. Use of the Suns-Voc for diagnosing outdoor arrays modules. In 2014 IEEE 40th Photovoltaic Specialist Conference (PVSC), pages 1928–1931, June 2014. ISSN: 0160-8371.

[31] Siyu Guo, Eric Schneller, Joe Walters, Kristopher O. Davis, and Winston V. Schoenfeld. Detecting loss mechanisms of c-Si PV modules in-situ I-V measurement. In Reliability of Photovoltaic Cells, Modules, Components, and Systems IX, volume 9938, page 99380N. International Society for Optics and Photonics, September 2016.

[32] Xingshu Sun, Raghu Vamsi Krishna Chavali, and Muhammad Ashraful Alam. Real-time monitoring and diagnosis of photovoltaic system degradation only using maximum power pointthe Suns-Vmp method. Progress in Photovoltaics: Research and Applications, 27(1):55–66, 2019. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pip.3043.

[33] D. C. Jordan and S. R. Kurtz. Photovoltaic Degradation Rates - An Analytical Review. Progress in Photovoltaics: Research and Applications, 21(1, January 2013), January 2013. Institution: National Renewable Energy Lab. (NREL), Golden, CO (United States).

[34] Jeha Kim. Outdoor Testing and Degradation of EVA and POE Encapsulated Photovoltaic Modules. Journal of the Korean Institute of Electrical and

Electronic Material Engineers, 29(12):847–852, 2016. Publisher: The Korean Institute of Electrical and Electronic Material Engineers.

[35] IEC. IEC 61215-2:2016, Terrestrial photovoltaic (PV) modules - Design qualification and type approval - Part 2: Test procedures. International Standard, International Electrotechnical Commission, 2016.

[36] Andrew W. Blakers, Aihua Wang, Adele M. Milne, Jianhua Zhao, and Martin A. Green. 22.8% efficient silicon solar cell. Applied Physics Letters, 55(13):1363–1365, September 1989. Publisher: American Institute of Physics.

[37] M. A. Green, A. W. Blakers, J. Zhao, A. M. Milne, A. Wang, and X. Dai. Characterization of 23-percent efficient silicon solar cells. IEEE Transactions on Electron Devices, 37(2):331–336, February 1990.

[38] Thorsten Dullweber. High-Efficiency Industrial PERC Solar Cells for Monofacial and Bifacial Applications. In High-Efficient Low-Cost Photovoltaics, volume 140, pages 65–94. Springer, November 2019.

[39] S. W. Glunz, R. Preu, and D. Biro. 1.16 - Crystalline Silicon Solar Cells: State-of-the-Art and Future Developments. In Ali Sayigh, editor, Comprehensive Renewable Energy, pages 353–387. Elsevier, Oxford, January 2012.

[40] Friederike Kersten, Peter Engelhart, Hans-Christoph Ploigt, Andrey Stekolnikov, Thomas Lindner, Florian Stenzel, Matthias Bartzsch, Andy Szpeth, Kai Petter, Johannes Heitmann, and Jrg W. Mller. Degradation of multicrystalline silicon solar cells and modules after illumination at elevated temperature. Solar Energy Materials and Solar Cells, 142:83–86, November 2015.

[41] Daniel Chen, Moonyong Kim, Bruno V. Stefani, Brett J. Hallam, Malcolm D. Abbott, Catherine E. Chan, Ran Chen, David N. R. Payne, Nitin Nampalli, Alison Ciesla, Tsun H. Fung, Kyung Kim, and Stuart R. Wenham. Evidence of an identical firing-activated carrier-induced defect in monocrystalline and multicrystalline silicon. Solar Energy Materials and Solar Cells, 172:293–300, December 2017.

[42] Namsu Kim, Hyunwoo Kang, Kyung-Jun Hwang, Changwoon Han, Won Sik Hong, Dajung Kim, Eonjoo Lyu, and Hwanyeon Kim. Study on the degradation of different types of backsheets used in PV module under accelerated conditions. Solar Energy Materials and Solar Cells, 120:543–548, January 2014.

[43] Yingbin Zhang, Jianmei Xu, Jing Mao, Jiahua Tao, Hui Shen, Yifeng Chen, Zhiqiang Feng, Pierre J. Verlinden, Pingxiong Yang, and Junhao Chu. Long-term reliability of silicon wafer-based traditional backsheet modules and double glass modules. RSC Advances, 5(81):65768–65774, July 2015.

[44] Sergiu Spataru, Peter Hacke, and Michael Owen-Bellini. Combined-Accelerated Stress Testing System for Photovoltaic Modules. In 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC), pages 3943–3948, June 2018.

[45] John H. Wohlgemuth and Sarah Kurtz. Using accelerated testing to predict module reliability. In 2011 37th IEEE Photovoltaic Specialists Conference, pages 003601–003605, June 2011. ISSN: 0160-8371.

[46] Ismail Kaaya, Michael Koehl, Amantin Panos Mehilli, Sidrach de Cardona Mariano, and Karl Anders Weiss. Modeling Outdoor Service Lifetime Prediction of PV Modules: Effects of Combined Climatic Stressors on PV Module Power Degradation. IEEE Journal of Photovoltaics, 9(4):1105–1112, July 2019. Conference Name: IEEE Journal of Photovoltaics.

[47] Bechara Nehme, Nacer K. Msirdi, Aziz Namaane, and Tilda Akiki. Analysis and Characterization of Faults in PV Panels. Energy Procedia, 111:1020–1029, March 2017. ZSCC: 0000014.

[48] Bennet Meyers and Mark Mikofski. Accurate Modeling of Partially Shaded PV Arrays. In 2017 IEEE 44th Photovoltaic Specialist Conference (PVSC), pages 3354–3359, June 2017.

[49] C Deline and B Marion. A Performance and Economic Analysis of Distributed Power Electronics in Photovoltaic Systems. Renewable Energy, page 20, 2011. ZSCC: 0000002.

[50] Daniele Rossi, Martin Omaa, Daniele Giaffreda, and Cecilia Metra. Modeling and Detection of Hotspot in Shaded Photovoltaic Cells. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 23(6):1031–1039, June 2015. ZSCC: NoCitationData[s1] Conference Name: IEEE Transactions on Very Large Scale Integration (VLSI) Systems.

[51] Y. el Basri, Michael Bressan, Lionel Sguier, Hussain A. Alawadhi, and Corinne Alonso. A proposed graphical electrical signatures supervision method to study PV module failures. Solar Energy, June 2015.

[52] S. Silvestre, A. Boronat, and A. Chouder. Study of bypass diodes configuration on PV modules. Applied Energy, 86(9):1632–1640, September 2009. ZSCC: 0000439.

[53] F. Martnez-Moreno, J. Muoz, and E. Lorenzo. Experimental model to estimate shading losses on PV arrays. Solar Energy Materials and Solar Cells, 94(12):2298–2303, December 2010.

[54] Daniele Giaffreda, Paolo Magnone, Matteo Meneghini, Marco Barbato, Gaudenzio Meneghesso, Enrico Zanoni, Enrico Sangiorgi, and Claudio Fiegna. Local Shunting in Multicrystalline Silicon Solar Cells: Distributed Electrical Simulations and Experiments. IEEE Journal of Photovoltaics, 4(1):40–47, January 2014. ZSCC: NoCitationData[s1] Conference Name: IEEE Journal of Photovoltaics.

[55] Dirk Stellbogen. Use of PV Circuit Simulation For Fault Detection in PV Array Field. Conference Record of the Twenty Third IEEE Photovoltaic Specialists, May 1993.

[56] K. Sinapis, C. Tzikas, G. Litjens, M. van den Donker, W. Folkerts, W. G. J. H. M. van Sark, and A. Smets. A comprehensive study on partial shading response of c-Si modules and yield modeling of string inverter and module level power electronics. Solar Energy, 135:731–741, October 2016.

[57] Chris Deline, Jenya Meydbray, and Matt Donovan. Photovoltaic Shading Testbed for Module-Level Power Electronics: 2016 Performance Data Update. Technical Report NREL/TP-5J00-62471, National Renewable Energy Lab. (NREL), Golden, CO (United States), September 2016.

[58] Lucia Serrano-Lujn, Jose Manuel Cadenas, Juan Faxas-Guzmn, and Antonio Urbina. Case of study: Photovoltaic faults recognition method based on data mining techniques. Journal of Renewable and Sustainable Energy, 8(4):043506, July 2016. ZSCC: 0000008 Publisher: American Institute of Physics.

[59] . Bognr, R. C. G. M. Loonen, R. M. E. Valckenborg, and J. L. M. Hensen. An unsupervised method for identifying local PV shading based on AC power and regional irradiance data. Solar Energy, 174:1068–1077, November 2018. ZSCC: 0000004.

[60] Elyes Garoudja, Fouzi Harrou, Ying Sun, Kamel Kara, Aissa Chouder, and Santiago Silvestre. Statistical fault detection in photovoltaic systems. Solar Energy, 150:485–499, July 2017.

[61] K. Steven Firth, Kevin Lomas, and Rees Simon. A simple model of PV system performance and its use in fault detection. Solar Energy, August 2009.

[62] V. Quaschning and R. Hanitsch. Influence of shading on electrical parameters of solar cells. In Conference Record of the Twenty Fifth IEEE Photovoltaic Specialists Conference - 1996, pages 1287–1290, May 1996. ZSCC: 0000081 ISSN: 0160-8371.

[63] Masaki Miwa, Sanshiro Yamanaka, Hajime Kawamura, Hideyuki Ohno, and Hideaki Kawamura. Diagnosis of a Power Output Lowering of PV Array with a (-dI/dV)-V Characteristic. In 2006 IEEE 4th World Conference on Photovoltaic Energy Conference, volume 2, pages 2442–2445, May 2006. ZSCC: 0000030 ISSN: 0160-8371.

[64] Arar Hemza, Haouam Abdeslam, Chenni Rachid, and Nouar Aoun. Simplified methods for evaluating the degradation of photovoltaic module and modeling considering partial shading. Measurement, 138:217–224, May 2019.

[65] R. Teodorescu, T. Kerekes, D. Sera, and S. Spataru. Monitoring and Fault Detection in Photovoltaic Systems Based On Inverter Measured String I-V Curves. 31st European Photovoltaic Solar Energy Conference and Exhibition, pages 1667–1674, November 2015. ZSCC: 0000002 ISBN: 9783936338393 Publisher: WIP.

[66] ET Instrumente. Series ESL-Solar 500, 2018.

[67] Franz Rubel, Katharina Brugger, Klaus Haslinger, and Ingeborg Auer. The climate of the European Alps: Shift of very high resolution Kppen-Geiger climate zones 18002100. Meteorologische Zeitschrift, 26(2):115–125, April 2017.

[68] Chelsey Bryant, Nicholas R. Wheeler, Franz Rubel, and Roger H. French. kgc: Koeppen-Geiger Climatic Zones, November 2017.

[69] Wladimir Kppen. The thermal zones of the Earth according to the duration of hot, moderate and cold periods and to the impact of heat on the organic world. Meteorologische Zeitschrift, 20(3):351–360, 2011. Place: Stuttgart, Germany Publisher: Schweizerbart Science Publishers.

[70] Wei-Heng Huang, Xuan Ma [aut, cre, Jiqi Liu, Menghong Wang, Alan J. Curran, Justin S. Fada, Jean-Nicolas Jaubert, Jing Sun, Jennifer L. Braid, Jenny Brynjarsdottir, and Roger H. French. ddiv: Data Driven I-v Feature Extraction, September 2018.

[71] Ernest Hasselbrink, Mike Anderson, Zoe Defreitas, Mark Mikofski, Yu-Chen Shen, Sander Caldwell, Akira Terao, David Kavulak, Zach Campeau, and David DeGraaff. Validation of the PVLife model using 3 million module-years of live site data. In 2013 IEEE 39th Photovoltaic Specialists Conference (PVSC), pages 0007–0012, June 2013.

[72] A. J. Curran, T. J. Burleyson, S. Lindig, D. Moser, and R. H. French. PVplr: SDLE Performance Loss Rate Analysis Pipeline, 2020.

[73] J. A. Gow and C. D. Manning. Development of a photovoltaic array model for use in power-electronics simulation studies. IEE Proceedings - Electric Power Applications, 146(2):193–200, March 1999. Publisher: IET Digital Library.

[74] Vito M. R. Muggeo. Estimating regression models with unknown break-points. Statistics in Medicine, 22(19):3055–3071, October 2003.

[75] Vito MR. Muggeo. Segmented: An R package to fit regression models with broken-line relationships. R news, 8(1):20–25, 2008.

[76] Hans W. Borchers. pracma: Practical Numerical Math Functions, December 2019.

[77] Benoit Thieurmel (R interface) and Achraf Elmarhraoui (R interface). suncalc: Compute Sun Position, Sunlight Phases, Moon Position and Lunar Phase, April 2019.

[78] Ronald Alan Sinton and Andres Cuevas. A quasi-steady-state open-circuit voltage method for solar cell characterization. In 16th European Photovoltaic Solar Energy Conference, Glasgow, UK, May 2000.

[79] M. Wang, T. J. Burleyson, J. Liu, A. J. Curran, E. J. Schneller, K. O. Davis, J. L. Braid, and R. H. French. SunsVoc: Constructing Suns-Voc from Outdoor Time-series I-V Curves, May 2020.

[80] Menghong Wang, Tyler J. Burleyson, Jiqi Liu, Alan J. Curran, Eric J. Schneller, Kristopher O. Davis, Jennifer L. Braid, and Roger H. French. SunsVoc: Constructing Suns-Voc from Outdoor Time-Series I-V Curves, June 2020.

[81] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An Introduction to Statistical Learning: With Applications in R. Springer Texts in Statistics. Springer, New York, 1st ed. 2013, corr. 5th printing 2015 edition edition, August 2013.

[82] R. Loudon and C. Baxter. Contributions of John Henry Poynting to the understanding of radiation pressure. Proceedings. Mathematical, Physical, and Engineering Sciences / The Royal Society, 468(2143):1825–1838, July 2012.

[83] H. Beyer. Tukey, John W.: Exploratory Data Analysis. Addison-Wesley Publishing Company Reading, Mass. Menlo Park, Cal., London, Amsterdam, Don Mills, Ontario, Sydney 1977, XVI, 688 S. Biometrical Journal, 23(4):413–414, 1981. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bimj.4710230408.

[84] Chuanxiao Xiao, Peter Hacke, Steve Johnston, Dana B. Sulas-Kern, Chun-Sheng Jiang, and Mowafak Al-Jassim. Failure analysis of field-failed bypass diodes. Progress in Photovoltaics: Research and Applications, n/a(n/a), June 2020.

[85] Kashif Ishaque, Zainal Salam, and Syafaruddin. A comprehensive MATLAB Simulink PV system simulator with partial shading capability based on two-diode model. Solar Energy, 85(9):2217–2227, September 2011.

[86] Yunlin Sun, Siming Chen, Liying Xie, Ruijiang Hong, and Hui Shen. Investigating the Impact of Shading Effect on the Characteristics of a Large-Scale Grid-Connected PV Power Plant in Northwest China. International Journal of Photoenergy, 2014.

[87] Philip Ingenhoven, Giorgio Belluardo, and David Moser. Comparison of Statistical and Deterministic Smoothing Methods to Reduce the Uncertainty of Performance Loss Rate Estimates. IEEE Journal of Photovoltaics, 8(1):224–232, January 2018. Conference Name: IEEE Journal of Photovoltaics.

[88] George Makrides, Bastian Zinsser, Markus Schubert, and George E. Georghiou. Performance loss rate of twelve photovoltaic technologies under field conditions using statistical techniques. Solar Energy, 103:28–42, May 2014.

[89] Julián Ascencio-Vásquez, Ismail Kaaya, Kristijan Brecl, Karl-Anders Weiss, and Marko Topič. Global Climate Data Processing and Mapping of Degradation Mechanisms and Degradation Rates of PV Modules. Energies, 12(24):4749, January 2019.

[90] Archana Sinha, Hamsini Gopalakrishna, Arun Bala Subramaniyan, Deepak Jain, Jaewon Oh, Dirk Jordan, and GovindaSamy TamizhMani. Prediction of Climate-Specific Degradation Rate for Photovoltaic Encapsulant Discoloration. IEEE Journal of Photovoltaics, 10(4):1093–1101, July 2020. Conference Name: IEEE Journal of Photovoltaics.

[91] A. M. Karimi, J. S. Fada, M. A. Hossain, S. Yang, T. J. Peshek, J. L. Braid, and R. H. French. Automated Pipeline for Photovoltaic Module Electroluminescence Image Processing and Degradation Feature Classification. IEEE Journal of Photovoltaics, pages 1–12, 2019.

[92] Ahmad Maroof Karimi, Justin S. Fada, Nicholas A. Parrilla, Benjamin G. Pierce, Mehmet Koyutürk, Roger H. French, and Jennifer L. Braid. Generalized and Mechanistic PV Module Performance Prediction From Computer Vision and Machine Learning on Electroluminescence Images. IEEE Journal of Photovoltaics, pages 1–10, 2020.