

Performance and stability of solar cells and modules: From laboratory characterization to field data analysis

Timon Sebastian Vaas

Energie & Umwelt / Energy & Environment

Band / Volume 685

ISBN 978-3-95806-871-1

Forschungszentrum Jülich GmbH
Institute of Energy Materials and Devices (IMD)
Photovoltaik (IMD-3)

Performance and stability of solar cells and modules: From laboratory characterization to field data analysis

Timon Sebastian Vaas

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment

Band / Volume 685

ISSN 1866-1793

ISBN 978-3-95806-871-1

Bibliografische Information der Deutschen Nationalbibliothek.
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte Bibliografische Daten
sind im Internet über <http://dnb.d-nb.de> abrufbar.

Herausgeber
und Vertrieb: Forschungszentrum Jülich GmbH
 Zentralbibliothek, Verlag
 52425 Jülich
 Tel.: +49 2461 61-5368
 Fax: +49 2461 61-6103
 zb-publikation@fz-juelich.de
 www.fz-juelich.de/zb

Umschlaggestaltung: Grafische Medien, Forschungszentrum Jülich GmbH

Druck: Grafische Medien, Forschungszentrum Jülich GmbH

Copyright: Forschungszentrum Jülich 2025

Schriften des Forschungszentrums Jülich
Reihe Energie & Umwelt / Energy & Environment, Band / Volume 685

D 82 (Diss. RWTH Aachen University, 2025)

ISSN 1866-1793
ISBN 978-3-95806-871-1

Vollständig frei verfügbar über das Publikationsportal des Forschungszentrums Jülich (JuSER)
unter www.fz-juelich.de/zb/openaccess.



This is an Open Access publication distributed under the terms of the [Creative Commons Attribution License 4.0](https://creativecommons.org/licenses/by/4.0/),
which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

The constant growth of the world population and the associated growing global energy demand represents a major challenge. The need to abstain nuclear and coal-fired power plants requires the rapid expansion of renewable energies and thus the expansion of photovoltaics (PV). In addition to increasing production capacity, improving the efficiency and sustainability of PV modules leads to higher yields. The resulting shorter payback times and the lower space required leads to increased economic interest and thus to a faster expansion of the PV industry.

This thesis deals with several topics in the area of efficiency and sustainability of PV modules. Research to improve the efficiency of solar cells and solar modules mainly focuses on the cell efficiency. For the purpose of comparability, this is verified conventionally at a fixed temperature and irradiation. The defined standard test conditions (STC) ensure the comparability of efficiency measurements in different laboratories around the world, but are not representative for the expected efficiency or expected yield of a module in operation consisting of several cells. The first topic addressed in this thesis is the extrapolation of the expected yield of a PV module from the laboratory-measured efficiency of a single CIGS solar cell. On the root of the extrapolation model stand temperature- and irradiation-dependent IV measurements. Based on realistic assumptions, a corresponding module characteristic of a 100 W_p module is extrapolated to each measured cell IV characteristic. To determine the yield of this module in realistic operating conditions, standard climatic reference profiles are used, which are defined in the IEC61853 standard. In addition, the extrapolation model can be used to determine how much yield potential there is in individual improvements to the cell performance.

In addition to the temperature and irradiation dependent efficiency of solar modules, efficiency stability is crucial for the yield of a solar module that is in operation for several years. Long-term experiments of PV modules in operation cover all possible influences affecting the performance. However, the superposition of these many influences also means that individual influences are difficult to separate from one another. In order to investigate the effect of individual influences, laboratory tests are required in which the operating conditions can be adjusted and controlled. One influence on CIGS modules in operation is the illumination with (sun) light. To investigate the separate influence of illumination in more detail, an accelerated degradation experiment was set up in which a total of 24 flexible CIGS solar cells were illuminated

under different operating conditions for approximately 1170 h. The operating temperature was varied between 25°C and 70°C and the irradiation intensity was varied between 0 Wm⁻² and 1000 Wm⁻². In addition, the influence of the operating bias was examined, with 12 cells kept in open circuit and 12 cells in short circuit. The results of the experiment show a strong dependency of the sustainability on the operating bias, with cells in short circuit showing high degradation rates. The process that leads to degradation is temperature-dependent (higher effect at higher temperatures) and mostly independent of the irradiation intensity. In addition, the analysis of the *IV* measurement data with the one-diode model indicates an increased probability of a recombination process with a high ideality factor. A possible recombination process that would explain this development is Shockley-Read-Hall (SRH) recombination in the quasi-neutral region of the p-n junction.

While laboratory experiments can lead to valuable insights on the effect of accelerated conditions, long-term experiments under realistic operating conditions are ultimately the best way to depict all influences affecting a PV modules efficiency in outdoor operation. In addition to measuring the module performance (often *IV* characteristics), this also requires the collection of meteorological data (e.g. module temperature and irradiation). Since the environmental influences are not controlled in these long-term experiments, the data structures recorded in this way are prone to errors. To this end, an analysis of the data requires a classification of the data quality and an appropriate filter before the analysis. In this work, two possible filter methods are developed and discussed. The first approach evaluates on the plausibility, which is classified using physical models describing the relationships between the different dimensions of the data recorded. The correlation of several occurring deviations is weighted using the Mahalanobis distance. The second filter method specifically aims to address a systematic error in the data often present in outdoor data, which occurs due to partial shading of the module and/or the irradiation sensor. Irradiation and operating temperature dependencies of the short circuit current of a PV module are described using several Gaussian process regressions (GPRs, a statistical method for interpolating data) and outliers from measurement and expectation are iteratively filtered out.

After the data quality has been assessed and an appropriate filter has been applied, the data can be further analysed. The two biggest challenges are the dimensionality and size of the data structures as well as the irradiation and temperature dependencies, which needs to be taken into account rating the performance. To reduce the dimensionality of the data, I use the so-called extended solar cell parameters (ESPs), an extension of the standard solar cell parameters I_{SC} , V_{OC} , I_{MPP} and V_{MPP} on a total of 10 parameters, which describe the complete form of an *IV* characteristic. Furthermore, a principal component analysis (PCA) is used, in which a new coordinate system is chosen for the 10 ESPs and linear correlations between the ESPs are thus separated. The 10 principal components (PCs) determined in this way are linearly uncorrelated and are described with several Gaussian processes that use the module temperature, irradiation and the time as input. In order to be able to process the amount of data, the data set for each individual PC is divided into monthly subsets and individual Gaussian process regressions are trained.

The model can thus reproduce the 10 PCs, consequently the 10 ESPs and the complete IV characteristic for any input of temperature, irradiation and time. In addition, the Gaussian process regressions provide information about the uncertainty of the output, which arises from the measurement uncertainty and the distance (in time, irradiation and temperature) to the actual measurements. The versatility of possible applications of the model is underlined with 3 examples (a simple representation of the time series to visualize seasonality, long-term degradation and acclimatization, the comparison to a classical performance ratio analysis and an application of the one-diode model based on the output of the GPRs). In addition, an analysis of the model's uncertainty shows high accuracy and good agreement between real and predicted error distributions.

Kurzfassung

Das stetige Wachstum der Weltbevölkerung sowie der damit verbundene wachsende globale Energiebedarf stellt eine große Herausforderung dar. Die Notwendigkeit des Verzichts auf Atom- und Kohlekraftwerke bedingt dabei den schnellen Ausbau der erneuerbaren Energien und damit den Ausbau der Photovoltaik. Neben der Steigerung der Produktionskapazität, führt die Verbesserung der Effizienz und Effizienzstabilität von PV Modulen zu höheren Erträgen. Zudem führen die resultierenden kürzeren Amortisationszeiten und der geringere Flächenbedarf zu einem gesteigertem wirtschaftlichem Interesse und damit zu einem schnelleren Ausbau der PV Branche.

Diese Doktorarbeit beschäftigt sich mit mehreren Thematiken im Bereich der Effizienz und Effizienzstabilität von PV Modulen. Forschung zur Verbesserung der Effizienz von Solarzellen und Solarmodulen konzentriert sich hauptsächlich auf die Zelleffizienz. Diese wird zwecks Vergleichbarkeit konventionell bei einer festen Temperatur und Einstrahlung verifiziert. Die definierten Standardtestbedingungen sorgen für die Vergleichbarkeit von Effizienzmessungen in unterschiedlichen Laboren auf der Welt, sind jedoch nicht repräsentativ für die erwartete Effizienz beziehungsweise den erwarteten Ertrag eines aus mehreren Zellen bestehenden Moduls im Betrieb. Das erste Thema, welches in dieser Doktorarbeit behandelt wird ist daher die Extrapolation eines erwarteten Ertrags eines PV Moduls von der im Labor gemessenen Effizienz einer einzelnen CIGS Solarzelle. Die Grundlage des Extrapolationsmodells bilden temperatur- und einstrahlungsabhängige Strom-Spannungs (IU) Messungen. Auf Basis realistischer Annahmen wird zu jeder vermessenen IU Zellkennlinie eine entsprechende Modulkennlinie eines 100 W_p -Moduls extrapoliert. Um den Ertrag dieses Moduls in realistischen Betriebsbedingungen zu bestimmen werden Standardklimareferenzprofile verwendet, welche in der IEC61853 Norm definiert sind. Darüber hinaus kann mit dem Extrapolationsmodell bestimmt werden, wie viel Ertragspotential in einzelnen Verbesserungen der Zelleffizienz stecken.

Neben der temperatur- und einstrahlungsabhängigen Effizienz von Solarmodulen, ist die Effizienzstabilität entscheidend für den Ertrag eines Solarmoduls, welches über mehrere Jahre in Betrieb ist. Realistische Langzeitversuche an Solarmodulen im Betrieb decken am besten alle möglichen Einflüsse ab. Die Überlagerung dieser vielen Einflüsse hat jedoch auch die Folge, dass individuelle Einflüsse schwer voneinander getrennt werden können. Um den Effekt einzelner Einflüsse zu untersuchen bedarf es daher Laborversuche, bei welcher die Betriebsbedingungen kontrolliert und eingestellt werden können. Ein Einfluss auf

CIGS Solarmodule in Betrieb ist die Beleuchtung mit (Sonnen-)Licht. Um den separierten Einfluss der Beleuchtung genauer zu untersuchen wurde ein beschleunigter Degradationsversuch durchgeführt, bei dem insgesamt 24 flexible CIGS Solarzellen unter unterschiedlichen Betriebsbedingungen für ungefähr 1170 h beleuchtet wurden. Die Betriebstemperatur wurde dabei zwischen 25°C und 70°C und die Einstrahlungsstärke zwischen 0 Wm^{-2} und 1000 Wm^{-2} variiert. Zudem wurde der Einfluss des Betriebspunktes untersucht, wobei 12 Zellen in offener Klemmspannung und 12 Zellen in Kurzschluss gehalten wurden. Die Ergebnisse des Versuches zeigen eine starke Abhängigkeit der Effizienzstabilität vom Betriebspunkt, wobei Zellen in Kurzschluss hohe Degradationsraten aufzeigen. Der Prozess, welcher zur Alterung führt ist dabei temperaturabhängig (größer bei höheren Temperaturen) und weitestgehend unabhängig von der Einstrahlungsstärke. Zudem weist die Analyse der *IU* Messdaten mit dem Ein-Dioden Modell auf eine erhöhte Wahrscheinlichkeit eines Rekombinationsprozesses mit hohem Idealitätsfaktor hin. Ein möglicher Rekombinationsprozess, welcher diese Entwicklung erklären würde ist Shockley-Read-Hall (SRH) Rekombination im quasi-neutralen Bereich des p-n Übergangs.

Während Laborexperimente wertvolle Erkenntnisse über die Auswirkung beschleunigter Degradationsbedingungen liefern können, sind Langzeitexperimente unter realistischen Betriebsbedingungen letztlich der beste Weg, um alle Einflüsse abzubilden, die sich auf die Effizienz eines PV-Moduls im Außenbetrieb auswirken. Dazu bedarf es neben der Messung der Modulperformance (oft *IU* Charakteristiken) auch die Erfassung meteorologischer Daten (z.B. Modultemperatur und Einstrahlung). Da die Umwelteinflüsse in diesen Langzeitversuchen nicht kontrolliert sind, sind die so erfassten Datenstrukturen fehleranfällig. Eine Analyse der Daten bedarf demnach eine Einordnung der Datenqualität und einen der Analyse vorgeschalteten Datenfilter. In dieser Arbeit werden zwei mögliche Filtermethoden erarbeitet und besprochen. Der erste Ansatz untersucht die Plausibilität der Zusammenhänge der unterschiedlichen Datentypen, welche mit physikalischen Modellen eingeordnet werden. Dabei wird die Korrelation mehrerer auftretender Abweichungen mit der Nutzung der Mahalanobisdistanz nicht übermäßig gewichtet. Die zweite Filtermethode richtet sich gezielt auf einen oft auftretenden systematischen Fehler in den Daten, welcher durch teilweise Abschattung des Moduls und oder des Einstrahlungssensors auftritt. Einstrahlungs- und Betriebstemperaturabhängigkeiten des Kurzschlussstroms eines PV Moduls werden mit mehreren Gaussprozessregressionen (einer statistischen Methode zur Interpolation von Daten) beschrieben und Ausreißer von Messung und Erwartung iterativ herausgefiltert.

Nachdem die Datenqualität bewertet ist und ein entsprechender Filter angewendet wurde, können die Daten weiter analysiert werden. Die zwei größten Herausforderung bestehen dabei durch die Dimensionalität und Größe der Datenstrukturen sowie durch die Einstrahlungs- und Temperaturabhängigkeiten, welche zur Einschätzung der Performance berücksichtigt werden müssen. Um die Dimensionalität der Daten zu reduzieren verwende ich dabei die sogenannten ESPs (extended solar cell parameters), eine Erweiterung der Standardsolarzellparameters I_{SC} , V_{OC} , I_{MPP} und V_{MPP} auf insgesamt 10 Parameter, welche die komplette Form einer *IU* Kennlinie beschreiben. Des Weiteren wird eine PCA (principal component analysis) angewendet, bei welcher ein neues Koordinatensystem für

die 10 ESPs gewählt wird und so lineare Korrelationen zwischen den ESPs getrennt werden. Die so bestimmten 10 PCs (principal components) sind linear unkorreliert und werden mit mehreren Gaussprozessen beschrieben, welche die Modultemperatur, Einstrahlung und den Zeitpunkt der Messung verwenden. Um die Datenmenge prozessieren zu können, wird der Datensatz jeder einzelnen PC in monatliche Teildatensätze geteilt und individuelle Gaussprozessregressionen trainiert. Das Modell kann so die 10 PCs, folglich die 10 ESPs und so die komplette *IU* Charakteristik für einen beliebigen Input aus Temperatur, Einstrahlung und Zeit reproduzieren. Darüber hinaus geben die Gaussprozessregressionen Aufschluss über die Unsicherheit des Outputs, welcher sich über die Messunsicherheit und die Distanz (in Zeit, Einstrahlung und Temperatur) zu den tatsächlichen Messungen ergibt. Die vielseitige Einsatzmöglichkeit wird anhand von 3 Beispielen (einer einfachen Darstellung der Zeitreihen zur Visualisierung von Saisonalität, Langzeitdegradation und Akklimatisierung, dem Vergleich zu einer klassischen performance ratio Analyse und einer aufbauenden Anwendung des Ein-Dioden Modells) aufgezeigt. Zudem zeigt eine Analyse der Genauigkeit des Modells lediglich geringe Abweichungen und eine gute Übereinstimmung realer und vorhergesagter Fehlerverteilungen.

Contents

Abstract	i
Kurzfassung	v
List of Figures	xi
Notations and Abbreviations	xv
List of Publications	xvii
1 Introduction	1
1.1 Predicting Outdoor PV Performance	1
1.2 Addressing Long-Term Degradation	2
1.3 Challenges in Long-Term Studies	3
1.4 Scope and Outline of this Thesis	3
2 Theoretical Background	5
2.1 Mathematical methods	7
2.1.1 Principal component analysis (PCA)	7
2.1.2 Mahalanobis distance	8
2.1.3 Gaussian process regression (GPR)	10
2.2 Solar cells and PV modules	14
2.2.1 Semiconductor solar cells	14
2.2.2 PV technologies	19
2.3 <i>IV</i> characteristics	25
2.3.1 The one-diode equivalent circuit model	25
2.3.2 The IEC60891 norm	29
2.3.3 Describing the shape of (measured) <i>IV</i> characteristics . .	30
2.4 Characterization methods and set-ups	36
2.4.1 Sun simulator set-up	36
2.4.2 Electro- and Photoluminescence set-up	36
2.4.3 Light induced degradation set-up	37
2.5 Outdoor Performance	38
2.5.1 Outdoor data	38
2.5.2 Performance ratio analysis and degradation estimation .	40
2.5.3 POA Modelling	42

2.5.4	Faiman module temperature model	42
2.5.5	CSER (IEC61853)	42
3	Extrapolation of Module Performance	45
3.1	Introduction and scientific context	46
3.2	Sample preparation	47
3.3	From cell performance to module performance	47
3.3.1	TCO series resistance	47
3.3.2	P2 series resistance	48
3.3.3	TCO Transmission Correction	49
3.3.4	Extrapolation of module IV characteristics	50
3.3.5	Resulting Module Performance	50
3.4	From module performance to module yield	51
3.5	Impact of performance changes on module yield	52
3.6	Summary	56
4	Light induced Degradation of CIGS solar cells	59
4.1	Introduction	60
4.2	CIGS Degradation, Accelerated Testing and Metastabilities	61
4.3	Experimental	63
4.3.1	Sample preparation	64
4.3.2	ex situ measurements	65
4.3.3	Light induced degradation set-up and in situ measurements	65
4.4	Results	66
4.4.1	Performance before LID	66
4.4.2	Degradation of the SSPs (in situ)	67
4.4.3	Degradation of the SSPs (ex situ)	69
4.5	Discussion	71
4.6	Summary and Outlook	75
5	Filtering of Outdoor data	77
5.1	Introduction	78
5.2	Mahalanobis distance filter	79
5.2.1	Concept	79
5.2.2	Mahalanobis filtering results	80
5.3	Gaussian Process Regression filter	85
5.3.1	Concept	85
5.3.2	GPR filtering results	87
5.4	Summary and Outlook	91
6	Analyzing PV Outdoor data	93
6.1	Introduction	94
6.2	A Gaussian process regression IV model	97
6.2.1	Filtering	97
6.2.2	ESPs	98
6.2.3	PCA	98
6.2.4	Mixture of Experts GPR	99
6.3	Results	101
6.3.1	Timeseries analysis	101

6.3.2	PR analysis	108
6.3.3	Physical analysis using the one-diode model	112
6.4	Validation	116
6.5	Summary and Outlook	123
7	Summary	125
7.1	Outlook	127

List of Figures

2.1.1 Example for a PCA of a 2D dataset	8
2.1.2 Euclidean and Mahalanobis distance of a 2D dataset	9
2.1.3 Example for predictions made with a GPR	10
2.1.4 Example for an optimized GPR and its limitations	12
2.1.5 Example for a failed optimization of a GPR	12
2.2.1 Fermi level of p and n doped semiconductor and pn junction . .	16
2.2.2 Silicon solar cell structures	20
2.2.3 Substrate and superstrate structure of thin-film solar cells . . .	22
2.2.4 P1P2P3 scribing line architecture	22
2.3.1 One-diode equivalent circuit model	25
2.3.2 J/V characteristic of a solar cell	26
2.3.3 Influence of parallel resistance on the JV characteristic	27
2.3.4 Influence of series resistance on the JV characteristic	28
2.3.5 Influence of the dark saturation current on the JV characteristic	29
2.5.1 The six standard climatic reference profiles provided in part 4 of the IEC61853 norm	43
3.3.1 Comparison of cell JV and module IV	51
3.4.1 Extrapolated ≈ 100 Wp CIGS module performance	53
3.5.1 Yield change due to a 10% gain in the six correction parameters	54
3.5.2 Yield gain in the six standard reference climates due to a 10% gain in V_{OC}	56
3.5.3 Extrapolated ≈ 100 Wp module performance gain with a 10% gain in V_{OC}	57
4.3.1 Transmittance of the ND filters used in LID experiments	64
4.4.1 JV characteristics of the 24 used CIGS solar cells at STC	67
4.4.2 Absolute change in in situ open circuit voltage ΔV_{OC} due to LID	68
4.4.3 Absolute change in ex situ open circuit voltage ΔV_{OC} at STC due to LID	70
4.4.4 Absolute change in ex situ short circuit current density ΔJ_{SC} at STC due to LID	71
4.4.5 Absolute change in ex situ efficiency $\Delta \eta$ at STC due to LID . .	71
4.5.1 Series resistance R_s before and after LID as well as after two weeks of DS	73

4.5.2 Dark saturation current density J_0 before and after LID as well as after two weeks of DS	74
4.5.3 Diode ideality factor n_{id} before and after LID as well as after two weeks of DS	74
4.5.4 Recombination current density J_x before and after LID as well as after two weeks of DS	75
5.2.1 Mahalanobis distance distribution for the Eugene mSi0166 dataset	81
5.2.2 Modelled versus measured G_{POA} for the unfiltered and filtered Eugene mSi0166 dataset	82
5.2.3 Corrected I_{SC} versus measured G_{POA} for the unfiltered and filtered Eugene mSi0166 dataset	83
5.2.4 Corrected V_{OC} versus T_{Mod} for the unfiltered and filtered Eugene mSi0166 dataset.	83
5.3.1 I_{SC} versus G_{POA} for the unfiltered and with one iteration filtered Italy CdTe1 dataset	89
5.3.2 I_{SC} versus G_{POA} for the with two and four iterations filtered Italy CdTe1 dataset	89
6.1.1 Illustration of the GPR IV model and its PV outdoor data processing steps	96
6.3.1 GPR PC timeseries prediction for the Italy CIGS4 dataset . . .	103
6.3.2 Direction of PC1 and PC2 for the Italy CIGS4 dataset	104
6.3.3 GPR ESP timeseries prediction for the Italy CIGS4 dataset . .	107
6.3.4 GPR PR prediction and classical PR estimate using a narrow irradiation band filter	110
6.3.5 GPR PR prediction and classical PR estimate using a broad irradiation band filter	111
6.3.6 GPR timeseries prediction fit of one diode model parameters for the Italy CIGS4 dataset	114
6.3.7 GPR timeseries prediction of the product $Nn_{id}E_a$ for the Italy CIGS4 dataset	115
6.4.1 GPR performance validation for ESP prediction reconstructed from 10 PCs for the Italy CIGS4 dataset	118
6.4.2 GPR performance validation for ESP prediction reconstructed from 7 PCs for the Italy CIGS4 dataset	118
6.4.3 GPR deviation and predicted deviation of the ESP prediction reconstructed from 7 PCs for the Italy CIGS4 dataset	121

Notations and Abbreviations

1D	one Dimensional
2D	two Dimensional
3D	three Dimensional
5D	five Dimensional
AC	Alternating Current
AM1.5G	Air Mass 1.5 Global Spectrum
ARIMA	AutoRegressive Integrated Moving Average
BIPV	Building Integrated PhotoVoltaic
CDF	Cumulative Distribution Function
CSER	Climatic Specific Energy Rating
DC	Direct Current
DFR	Design For Recycability
DH	Damp Heat
DHI	Diffuse Horizontal Irradiation
DLIT	Dark Lock-In Thermography
DNI	Direct Normal Irradiation
DS	Dark Storage
DUT	Device Under Test
EL	ElectroLuminescence
EoL	End of Life
ESP	Extended Solar cell Parameter
EVA	Ethylene Vinyl-Acetate
GHI	Global Horizontal Irradiation
GPR	Gaussian Process Regression
GPU	Graphics Processing Unit
IEC	International Electrotechnical Commission
ILIT	Illuminated Lock-In Thermography
IRT	InfraRed Thermography
ITO	Indium Tin Oxide
IQR	Inter-Quartile Range
L-BFGS	Limited-memory Broyden–Fletcher–Goldfarb–Shanno
LED	Light Emitting Diode
LIT	Lock-In Thermography
LID	Light Induced Degradation
LS	Light Soaking
MPP	Maximum Power Point

ND	N eutral D ensity
NOCT	N ominal O perating C ell T emperature
NREL	N ational R enewable E nergy L aboratory
OC	O pen C ircuit
PC	P rincipal C omponent
PCA	P rincipal C omponent A nalysis
PDT	P ost D eposition T reatment
PERC	P assivated E mitter and R ear C ontact
PID	P otential I nduced D egradation
PL	P hoto L uminescence
POA	P oint O f A rray
PPC	P ersistent P hoto C onductivity
PR	P erformance R atio
PV	P hoto V oltaic
RBF	R adial B asis F unction
RMS	R oot M ean S quare
RMSE	R oot M ean S quared E rror
SC	S hort C ircuit
SCR	S pace C harge R egion
SHJ	S ilicon H etero J unction
SMU	S ource M easure U nit
SRH	S hockley- R ead- H all
SSDP	S imple S ky D ome P rojector
SSP	S tandard S olar cell P arameter
STC	S tandard T est C onditions
TCO	T ransparent C onducting O xide
UV	U ltra V iolet
UVF	U ltra V iolet F luorescence
VIPV	V ehicle I ntegrated P hoto V oltaic
ZSW	Z entrum für S onnenenergie- und W asserstoff-Forschung

List of Publications

- Vaas, T. S., and Pieters, B. E. "Application of Extended Solar Cell Parameters to monitoring of IV characteristics." 2021 IEEE 48th Photovoltaic Specialists Conference (PVSC). IEEE, 2021.
- Vaas, T. S., Pieters, B. E., Friedlmeier, T. M., Gerber, A., and Rau, U. "Extrapolating CIGS Module Performance From Laboratory Cell Performance Using IEC61853 Standard Reference Climatic Profiles." IEEE Journal of Photovoltaics, 2023.
- Vaas, T. S., Körtgen, J., Sovetkin, E., Rau, U., and Pieters, B. E. "Plausibility Filtering of PV Outdoor Data." 2023 IEEE 50th Photovoltaic Specialists Conference (PVSC). IEEE, 2023.
- Vaas, T. S., Pieters, B. E., Roosen-Melsen, D., van den Nieuwenhof, M., Kingma, A., Theelen, M., Zahren, C., Gerber, A., and Rau, U. "Light induced degradation of CIGS solar cells." Solar Energy Materials and Solar Cells, 2024.
- Vaas, T. S., Pieters, B., Sovetkin, E., Gerber, A., and Rau, U. "A Gaussian process regression IV model for PV outdoor data." Progress in Photovoltaics: Research and Applications, 2025.

Introduction

The growth of the photovoltaic (PV) sector is critical to addressing the increasing global energy demand. Projections for 2050 estimate that the globally installed PV capacity could range between 14 and 80 terawatts (TW) within the next 25 years [1]. To meet this demand, not only an increase in overall production is necessary, but improvements in the efficiency and sustainability of PV modules are also essential. This thesis focuses on the "Analysis and Prediction of Outdoor PV Data", aiming to provide a comprehensive framework for predicting and understanding PV performance from the laboratory to real-world conditions.

1.1 Predicting Outdoor PV Performance

The journey begins with predicting outdoor PV performance based on laboratory solar cell data. Recent advancements in various solar cell technologies, including cadmium telluride (CdTe), copper indium gallium selenide (CIGS), perovskite, and silicon heterojunction cells, have shown remarkable efficiency improvements [2]. However, the performance of individual cells under standard test conditions (STC) does not directly translate to the performance of complete PV modules in outdoor environments.

To bridge this gap, it is essential to extrapolate module performance from cell-level data. This involves understanding the interplay between cell characteristics and module configuration, accounting for factors such as cell and module size, interconnection methods, dead area losses, and series resistances. Standard test conditions, defined as an irradiance of $G_{\text{STC}} = 1000 \text{ W/m}^2$ and a cell temperature of $T_{\text{STC}} = 25^\circ\text{C}$, provide a baseline for comparison. However, these conditions rarely reflect real-world operating environments.

A more realistic approach involves using Nominal Operating Cell Temperature (NOCT) conditions, which consider medium-high irradiance ($G_{\text{NOCT}} = 800 \text{ W/m}^2$) and moderate ambient temperature ($T_{\text{amb,NOCT}} = 20^\circ\text{C}$). Addition-

ally, the IEC61853 standard offers a reference for outdoor conditions across various climates [3], facilitating more accurate predictions of PV module performance in different environments. This approach allows for the calculation of the Climatic Specific Energy Rating (CSER), which estimates the expected yield of PV modules under diverse climatic conditions and compares the estimated outdoor performance with the STC efficiency [4].

1.2 Addressing Long-Term Degradation

Extrapolating from laboratory cells to outdoor performance can predict the performance of a PV module measured at a specific point in time. However, this prediction does not account for long-term degradation effects, which are influenced by numerous environmental factors. Degradation mechanisms in PV modules are complex, involving various overlapping processes that contribute to performance loss over time. Given the complexity and lengthy timescales (up to 20 years), this thesis does not provide a comprehensive approach to predicting long-term degradation.

On a laboratory scale, stability can be characterized to some extent. Due to the high costs of laboratory infrastructure, it is generally not feasible to conduct stability experiments on cells and mini-modules over many years. Instead, laboratory characterization of cell stability is performed using accelerated ageing tests, such as damp heat (DH) and potential induced degradation (PID), which are typically conducted over weeks to months. The primary purpose of these tests was not initially to gain deeper insights into the underlying mechanisms and causes of degradation. In both accelerated ageing tests and long-term outdoor studies, multiple degradation effects might overlap due to the simultaneous influence of various factors on the PV module. Laboratory experiments at the cell scale are more useful for decomposing these influences. Under laboratory conditions, parameters such as humidity, bias, temperature, and irradiance can be tuned and controlled. Furthermore, extensive ex-situ and in-situ characterization is less tedious at the cell level.

However, accelerated testing procedures cannot replace actual long-term degradation studies. Long-term studies on PV modules have the clear advantage of best depicting real operating conditions and outdoor influences. Thus, such long-term studies are crucial to test and validate the results of accelerated testing procedures. Studies of PV module performance, such as those conducted by Schweiger *et al.* [5] and Marion *et al.* [6], provide valuable insights into degradation patterns. These studies, along with performance ratio analyses of operational PV systems, are essential for assessing the expected lifetime of PV

modules.

1.3 Challenges in Long-Term Studies

Long-term outdoor studies are hampered by several factors. The first limitation is the extended time over which such experiments must be operated. Furthermore, outdoor conditions are explicitly not controlled, and they should not be controlled to reflect realistic conditions. Consequently, such datasets, typically comprising electrical performance measures and meteorological data, often suffer from disturbances leading to measurement errors, and missing or otherwise invalid data. These datasets must thus be rigorously analysed to ensure data quality. Filtering the data appropriately ensures that the dataset remains representative and is a crucial step for further analysis.

1.4 Scope and Outline of this Thesis

In this work, a model to extrapolate CIGS PV module yield from cell performance is developed and laboratory degradation experiments on CIGS solar cells are presented. Furthermore, filtering methods and methodologies for further analysis of module performance and degradation over time are developed. This type of analysis is crucial to validate and test the completeness of degradation mechanisms studied in laboratory settings, and thus allows for a more comprehensive understanding of laboratory cell performance to long-term outdoor performance.

In Chapter 2 the theoretical background needed throughout the thesis is discussed. After a quick summary of the used mathematical concepts and tools, the theory of semiconductor solar cell physics is summarized and common PV technologies are discussed. Since *IV* characteristics are used throughout the whole thesis as the most important performance measure of solar cells and PV modules, descriptive and physical models for *IV* characteristics, temperature and irradiation dependencies as well as fitting algorithms for *IV* characteristics are discussed. Afterwards the used laboratory set-ups are presented. In the last section of the first chapter outdoor data is discussed, emphasizing common, used and reference data profiles, data issues and challenges, common performance ratio analysis approaches as well as irradiance and module temperature modelling. Chapter 3 presents a model to extrapolate CIGS module performance from cell performance as well as to extrapolate module yield in different climates from module performance. Furthermore, the effect of potential improvements on cell level on the yield is quantified. Chapter 4 presents laboratory experiments,

studying light induced degradation of CIGS solar cells as well as the dependency on temperature, irradiation intensity and applied bias condition. Chapter 5 introduces two different approaches for data quality assessment and filtering. In Chapter 6 a model for statistically analysing PV outdoor data is presented. Finally, in Chapter 7 the presented work is summarized.

Theoretical Background

2.1	Mathematical methods	7
2.1.1	Principal component analysis (PCA)	7
2.1.2	Mahalanobis distance	8
2.1.3	Gaussian process regression (GPR)	10
2.2	Solar cells and PV modules	14
2.2.1	Semiconductor solar cells	14
2.2.2	PV technologies	19
2.3	<i>IV</i> characteristics	25
2.3.1	The one-diode equivalent circuit model	25
2.3.2	The IEC60891 norm	29
2.3.3	Describing the shape of (measured) <i>IV</i> characteristics . .	30
2.4	Characterization methods and set-ups	36
2.4.1	Sun simulator set-up	36
2.4.2	Electro- and Photoluminescence set-up	36
2.4.3	Light induced degradation set-up	37
2.5	Outdoor Performance	38
2.5.1	Outdoor data	38
2.5.2	Performance ratio analysis and degradation estimation .	40
2.5.3	POA Modelling	42
2.5.4	Faiman module temperature model	42
2.5.5	CSER (IEC61853)	42

This chapter provides a detailed introduction to the key theoretical concepts required for the thesis. Section 2.1 lays the foundation by introducing the mathematical concepts applied in the thesis. Section 2.2 delves into the physics behind solar cells, providing an overview of c-Si, CdTe, and CIGS PV technologies. The emphasis shifts to *IV* characteristics in Section 2.3, highlighting them as a crucial performance indicator for solar cells and PV modules. In Section 2.4 the experimental setups used in the thesis are described, including the sun simulator, electroluminescence (EL) and photoluminescence (PL) setup, and the light induced degradation (LID) setup. Finally, Section 2.5 focuses on the PV outdoor data used, explores common analysis

techniques for the performance ratio (PR) and the degradation of PV modules, and introduces concepts such as Plane of Array (POA) modelling, PV module temperature modelling, and outdoor performance rating.

2.1 Mathematical methods

This section covers the key mathematical methods employed in the thesis. It includes a succinct overview of principal component analysis, the Mahalanobis norm, and Gaussian process regressions.

2.1.1 Principal component analysis (PCA)

Principal Component Analysis (PCA) essentially transforms the coordinate system of a dataset. This method first introduced by Pearson aims to identify the directions where the data spreads out the most, revealing the underlying patterns of a set of n -dimensional data points [7]. In simpler terms, PCA finds new axes (principal components) for the dataset such that each new axis accounts for as much variability in the data as possible. Here's how it works:

- The first principal component (PC) is the direction that best captures the variance in the data, akin to finding the best straight line that fits through a cloud of n -dimensional data points.
- Subsequent PCs identify the next best directions of variance that are orthogonal (at right angles) to all previously determined PCs, ensuring a lack of redundant information.
- This process redefines the dataset's basis into an orthonormal one, where the axes (PCs) are linearly uncorrelated and orthogonal to each other, providing a comprehensive description of the n -dimensional space.

Figure 2.1.1 shows an example for a PCA applied to a 2D dataset of correlated normal distributed random variables. Note that the axes of x and y are normalized with the standard deviation in x respectively y direction. The red and blue arrow depict the direction of PC1 and PC2, respectively, as well as the variance of the dataset in the respective direction (by the length of the arrows).

For this example one finds the approximate directions

$$\begin{pmatrix} \text{PC1}_x \\ \text{PC1}_y \end{pmatrix} = \begin{pmatrix} 0.718 \\ 0.696 \end{pmatrix}$$

and

$$\begin{pmatrix} \text{PC1}_x \\ \text{PC1}_y \end{pmatrix} = \begin{pmatrix} -0.696 \\ 0.718 \end{pmatrix}.$$

PC1 points in the direction of highest variance in the dataset, while PC2 is orthogonal to PC1. Note that the two directions form a orthonormal base, where the two vectors are orthogonal to each other and have a (Euclidean) length of 1.

Expressing all 2D datapoints $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ in terms of the directions of the PCs results in a basis change. Formally the basis change is performed by multiplying each point \vec{x} by a translation matrix \mathbb{P} to compute the representation of the

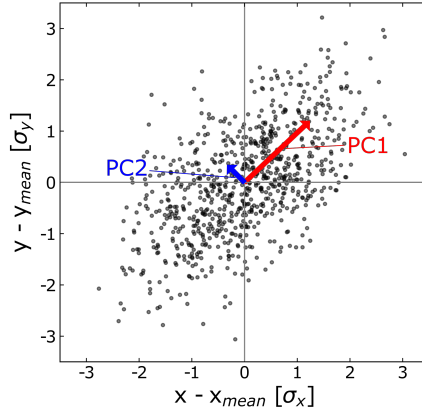


Figure 2.1.1: Example for a PCA respectively the PCs of a 2D dataset of correlated normal distributed random variables. The direction of the red (blue) arrow depict the first (second) PC, while the length depicts the variance found for the respective direction.

point in terms of the PCs

$$\overrightarrow{\text{PC}} = \begin{pmatrix} \text{PC1} \\ \text{PC2} \end{pmatrix} = \begin{pmatrix} \text{PC1}_x & \text{PC1}_y \\ \text{PC2}_x & \text{PC2}_y \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \mathbb{P}\vec{x}. \quad (2.1)$$

For this example with a 2D dataset, the translation matrix \mathbb{P} has the dimensions of 2×2 . In general, for a n -dimensional dataset, a PCA results in the determination of a $n \times n$ translation matrix, where the i -th row reflects the direction of the i -th PC.

A key advantage of a PCA is its ability to streamline the data by removing linear correlations, since the PCs are uncorrelated. It effectively reduces the dataset's dimensionality if it can be accurately represented in fewer dimensions. Thus, PCA not only clarifies which directions in the data are the most informative but also identifies if the dataset can be simplified into a lower-dimensional space, shedding light on its intrinsic structure.

2.1.2 Mahalanobis distance

The Mahalanobis distance was first introduced by Mahalanobis as a measure of the distance between a point \vec{x} and the mean $\vec{\mu}$ of a distribution \mathbb{Q} . A reprint of the work published first in 1936 can be found in [8]. With a given distribution \mathbb{Q} on \mathbb{R}^n , the Mahalanobis distance of \vec{x} from \mathbb{Q} is

$$d_M(x, \mu) = \sqrt{(\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \vec{\mu})}, \quad (2.2)$$

where Σ is the non negative definite covariance matrix of \mathbb{Q} . The Mahalanobis distance effectively measures how far a point \vec{x} is from the mean $\vec{\mu}$ of a dataset \mathbb{Q} . However, instead of calculating this distance in the usual way, it considers the variance of \mathbb{Q} in the direction of the distance between \vec{x} and $\vec{\mu}$. This means

it takes into account how spread out the dataset is in the direction of interest, providing a distance that is adjusted for the dataset's overall variability.

Figure 2.1.2 gives an example for computing Euclidean and Mahalanobis distances for the 2D dataset of correlated normal distributed random variables also used in Section 2.1.1. Note that the axes of x and y are normalized with the standard deviation in x respectively y direction. The red, orange and yellow circles in Fig. 2.1.2 (a) depict a Euclidean distance of one, two and three standard deviations from the mean of the 2D dataset. Incorporating the covariance of the 2D dataset in form of computing the Mahalanobis distance (see Fig. 2.1.2 (b)) results in the red, orange and yellow ellipses marking distances of one, two and three standard deviations, respectively.

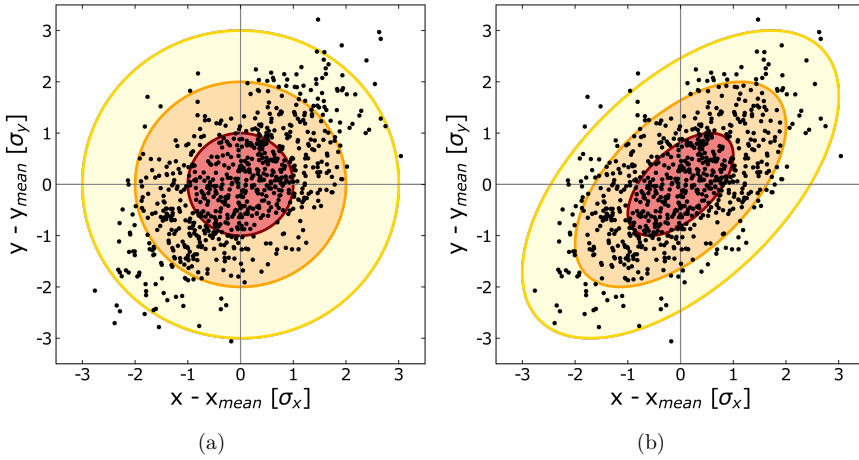


Figure 2.1.2: (a)Euclidean and (b) Mahalanobis distances of a 2D dataset of correlated normal distributed random variables. The red, orange and yellow circles respectively ellipses indicate a Euclidean respectively Mahalanobis distance of one, two and three standard deviations.

Illustratively, the Mahalanobis distance describes the distance of a point \vec{x} from the mean $\vec{\mu}$ of a dataset \mathbb{Q} incorporating (linear) correlations in the dataset \mathbb{Q} . The Mahalanobis distance is closely related to a PCA. Computing the Mahalanobis distance is equivalent to changing the basis according to a PCA, subsequently normalizing the data to the variance in the direction of the PCs and computing the Euclidean distance in the new normalized basis.

For normal distributed n -dimensional random variables, the Mahalanobis distance can further be used to define a threshold Mahalanobis distance, where a given quantile of the dataset has a Mahalanobis distance to the mean $\vec{\mu}$ smaller or equal to the threshold. Assuming \mathbb{Q} as a normal distribution in \mathbb{R}^n , d_M^2 follows the χ^2 -distribution with n degrees of freedom [9]. Equating the cumulative distribution function (CDF) of the χ^2 -distribution with a given quantile q gives

$$F_{\chi^2}(n, d_{M, \text{threshold}}) = q, \quad (2.3)$$

where $F_{\chi^2}(n, x)$ is the CDF of the χ^2 -distribution with n degrees of freedom,

and $d_{M,\text{threshold}}$ is the Mahalanobis distance threshold for which only $1 - q$ of the data points have a higher Mahalanobis distance than $d_{M,\text{threshold}}$.

2.1.3 Gaussian process regression (GPR)

Gaussian Process Regression (GPR) serves as a probabilistic, non-parametric approach for predicting outcomes for inputs that have not been previously observed. By leveraging the inherent randomness in data measurements, GPR excels in not only interpolating desired outputs but also in providing a measure of certainty (or uncertainty) about these predictions. This section is designed to capture the core principles of GPRs, elucidating their application and significance within the scope of this thesis. Furthermore, the strengths and limitations of GPRs are demonstrated with a simplified example.

For a thorough examination of GPR, including discussions on theoretical underpinnings, model selection strategies, the challenges posed by large datasets, and the implementation of approximate GPR methods, the work of Williams and Rasmussen ([10]) is strongly recommended. This resource offers extensive insights into GPR, making it an invaluable reference for those seeking to deepen their understanding of these models.

Given a discrete set of observations $y(x_i)$, that are assumed to depend on the measurements x_i , a Gaussian process regression predicts y at the unsampled inputs \tilde{x}_j . Figure 2.1.3 visualizes a GPR $f_y : \mathbb{R} \rightarrow \mathbb{R}$ trained with the observations y (red) given for the three discrete measurements x_i . Here, the data y follows the underlying function $y = f(x) = x^2 \sin(x)$ (red dashed line), which is in a real case unknown, respectively to be approached with the predictions. For simplicity, the modelled data y exhibit no noise. In the given example the goal of the GPR is to determine the best guess for \tilde{y} (blue) as well as the uncertainty (95% confidence interval in light blue) at the unsampled locations $\tilde{x} \in (0, 10)^1$.

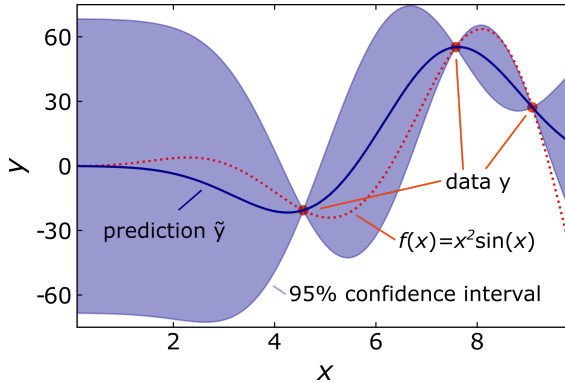


Figure 2.1.3: Example for a Gaussian Process Regression aiming to predict $\tilde{y}(\tilde{x})$ from the given discrete observations y at x_i .

To predict $\tilde{y}(\tilde{x}_j)$ one describes $\tilde{y}(\tilde{x}_j)$ as a weighted sum of $y(x_i)$. The weights

¹ \tilde{y} is modelled for a finite number of \tilde{x} with regular spacing, since $\tilde{x} \in (0, 10)$ contains a infinite number of real numbers. The output of a GPR prediction is always a finite set of values, since a GPR cannot determine the underlying function.

are defined via a correlation matrix $K_y = K_y(x, \tilde{x})$, where $\tilde{y}(\tilde{x}_j)$ is close to $y(x_i)$, if \tilde{x}_j is close to x_i . Beside the discrete x_i and \tilde{x}_j a GPR introduces a set of hyperparameters Θ , that need to be optimized during regression, to define the correlation matrix $K_y = K_y(x, \tilde{x}, \Theta)$. The correlation respectively covariance matrix is also called kernel and defines the family of functions $y(x_i)$ used to predict $\tilde{y}(\tilde{x}_j)$, i.e. it defines the shape of the prediction (e.g. smoothness and periodicity). Since the kernel reflects the correlation between a discrete set of n data points, it is $n \times n$ dimensional and in general positive definite as well as symmetric.

The optimization of the hyperparameters aims to maximize the probability (or marginal likelihood) $p(y|x, \Theta)$ of observing y with given input x and Θ . With the assumption of y being a set of random variables, any finite number of which have a joint Gaussian distribution, the probability respectively the logarithm of this probability (the log marginal likelihood) reads:

$$\log p(y|x, \Theta) = -\frac{1}{2}y^\top K_y^{-1}y - \frac{1}{2}\log |K_y| - \frac{n}{2}\log 2\pi. \quad (2.4)$$

Here, the term $-\frac{n}{2}\log 2\pi$ is a normalization constant depending on the number n of observations y and the term $-\frac{1}{2}\log |K_y|$ reflects the complexity penalty arising due to the choice of the covariance function as well as from the inputs. Beside its complexity the kernel or covariance matrix choice is crucial for a GPRs performance. One of the most common choices is the radial basis function (RBF) kernel, introducing an exponential decline of the correlation with increasing distance defined by the lengthscale hyperparameter l . The RBF kernel is defined as

$$K(x, \tilde{x}) = \exp\left(-\frac{\|x - \tilde{x}\|^2}{2l^2}\right). \quad (2.5)$$

The use of the RBF kernel implies the assumption, that the output functions to predict are smooth, i.e. the family of functions $y(x)$ used to predict $\tilde{y}(\tilde{x}_j)$ are smooth with varying x . Figure 2.1.4 shows the extended example (w.r.t. Fig. 2.1.3) of an optimized GPR trained with seven discrete observations y . Here, the GPR is trained with a one dimensional RBF kernel. The optimization of the hyperparameter based on the seven observations results in $l = l_x = 1.03$. One can see, that the predictions $\tilde{y}(\tilde{x}_j)$ are smooth with varying x and close to the expected underlying function $f(x)$ as long as \tilde{x}_j is close to a given x_i . Furthermore, it is clear, that the predicted uncertainty is low at \tilde{x}_j close to x_i and rises with rising distance. Since the example introduces a noise free case, the arising predicted uncertainty reflects solely the lack of information at the respective points \tilde{x}_j .

In Fig. 2.1.4 the main limitation of GPRs is clearly visible. Since a GPRs prediction relies on the correlation of x_i and \tilde{x}_j it is unable to make reliable predictions far away from the observations, i.e. a GPR is not suitable to extrapolate a discrete dataset. This is reflected in an enhanced uncertainty on the prediction far away from the observations. Far away from the observations a GPRs predictions tend to 0. In real regression scenarios the mean of the n observations μ_y , given by

$$\mu_y = \sum_{i=0}^n \frac{y(x_i)}{n}, \quad (2.6)$$

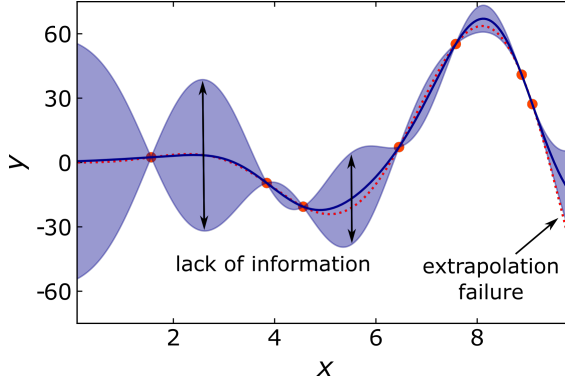


Figure 2.1.4: Example for an optimized Gaussian Process Regression with a lengthscale of $l_x = 1.03$ and its limitations with a discrete set of given observations

is often subtracted from the data. In the given example the underlying function oscillates around 0 and the mean of the observations is not subtracted. The optimized lengthscale $l_x = 1.03$ determines how fast the prediction tends towards 0 with rising distance from the observations. Here, and in general, the optimization of the hyperparameters is crucial for a GPRs performance. Figure 2.1.5 visualizes a GPR trained with the same data as in Fig 2.1.4, but with the unoptimized fixed hyperparameter $l_x = 0.4$. The performance of the GPR close to the observations remains good, while the lower lengthscale leads to an exaggerated tendency towards 0 in between observations.

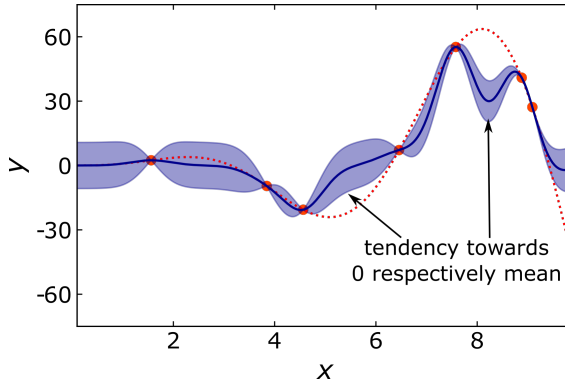


Figure 2.1.5: Example for a Gaussian Process Regression with a fixed lengthscale of $l_x = 0.4$ reflecting an unoptimized GPR respectively a GPR with an inaccurate result of the optimization routine.

In the limit of a lengthscale approaching 0 the resulting prediction $\tilde{y}(\tilde{x}_j)$ would still be accurate at the exact location of the observations ($\tilde{x}_j \in x_i$) but equal to 0 at any other point ($\tilde{x}_j \notin x_i$). The other extreme case of a very high lengthscale would result in a prediction equal to the mean of the data, since for every \tilde{x}_j the influence of all observations y is assumed to be equal with a constant correlation matrix.

For datasets, that are assumed to exhibit significant noise, a white noise kernel $K_{\text{WN}} = \sigma_n^2 I$ (I denotes the identity matrix) is often added to describe Gaussian noise with variance σ_n^2 in the observations. Referring to the given example this allows for deviations of the predictions from the examples, where the distribution of deviations should approach a Gaussian for high n (number of observations). In the context of introduced noise, one can further interpret a too high (low) lengthscale in a GPR with RBF kernel with underfitting (overfitting) the data leading to a very high (low) predicted noise level.

A closer look on the first term of 2.4, that is influenced by the observations y , introduces the second main limitation of the usability of GPRs. Since the computation time to invert a positive definite symmetric $n \times n$ matrix scales with n^3 , approximation methods might be necessary for large datasets. In this thesis I use a “mixture of experts” approach to deal with the high computation time arising to dataset sizes of $n \approx 60000$. The “mixture of experts” approach is based on splitting the dataset into m subsets, training individual GPR on subsets and overlaying their predictions. Each individual GPR can be interpreted as an “expert” on the subspace spanned by the used subset of inputs, i.e. each GPR will result in predictions with lower uncertainties on the subspace covered by the inputs used for training and predictions with higher uncertainties elsewhere. The accuracy of “mixture of experts” approaches is dependent on the choice how to split the dataset and the how to overlay the respective predictions of the multiple GPRs. Using a “mixture of experts” approach reduces the computation time by a factor of up to approximately $\frac{1}{m^2}$ with m being the number of subsets, since instead of inverting one $n \times n$ matrix (computation scales with n^3), m matrices of size $\frac{n}{m} \times \frac{n}{m}$ (computation time scales with $m \frac{n^3}{m^3} = \frac{n^3}{m^2}$) need to be inverted during training of the GPRs. Note, that a “mixture of experts” approach has been shown to be, in general, a suitable way to overcome challenges arising with big datasets [11].

2.2 Solar cells and PV modules

This section provides an overview of the theoretical foundation related to solar cell physics and highlights key differences among prevalent photovoltaic (PV) technologies.

2.2.1 Semiconductor solar cells

Absorption of light and generation of electron-hole pairs

The principle of solar cells is based on the generation of electrical energy through the absorption of (sun) light. Every atom has specific energetic states that can be occupied by electrons. If the photon energy is equal to the energy difference of two possible energetic states the absorption of the photon can lead to the excitation of an electron to a higher energetic state. For single atoms, there are only a few possible states, leading to distinctive absorption lines (i.e. absorption spectra featuring a few clearly defined and narrow peaks).

For condensed matter, the many different energetic or electronic states become indistinguishable and form so called energy bands. The Fermi level E_F defines the characteristic energy of electrons within a material. Apart of (thermally) excited electrons all electronic states below E_F are occupied and all electronic states above E_F are unoccupied (or empty). For metals E_F lies within an energy band, leading to a high density of electronic states near the Fermi level, implying a high mobility of electrons. In semiconductors and insulators E_F lies within two energy bands, the valence and the conduction band. These bands are separated by a band gap E_g , allowing no, or only very few, possible electronic states in between the two energy bands. At 0K the states in the conduction band are completely empty and the states in the valence band are completely filled with electrons. Unlike insulators, semiconductors have a smaller energy gap between E_F (the Fermi level) and the energy bands. This smaller gap allows for electrons to be thermally excited to the conduction band. As the conductivity depends on the number of electrons in the bands, and at any given temperature, only a certain number of electrons get excited to this level, these materials are called 'semiconductors', i.e. they conduct under circumstances.

Beside thermal excitation, also the absorption of a photon can lead to the excitation of an electron to the conduction band. This process leaves behind an "empty state" in the valence band, known as a "hole." Essentially, this hole is a quasi-particle that represents the absence of an electron. Therefore, when an electron is excited to the conduction band, an electron-hole pair is created. To generate an electron-hole pair through photon absorption, the photons energy needs to be at least equal to the energy of the band gap.

Semiconductors are effective absorbers for photons with an energy, that exceeds the semiconductors band gap energy E_g ². Considering the AM1.5G spectrum

²For indirect semiconductors the maximum of the valence and the minimum of the conduction band are shifted in momentum space. Due to the conservation of momentum, the additional participation of a phonon is required. The photon energy required for absorption is modified by the energy of the phonon. As a result also a small share of photons with a lower energy than E_g can be absorbed and the increase in the absorption coefficient is less steep with increasing photon energy.

of the sunlight, i.e. the spectrum of the sun after a passage through 1.5 earth atmospheres, a considerable share of the photons can be absorbed by various types of semiconductors. In contrast, insulators do not serve well as materials for absorbing sunlight, because the energy needed to overcome the band gap in insulators is higher than the energy of most photons in the solar spectrum.

Thermalisation and recombination

Once an electron (hole) is excited to the conduction (valence) band, it relaxes to the band edge, since the band edge is energetically most favourable inside the energy band. This leads to the creation of phonons (quasi particles describing quantized lattice vibrations) and such to thermalisation of the semiconductor. The relaxation to the band edge is a fast process, since within the energy band the possible energetic states are (nearly) continuous, enhancing the probability of the process.

Once electron and hole have reached the energy band edges, the energetic separation of valence and conduction band leads to a dramatically decreased probability of further thermalisation. However, there are many different mechanisms, that lead to a limited lifetime of generated electron-hole pairs. These mechanisms are called recombination mechanisms, since they describe the recombination of an electron and a hole.

Selective contacts

Beside the absorption of the photon energy and the creation of an electron-hole pair, the separation and extraction of the charge carriers is essential to create an efficient solar cell. First of all, a considerable lifetime for the generated electron-hole pairs is necessary. Second, the separation of the charge carriers needs to be fast enough to ensure, that only a minor share will recombine directly after generation and the charge carriers need to be extracted, i.e. a solar cell needs to have selective contacts. In semiconductor solar cells the separation of electrons and holes as well as the extraction of the charge carriers via selective contacts is established via a pn junction³.

The pn junction

To understand the charge carrier separation in semiconductor solar cells via a pn junction, a deeper look on the electrical properties of semiconductors is necessary. The charge carrier densities of the holes in the valence band n_h and the electrons in the conduction band n_e are of central importance. The charge carrier densities increase with temperature and can be expressed using the Boltzmann approximation to

$$n_h = N_v \exp\left(\frac{E_v - E_{fh}}{kT}\right) \quad (2.7)$$

³Note, that there are in principal other ways to establish selective contacts in a solar cell, but the most common realization in semiconductor solar cells is via a pn junction.

and

$$n_e = N_c \exp\left(\frac{E_{fe} - E_c}{kT}\right), \quad (2.8)$$

where k is the Boltzmann constant [12], E_{fh} and E_{fe} are the quasi Fermi levels of the holes and electrons and N_c and N_v are the effective density of states at the conduction band edge E_c and the valence band edge E_v , respectively.

The quasi Fermi levels are in equilibrium equal to the Fermi level E_F and approximately in the middle of the band gap in intrinsic semiconductors. An example for an intrinsic semiconductor is silicon. Defects in the crystal structure of a semiconductor can create electronic states within the band gap. Shallow defects (defects which are energetically close to the valance or conduction band) may lead to an effect called doping. Doping can be of two types: n-type and p-type. In n-type doping, a defect state close to the conduction band donates an electron to the conduction band, increasing the electron concentration. Conversely, in p-type doping, a defect state close to the valence band accepts an electron from the valence band, increasing the hole concentration. Doping is often intentionally used to modify the electrical properties of a semiconductor. To achieve an n-type doping, atoms of the semiconductor lattice are replaced by atoms containing more outer electrons. The increase in electron density results in $n_e > n_h$ equivalent to a shift of the Fermi level closer to the conduction band. Likewise, p-type doping is achieved by substituting some semiconductor atoms in the lattice with atoms with fewer outer electrons, thus leading to $n_h > n_e$, and the Fermi level is shifted closer to the valence band. Common dopants for n and p-type silicon are f.e. phosphorus and boron atoms, introducing one access electron to the conduction band and one access hole to the valence band, respectively. Figure 2.2.1 (a) illustrates this shift of Fermi level in n and p-type semiconductors.

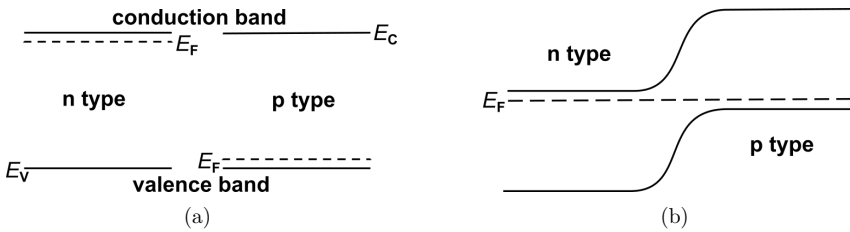


Figure 2.2.1: (a) Position of the Fermi level E_F within the band gap of a n and p-type semiconductor and (b) illustration of the alignment of the Fermi level in a pn homo junction.

In most semiconductor solar cells, n and p doped materials are brought into contact, creating a pn junction. Due to alignment of the Fermi levels, the conduction and valence band are bended. Figure 2.2.1 (b) illustrates the alignment of the energy bands for a pn homo junction, where the n-type and p-type

material consist of the same semiconductor ⁴. In a pn junction, the difference in the charge carrier densities of the two doped materials results in a diffusion current of charge carriers. The accumulation of holes in the n-type material and electrons in the p-type material induces a rising built-in electric field, generating a drift current in the opposite direction. Once thermal equilibrium is reached, the diffusion and drift currents balance each other, leading to no net current flow. Such, the space charge region (SCR) is created, where the electric field pulls all charge carriers to one or the other side.

In order to enable a net current flow through a pn junction in the absence of generation, holes and electrons must recombine. The strong difference of charge carrier concentration in doped semiconductors causes the net current to be carried mostly by the majority charge carriers, i.e. by electrons in n-type and by holes in p-type semiconductors. Due to the internal electric field, an effective energy barrier is built for the majority charge carriers. With an applied external positive voltage this energy barrier is reduced and the probability for recombination is enhanced. The linear reduction of the internal electric field with applying an increasing positive voltage results in an exponential increase in recombination current.

Applying a (small) negative external voltage enhances the internal electric field of a pn junction. This causes the enlargement of the space charge region, i.e. the majority carriers are pulled away from each other. Thus, (in the absence of optical generation) only a small net generation current of thermally excited carriers occurs. This net current is called dark saturation current. Since a pn junction limits the current for negative voltages and allows the current to flow for positive voltages it behaves as a diode (i.e. exhibits an asymmetric conductivity, where the conduction is much higher in one direction than in the other).

Recombination mechanisms

In solar cells different mechanisms of recombination occur. The reverse process of the generation of an electron-hole pair by the absorption of a photon is called radiative recombination. It describes the process of recombination of an electron of the conduction band with a hole of the valence band under emission of a photon. The rate of radiative recombination depends on the electron and hole concentration and is given by

$$R_{\text{rad}} = Bn_e n_h, \quad (2.9)$$

where B is the material dependent radiative recombination constant. Radiative recombination is generally not the dominant type of recombination in semiconductor solar cells ⁵. Nevertheless, radiative recombination may contribute to a performance loss, when electrons and holes are not separated fast enough.

⁴In a pn hetero junction the semiconductor of the n-type and the semiconductor of the p-type differ. Due to the difference of the two materials band gaps, a shift in the energy level of conduction and valence band edge at the interface of the two regions arises. Furthermore, introducing an intrinsic semiconductor between n-type and p-type semiconductor creates a p-i-n junction, leading to a more gradual transition between p-type and n-type band structure.

⁵Especially in indirect semiconductors the radiative recombination constant is very low due to the necessity of a phonon in the recombination process due to conservation of momentum analogous to the generation of an electron-hole pair in an indirect semiconductor

Furthermore, the process of radiative recombination is used to validate a solar cells or PV modules performance with electro- and photoluminescence (EL and PL) measurements. More details on the measurement technique can be found in Section 2.4.2.

A more decisive recombination mechanism in solar cells is Shockley-Read-Hall (SRH) recombination. It describes the recombination of an electron and a hole via a defect state within the band gap. Since the energy difference is reduced, the process is more likely than radiative recombination. The SRH recombination rate can be described via

$$R_{\text{SRH}} = \frac{n_e n_h - n_i^2}{n_e \tau_h + n_h \tau_e}, \quad (2.10)$$

where n_i is the intrinsic charge carrier density and τ_e respectively τ_h is the lifetime of an electron respectively a hole in a defect state. Assuming that the lifetimes of the charge carriers are similar, the SRH recombination rate scales with the minority charge carrier density. The minority charge carrier density is exponentially dependent on the change of the minority charge carrier quasi Fermi level (see 2.7 and 2.8) and to this end also exponentially dependent on the voltage. Consequently, an exponential dependence on the voltage also results for the recombination current density through a pn junction

$$J_{\text{rec}} = J_0 \left[\exp \left(\frac{eV}{n_{\text{id}} k T} \right) - 1 \right], \quad (2.11)$$

where e denotes the elementary charge and J_0 describes the dark saturation current of recombining electron-hole pairs. Equation 2.11 further introduces the ideality factor n_{id} . The ideality factor accounts for different recombination mechanisms (in different areas of the solar cell) with different exponential dependencies on the applied voltage. For SRH recombination in the bulk the ideality factor equals 1. The quasi Fermi level of the majority charge carrier is fixed and the influence of the voltage results in a shift of the quasi Fermi level of the minority charge carrier. In the SCR the ideality factor for the SRH recombination can be derived to 2, since the voltage effects the quasi Fermi level of both charge carriers, leading to the quasi Fermi level of electrons and holes is shifted by half of the amount. Furthermore, most recombination mechanisms are thermally activated. Putting this into context, it is suitable to express J_0 with an activation energy E_a to

$$J_0 = J_{00} \left(\frac{T}{T_0} \right)^3 \left[\exp \left(\frac{-E_a}{n_{\text{id}} k T} \right) - 1 \right], \quad (2.12)$$

where T_0 is a material specific constant often assumed to be in the range of T , enabling to neglect the term $\left(\frac{T}{T_0} \right)^3$. For SRH recombination in the SCR and in the bulk the activation energy equals the band gap energy ($E_a = E_g$).

In addition to radiative and SRH recombination, two important recombination mechanisms are Auger recombination and interface recombination. Auger recombination is a three body process with the additional participation of a hole or an electron, which absorbs the energy of the recombining electron-hole pair

and relaxes to the band edge exciting phonons, i.e. leading to thermalisation of the lattice. Interface recombination refers to recombination processes at the pn interface. Due to band bending or a pn hetero interface, i.e. an interface between a p and a n-type semiconductor from different materials, the local band gap might be reduced. This leads to a decreased activation energy $E_a = \Phi_b < E_g$. With a similar argument as for SRH recombination in the bulk one can derive an ideality factor of $n_{id} = 1$ for interface recombination. Furthermore, tunnelling enhanced recombination [13], thermionic field emission [14], potential fluctuations [15] as well as trap assisted tunnelling [16] may lead to contributions of the recombination current with ideality factors $n_{id} > 2$.

Performance limitations in solar cells

Recombination is unavoidable at temperatures beyond absolute 0 and limits the performance of a solar cell. Another unavoidable loss factor of solar cells arise directly from the band structure of semiconductors. With a given band gap, only photons with an energy higher than the band gap can produce electron-hole pairs and thus be electrically used. On the other hand, if the energy of the absorbed photon is larger than the band gap energy, the hole and the electron relax to the band edge under step-wise excitation of phonons, i.e. resulting in thermalisation of the lattice. Thus, only the band gap energy can be used. If one considers the AM1.5G solar spectrum, a maximum of the theoretically usable energy of the irradiation of the sun of 33.2% results for a band gap of 1.34 eV. This maximum is called Shockley Queisser limit.

In addition to the inevitable loss mechanisms, there are a few other factors that further influence the performance of solar cells. Incomplete absorption due to a finite width of the absorber, resistive losses due to finite mobilities in the semiconductor and contact resistances, a finite parallel resistance as well as parasitic absorption in the packaging material and transmissive layers and reflective losses at interfaces may all contribute to a decreased efficiency.

To measure a solar cells efficiency, illuminated IV characteristics are used. Section 2.3 summarizes the one-diode model often used to describe the shape of an IV characteristic, discusses how various loss mechanisms affect the IV characteristic and the efficiency and introduces models used to describe the shape of IV characteristics as well as equations, that quantify temperature and irradiation dependencies of the IV characteristic.

2.2.2 PV technologies

While c-Si wafer based PV technologies dominate the PV market (95% market share [17]), thin-film technologies like CIGS and CdTe are important technologies regarding the rising market of integrated PV applications (e.g. building and vehicle integrated PV). Their possibility for deposition on flexible substrates open a wide range of flexible and light-weight PV applications. In the following section c-Si, CIGS and CdTe PV technologies are discussed.

c-Si

The combination of high module efficiency and the possibility for high throughput and low cost production makes c-Si PV the obvious choice for conventional PV systems. The most common technology used in c-Si modules is the passivated emitter and rear contact (PERC) cell structure shown in Fig. 2.2.2 (a). The p-type doped c-Si absorber is passivated at the front side utilizing silicon oxide (SiO_Y) and silicon nitride (SiN_X). Passivation layers fulfil the purpose of reduced interface recombination as they reduce the probability of SRH recombination over defect states. Furthermore, the SiO_Y and SiN_X act as anti-reflective layers incorporating more light into the absorber. Near the front surface and near the aluminium contacts a phosphorous diffusion process creates a n^+ respectively n^{++} doped silicon region establishing the pn homo junction. The back side of the PERC solar cell is passivated using an aluminium oxide (Al_2O_3) and SiN_X layer. The area near the aluminium back contact is considered heavily doped (p^{++} region). The advantages of the PERC structure mainly arise due to their well established, reliable and cost-effective production processes. The disadvantage, however, arises from the structure itself. Auger recombination in the p^{++} region and interface defects at the metal semiconductor interface leading to performance limitations. Note, that typically there are much more contacts at the back side ensuring a low series resistance contribution. At the front side all contact areas contribute to I_{SC} losses since no light is incorporated in the solar cell through the contact areas.

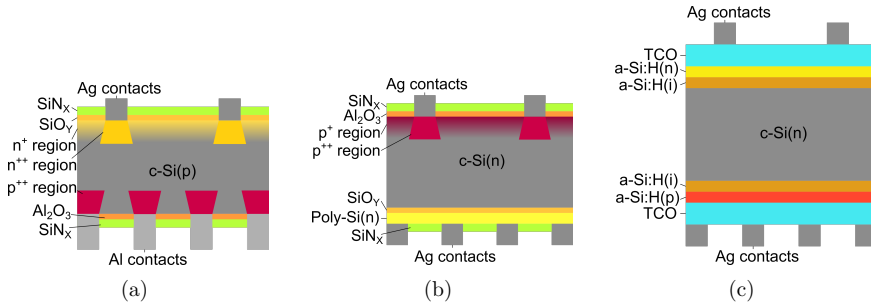


Figure 2.2.2: Typical structure of (a) passivated emitter and rear contact (PERC) (b) tunnel oxide passivated contact (TOPCon) and (c) silicon heterojunction (SHJ) solar cells.

Passivated contact silicon solar cell structures avoid the direct metal semiconductor interface, exhibiting a higher potential in efficiency. The two most prominent passivated contact structures are the tunnel oxide passivated contact (TOPCon) and silicon heterojunction (SHJ) cell structure. The TOPCon solar cell structure is a direct evolution of the PERC cell structure shown in Fig. 2.2.2 (b), implementing contact passivation layers at the back contact. The absorber is n-type doped c-Si. The front contacts are established using the passivating Al_2O_3 and SiN_X layers and a direct contact of aluminium and p^{++} doped c-Si. The back contact is passivated with a SiO_Y layer and an additional poly-crystalline Si layer providing a low contact resistance due to increased charge carrier mobility.

The use of the poly-crystalline Si layer is only possible on the back side due to parasitic absorption of light. Since the TOPCon structure is an evolution from the PERC structure the production processes are fundamentally compatible with implemented production processes. Furthermore, the reduced necessity for big contact areas on the back-side makes the TOPCon solar cell more suitable for the use in bifacial PV modules.

The SHJ solar cell structure is shown in Fig. 2.2.2 (c). The approach is fundamental different from the PERC and TOPCon solar cell structures using amorphous hydrogenated silicon (a-Si:H) layers for passivation and a transparent conducting oxide (TCO) layer to ensure low contact resistances. As absorber the SHJ structure uses n-type c-Si. The contact and passivating layer structure is symmetric on both sides using an intrinsic a-Si:H, a doped a-Si:H (typically n-type at the front and p-type at the back side) and a TCO layer. SHJ solar cells exhibit much less recombination due to the complete avoidance of direct semiconductor metal contacts. To this end, SHJ solar cells exhibit the best V_{OC} potential. Furthermore, they are inherently suitable for bifaciality of PV modules. However, the substantially different manufacturing process, w.r.t. the established PERC solar cell structure, leads to comparably high production cost and parasitic absorption of light in the TCO and a-Si:H layers might lead to a lower I_{SC} potential.

A more detailed summary about the three presented concepts as well as about their advantages and disadvantages can be found in [18].

thin-film PV

Apart of the conventional and emerging c-Si based solar cell technologies, CIGS and CdTe thin-film PV technologies have the largest market share. In contrast to the wafer based technologies thin-film solar cells use direct band gap semiconductor absorbers. The direct band gap leads to much higher absorption coefficients and enables up to two orders of magnitude thinner absorber layers. Thin-film solar cells are deposited either on a sub- or superstrate. The difference is the direction from where the light enters the solar cell. Figure 2.2.3 (a) shows the substrate configuration of a thin-film solar cell. The substrate provides mechanical stability, the back contact layer provides a high electrical conductivity and ensures a good charge extraction, the n- and p-type semiconductor build the pn junction and the front contact ensures a good electrical contact as well as a good incorporation of light. In a superstrate configuration (Fig. 2.2.3 (b)) the conductive and transparent front contact is deposited on the superstrate followed by the solar cell stack (pn junction and passivation layers) and the back contact. Beside the mechanical stability, the superstrate needs to provide a good transparency, since the light enters the solar cell through the superstrate. The depicted cell structure in Fig. 2.2.3 presents the minimum number of layers needed to build a thin-film solar cell. The solar cell stack is depicted by n- and p-type semiconductor, but might consist also of additional buffer layers used for contact passivation. Furthermore, for organic solar cells the absorber itself is not a semiconductor.

Thin film modules consist of several cells connected in series and or parallel. A common structure in commercial thin-film modules is the so called “monolithic interconnection”, where the series connection of multiple solar cells is established

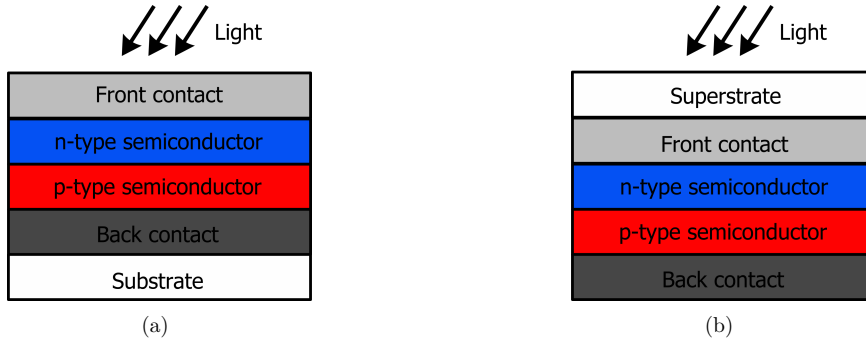


Figure 2.2.3: (a) Substrate and (b) superstrate structure of thin-film solar cells consisting of sub- respectively superstrate, front contact, pn junction and back contact.

via scribing lines. It is called monolithic as it does not involve external wires or soldering. Figure 2.2.4 shows the common P1P2P3 scribing line architecture in a substrate configuration, which can be implemented in the deposition steps of the layers. The scribing line P1 is typically implemented using laser-scribing after the deposition of the back contact on the substrate. In CIGS, the P2 and P3 are usually performed mechanically by a graver after the deposition of the solar cell stack respectively after the deposition of the TCO layer. The P1 and P3 lines separate two adjacent cells, whereas the P2 line connects the front contact of one cell with the back contact of the next cell.

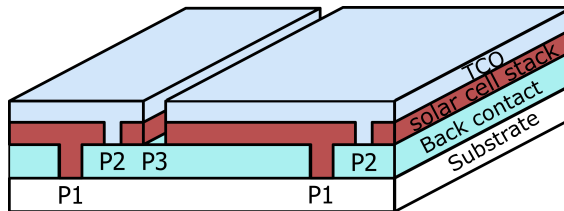


Figure 2.2.4: P1P2P3 scribing line architecture often used in commercial thin-film PV modules. The scribing line architecture is incorporated during the substrate deposition process.

The series connection of multiple cells via scribing lines in thin-film solar cells has the advantage of the possibility to implement the scribing line processes into the deposition steps. This allows for production of large scale PV modules within a comparably simple production line. The implementation of the scribing lines can, on the other side, cause irregularities within thin-film modules. The laser beam of the P1 scribing process can be interrupted, reflected or scattered by irregularities in the substrate or soiling of the surface of the substrate. This can cause scribing line interruptions. While scribing of P2 and P3 mechanical wear can cause unsmooth edges and can also lead to scribing line interruptions. In addition, the deposition of different layers can lead to variations in layer thicknesses and implemented defects. All of these irregularities affect the electronic structure of

parts of the module, leading to a spread in the performance of individual cells. To this end, the already challenging up-scaling of deposition processes from small area cells to large area thin-film modules are additionally impeded. Beside the conventional scribing line architecture presented, the serial connection of several cells in thin-film modules can further be established using a front grid connected to the back contact of the next cell.

In conventional thin-film modules glass is used as substrate and an ethylene vinyl-acetate (EVA) foil encapsulation avoids external influences on the solar cell. Furthermore, a transmissive front glass provides mechanical stability. Typically, the encapsulation foil is approximately 150 μm thick, front and substrate glass are approximately 3 mm thick.

Beside the conventional glass structure of thin-film modules, the deposition of the layers is also possible on flexible substrates. This allows for the application of thin-film PV modules on uneven respectively curved surfaces and in applications, where a light-weight PV module is a necessity. This leads to thin-film PV being a suitable solution for multiple integrated PV applications.

CIGS

Conventional CIGS solar cells consist of five layers which can be deposited on glass or flexible substrates. The back contact is typically made of a 1 μm thick molybdenum (Mo) layer and is highly conductive ($\rho_{\text{Mo}} = 54 \cdot 10^{-9} \Omega\text{m}$). The absorber consists of a typically 2 μm thick p-type copper gallium indium diselenide $\text{Cu}(\text{In}_{1-x}\text{Ga}_x)\text{Se}_2$ layer. Variation of the proportion of gallium (Ga) and indium (In) content allows for an adjustment of the band gap energy. The highest efficiencies of CIGS solar cells are achieved with a share of gallium of $x \approx 0.3$ resulting in a band gap $E_{\text{g, CIGS}} \approx 1.15 \text{ eV}$ [19]. A 50 nm thick layer of cadmium sulphide (CdS) acts as the n-type semiconductor of the pn-hetero junction with a band gap of $E_{\text{g, CdS}} \approx 2.42 \text{ eV}$. A combination of an intrinsic zinc oxide (i-ZnO) layer ($\approx 50 \text{ nm}$) and an aluminium doped zinc oxide (ZnO:Al) ensure a good passivation as well as high transmission and high conductivity ($\rho_{\text{ZnO:Al}} = 9 \cdot 10^{-6} \Omega\text{m}$). The thickness of the ZnO:Al layer depends on the design of the PV module. As already mentioned, conventional thin-film PV modules use the P1P2P3 scribing line architecture, requiring a high conductive respectively thicker TCO layer. With an architecture implementing a front contact metal grid a less conductive, thinner, ZnO:Al layer can be used. Except from the substrate the CIGS solar cell stack is approximately 4 μm thick.

CdTe

Conventional CdTe solar cells consist of five layers deposited on glass or flexible substrates respectively superstrates. An approximately 240 nm thick Indium tin oxide (ITO) and an approximately 30 nm thick tin(IV) oxide (SnO_2) layer act as transmissive front contact. The tin(IV) oxide supports the conductivity and provides essential adhesion and passivation benefits. The pn junction comprises a n-type cadmium sulphide (CdS) layer, typically around $\approx 150 \text{ nm}$ thick, paired with a substantially thicker p-type cadmium telluride (CdTe) absorber layer, generally with a thickness about $\approx 6 \mu\text{m}$. The back contact is established using different materials like copper telluride (Cu_2Te) or tin telluride (Sb_3Te_2). A

more detailed overview of thin-film solar cells used in commercial thin-film PV applications can be found in [20].

2.3 IV characteristics

In this section the theoretical background regarding *IV* respectively *JV* characteristics is given.

2.3.1 The one-diode equivalent circuit model

The current density-voltage respectively current-voltage characteristic (*JV* respectively *IV* characteristic) is from central importance for characterizing solar cells and PV modules. As discussed in Section 2.2.1 the pn junction of a solar cell is a diode. The one-diode equivalent circuit model represented in Fig. 2.3.1 can be used to determine the relationship between voltage and current in a solar cell. It takes the diode character of the pn junction, the generation of electron-hole in the absorber reflected as a current source (J_{Ph}), a series resistance R_{S} and a parallel resistance R_{P} into account.

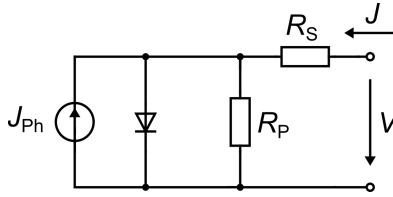


Figure 2.3.1: One-diode equivalent circuit model representing a solar cells equivalent circuit. The pn junction is depicted as a diode and a current source connected in parallel reflects the generated photo current due to absorption of light in the absorber of the solar cell. A finite parallel and shunt resistance reflect resistive losses.

Applying Kirchhoff's circuit laws results in the equation

$$J = J_0 \left[\exp \left(\frac{e(V - JR_{\text{S}})}{n_{\text{id}} kT} \right) - 1 \right] + \frac{V - JR_{\text{S}}}{R_{\text{P}}} - J_{\text{Ph}}, \quad (2.13)$$

where J_{Ph} denotes the photo current density. Numerically, the implicit equation can be solved for given values of R_{S} , R_{P} , J_{Ph} , n_{id} and J_0 . Figure 2.3.2 gives an example for a numerical solution of a *JV* characteristic with the parameters $R_{\text{S}} = 0.5 \Omega \text{cm}^2$, $R_{\text{P}} = 10^5 \Omega \text{cm}^2$, $J_0 = 3.0 \cdot 10^{-9} \text{mA/cm}^2$, $n_{\text{id}} = 1.2$ and $J_{\text{ph}} = 32.0 \text{mA/cm}^2$ at $T = 25^\circ \text{C}$.

The short circuit current density J_{SC} is defined as the current density when the applied voltage to the device is $V = 0$ (i.e. the short circuit condition). For reasonable assumptions of series and parallel resistances in functional solar cells, J_{SC} is directly given by J_{Ph} . One can interpret J_{SC} as a measure of the amount of generation of electron-hole pairs. Note, that J_{Ph} is approximated to be voltage independent in the one-diode model. In real solar cells a slight decrease of the photo current with increasing system voltage is often observed.

The open circuit voltage V_{OC} is defined as the voltage where the current through the device is $J = 0$ (i.e. the open circuit condition). At V_{OC} the recombination current through the diode needs to compensate the photo current. Consequently,

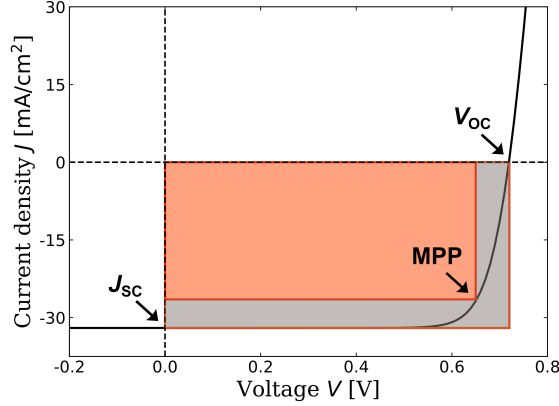


Figure 2.3.2: Numerical solution of the JV characteristic of the one-diode equivalent circuit model using the parameters $R_S = 0.5 \Omega \text{cm}^2$, $R_P = 10^5 \Omega \text{cm}^2$, $J_0 = 3.0 \cdot 10^{-9} \text{mA/cm}^2$, $n_{id} = 1.2$ and $J_{ph} = 32.0 \text{mA/cm}^2$ at $T = 25^\circ\text{C}$. J_{SC} , V_{OC} as well as MPP are marked and the orange and grey areas give a visual interpretation of the definition of the FF in 2.14.

the open circuit voltage depends on the photo current density J_{ph} , on the dark saturation current density J_0 and on the diode ideality factor n_{id} . J_0 and n_{id} reflect the combined influence of different recombination processes on the recombination current through the pn junction. Such, one can interpret V_{OC} as a measure for the amount of recombination in the solar cell.

Beside J_{SC} and V_{OC} the maximum power point (MPP) is marked in Fig. 2.3.2. The MPP describes the point of the maximum power output P_{MPP} of the solar cell and thus in an ideal case also its operating point. Another important parameter of solar cells regarding the performance is the fill factor FF. It is given by the quotient of P_{MPP} and the product of the V_{OC} and the short circuit current I_{SC} to

$$\text{FF} = \frac{P_{MPP}}{V_{OC}I_{SC}} = \frac{V_{MPP}J_{MPP}}{V_{OC}J_{SC}}. \quad (2.14)$$

The orange and grey areas in Fig. 2.3.2 can be seen as a visual interpretation of the FF. The orange area is equal to the maximum power density $p_{MPP} = P_{MPP}/\text{Area}$ and orange and grey area together are equal to the product $V_{OC}J_{SC} = V_{OC}I_{SC}/\text{Area}$. Illustratively, the FF equals the ratio of the small rectangle to the big rectangle.

The efficiency is the most important performance measure of a solar cell or PV module and is given via the quotient of the power output at MPP and the power input of the irradiation on the solar cells or PV modules area:

$$\eta = \frac{P_{MPP}}{P_{in}} = \frac{V_{MPP}I_{MPP}}{P_{in}} = \frac{\text{FF}V_{OC}I_{SC}}{P_{in}}. \quad (2.15)$$

For the determination of P_{MPP} and, with this, for the determination of the efficiency of solar cells and PV modules illuminated JV respectively IV characteristics need to be measured. In order to guarantee a good comparability of

illuminated *IV* characteristic measurements under laboratory conditions, the AM1.5G solar spectrum is often used and simulated with prescribed accuracy in sun simulators. The AM1.5G solar spectrum is a defined spectrum of sunlight passing through 1.5 earth atmospheres. Varying the cell temperature and irradiation intensity depicts further insights on a solar cells performance, since different performance limiting mechanisms might be dominant under different conditions. Section 2.4.1 summarizes the specifications of the sun simulator used in this thesis to validate a solar cells performance in terms of illuminated *IV* characteristics. Beside measurements under defined laboratory conditions, *IV* characteristics are often used to verify a PV modules performance in field under varying outdoor conditions. Environmental data, like the Plane Of Array (POA) irradiation, G_{POA} , and module temperature, T_{Mod} , is required for evaluation of the measurements.

In addition to the determination of the efficiency, the measurement of the illuminated *JV* characteristic provides more detailed information on a solar cells or PV modules performance. The influence of various effects on the shape of the *JV* characteristic are not always intuitive. Using the numerical solution of the one-diode model presented in Fig. 2.3.2, the variation of the parameters gain insight how, f.e. the series and parallel resistance as well as the recombination affects the shape of an *JV* characteristic.

Figure 2.3.3 shows the numerical solution of the *JV* characteristic varying the parallel resistance. Low parallel resistances show the effect of shunt paths in the solar cell.

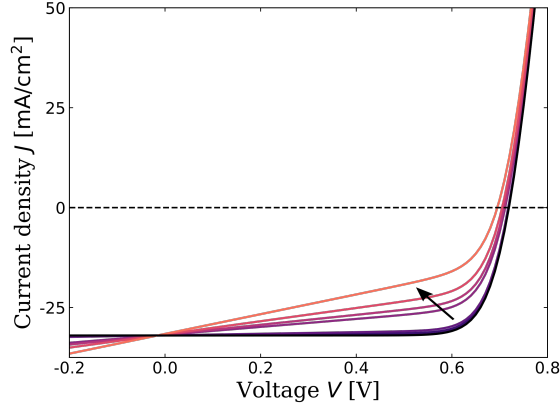


Figure 2.3.3: Numerical solution of the *JV* characteristic for varying parallel ($R_P = 40, 60, 80, 100, 500, 10^3, 10^4, 10^5 \Omega\text{cm}^2$) resistance. The black line shows the solution for the parameters in Fig. 2.3.2. The arrows show into the direction of decreasing parallel resistance and, with this, in the direction of decreasing performance.

Figure 2.3.4 shows the numerical solution of the *JV* characteristic varying the series resistance. High series resistances are usually connected to issues regarding the contacts or high sheet resistances of a TCO layer.

Figure 2.3.5 shows the numerical solution of the one-diode model for varying dark saturation current densities J_0 . Since the diode ideality factor n_{id} is kept

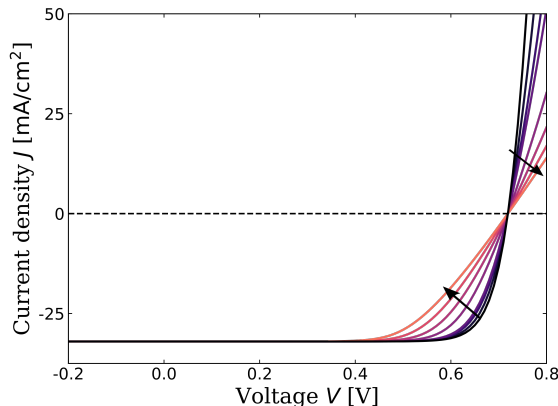


Figure 2.3.4: Numerical solution of the JV characteristic for varying series ($R_s = 0.2 - 5 \Omega\text{cm}^2$) resistance. The black line shows the solution for the parameters in Fig. 2.3.2. The arrows show into the direction of increasing series resistance and, with this, in the direction of decreasing performance.

constant this variation reflects increasing/ decreasing recombination rates, while the recombination mechanism remains the same.

The one-diode model represents, all recombination processes with only a single diode ideality factor and dark saturation current density, i.e. with one diode. In real solar cells many recombination mechanisms might contribute as discussed in Section 2.2.1. Furthermore, which recombination mechanism is dominant and to this end limiting the solar cells performance might be voltage dependent. One way to overcome the limitations of describing the pn junctions recombination with one diode is the usage of the two-diode model, introducing a second recombination current contribution with differing diode ideality factor and dark saturation current density in parallel. Fitting the one- as well as the two-diode model to a measured IV characteristic comes with implications, that are further discussed in Section 2.3.3

Beside illuminated IV characteristics, dark IV characteristics can provide valuable information about a solar cells performance. Since only a possibility to darken the sample and a source measurement unit are necessary, dark IV measurements are less complex to perform. Since the illuminated IV characteristic is not necessarily just the superposition of photo current and dark IV characteristic, the dark IV characteristic is not sufficient to predict the efficiency of a solar cell. Nevertheless, dark IV measurements are often used to evaluate the device's recombination properties, series, and parallel (shunt) resistances.

IV respectively JV characteristics yield much information on basic properties of solar cells. However, if several solar cells are measured in series and/ or parallel connection as it is the case for a PV module, the interpretation of various influences can become more complex. The result of such a measurement is in the easiest case the superposition of several IV characteristics. In general the solar cells can affect each others performances. Since the same current flows through all cells connected in series, cells providing a high current might be limited by cells exhibiting a reduced current.

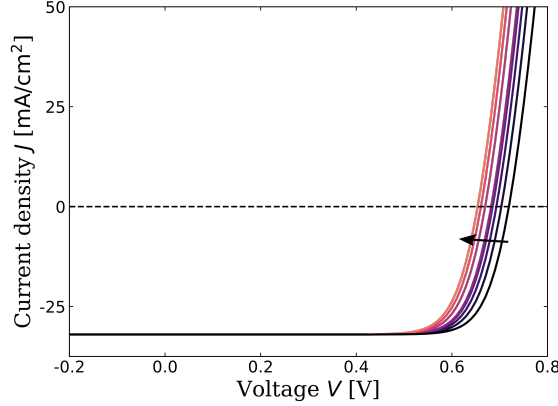


Figure 2.3.5: Numerical solution of the JV characteristic of the one-diode equivalent circuit model with varied dark saturation current density between $J_0 = 3.0 - 25.0 \cdot 10^{-9} \text{ mA/cm}^2$. The black line shows the solution for the parameters in Fig. 2.3.2. The arrow points in the direction of increasing dark saturation current density and, with this, increasing recombination rates.

2.3.2 Temperature and Irradiation dependencies according to the IEC60891

The irradiation and temperature dependencies of IV characteristics vary throughout different technologies and producers. Nevertheless, some similarities can be observed. To describe the irradiation and temperature dependencies of various PV modules and solar cells, many methods are proposed in literature. The International Electrotechnical Commission (IEC) published the IEC60891 norm, which provides several standardized procedures to translate IV characteristics from one condition to another. In this work, only procedure 2 of the IEC60891 norm is used. The IEC60891:2009 norm [21] describes procedure 2 as:

$$I_2 = I_1 (1 + \alpha(T_2 - T_1)) \frac{G_2}{G_1}, \quad (2.16)$$

$$V_2 = V_1 + V_{OC1}(\beta(T_2 - T_1) + a \ln(\frac{G_2}{G_1})) - R_s(I_2 - I_1) - \kappa I_2(T_2 - T_1), \quad (2.17)$$

where current I_1 and voltage V_1 are a current and voltage pair from the IV characteristic at temperature T_1 and irradiation G_1 . The current-voltage pair is translated to current I_2 and voltage V_2 at temperature T_2 and irradiation G_2 . The procedure uses 5 coefficients:

- α The temperature coefficient of current [K^{-1}]
- β The temperature coefficient of the voltage [K^{-1}]
- a The irradiation correction factor of the voltage [V]
- R_s The series resistance of the device [Ω]

- κ The temperature coefficient of R_s [ΩK^{-1}]

By translating the points in an IV curve at T_1 and G_1 one by one to new current-voltage coordinates a translated IV characteristics is obtained corresponding to temperature T_2 and irradiation G_2 .

In 2021 the IEC60891 norm was updated and procedure 2 was adjusted. The transformation equations of procedure 2 from the IEC60891:2021 norm [22] are

$$I_2 = I_1 \frac{G_2}{G_1} \frac{1 + \alpha(T_2 - 25^\circ C)}{1 + \alpha(T_1 - 25^\circ C)}, \quad (2.18)$$

$$\begin{aligned} V_2 = & V_1 - R_{S1}(I_2 - I_1) - \kappa I_2(T_2 - T_1) \\ & + V_{OC,STC} \beta [f(G_2)(T_2 - 25^\circ C) - f(G_1)(T_1 - 25^\circ C)] \\ & + V_{OC,STC} \beta \left[\frac{1}{f(G_2)} - \frac{1}{f(G_1)} \right], \end{aligned} \quad (2.19)$$

where

$$f(G) = B_2 \ln^2 \left(\frac{1000 \text{ W m}^{-2}}{G} \right) + B_1 \ln \left(\frac{1000 \text{ W m}^{-2}}{G} \right) + 1, \quad (2.20)$$

$$R_{S1} = R_S + \kappa(T_1 - 25^\circ C), \quad (2.21)$$

where the correction factor a from the previous version was substituted by a non-linear description of the irradiation dependency of the voltage with 2 correction coefficients:

- B_1 linear irradiation correction factor of the voltage [-]
- B_2 quad. irradiation correction factor of the voltage [-]

Beside considering the non-linear irradiation dependency of the voltage, the updated procedure 2 avoids inconsistencies regarding back and forth translation of two different conditions. While for the 2009 version the result of translating $IV(G_1, T_1) \rightarrow IV(G_2, T_2) \rightarrow \tilde{IV}(G_1, T_1)$ is in general not consistent, the updated version guarantees that $IV(G_1, T_1) = \tilde{IV}(G_1, T_1)$. Furthermore, the updated procedure is more referenced to STC conditions. This increases the comparability, since a PV module's performance is in general verified at these standardized conditions.

The performance of procedure 2 of the IEC60891:2021 norm is verified in literature [23]. The compact description and the verified good performance of describing temperature and irradiation dependencies leads to widespread use of procedure 2 of the IEC60891 norm in the PV community.

2.3.3 Describing the shape of (measured) IV characteristics

This section introduces to concepts used to describe the shape of (illuminated) IV characteristics. A (measured) IV characteristic consists of several current-voltage (I/V) pairs. As for every measurement, all I/V pairs are subject to a certain degree of measurement uncertainty or noise level. For laboratory 4 quadrant measurements, performed under stable irradiation conditions with

suitable source measure units (SMUs), the signal to noise ratio is usually small and does not impede the data analysis. A common problem, however, is the solar cells, respectively PV modules temperature during a measurement. PV modules and solar cells heat up during illumination, such that the cell respectively module temperature might differ at the start and beginning of a measurement. As discussed in Section 2.3.2, the temperature has a noticeable influence on the *IV* characteristic and temperature differences might need to be considered.

For outdoor measurements the data quality can be influenced by several factors. Since the conditions are not controlled, unstable irradiation conditions during the measurement, partial shading, unsteady module temperature and wind conditions can influence the measurement. Outdoor data is often acquired with programmable loads, leading to incomplete *IV* characteristics, i.e. such equipment can only measure the active range of the PV module and thus the short and open circuit points are not directly measured and cannot be interpolated from the measurement either. Furthermore, malfunction of a SMU or programmable load might cause an incomplete *IV* characteristic, where parts of the current or voltage range are missing.

Beside implications arising due to incomplete or noisy measured *IV* characteristics, the serial and parallel connection as well as the use of bypass diodes in conventional PV modules might impede fitting procedures. When individual cells or (sub-) strings inside of a PV module or array provide considerably less current as the remaining cells or (sub-) strings, bypass diodes maximize the power output. The *IV* characteristic, however, looks irregular in these situations, i.e. a "S-shape" is often observed.

The standard solar cell parameters (SSPs)

The most widely used feature vector of describing an *IV* characteristic are the standard solar cell parameters (SSPs). V_{OC} , I_{SC} , V_{MPP} and I_{MPP} describe three key points of the *IV* characteristic, namely the open and short circuit as well as the maximum power point. While the SSPs depict valuable and the most important information an *IV* characteristic, the reduction of the complete *IV* characteristic to 4 parameters implies information loss. The *IV* characteristic cannot be adequately reconstructed using solely the SSPs and f.e. an *IV* characteristic with "S-shape" and an *IV* characteristic with strong series resistance are undistinguishable using only the SSPs.

To extract the SSPs from a measured *IV* characteristic a suitable interpolation approach as well as a noise robust approach to extrapolate the *IV* to open respectively short circuit might be needed. Nevertheless, extracting the SSPs is in any case the first step of analysing an *IV* characteristic, since using these easy and standardized parameters allows for meaningful interpretation and comparison of solar cells and PV modules performance.

One- and two-diode model fits

Fitting the one-diode model (see Section 2.3.1) or the two-diode model are common approaches to extract features of an *IV* characteristic [24–26]. Their advantage is the representation of the *IV* characteristic in terms of physically meaningful parameters. However, the parametrization of the two equations is

quite challenging, which is reflected by a vast number of papers to this subject. A detailed overview of proposed methods is given by the review papers in [24–26]. The one-diode model introduces five (free) parameters to describe the *IV* characteristic. In comparison to the SSPs, the one-diode model can track more information, inherited by the shape of the *IV* characteristic. The two major downsides of using the one-diode model equation to describe an *IV* characteristic are the challenging parametrization as well as the correlation of the diode model parameters. A small change of the *IV* might result in large differences in the parameters, if the parametrization procedure is ill-conditioned. Furthermore, the one-diode model is limited to *IV* shapes, that can be described by a diode, i.e. "S-Shapes" can f.e. not be described.

The diode ideality factor in the one-diode model can be seen as an effective ideality factor arising from the joint influence of different recombination processes. In the two-diode model the two different ideality factors are introduced, by adding a second diode character describing the recombination current. The two ideality factors are often fixed to $n_{id,1} = 1$ and $n_{id,2} = 2$, since most recombination processes can physically be described with an ideality of 1 respectively 2. The dark saturation currents $I_{0,1}$ and $I_{0,2}$ then reflect the contributions of the respective mechanisms to the overall recombination current. Fixing the two ideality factors, however, excludes recombination mechanisms with $n_{id} \neq 1, 2$. To this end also two-diode models with free ideality factors and dark saturation currents are used in literature describing the *IV* characteristic with 7 free parameters. On the first look, this seems to be an advantage over the one-diode model and the two-diode model with fixed ideality factors, since the model is capable to describe a more complex diode character often present in real solar cells and PV modules. On the second look, adding 2 parameters to describe the diode character in more detail extends the problems of the correlated parameter space and the tendency of ill-conditioned parameterization procedures as mentioned for the one-diode model. Often, an *IV* characteristic can be adequately described using multiple combinations of the two dark saturation currents, $I_{0,1}$ and $I_{0,2}$, and the two diode ideality factors, $n_{id,1}$ and $n_{id,2}$. Therefore, an optimization procedure to find the best parameterization might face the challenge of many local minima in the parameter space and the result of the parameterization is less meaningful.

The ESP model

One- and two-diode model are based on device physics and their fit to an *IV* characteristic can result in meaningful parameters, gaining a better understanding of the PV modules or solar cells performance. The physics based approach leads, on the other side, to the restriction, that the measured *IV* characteristic can be described with the fundamental physical assumptions of the model. More complex *IV* shapes, that originate from the interplay of many cells in a module, cannot be described adequately. Furthermore, in the presence of noisy datasets the parameterization is further impeded. Another approach to track the information of an *IV* characteristic is a descriptive approach.

Pieters introduces a descriptive model for the complete shape of the *IV* characteristic based on the extension of the SSPs in terms of extended solar cell parameters (ESPs) [27]. The ESPs extend the SSPs with the slopes of the *IV* at

short and open circuit (G_{SC} and R_{OC})⁶ as well as with two additional key points (‘upper quasi maximum power point’ [I_{qmp+} , V_{qmp+}] and ‘lower quasi maximum power point’ [I_{qmp-} , V_{qmp-}]) to comprehensively describe the complete shape of the IV characteristic.

The lower and upper quasi maximum power points incorporate information on the shape of the IV characteristic between short circuit (SC) and maximum power point (MPP) and between MPP and open circuit (OC), respectively. These points are defined using the power curve ($P(V)$) associated with the IV . V_{qmp+} is defined as the voltage between MPP and OC, where the power curve exhibits the largest difference from a straight line drawn between the two coordinates [V_{MPP}, P_{MPP}] and [$V_{OC}, 0$]. I_{qmp+} is then the associated current of the IV at $V = V_{qmp+}$. Analogously, V_{qmp-} is defined as the voltage, where the power curve exhibits the largest difference from a straight line drawn between the two coordinates [$0, I_{SC}$] and [V_{MPP}, P_{MPP}] and I_{qmp-} is the associated current at $V = V_{qmp-}$. Thus, the ESPs reduce the shape of the IV characteristic to 10 parameters.

The choice of the 2 key points has benefits regarding the additional information, that they inherit as well as benefits regarding the parameterization of the model. Compared to other parametrization routines, the ESPs result in less information loss, since also more subtle features of the IV characteristic are tracked. By reconstructing the IV characteristics Pieters demonstrated on a large dataset of 2.2 million IV s, covering different PV technologies, that the ESPs provide an accurate description of the IV showing for 90% of the IV characteristics an root mean squared error (RMSE) w.r.t. the original measurements below 0.2%. For comparison, for the one diode model this is about 1% and for the Karmalkar-Haneefa model [29] 1.2% [27].

Another benefit, is that the 10 ESPs describe 5 key points on the IV as well as their slopes. Thus, also the irradiation and temperature corrections discussed in Section 2.3.2 can be applied to directly translate the ESPs to different conditions. More detailed explanations to the ESPs can be found in [27].

The IV Bézier model

Pieters introduces various interpolation models to reconstruct an IV characteristics from the ESPs. A particularly effective model is the so called “PV Bézier model”, based on a cubic Bézier interpolation[27]. Pieters describes a method to parameterize the PV Bézier model directly from the ESP. Note that for Bézier curves robust and efficient fitting algorithms exists, see for example [30]. Here I introduce the PV Bézier model and its parameters as an alternative representation of the ESPs. The PV Bézier parameters also entail the SSPs, V_{OC} , I_{SC} , V_{MPP} and I_{MPP} .

The model uses cubic Bézier curves - parametric curves between a start coordinate P_0 and an end coordinate P_3 . In between the start and end coordinate the shape is determined by two control points P_1 and P_2 . A cubic Bézier curve can be

⁶Savitzki-Golay filters are used to obtain noise robust estimators for the slopes at short and open circuit [28].

written as:

$$B(t) = (1 - t)^3 P_0 + 3t(1 - t)^2 P_1 + 3t^2(1 - t) P_2 + t^3 P_3, \quad (2.22)$$

where the time t ranges from 0 to 1 and $B(t)$ is the coordinate at time t . At $t = 0$ it results, that $B(t)|_{t=0} = P_0$ and at $t = 1$ it results, that $B(t)|_{t=1} = P_3$. $B(t)$ can be seen as a trajectory, where at each point the curve has a derivative and a velocity w.r.t. the time variable t . The control points P_1 and P_2 encode the slopes at the start and end coordinates, which becomes markable, considering the derivative of the trajectory

$$B'(t) = 3(1 - t)^2(P_1 - P_0) + 6t(1 - t)(P_2 - P_1) + 3t^2(P_3 - P_2). \quad (2.23)$$

At the start point P_0 it follows, that $B'(t)|_{t=0} = 3(P_1 - P_0)$ and at the end point P_3 one can derive $B'(t)|_{t=1} = 3(P_3 - P_2)$. The trajectory in the start point (end point), thus follows the line segment $P_0 - P_1$ ($P_2 - P_3$) and the velocity is proportional to the length of the line segment.

To describe a complete *IV* characteristic with cubic Bézier curves the *IV* characteristic is divided in two sections, one from SC to MPP and one from MPP to OC. The model adds to the SSPs a differential short circuit point $[V_{\text{dsc}}, I_{\text{dsc}}]$, which encodes the short circuit slope and effectively how long this slope dominates the shape of the *IV* characteristics. Similarly the model has a differential open circuit point $[V_{\text{doc}}, I_{\text{doc}}]$. Since the slope at MPP equals $-I_{\text{mpp}}/V_{\text{mpp}}$, the analogously defined differential MPP point $[V_{\text{dmpp}}, I_{\text{dmpp}}]$ effectively describes how long the slope at MPP dominates the *IV* characteristic in both directions. Thus, the PV Bézier parameters also reduce the shape of the *IV* characteristic to 10 parameters. As for the ESPs, the performance of the model tested on *IV* characteristics of different module types operated in different climates is discussed in [27].

The PV-CRAZE library

To analyze *IV* curves throughout this thesis I use the open source PV-CRAZE library [31]. The PV-CRAZE library implements, beside others, a fast and noise-robust parameterization of the ESP model, the one-diode model as well as of the PV Bézier model. Furthermore, as a part of the extraction of the ESPs, PV-CRAZE detects various problems with *IV* characteristics. In particular it will flag data-points when

- it is not at all possible to determine all ESPs (e.g. large parts of an *IV* are missing)
- it is not possible to reliably determine the ESPs, (e.g. too few data to make a reliable fit)
- the *IV* has unexpected features (e.g. multiple MPPs, or non-monotonous *IV* characteristics)

Throughout the thesis, I use the library as a first filtering step for *IV* data, excluding *IV*s that are flagged as unreliable. Furthermore, the PV-CRAZE library is used for fitting of the one-diode model to *IV* characteristics. The

fitting procedures for the diode model are rather involved as it is notoriously hard to obtain a robust fitting algorithm [27]. As some of the used optimization algorithms in PV-CRAZE are non-deterministic, and fitting with PV-CRAZE is generally fast, I used multiple fits for each data-point to select the best fit. More details about the library, its fitting performance as well as the processing time for large data sets can be found in in [27].

2.4 Characterization methods and set-ups

In this section the main characterization methods used to verify solar cell's performance as well as the used set-ups are discussed.

2.4.1 Sun simulator set-up

As introduced in Section 2.2.1 sun simulators are one of the most important characterization tools for solar cells and PV modules. For the characterization of solar cells with illuminated and dark *IV* characteristics I use a steady state WACOM WXS-140S-Super sun simulator with class A accuracy of the AM1.5G solar spectrum. The sun simulator uses a tungsten and a xenon lamp to reproduce the AM1.5G solar spectrum. Various optics ensure a steady and homogeneous irradiation on an area of $10 \times 10 \text{ cm}^2$ with inhomogeneities of less than 2%. The cell temperature control is ensured via an external water cooling system, that can be connected to various cell holders. Measurements at lower irradiation intensities can be performed using neutral density filters.

2.4.2 Electro- and Photoluminescence set-up

With the same principle as light emitting diodes, injected charge carriers can cause radiative recombination within a solar cell leading to light emission. Electrically induced emission of photons is called electroluminescence (EL). When the charge carriers are not injected due to an applied voltage but created due to the absorption of light, the emission of radiation is called photoluminescence (PL).

The emission signal can be detected with a camera, whose detector is sensitive in the wavelength region around the band gap energy. The measured EL intensity increases exponentially with the local voltage, since the recombination current is proportional to $\exp(\frac{eV}{kT})$. The measured PL intensity increases linearly with the illumination intensity applied.

EL and PL measurements can be used to verify local differences in a solar cells or PV modules performance. In EL images, inactive and poorly contacted areas as well as shunts show up as dark areas respectively dark spots. Due to the difference of charge carrier origin, variations in series resistances over the solar cell do not effect the luminescence intensity in PL images. Furthermore, the signal strength is dependent on the non-radiative recombination probability of charge carries, since non-radiative recombination does not contribute to the EL respectively PL signal. EL imaging, PL imaging and especially the combination of both imaging techniques can yield valuable insights on a solar cells or PV modules performance.

In this thesis all EL and PL measurements are performed on the same set-up. The set-up uses a Princeton instruments Nirvana 640 Indium Gallium Arsenide (InGaAs) camera with a 1050 nm band-pass filter with a full width half maximum of 30 nm to measure the luminescence signal. The EL images are acquired using a Keithley 238 high current SMU. The camera position can be changed with an Isel X-Y moving stage, enabling EL image acquisition for DUTs of a size up to $32 \times 25 \text{ cm}^2$.

The PL images are acquired using a 808 nm laser as a light source and a beam homogenizer ensuring a homogeneity of $\pm 2.5\%$ on a $12 \times 12 \text{ cm}^2$ plane and a homogeneity of $+2.5\% - 6.5\%$ on a $17 \times 17 \text{ cm}^2$ plane. The light intensity is calibrated using a silicon hetero junction solar cell to an effective one sun equivalent irradiation in the measurement plane.

2.4.3 Light induced degradation set-up

For the laboratory experiments discussed in Chapter 4 I use a light induced degradation set-up. The three main components of the set-up are a light source, three water cooling systems and a SMU. This way the set-up allows to study the effect of light and temperature on the long-term performance of multiple device under test (DUT), verified with in-situ *IV* measurements. As light source a metal halide lamp HMI 4000 W/DXS - Osram 6000K is used. A long pass filter of 390 nm blocks ultraviolet (UV) light and ensures no overlaid effect of UV degradation. The set-up is shielded from external light sources and the irradiation level can be tracked with a silicon photodiode in the middle of the illuminated area. The homogeneity of the irradiation can be verified in the measurement plane using two reference cells (PRC Krochmann GmbH PRC911214/1 and PRC911214/2 with a BG17 filter of a size by $2 \text{ cm} \times 2 \text{ cm}$). The two reference cells measure the total as well as the filtered short wavelength intensity.

The DUT temperature can be controlled to three different temperatures with the three water cooling systems. Furthermore, within the set-up, the DUTs are thermally decoupled from the set-up, ensuring a well defined temperature. The temperature of the refrigerated/heated water systems as well as the room temperature can be tracked with PT100 temperature sensors. Furthermore, the working point of the DUT can be set to OC or SC in between *IV* measurements using relays.

2.5 Outdoor Performance

This section explores prevalent patterns in PV outdoor data and introduces the specific datasets utilized in this study. It delves into crucial methodologies for modelling PV module temperature and in-plane irradiation, fundamental for understanding module performance under varying environmental conditions. Additionally, I provide an overview of the approach to assessing PV module outdoor performance as incorporated in the IEC61853 standard. This standard outlines procedures for rating the performance of PV modules in differing climates and operational contexts, offering a comprehensive framework relevant to our analysis.

2.5.1 Outdoor data

PV outdoor data consists in general of performance measures of PV modules or strings and additional meteorological measurements, that set the performance measure into the context of the condition the performance is measured at.

Meteorological data

One part of the meteorological data are irradiation measurements. The most important irradiation measure is the point of array irradiation G_{POA} measuring the irradiation in the plane of the installed PV module. In addition global horizontal irradiation (GHI), diffuse horizontal irradiation (DHI) and direct normal irradiation (DNI) can provide additional information about the effective irradiation reaching a PV module. GHI can further be determined from DNI and DHI using

$$\text{GHI} = \text{DNI} \cos \Theta + \text{DHI}, \quad (2.24)$$

where Θ is the zenith angle. With sufficient knowledge about the topology of the surrounding or the assumption of no topology causing shading, G_{POA} can be modelled using GHI, DHI, location, elevation and time. The used modelling approach in this thesis is discussed in Section 2.5.3. In addition the spectrum of the respective irradiation measurements or the average photon energy can provide more information on the measured irradiation parameters.

The most important temperature measure is the module temperature T_{Mod} as it directly relates to a PV modules performance. In addition the ambient air temperature T_{amb} , wind speed v and wind direction as well as rainfall and humidity can provide valuable insights about the operating conditions and might be used to predict T_{Mod} in case the parameter is not available.

Electrical performance measures

Performance measures of PV modules and arrays vary significantly regarding their complexity. While for PV arrays often only the AC (alternating current) or DC (direct current) power is tracked, some systems track AC respectively DC voltage and current. The DC performance reflects the direct output of the PV system respectively string and is a direct performance measure of the system respectively string, the AC performance incorporates inverter losses and is more

relevant evaluating the complete system, rather than rating the PV modules solely.

As discussed in Section 2.3, *IV* characteristics are the most important and most frequently used performance measure of solar cells and PV modules. Since an *IV* measurement takes time and require suitable equipment, regular *IV* measurements are not common for monitoring commercially operated PV systems. Regular *IV* measurements of operated PV modules can, however, yield valuable insights about performance degradation and seasonality and are often subject of research projects observing the PV module performance over time.

Common challenges

Before processing, the acquired data quality needs to be evaluated. Beside missing data and repetitive data, common issues are sensor alignment, sensor shading, sensor drift, data and time shift, inverter clipping calibration problems, data synchronization and data aggregation. A good summary of common data quality issues is given by Lindig *et al.* [32]. After the evaluation of the data quality, filtering approaches are applied, to exclude systematic errors and reduce the dataset to e.g. clear-sky instances or a range of in-plane irradiation or temperature conditions. Data filtering approaches are further discussed in Chapter 5, where I further present two proposed filtering approaches.

Imaging techniques

Further performance measures for PV modules and systems in operation are imaging techniques often used to monitor large scale PV power plants. The decreased performance of single modules affects the whole string and, such, the performance of the complete system. Electrical performance measures like the string output power provide insight on the strings performance, but are not suitable to rate the performance of individual PV modules being part of the string. Infrared thermography (IRT), EL and PL as well as ultraviolet fluorescence (UVF) imaging can be used to identify individual malfunctioning PV modules.

A good overview of common imaging techniques used for inspection and monitoring of large scale PV systems is given by Høiaas *et al.* [33]. EL and PL are already discussed in Section 2.4.2. In field, PL images can be acquired using an external light source, i.e. using a laser or light emitting diodes (LEDs) or the sun light, disconnecting a string or individual modules to open circuit. EL image acquisition in field requires a portable power supply and is often done at night to ensure a better signal to noise ratio. Furthermore a high signal to noise ratio for measurements during the day can be achieved using lock-in techniques. IRT is possible without disconnecting any modules or strings and is the most common imaging technique used for PV system monitoring. Disconnected modules or cells, bypassed cells or substrings inside a module as well as shunted, delaminated or otherwise defected cells and parts of the module can be identified using infrared thermography. UVF imaging is an emerging technique in PV system monitoring. PV modules are excited using UV light and the resulting fluorescence is usually capture common RGB cameras. UVF imaging can detect cracks in single cells, not recognizable with optical inspection and IRT. Such,

UVF imaging can help to detect less pronounced causes for power reduction within a PV system.

TÜV data

Among the used data in this work I use a subset of PV outdoor data acquired by TÜV Rheinland. Apart from the Thuwal dataset (which was not yet available at that time, but is acquired with the same module types and measurement devices) the data is published in [5]. For this dataset periodic IV measurements were performed on industrial produced PV modules of different technologies at 5 different test locations in different climate zones over up to nearly three years. The operated modules cover an amorphous/microcrystalline silicon tandem module, a conventional PERC silicon module, three CdTe modules and four CIGS modules from different producers. Note that the data was collected approximately ten years ago and therefore reflects the PV technologies that were commercially available at that time. Similar modules of all 9 types were operated in Ancona (Italy), Phoenix (USA), Cologne (Germany), Chennai (India) and Thuwal (Saudi Arabia). Depending on the module type and the location the operation period varies between 17 and 35 months resulting in dataset sizes between ≈ 29.000 and ≈ 65.000 IV characteristics. The individual size of the datasets is listed in Tab. 5.2. In total the datasets consist of approximately 2.2 million individual IV characteristics. In addition I use the simultaneously measured G_{POA} and T_{Mod} . The IV characteristics are measured using programmable loads in intervals of ten minutes. The module temperature T_{Mod} reflects the average of two measurements with Pt100 temperature sensors mounted on the back of each PV module. The irradiation G_{POA} is measured in plane by a ventilated pyranometer.

NREL data

As a second PV outdoor dataset I use a subset of data provided by the National Renewable Energy Laboratory (NREL) described in [6]. From the dataset I use data from 11 different modules, operated for 13 months in Cocoa, Florida, and another 13 months in Eugene, Oregon. The different modules cover various thin-film as well as wafer based technologies (mono-crystalline Si, multi-crystalline Si, CdTe, CIGS, amorphous/crystalline Si heterojunction, amorphous/microcrystalline Si tandem, amorphous Si tandem and amorphous Si triple junction). The dataset sizes vary between ≈ 34.000 and ≈ 43.000 IV characteristics. The individual size of the datasets is listed in Tab. 5.1 summing up to a total of ≈ 880.000 individual IV characteristics. In this work I only use the SSPs rather than the complete IV as well as the simultaneously to the IV measured G_{POA} and T_{Mod} . The dataset is publicly available enabling, that the performance of the filtering approach presented in Section 5.3 can be compared with other approaches aiming for a similar outcome.

2.5.2 Performance ratio analysis and degradation estimation

Crucial for the rating of a PV system is the performance respectively the performance over time. A common way to depict the performance is the

performance ratio (PR) defined in the IEC61724 norm [34]. For the AC/DC output performance the PR is given by

$$\text{PR}_{\text{AC/DC}} = \frac{E_{\text{AC/DC}}/P_{\text{nom}}}{H_{\text{POA}}/G_{\text{STC}}} = \frac{E_{\text{AC/DC}}G_{\text{STC}}}{H_{\text{POA}}P_{\text{nom}}}, \quad (2.25)$$

where P_{nom} is the nominal AC respectively DC output power at STC conditions, $G_{\text{STC}} = 1000 \text{ Wm}^{-2}$ is the STC irradiation and H_{POA} is the normalized integrated in-plane irradiation for the considered time span. The PR is defined on a time interval, but is often generalized for the DC output power to a discrete time using

$$\text{PR} = \frac{P_{\text{DC}}G_{\text{STC}}}{G_{\text{POA}}P_{\text{STC}}}, \quad (2.26)$$

where I substituted $P_{\text{nom}} = P_{\text{STC}}$ for the DC situation. As discussed in Section 2.3.2 the performance of a PV module and, thus, the performance of a string of modules or a PV system is dependent on the module temperature T_{Mod} . The temperature corrected performance ratio PR_T includes the temperature dependency of the output power using

$$\text{PR}_T = \frac{P_{\text{DC}}G_{\text{STC}}}{G_{\text{POA}}P_{\text{STC}}(1 + \gamma(T_{\text{Mod}} - T_{\text{STC}}))}. \quad (2.27)$$

Here, the STC performance is corrected to the performance at the measured module temperature using a linear relationship described with the correction coefficient γ similar to the linear correction of the currents in an IV according to the IEC60891 norm in 2.18 using the correction coefficient α . Beside the temperature dependency, a PV modules performance is further influenced by the overall irradiation intensity ⁷ as well as by the irradiation spectra, incidence angle and soiling [5].

Degradation estimates of PV module or system performance are often based on the determination of the PR. The most common approach is a linear fit of the PR respectively PR_T , to determine a relative performance loss (given in %). Especially for the not temperature corrected PR the linear fit is further often supplemented by a periodic term describing the yearly seasonal variation of the PR.

The estimate of the PR and degradation rates is highly dependent on the chosen filtering approach. For simplicity many approaches are based on using only a (small) subset of the available data, including only measurements, obtained f.e. at irradiation intensities close to G_{STC} ($800 - 1200 \text{ Wm}^{-2}$, [36]), at clear sky instances [32] or at high angle of incidence [37]. Which filtering approach is suitable is dependent on the data specific challenges and is crucial for PR and degradation rate estimates. More on data filtering approaches can be found in Chapter 5.

⁷Typically the low light efficiency of PV modules is lower, than the STC efficiency due to the finite shunt resistance [35]

2.5.3 POA Modelling

To model the POA irradiation I use the open-source simple sky dome projector library (SSDP) [38]. The library implements the Perez All-Weather sky model [39] and accounts for diffusive and direct irradiation contributions. As input, the model uses the location (latitude, longitude and elevation), orientation and tilt of the PV module as well as the time and local effective albedo of the surrounding and measured global horizontal direct and diffusive irradiation GHI and DHI. Optionally a local topography may be provided to simulate shading from the surrounding structures and terrain. More information on SSDP can be found in [40].

2.5.4 Faiman module temperature model

The Faiman module temperature model aims to predict the module temperature T_{Mod} from the ambient air temperature T_{amb} , the in-plane irradiation G_{POA} and the wind speed v . The equation to determine T_{Mod} according to [41] is given by

$$T_{\text{Mod}} = T_{\text{amb}} + \frac{G_{\text{POA}}}{u_0 + u_1 v}, \quad (2.28)$$

where u_0 and u_1 are module specific parameters, that need to be determined experimentally. With the Faiman model accuracies with a bias error below 1 K can be achieved. An evaluation of the accuracy of the Faiman model as well as on the sensitivity of the performance with varying u_0 and u_1 can be found in [42].

2.5.5 CSER (IEC61853)

The IEC61853 norm implements a procedure to determine the Climatic Specific Energy Rating (CSER) for 6 different standard reference climatic profiles. Descriptively, the CSER determines the ratio between overall outdoor performance and Standard Test Condition (STC) efficiency η_{STC} . The 6 reference climates and the procedure how to determine the CSER are presented shortly hereinafter. Part 4 of the IEC61853 norm [3] provides the 6 standard reference climatic profiles. Figure 2.5.1 summarizes the 6 reference climates regarding the monthly integrated in-plane energy input $H_{p,\text{month}}$ and the maximum, mean and minimum ambient temperatures. The six climates cover tropical humid, subtropical arid, subtropical coastal, temperate coastal and temperate continental climates as well as a high elevation reference climate. The annual in-plane energy input $H_{p,\text{year}}$, the latitude and the annual average ambient temperature of the 6 reference climates are summarized in Tab. 2.1. Beside month, day, hour, ambient temperature, wind speed at module height, sun elevation and sun incidence angle, the reference profiles of part 4 of the IEC61853 norm provide global and direct horizontal as well as global and direct in-plane irradiation and the global in-plane irradiation for a set of discrete wavelength bands.

Part 3 of the IEC61853 norm [4] describes the equations and steps to determine the CSER. The procedure can be split into two steps. The first step is determining the effective in-plane irradiation G_{eff} and module temperature T_{mod} at every

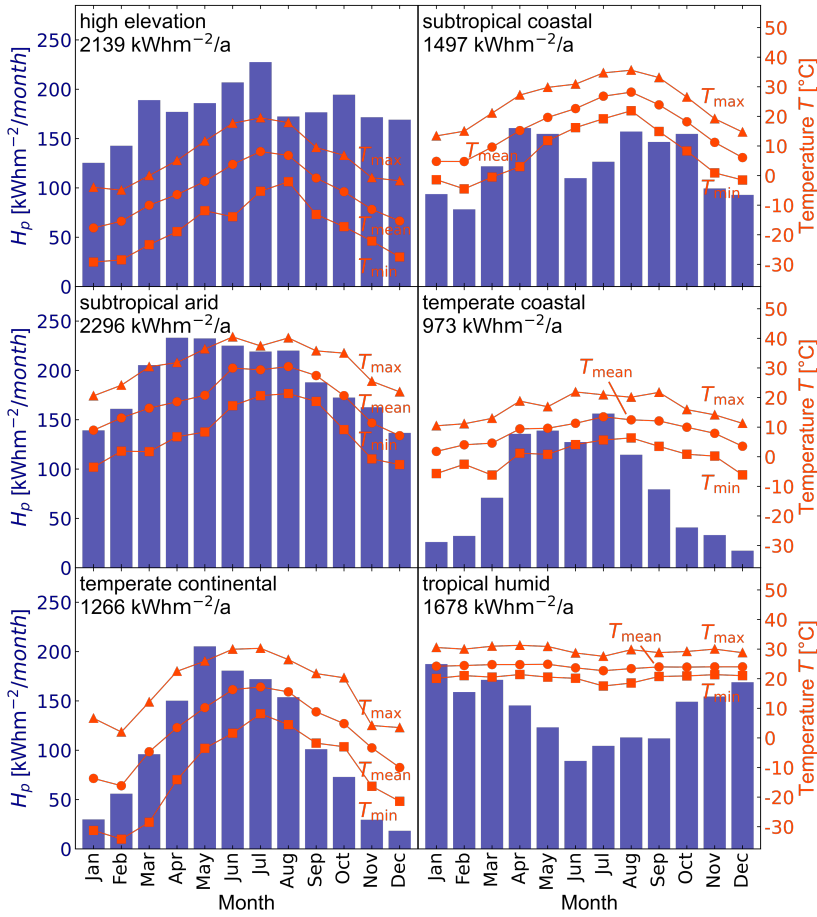


Figure 2.5.1: Monthly energy input H_p and maximum T_{\max} , minimum T_{\min} and mean T_{mean} temperature in the six standard climatic reference profiles provided in part 4 of the IEC61853 norm [3].

Table 2.1: Annual energy input $H_{p,\text{year}}$, latitude and annual average temperature $T_{\text{ave},\text{year}}$ of the IEC61853 standard reference climates [3].

reference climate	$H_{p,\text{year}}$	latitude	$T_{\text{ave},\text{year}}$
high elevation	2139 kWhm ⁻² /a	34°0' N	-5.4 °C
subtropical coastal	1497 kWhm ⁻² /a	33°3' N	16.0 °C
subtropical arid	2296 kWhm ⁻² /a	33°22' N	19.5 °C
temperate coastal	973 kWhm ⁻² /a	56°0' N	8.4 °C
temperate continental	1266 kWhm ⁻² /a	57°0' N	2.5 °C
tropical humid	1678 kWhm ⁻² /a	1°0' S	24.1 °C

given time t_j . The effective irradiation $G_{\text{eff},j}$ is determined using direct and diffuse⁸ irradiation components. Furthermore, the procedure for determining $G_{\text{eff},j}$ adapts for the second to third order (w.r.t. their influence on the overall yield) angular and spectral effects. $T_{\text{mod},j}$ is calculated using the Faïman model [41] according to 2.28.

In the second step the actual yield is computed. For this a performance matrix of the PV module is needed, i.e. the Maximum Power Point (MPP) output power at a set of effective in-plane irradiation G_{eff} and module temperature T_{mod} values. For every time step j of the hourly standard reference climatic profile data, the module output power $P_{\text{mod},j}$ is then calculated using linear interpolation. Afterwards, the hourly output power $P_{\text{mod},j}$ is integrated to an energy output $E_{\text{mod},\text{period}}$ of a given period. Conventionally, the period is one year such that $E_{\text{mod},\text{year}}$ is the annual energy output of a module placed in the theoretical standard reference climate. The CSER w.r.t. one year is given by

$$CSER = \frac{E_{\text{mod},\text{year}} G_{\text{ref}}}{H_{p,\text{year}} P_{\text{max,STC}}}. \quad (2.29)$$

Here, $G_{\text{ref}} = 1000 \text{ Wm}^{-2}$ is the reference irradiation at STC, $P_{\text{max,STC}}$ is the module output power at MPP under STC and $H_{p,\text{year}}$ is the sum of hourly global in-plane irradiation for the reference climatic profile in the given period (i.e. the integrated annual energy input). Inserting $\eta_{\text{year}} = E_{\text{mod},\text{year}}/H_{p,\text{year}}$ and $\eta_{\text{STC}} = P_{\text{max,STC}}/G_{\text{ref}}$, the definition of the CSER can be written as

$$CSER = \frac{\eta_{\text{year}}}{\eta_{\text{STC}}}. \quad (2.30)$$

Thus, the CSER gives the performance ratio between the annual average module efficiency at the reference climatic profile and the STC efficiency. Note, however, that the CSER can be determined for any given period (e.g. for every month of the standard climatic profiles).

⁸The diffuse component of irradiation D_j is given by the difference of global irradiation G_j and direct component B_j ($D_j = G_j - B_j$).

Extrapolation of Module Performance

3.1	Introduction and scientific context	46
3.2	Sample preparation	47
3.3	From cell performance to module performance	47
3.3.1	TCO series resistance	47
3.3.2	P2 series resistance	48
3.3.3	TCO Transmission Correction	49
3.3.4	Extrapolation of module <i>IV</i> characteristics	50
3.3.5	Resulting Module Performance	50
3.4	From module performance to module yield	51
3.5	Impact of performance changes on module yield	52
3.6	Summary	56

In this chapter an extrapolation method to estimate the module yield under various climatic conditions from the laboratory CIGS cell performance is presented. As input for this extrapolation method, measured cell *JV* characteristics at various irradiation intensities and temperatures are used. A model to estimate the module *IV* characteristic based on the measured cell *JV* characteristic is developed. As a result, module *IV* characteristics at various irradiation intensities and temperatures are obtained. By fitting appropriate temperature and irradiation coefficients, procedure 2 of the IEC60891 norm (see section 2.3.2) is then employed to obtain a compact, temperature- and irradiation-dependent model for the module *IV* characteristic. Finally, the IEC61853 norm (see section 2.5.5), which defines a set of standardized reference climatic profiles is implemented. In combination with the Faïman model for the module temperature (see section 2.5.4), the module performance under the six standardized climates of the IEC61853 is estimated. Furthermore, the impact on the climate-specific yield of various cell optimizations is studied.

3.1 Introduction and scientific context

Optimization efforts of Copper Indium Gallium Diselenide Cu(In,Ga)Se_2 (CIGS) solar cells has resulted in a current record efficiency of 23.6% [2]. The performance of high efficiency CIGS solar cells is validated under Standard Test Conditions (STC) [43], i.e. under the reference AM1.5G spectrum at $G_{\text{STC}} = 1000 \text{ Wm}^{-2}$ and a reference cell temperature of $T_{\text{STC}} = 25^\circ\text{C}$. Furthermore, also efforts on optimization simulations and theoretical improvements most often refer to cell performance at STC [44–46].

Similarly, the performance of PV modules is most often verified at STC respectively at Nominal Operating Cell Temperature (NOCT) conditions ($G_{\text{NOCT}} = 800 \text{ Wm}^{-2}$, $T_{\text{amb,NOCT}} = 20^\circ\text{C}$) [47]. These standardized conditions guarantee a comparability of measurements performed all over the world. However, there has always been a discussion in the PV community about whether verifying the performance of solar cells or modules at only one or a few operating conditions is sufficient. As a result, various approaches of rating or modeling PV module as well as PV array outdoor performance under realistic operating conditions are described in literature [4, 47–50].

While a share of the PV community already agrees on the necessity to verify PV module performance under realistic operating conditions, the efficiency of CIGS solar cells as well as cell optimization approaches still refer to STC performance. Zinßer *et al.* showed that cell optimization w.r.t. STC can result in less yield compared to cell optimization w.r.t. differing conditions based on typical meteorological year’s irradiation data in different locations [51]. This finding clearly underlines that a single measurement at STC can not reflect how a PV module consisting of multiple cells would perform under different conditions occurring during operation in various climates.

To verify cell performance in terms of expected yield on module level, a first step is to extrapolate module performance from cell performance. Yousuf *et al.* developed a method to extrapolate module from cell performance in case of Si-based PV modules [52]. For Si-based technologies, cell-to module losses arise mainly due to resistive losses caused by the cell-to-cell interconnection [53]. The standard in CIGS glass-based modules is a monolithic cell to cell interconnection. The extrapolation of CIGS module performance from cell performance requires realistic assumptions of TCO properties and the P1-P2-P3 scribing line architecture [54, 55]. Besides additional dead area and resistive losses, the cell-to-module efficiency gap is further influenced by lateral inhomogeneities (up-scaling of deposition processes) [53].

The second step is the translation of module performance to location respectively climate specific conditions. The IEC61853 norm implements six standardized climatic reference locations [3] depicting typical operating conditions of locations all over the world. With a suitable way to describe temperature and irradiance dependencies as well as with a model to describe the effective irradiance and module temperature one can extrapolate the yield of a PV module operated in different climates [4]. The standardized nature of the defined conditions ensures the comparability of the extrapolated yield, setting a reference for yield calculations.

This chapter presents a model, that extrapolates PV module outdoor performance

respectively yield in various climates from CIGS cell performance. The presented results are published in [56].

3.2 Sample preparation

The analysed CIGS solar cell was produced with an alkali-fluoride post-deposition treatment [57] and the standard Mo/CIGS/CdS/i-ZnO/Al:ZnO architecture at ZSW (Zentrum für Sonnenenergie- und Wasserstoff-Forschung) in Stuttgart. The cell has a total TCO thickness of 150 nm and an associated TCO sheet resistance of $66.7 \Omega/\text{sq}$. The front grid covers 2.78% of the cell area $A_{\text{cell}} = 0.51 \text{ cm}^2$. The cell is characterized with JV characteristics at 9 different temperatures ranging from 4°C to 70°C and 7 different irradiation intensities ranging from 0 to 1300 Wm^{-2} . There, the cell is contacted with metal tips on the front grid and on the Mo back contact layer next to the cell area. Using procedure 2 of the IEC60891 norm I compute $J_{\text{SC,STC}} = 31.2 \text{ mAcm}^{-2}$, $V_{\text{OC,STC}} = 0.72 \text{ V}$, $J_{\text{MPP,STC}} = 27.2 \text{ mAcm}^{-2}$ and $V_{\text{MPP,STC}} = 0.60 \text{ V}$. The cell has a fill factor of 71.9% and an efficiency of 16.24% at STC.

3.3 From cell performance to module performance

In the following, the model to estimate the module performance from lab cell is discussed. For simplicity, the focus lies on first order effects, resulting in relatively simple models. The CIGS module design used here is based on a commercial CIGS module. The basic design parameters (i.e. geometric properties and the TCO sheet resistance) are listed in Tab. 3.1.

Table 3.1: Module design parameters based on a commercial CIGS module. Here w_{cell} and l_{cell} are the width and length of a cell stripe, respectively, w_{da} is the dead-area between cells, and N_{cell} the number of series connected cells. The overall module dimensions are given by the width W_{module} and the length L_{module} . The edge removal widths are E_{short} and E_{long} , along the short and long edges, respectively, and R_{TCO} is the TCO sheet resistance.

w_{cell}	3.0495	mm	l_{cell}	117.46	cm
w_{da}	0.2605	mm	N_{cell}	175	-
W_{module}	60.0	cm	L_{module}	120.0	cm
E_{short}	1.0375	cm	E_{long}	1.27	cm
R_{TCO}	25.0	Ω/sq			

3.3.1 TCO series resistance

As the laboratory cell is a small cell contacted with a grid, the series resistance originating from the electrodes is low. In the monolithically series-connected thin-film modules, the series resistance originating from the TCO electrode is higher. In general the design of monolithically series-connected thin-film

modules is adapted to the TCO sheet resistance as an optimum exists as a trade off between dead area losses and series resistance losses [58]. In a first-order approximation the series resistance originating from the TCO electrode of a cell stripe with width w in the x direction and length l in the y dimension is derived. It is assumed, that the metal electrode is much more conductive than the TCO, allowing to neglect this series resistance contribution. The series connection in the module is achieved such that the TCO electrode is connected to the back electrode of an adjacent cell at $x = w$, over the complete length of the cell stripe. On the other side at $x = 0$ the TCO electrode is isolated from the adjacent cell. Homogeneous cell and electrode properties are assumed as well as, that the current density flow through the cell is uniform over the cell area. In this case the potential is independent of the y coordinate and the potential can be described with a one dimensional model. The lateral current in the TCO electrode is

$$I(x) = J_c l x, \quad (3.1)$$

where I is the lateral current and J_c is the current density per unit area through the solar cell. Due to the electrode sheet resistance the electric field in the electrode is then

$$\begin{aligned} E(x) &= -\frac{I(x)R_{sq}}{l} \\ &= -J_c x R_{sq}, \end{aligned} \quad (3.2)$$

where R_{sq} is the electrode sheet resistance. The potential is obtained by integration of 3.2

$$V(x) = V_0 - \frac{J_c R_{sq}}{2} x^2, \quad (3.3)$$

where V_0 is an integration constant. To compute the effective series resistance, the potential difference between the contact at $x = w$ and the average cell voltage is considered. The average cell voltage equals

$$\begin{aligned} \bar{V} &= \frac{\int_0^w V(x) dx}{w} \\ &= V_0 - \frac{J_c R_{sq} w^2}{6}. \end{aligned} \quad (3.4)$$

The series resistance then follows as

$$\begin{aligned} R_s &= \frac{\bar{V} - V(w)}{I(w)} \\ &= \frac{R_{sq} w}{3l}, \end{aligned} \quad (3.5)$$

where $I(w)$ is the total cell current.

3.3.2 P2 series resistance

In the following the P2 series resistance is estimated. Yoon *et al.* measured a contact resistivity on the order of $5 \times 10^{-3} \Omega \text{cm}^2$ [59]. Thus, with a P2 scribing width in the order of 0.01 cm, the series resistance for a 1 cm long cell stripe

($l = 1$ cm), is on the order of $0.5, \Omega\text{cm}$. If a TCO resistance of $25 \Omega/\text{sq}$ and a cell width $w = 0.4$ cm is assumed, a series resistance of $3.33 \Omega\text{cm}$ is obtained using 3.5. Thus, for the module design the P2 contribution can be considered low compared to the TCO sheet resistance contribution and, for simplicity, this contribution is further neglected.⁹

3.3.3 TCO Transmission Correction

As listed above in section 3.2, the sheet resistance of the used TCO in the laboratory cell is $66.7 \Omega/\text{sq}$, and the thickness is 150 nm. As a more conductive TCO is needed for a monolithic interconnection, the additional absorption losses in the TCO which would result from a thicker TCO need to be estimated. The TCO sheet resistance is assumed to be inversely proportional to the TCO thickness. Thus, aiming for a TCO sheet resistance of $25 \Omega/\text{sq}$, as listed in Tab. 3.1, a factor 2.668 thicker TCO, i.e. an approximately 400 – nm thick TCO layer is needed. The transmission coefficient is estimated using a simple Lambert-Beer law

$$T_{\text{TCO}}(\lambda) = \exp[-\alpha(\lambda)d], \quad (3.6)$$

where $\alpha(\lambda)$ is the wavelength-dependent absorption coefficient, and d the thickness of the TCO. Thus the transmission ratio equals

$$\hat{T}_{\text{TCO}}(\lambda) = \exp[-\alpha(\lambda)(d_{\text{module}} - d_{\text{cell}})], \quad (3.7)$$

where d_{module} and d_{cell} are the TCO thicknesses in the module and cell, respectively ($d_{\text{module}} = 400$ nm and $d_{\text{cell}} = 150$ nm). To estimate the effective transmission ratio the weighted average over the relevant range of the AM1.5g spectrum is computed.

$$\hat{T}_{\text{TCO,eff}} = \frac{\int_{\lambda_0}^{\lambda_1} \hat{T}_{\text{TCO}}(\lambda) F_{\text{AM1.5g}}(\lambda) d\lambda}{\int_{\lambda_0}^{\lambda_1} F_{\text{AM1.5g}}(\lambda) d\lambda}, \quad (3.8)$$

where $F_{\text{AM1.5g}}(\lambda)$ is the spectral photon flux for the AM1.5g spectrum, λ_0 and λ_1 mark the start and end of the relevant spectral range of $\lambda_0 = 400$ nm and $\lambda_1 = 1100$ nm. To obtain a TCO with a sheet resistance of $25 \Omega/\text{sq}$ the effective transmission results in $\hat{T}_{\text{TCO,eff}} = 0.9708$, i.e. the thicker TCO in the module will effectively reduce the short circuit current by approximately 3%.

Finally, the area of the metal grid is considered. As mentioned before, the grid on the cell covers 2.78 % of the cell area. Hence, most of the additional TCO absorption losses are compensated by the absence of a grid in the module. The overall ratio in photo current density is thus expected to be $f_{\text{Ph}} = 0.9708 + 0.0278 = 0.9986$. Thus, the cell JV characteristics are corrected using

$$J_{\text{module}}(V) = J_{\text{cell}}(V) + J_{\text{SC,cell}}(f_{\text{Ph}} - 1), \quad (3.9)$$

⁹In reference [59], the P2 series resistance is deemed a large contribution with respect to the TCO resistance. However, Yoon *et al.* derived the expression for the TCO resistance $R_s = R_{\text{sq}} w^2 / 2$, which is not correct (compare with 3.5). Note that this erroneous series resistance expression has the unexpected consequence that the overall TCO resistance is independent of cell length.

where, $J_{\text{SC,cell}}$ is the measured cell short circuit current density.

3.3.4 Extrapolation of module IV characteristics

Having expressions to correct the cell current density and account for the module TCO series resistance, the cell JV can be extrapolated to a module IV . The measured cell JV is considered as a list of current value pairs with index i . Each cell current density is converted to a corresponding module current

$$I_{i,\text{module}} = w_{\text{cell}} l_{\text{cell}} J_{i,\text{cell}} f_{\text{Ph}}. \quad (3.10)$$

The corresponding module voltages are corrected for the number of series-connected cells and the added TCO series resistance

$$V_{i,\text{module}} = N_{\text{cell}} \left(V_{i,\text{cell}} + \frac{R_{\text{sq}} w_{\text{cell}}}{3 l_{\text{cell}}} I_{i,\text{module}} \right). \quad (3.11)$$

Note that the model discussed does not account for cell-to-module losses due to lateral inhomogeneities usually present due to the up-scaling of deposition processes. Furthermore, it does not incorporate an additional series resistance component of the P1P2P3 scribing line structure in addition to the TCO sheet resistance. To this end, the model does not incorporate all contributions of the cell-to-module efficiency gap in commercial fabricated CIGS modules and only focuses on unavoidable losses due to parasitic absorption in the TCO layer, resistive losses due to the TCO sheet resistance as well as dead area losses.

3.3.5 Resulting Module Performance

The extrapolation method is based on the performance matrix for the cell, where the cell JV was measured as a function of irradiation and temperature. Using the extrapolation method described in before, the corresponding module IV characteristics are computed.

Figure 3.3.1 shows a measured cell JV characteristic and the associated extrapolated module IV characteristic at $G = 1000 \text{ W m}^{-2}$ and $T = 28.5^\circ\text{C}$. The scaling adapts for the number of cells ($V_{\text{m}} = N_{\text{cells}} V_{\text{c}}$) and further for the module-to-cell area ratio ($I_{\text{m}} = A_{\text{module}}/N_{\text{cells}} \times J_{\text{c}}$) to visualize cell-to-module I_{SC} losses (reduced transparency of the TCO, dead-area losses).

To describe the temperature and irradiation dependencies of the module IV characteristic, procedure two of the IEC60891 norm is used. The six correction coefficients as well as the standard solar cell parameters (SSPs), i.e. I_{SC} , V_{OC} , I_{MPP} and V_{MPP} , at STC are determined from the extrapolated module IV performance matrix using the four SSPs, irradiation intensity and temperature of every IV characteristic and three least-squares fits. The correction coefficients are: $\alpha = -5.4 \times 10^{-4} \text{ K}^{-1}$, $\beta = -2.8 \times 10^{-3} \text{ K}^{-1}$, $B_1 = 4.3 \times 10^{-2}$, $B_2 = 9.4 \times 10^{-3}$, $R_{\text{S}} = 7.3 \Omega$ and $\kappa = -2.5 \times 10^{-5} \Omega \text{ K}^{-1}$. Note that the result for the temperature coefficient for the current (α) is negative, which is also observed for other CIGS solar cells [60].

The SSPs at STC of the extrapolated module are: $I_{\text{SC}} = 1.14 \text{ A}$, $V_{\text{OC}} = 126.7 \text{ V}$, $I_{\text{MPP}} = 0.98 \text{ A}$ and $V_{\text{MPP}} = 101.1 \text{ V}$. The module has a fill factor of 68.2% and an efficiency of 13.73% (total area) respectively 15.77% (active area) at STC.

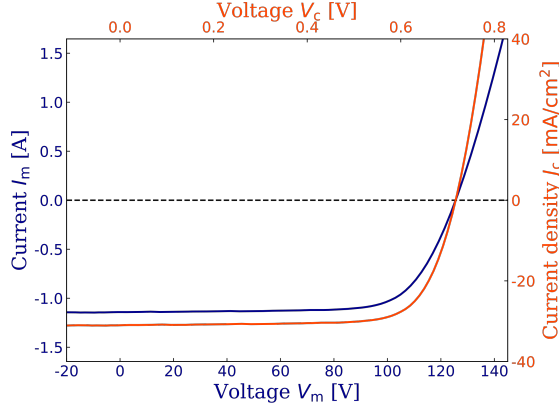


Figure 3.3.1: Comparison of cell JV (orange) and module IV (darkblue) at $G = 1000 \text{ Wm}^{-2}$ and $T = 28.5^\circ\text{C}$. Note that the scaling adapts for the module-to-cell area ratio ($I_m = A_{\text{module}}/N_{\text{cells}} \times J_c$) and the number of cells ($V_m = N_{\text{cells}}V_c$).

The maximum output power at STC is $P_{\text{MPP,STC}} = 98.83 \text{ W}$. The extrapolated module has a nominal output power of $P_{\text{nom}} \approx 100 \text{ Wp}$.

3.4 From module performance to module yield

To extrapolate the module performance respectively the expected yield in the 6 standard reference climates, I use the climate data given by part 4 of the IEC61853 norm (see section 2.5.5). I neglect angular and spectral effects on the effective in-plane irradiation and assume $G_{\text{eff},j} = G_j$. For the determination of the module temperature (T_{mod}) from wind speed (v), effective in-plane irradiation ($G_{\text{eff},j}$) and ambient temperature (T_{amb}) according to the Faiman model (see section 2.5.4), the module-dependent parameters u_0 and u_1 are needed. I use $u_0 = 29.8 \text{ WK}^{-1}\text{m}^{-2}$ and $u_1 = 4.79 \text{ WsK}^{-1}\text{m}^{-3}$ from a study about the Faiman model accuracy in different climates [42]. The used parameters u_0 and u_1 represent the average of the Faiman model fit on half-year module temperature data of commercial CIGS modules operated in five different climates. Barykina and Hammer further show that the determined parameters result in a high accuracy of the module temperature with a bias error less than 1 K and a root mean squared error less than 2.5 K for all 5 investigated locations [42].

With the hourly values of $G_{\text{eff},j}$ and $T_{\text{amb},j}$ and assuming an ideal MPP tracker I compute the expected yield and associated CSER for the 6 standard reference climates from the module STC performance and the 6 correction coefficients.

Figure 3.4.1 summarizes the monthly yield (blue bars) and monthly CSER (red dots). The annual energy output $E_{\text{mod,year}}$ as well as the overall CSER of the extrapolated $\approx 100 \text{ Wp}$ module for the IEC61853 standard reference climates are given in the top right corner of each plot and are further summarized in Tab. 3.2. The high elevation climate shows the highest yield (218 kWh/a) and the temperate coastal climate the lowest yield (95 kWh/a). The highest overall CSER can be found for the high elevation climate (103.2%). For all other

Table 3.2: Annual energy output $E_{\text{mod,year}}$ as well as overall Climatic Specific Energy Rating CSER of the extrapolated ≈ 100 Wp module in the IEC61853 standard reference climates.

reference climate	$E_{\text{mod,year}}$	CSER
high elevation	218 kWh/a	103.2 %
subtropical coastal	140 kWh/a	94.9 %
subtropical arid	209 kWh/a	92.3 %
temperate coastal	95 kWh/a	98.4 %
temperate continental	123 kWh/a	98.6 %
tropical humid	151 kWh/a	91.1 %

climates, the overall CSER is lower than 100% with the lowest CSER found for the tropical humid climate (91.1%).

From Fig. 3.4.1 one can further see, that the monthly yield shows a strong seasonality in temperate coastal and temperate continental climates. In the high elevation and subtropical arid climate the yield varies much less over the year. The CSER shows also a strong seasonality in high elevation, temperate continental as well as in the two subtropical climates. For the tropical humid and temperate coastal climate I observe a rather constant CSER over the year. Comparing Fig. 2.5.1 and Fig. 3.4.1 one can see, as expected, that the yield shows a strong correlation with the monthly energy input H_p . Furthermore, a strong correlation of mean ambient temperature T_{mean} with the monthly CSER can be seen. The overall difference in performance of 12.1% w.r.t. STC efficiency (difference between overall CSER of high elevation and tropical humid) shows clearly the necessity of climate-dependent performance considerations for CIGS (and in general PV) modules.

3.5 Impact of performance changes on module yield

If I now assume a potential gain in the solar cell parameters at STC or in the correction coefficients I can determine the associated gain in the overall yield in the six standard reference climates. I evaluated yield changes by varying all parameters individually in a range of $\pm 10\%$. I find that the gain or loss in yield is linear in the change of all parameters, i.e. linear in a change of V_{OC} and V_{MPP} , I_{SC} and I_{MPP} and linear in a change of all 6 correction coefficients, respectively (not shown). For simplicity, I further refer to yield changes with respect to a 10% gain of the individual parameters. Considering a gain in the photo current (I_{SC} and I_{MPP} at STC) of 10%, I find a rather constant yield gain between 10.20% (high elevation) and 10.44% (temperate coastal, not shown) for the six standard reference climates.

Next I assume a potential gain in the six correction coefficients, i.e. α , β , and κ are considered to be less negative, while B_1 and B_2 are considered to be less positive. For R_s I consider a more positive value as a gain, since the coefficient does not reflect the real series resistance of the device and needs to be interpreted as a measure of how the series resistance changes with a given condition change

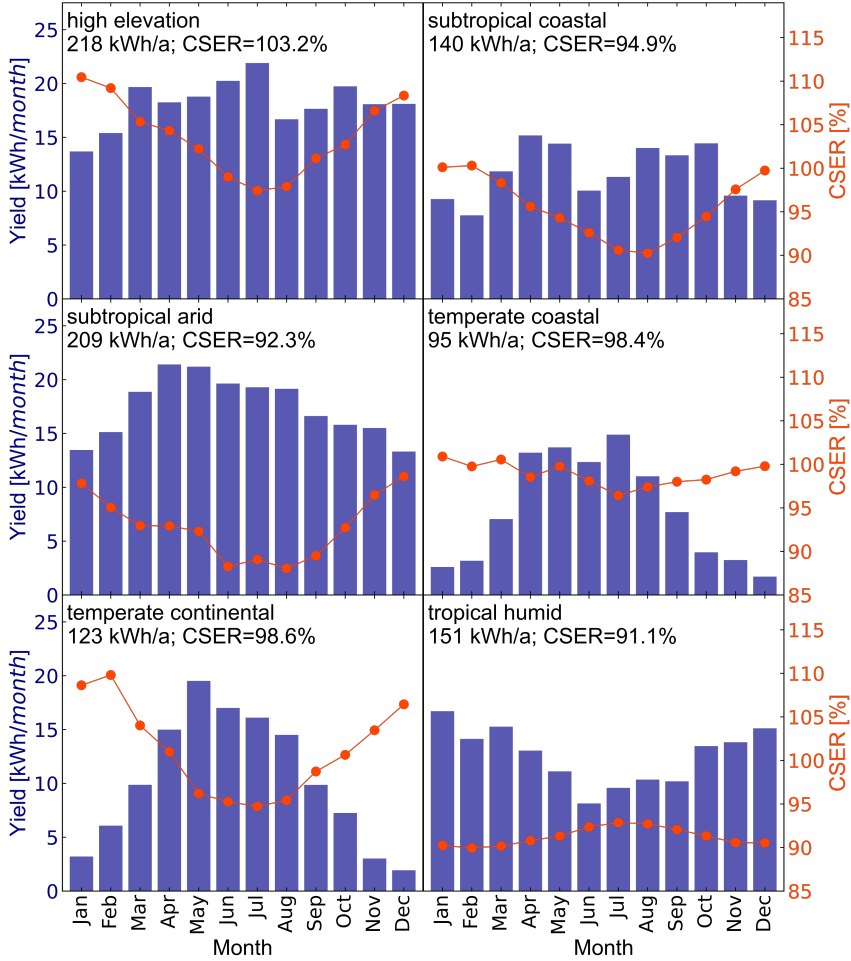


Figure 3.4.1: Extrapolated performance from CIGS laboratory cell towards an ≈ 100 Wp module for the six standard reference climates of the IEC61853 norm. The blue bars show the monthly module yield, while the orange dots present the associated monthly CSER. For the six reference climates the annual yield varies between 95 kWh/a (temperate coastal) and 218 kWh/a (high elevation). The overall (annual) CSER varies between 91.1% (tropical humid) and 103.2% (high elevation).

w.r.t. STC.

Figure 3.5.1 shows the yield gain for a 10% gain of the six correction coefficients. I find nearly no yield gain due to a 10% gain in κ independent of the reference climate. In contrast, a 10% gain in β is highly dependent on the climate and leads to the highest rise in overall yield in warm climates (0.73% in subtropical arid climate) or even to a yield loss in cold climates (-0.35% in the high elevation climate). The yield gain resulting from a 10% gain in the coefficients α , R_S , B_1 and B_2 is also climate dependent and correlates with the reference climates average temperature $T_{\text{ave,year}}(\alpha)$ respectively with the reference climates energy input $H_{p,\text{year}}(R_S, B_1 \text{ and } B_2)$ in Tab. 2.1.

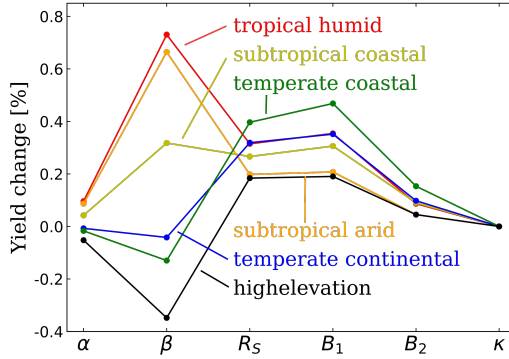


Figure 3.5.1: Yield change in the six standard reference climates due to a 10% gain in the 6 correction parameters α , β , R_S , B_1 , B_2 and κ implemented in the IEC60891 norm.

In the past 20 years, efforts on optimizing CIGS solar cell efficiency has resulted in a gain of more than 4% absolute, associated with a V_{OC} gain of more than 11% [2]. If I assume a potential increase in V_{OC} (and also in V_{MPP}) at STC on cell level, I have to consider that a potential gain in V_{OC} would imply a gain in the temperature coefficient of the voltages β . I can derive the temperature dependency of the open circuit voltage according to the IEC60891 norm. To this end $I_1 = I_2 = 0$, $T_2 = T$ and $T_1 = T_{STC}$ are substituted in 2.19 and, as I consider the STC reference irradiation, $G_1 = G_2 = 1000 \text{ Wm}^{-2}$ implying $f(G_1) = f(G_2) = 1$. From this it follows that

$$V_{OC}(T) = V_{OC}(T_{STC})[1 + \beta(T - T_{STC})], \quad (3.12)$$

where $T_{STC} = 25^\circ\text{C}$ is the STC reference temperature. I can compare this result to expressions derived from the standard one-diode equation. The standard one-diode equation is given in 2.13. Neglecting the influence of series and parallel resistances, accounting for the temperature dependency of the dark saturation current given in 2.3.5, rewriting the equation for the current instead of current density, inserting $I_{SC} = I_{Ph}$ and using $E_g = E_a$ as activation energy of the recombination mechanism (as e.g. the case in SRH recombination) results in

$$I(V) = I_{00} \left(\frac{T}{T_0} \right)^3 \exp \left(\frac{-E_g}{n_{id} k T} \right) \left[\exp \left(\frac{V}{n_{id} k T} \right) - 1 \right] - I_{SC}, \quad (3.13)$$

where I_{00} is a constant, T_0 is a material specific constant, E_g is the band gap, n_{id} is the diode ideality factor and k is the Boltzmann constant. By equating 3.13 to 0 and solving for the voltage, I find

$$V_{OC}(T) = E_g + n_{id}kT \left[\ln \left(\frac{I_{SC}}{I_{00}} \right) - 3 \ln \left(\frac{T}{T_0} \right) \right]. \quad (3.14)$$

Assuming the term T/T_0 is close to 1, 3.14 simplifies to

$$V_{OC}(T) = E_g + n_{id}kT \ln \left(\frac{I_{SC}}{I_{00}} \right) = E_g + AT. \quad (3.15)$$

Thus, I have two linear expressions for the temperature dependency of the V_{OC} in 3.12 and 3.15. From this it follows that

$$A = V_{OC}(T_0)\beta \quad (3.16)$$

and

$$E_g = V_{OC}(T_0)(1 - \beta T_0). \quad (3.17)$$

Thus, obtaining a device with an improved open circuit voltage, V'_{OC} , whilst keeping the band gap E_g the same ¹⁰, it follows from 3.17 that

$$\beta' = \frac{V'_{OC}(T_0) - V_{OC}(T_0)}{V'_{OC}(T_0)T_0} + \frac{V_{OC}(T_0)}{V'_{OC}(T_0)}\beta, \quad (3.18)$$

where β' is the new temperature coefficient for the voltages. To this end, a gain in V_{OC} is expected to result in a gain in the associated temperature coefficient β , i.e. β is expected to be less negative due to a gain in V_{OC} .

Figure 3.5.2 (a) shows the relative yield gain with a 10% gain in V_{OC} and V_{MPP} with (orange) and without (darkblue) taking the associated gain in β into account. The associated gain in β leads to an extra yield gain in warm climates of up to 1.61% (subtropical arid), whereas it leads to a loss in cold climates of up to -0.77% (high elevation).

Figure 3.5.2 (b) shows the absolute yield gain with a 10% V_{OC} gain with and without the associated gain in β . I find an absolute yield gain of up to 21 kWh/a (high elevation) not considering the gain in the temperature coefficient β (darkblue) and of up to 24 kWh/a (subtropical arid) considering the gain in β (orange). The lowest absolute yield gain I find for the temperate coastal climate (≈ 9 kWh/a for both cases).

The extrapolated module reflects an ≈ 100 Wp module. With this context, a gain of one kWh represents ten additional full load hours of operating the PV module. In Fig. 3.5.3 a monthly resolution of the yield gain due to a 10% gain in V_{OC} and the associated gain in β is shown. In addition, the absolute change in the overall CSER is shown. Here, a $\Delta CSER > 0$ depicts an increase in the CSER due to the V_{OC} gain. I find a strong correlation of the monthly absolute CSER change (red) and the monthly absolute yield gain (blue bars) for all six climates. The overall

¹⁰Note that this assumption also implies, that the dominant recombination mechanism and with this the activation energy remains the same in the temperature range under consideration.

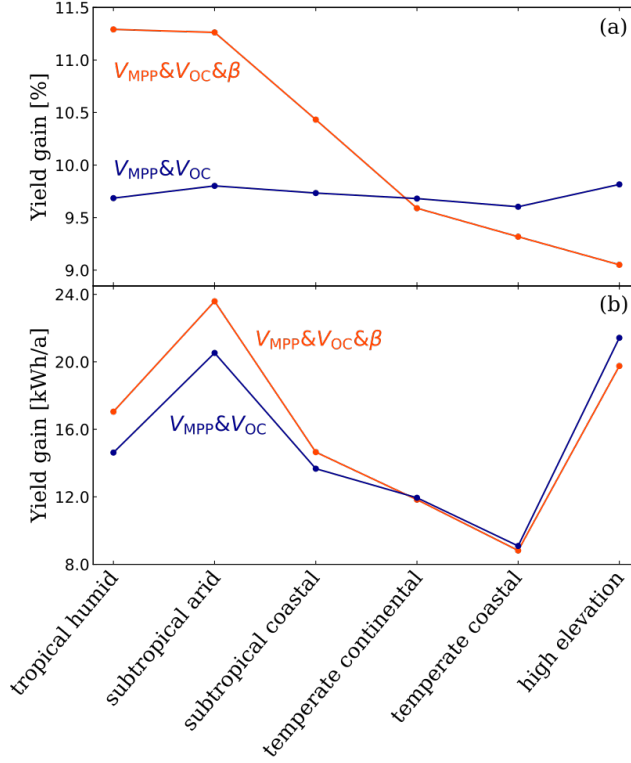


Figure 3.5.2: Relative (a) and absolute (b) yield gain in the six standard reference climates due to a 10% gain in V_{OC} (and V_{MPP}) with (orange) and without (darkblue) considering the associated gain in the temperature coefficient of the voltage β .

annual absolute CSER change is negative for cold climates with the highest change for the high elevation climate (-0.9%) and positive for warm climates with a change up to 1.1% absolute (subtropical arid and tropical humid). A V_{OC} gain respectively the associated gain in the temperature coefficient β leads to less variance in the overall CSER across different climates. The maximum difference in CSER respectively the maximum difference in performance w.r.t. STC conditions drops from 12.1% to 10.2% . To this end, a gain in V_{OC} or in the temperature coefficient β leads to a more stable performance of a PV module operated in different climates.

3.6 Summary

This chapter has introduced an extrapolation method to calculate module yield from laboratory cell performance. The presented model is based on fundamental temperature and irradiation dependent JV characteristics of a CIGS solar cell measured under laboratory conditions. I have introduced a model to extrapolate module IV characteristics incorporating associated cell-to-module losses due to

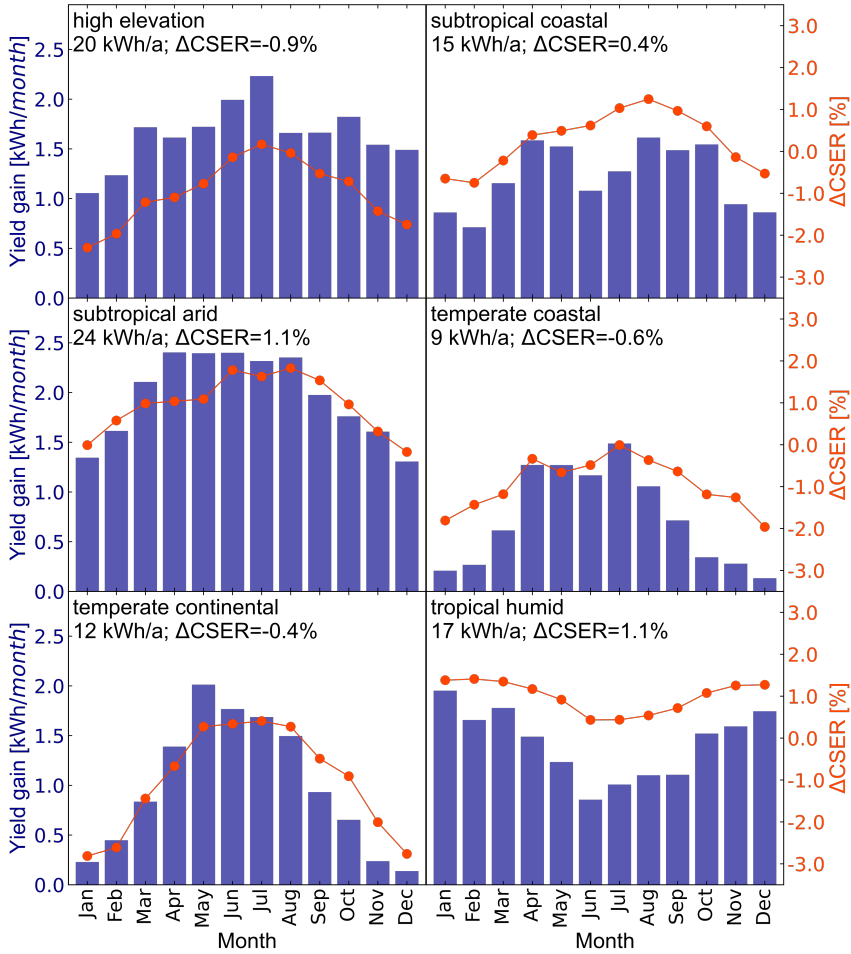


Figure 3.5.3: Extrapolated ≈ 100 Wp module performance gain with a 10% gain in V_{OC} and the associated gain in the temperature coefficient β in the IEC61853 standard reference climates. The blue bars show the monthly gain in the module yield, while the orange dots present the associated absolute CSER change $\Delta CSER$. The overall annual yield gain varies between 9 kWh/a (temperate coastal) and 24 kWh/a (subtropical arid). The overall annual absolute CSER change varies between -0.9% (high elevation) and 1.1% (subtropical arid and tropical humid).

a change of the TCO thickness as well as dead-area losses. In the next step I employed the temperature and irradiation correction equations from procedure 2 of the IEC60891 norm to determine the standard solar cell parameters (SSPs) at STC as well as the temperature and irradiation dependencies of the module IV characteristic. Using the standard reference climatic profiles provided in the IEC61853 norm, I extrapolated module performance at different climates, i.e. I determined the expected annual and monthly yield as well as the CSER (i.e. the overall outdoor performance w.r.t. STC performance) at the 6 reference climates. The expected annual yield varies between 95 and 218 kWh/a. The overall CSER varies between 91.1 and 103.2%. These differences of up to 12.1% absolute in the CSER in different climates underline the necessity to evaluate cell performance not only with one or a few IV characteristics at specific conditions. As a final step I quantified how a potential gain in the SSPs and in the 6 temperature and irradiation coefficients is reflected in the overall annual yield. Such, I quantified how a potential gain on cell level would influence the expected yield of a module consisting of many cells. For a potential gain in the open circuit voltage V_{OC} I considered an implied gain in the temperature coefficient of the open circuit voltage β . This implied gain leads to an additional yield gain in warm climates with an enhanced V_{OC} . With a suitable approach to extrapolate module performance from cell performance, the concept is in general also applicable to other technologies.

Light induced Degradation of CIGS solar cells

4.1	Introduction	60
4.2	CIGS Degradation, Accelerated Testing and Metastabilities . . .	61
4.3	Experimental	63
4.3.1	Sample preparation	64
4.3.2	ex situ measurements	65
4.3.3	Light induced degradation set-up and in situ measurements	65
4.4	Results	66
4.4.1	Performance before LID	66
4.4.2	Degradation of the SSPs (in situ)	67
4.4.3	Degradation of the SSPs (ex situ)	69
4.5	Discussion	71
4.6	Summary and Outlook	75

In this chapter, I will present the details and results of a light induced degradation (LID) experiment on CIGS solar cells. While in literature the focus of LID of CIGS modules and cells is on metastable changes analysed on time scales ranging from minutes to hours, the focus of this work is on time scales up to more than 1000 hours. Industrially produced encapsulated CIGS solar cells are exposed for approximately 1170 hours to light with varied intensity under varied temperature conditions. Such, I aim to study temperature and light intensity dependencies of the observed performance changes. Furthermore, I study the influence of applied bias by comparing LID at short and open circuit. I demonstrate that LID under short circuit (SC) conditions leads to V_{OC} degradation, while being temperature assisted and not dependent on the irradiance intensity. CIGS solar cells kept at open circuit (OC) conditions appear to be stable under illumination. Exploiting the one-diode model, I further connect the observed temperature assisted performance loss to enhanced recombination with lower ideality factor, in comparison to the dominant recombination process before degradation.

4.1 Introduction

Copper Indium Gallium Diselenide Cu(In,Ga)Se_2 (CIGS) solar cells achieved a current record efficiency of 23.6% [2] and have to this degree shown to be competitive to the more established wafer based technologies. Especially due to the feasibility of CIGS solar cells in flexible and light weight photovoltaic (PV) applications, the technology possesses a potential growth in the integrated PV market, e.g. in the vehicle integrated PV (VIPV) and building integrated PV (BIPV) market.

Considering warranted lifetimes of PV modules of typically 20-30 years, their performance stability is for all PV applications one of the most crucial factors regarding power generation and yield. For conventional glass-glass CIGS modules the performance degradation has been studied in great detail as comprehensively summarized in the two review articles from Theelen and Daume [61] respectively Kettle *et al.* [62]. Long-term studies on the performance under outdoor conditions, such as those performed in references [5, 63], are ultimately the best way to evaluate the stability of CIGS modules as the conditions are closest to realistic operation conditions. Under outdoor conditions conventional CIGS modules are shown to experience a performance drop in the beginning of operation (acclimatization) [5] overlaid with seasonal variations and a linear degradation [5, 63]. While the research on outdoor degradation shows a clear trend, the exact causes for the observed acclimatization, seasonality and degradation are hard to infer from such measured time series. Under outdoor conditions various factors like temperature, irradiance (long and short wavelength visible as well as UV), humidity, wind as well as soiling, may influence the modules performance. Furthermore, the conditions are highly correlated, e.g. PV modules heat up under illumination, leading to a clear correlation of irradiance and module temperature.

Laboratory experiments are more suited to study the isolated effects from the various influences by an accurate control of the degradation conditions, enabling to decompose effects, that are typically overlaid in outdoor experiments. Hence, laboratory experiments can be used to elaborate on the insights on degradation of CIGS modules gained in outdoor studies. Another advantage of laboratory experiments is the time scale. While outdoor studies have to be performed over years to make meaningful statements about the stability, laboratory experiments can provide more extreme conditions (e.g. high irradiance, high humidity, high temperature) which allow for much shorter time spans.

One aspect of outdoor degradation is the influence of illumination on the CIGS module performance, i.e. light induced degradation (LID). In literature, effects of illumination on the performance of CIGS solar cells is mainly focused on metastable performance changes [64–68] evaluating performance changes over short time periods, i.e. over minutes to hours. Walkons *et al.* describe a light soaking experiment of up to ≈ 135 h on differently produced CIGS solar cells¹¹ [69]. Cells from two different batches produced with different alkali treatments were subject to the study. Each cell type was exposed to 1000 W/m^2 of AM1.5 g light at different temperatures between 50 and 85°C and short circuit (SC) as well as open circuit (OC) conditions. Walkons *et al.* observe an increase in

¹¹Unfortunately it is unclear whether the used samples are encapsulated or not

V_{OC} upon light soaking under OC and a decrease under SC conditions. An increased degradation temperature seems to enhance the observed impact on V_{OC} . Furthermore, RbF Post deposition treatment (PDT) is shown to lead to more stable performance under illumination [69].

In this work I systematically study the dependency of light induced degradation (LID) of CIGS devices on the light intensity, temperature and bias condition on a time scale of more than 1000 hours. I expose a set of industrially produced and encapsulated CIGS solar cells to light under different illumination intensities, temperature and bias conditions. To gain insights on the interplay of light and temperature I expose the solar cells to three different temperatures between 25 and 70°C and four light intensities between 0 and 1000 W/m². All combinations of temperature and light intensity are studied, keeping one set of cells at SC and another set of cells at OC. With frequent in situ JV measurements as well as with a detailed ex situ characterization before and after ≈ 1170 h of light exposure I show, that the conditions during light exposure strongly influence the observed performance changes.

Before I discuss the experiments in detail I first give a small overview of laboratory (accelerated) degradation testing and the theory behind the observed degradation and the metastable changes in Section 4.2. In Section 4.3 I comprehensively explain the measurement procedure, describe the sample preparation and discuss the used set-ups. Section 4.4 displays the main results structured in the solar cells performance before LID, the in situ observed changes and the ex situ verified comparison of the solar cells performance before and after LID. In Section 4.5 I derive deeper insights on the observed changes employing the one-diode model. Finally, in Section 4.6 the major results are summarized and an outlook on future work is given.

This chapter presents temperature, bias and light intensity dependent laboratory LID experiments on encapsulated CIGS solar cells. The presented work is published in [70].

4.2 CIGS Degradation, Accelerated Testing and Metastabilities

Long-term degradation studies, such as those in references [5, 63], are both tedious and expensive. Ideally, the duration of such long-term degradation studies would be oriented towards the module lifetime itself, however, with a warranted lifetime of PV modules of typically 20-30 years this is unfeasible. Consequently, such studies are limited to one or a few years, which may be enough to study the mayor effects in long-term degradation. Arguably, the economic lifetime of PV modules is likely much shorter [71]. However, one of the reasons that the economic lifetime of PV modules is much shorter is the technological advances over time [71], which in turn implies a risk that the outcome of a long-term degradation experiment is outdated before it is completed. Research on degradation of CIGS modules (and PV modules in general) tends to focus on accelerated aging tests, evaluating the susceptibility of the modules to common causes of degradation in operation.

Damp Heat (DH) testing quantifies their susceptibility to humidity and tem-

perature [72–79]. In general a performance loss is observed under DH, which is associated with a decrease in open circuit voltage (V_{OC}), short circuit current (I_{SC}), fill factor (FF), and shunt resistance (R_{sh}), as well as an increase in series resistance (R_S). The performance loss ratios appear to be highly dependent on the packaging material, encapsulant as well as the solar cell stack itself (e.g. TCO and buffer layer). Suitable packaging and encapsulation materials can prevent high performance loss ratios due to humidity ingress leading to an enhanced series resistance, i.e. TCO [72–76] respectively molybdenum back contact [77–79] degradation. Furthermore, enhanced shunting due to DH treatment is observed [75, 76].

Potential induced degradation (PID) accelerated aging tests on CIGS modules quantify the susceptibility of PV modules to high system voltages common in medium to large PV systems. In literature several studies quantify the performance loss due to PID [80–87]. As for DH treatments the performance loss ratios due to PID are highly dependent on the design and production process of the CIGS module, where a performance drop to close to 0% as well as nearly no performance drop is observed after 50 h of PID for differently produced devices Fjällström *et al.*. Here, high performance loss rates are associated with a large drop in V_{MPP} associated with a large V_{OC} degradation. Further studies on PID find increased bulk and interface defect concentrations [81], a decreased bulk doping concentration [82] as well as an increase in the series resistance due to TCO degradation [84, 85]. Furthermore, PID appears to be dependent on the applied bias condition of the PV module, where faster degradation rates are found for short circuited modules w.r.t. open circuited modules [86]. Fjällström *et al.* further observe non-permanent PID with restored electrical performance close to initial performance after storage in darkness or reversing the electric field that led to the PID [87].

The observed reversibility of PID introduces another complexity. Observed performance changes of CIGS solar cells and modules have to be considered in the context of metastabilities. The metastable nature of CIGS solar cells has already been observed decades ago [64]. Phenomenological it has been found, that CIGS solar cells and thin films exhibit persistent photoconductivity (PPC) due to illumination [88]. Furthermore, under white light illumination (light soaking) a gain in V_{OC} , FF and capacitance has been observed [64–68], where the gain in capacitance as well as in conductivity is attributed to yellow light illumination (absorption in the CIGS absorber) [68, 89, 90] and the gain in FF attributed to blue light illumination (absorption in the CdS buffer layer) [91]. Under forward bias CIGS solar cells are shown to exhibit a gain in V_{OC} as well as in capacitance [64–66, 92–94].

Recent studies on metastabilities in CIGS solar cells have shown, that the amplitude of observed metastabilities is dependent on the production process [69, 95, 96]. Repins *et al.* examined CIGS metastabilities due to light soaking on devices produced with different buffer layers. The efficiency of the differently produced CIGS solar cells is found to be anti-correlated with the amplitude of metastable V_{OC} changes due to light exposure, where high-efficient solar cells exhibit a less metastable nature [95]. Walkons *et al.* showed, that, while temperature assisted (at 65°C) light soaking at open circuit has a beneficial effect on V_{OC} , temperature assisted light soaking at short circuit conditions has

a negative impact on the performance of different CIGS devices [69].

The most common theory in literature to explain the metastable behaviour of CIGS solar cells is based on a selenium copper divacancy defect complex ($V_{\text{Se}} - V_{\text{Cu}}$) in the CIGS absorber first proposed by Lany and Zunger [97]. According to Lany and Zunger the $V_{\text{Se}} - V_{\text{Cu}}$ defect complex has two configurations, that act as shallow donor and shallow acceptor respectively. Which configuration of the defect complex is dominant depends on the position of the Fermi-level within the bandgap. With a Fermi-level close to the valence band (as in typically p-type CIGS absorbers) the shallow donor configuration is dominant, although the metastable equilibrium of the defect configuration may be shifted by (photo-generated) free electrons. While the presence of such a metastable defect configuration seems to be consistent with literature it has been shown, that this explanation alone is not sufficient to explain the observed magnitude and timescale of observed metastabilities [98–100].

The high attention in literature on studying DH, PID and metastabilities of CIGS devices indicates that it is a challenging task to gain more insight into the underlying mechanisms. There are several links that copper-, alkali- and in general ion-migration into the device seems to play a crucial role regarding degradation [75, 76, 80–85] and metastable changes [101–103]. The crucial role of alkali elements is underlined by the recent improvements in efficiency achieved by incorporating alkali-fluoride (i.e. RbF) in post deposition treatments (PDT) [104, 105], leading to a substitution of sodium with heavier and less mobile rubidium atoms. Beside the increased efficiency, RbF PDT seems to reduce the amplitude of the metastable behaviour of CIGS solar cells under light exposure [69].

The analysis of ion migration and changed concentrations of for example sodium or copper due to DH, PID, bias or light exposure gains valuable insights on how the stoichiometry in CIGS solar cells affects its performance. However, the influence of e.g. increased sodium content on the band structure or on shallow and deep defect concentrations are not fully understood. Furthermore, a stoichiometric analysis always reflects only a small share of the solar cell not being able to track lateral variations. Because of the variety and complexity of the reactions during the production of CIGS thin films, the low comparability of different cells from different production lines used in different studies impedes the interpretation.

4.3 Experimental

In total 24 CIGS solar cells are exposed to light induced Degradation (LID). During the LID for 1170 h, 12 cells are kept in short circuit (SC) and 12 cells are kept in open circuit (OC) conditions. Both sets of 12 samples are degraded under 12 different combinations of 3 temperatures ($T_1 = 25^\circ\text{C}$, $T_2 = 50^\circ\text{C}$, $T_3 = 70^\circ\text{C}$) and 4 irradiation intensities ($G_0 = 0.0\%$, $G_1 = 13.5\%$, $G_2 = 47.3\%$ and $G_3 = 100.0\%$ of one sun equivalent irradiation). To create an irradiance level of 13.5 respectively 47.3% I use neutral density (ND) filters. The spectral transmittance of the used ND filters is depicted in Fig. 4.3.1. Using the AM1.5G photon flux, I find an effective transmittance of $T_{\text{eff,AM1.5G,1}} = 13.53\%$ and $T_{\text{eff,AM1.5G,2}} = 47.30\%$ for the two used ND filters.

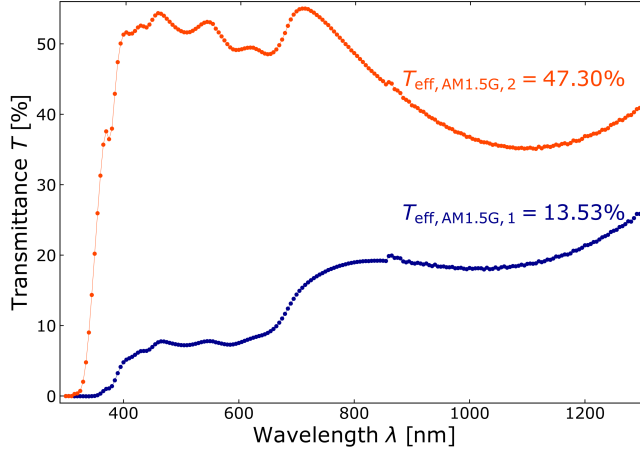


Figure 4.3.1: Transmittance of the used ND filters in the wavelength range between 300 and 1300 nm. Using the AM1.5G photon flux, I find an effective Transmittance of $T_{\text{eff,AM1.5G},1} = 13.53\%$ and $T_{\text{eff,AM1.5G},2} = 47.30\%$.

4.3.1 Sample preparation

The used samples are cut from two industrial produced flexible $308 \times 44 \text{ mm}^2$ CIGS solar cells from MiaSolé that were commercially acquired [106]. The solar cells are produced on a stainless steel substrate used as back contact and the front contact is established using MiaSolé's Ultrawire interconnect technology. From each solar cell I cut 16 similar solar cells with a width of $w_{\text{cell}} = 44.0 \text{ mm}$ and a length of $l_{\text{cell}} \approx 12.0 \text{ mm}$. Each solar cell is then contacted with aluminium busbars using a commercial isotropic conductive adhesive and encapsulated by vacuum lamination using a standard layer stack comprising all commercially available flexible frontsheet, flexible backsheet and polyolefin based encapsulants. OMEGA 5SC-TT-K-36-36 thermocouples are placed at the centre of the back contact of each cell and encapsulated within.

In total 32 CIGS solar cells were prepared. All 32 cells were characterized using EL and PL imaging as well as with a JV characteristic at standard test conditions (STC, $T = T_1 = 25^\circ\text{C}$ and $G = G_3$). Afterwards I selected 24 of the 32 cells for the light induced Degradation (LID) experiments aiming to have the most comparable performance throughout the used solar cells. As described the 32 cells are cut from two industrial produced cells. Since during the LID I keep 12 cells in SC and 12 cells in OC, I chose for each condition to use 12 cells, which are cut from the same solar cell. I selected 12 out of 16 cells for both conditions¹². I first excluded solar cells with visible defects in the EL respectively PL images. Furthermore, I chose the 12 samples with the most similar fill factor (FF) of around 73% at standard test conditions (STC).

¹²Note, that both used solar cells are from the same production line ensuring a comparability of the results obtained at OC and SC.

4.3.2 ex situ measurements

After the sample selection I installed the samples in the sample holder, consisting of a copper block (for tempering with a refrigerated/heated bath circulator) and an anodized aluminium block (for holding the ND filters and reducing stray lights). The solar cells are fixed in between the two blocks and soldered to four pin connectors to ensure a steady and low external series resistance during all electrical measurements. Afterwards the solar cells are kept in dark storage for one week till the characterization before the LID.

Directly before and after the LID as well as two weeks after the LID the CIGS solar cells are characterized with a steady state WACOM WXS-140S-Super sun simulator with class A accuracy of the AM1.5G solar spectrum (details on the used set-up can be found in Section 2.4.1). The JV characteristic of every cell is measured at 12 different conditions (at G_0 , G_1 , G_2 and G_3 as well as T_1 , T_2 and T_3). I start at T_1 and measure the JV characteristics at G_0 , G_1 , G_2 , G_3 and again at G_0 followed by measurements at T_2 and T_3 in the same order of irradiation levels. The irradiation levels are achieved using the same filters as during the LID with the transmissions shown in Fig. 4.3.1. In between the measurements at different temperatures the cells are tempered using a climate chamber.

After the JV characterization, the cells are further characterized with EL and PL images. To this end I use a Princeton instruments Nirvana 640 Indium Gallium Arsenide (InGaAs) camera. The camera is equipped with a 1050 nm band-pass filter with a full width half maximum of 30 nm to block the 808 nm laser bias illumination for the PL experiments. The integration time is set to 100 ms and for both EL and PL, a background image is subtracted. The electrical bias of 0.75 V during the EL experiments is provided by a Keithley 238 high current source measurement unit (SMU). To allow for the sample to reach steady state conditions, the image acquisition starts ≈ 25 s after the solar cell is biased. For more details on the used photo- and electroluminescence set-up see Section 2.4.2.

4.3.3 Light induced degradation set-up and in situ measurements

To ensure constant degradation temperatures the sample holders are thermally decoupled from the set-up and the temperature of the refrigerated/heated bath circulators as well as the room temperature are tracked with PT100 temperature sensors. As light source I use a metal halide lamp HMI 4000 W/DXS - Osram 6000K with a long pass filter of 390nm that blocks ultraviolet (UV) light. The set-up is shielded from external light sources and the irradiation level is tracked with a silicon photodiode in the middle of the illuminated area. The homogeneity of the irradiation is verified at the positions of the cells with two reference cells (PRC Krochmann GmbH PRC911214/1 and PRC911214/2 with a BG17 filter of a size by 2 cm x 2 cm) before and directly after the LID to deviations below $\pm 3.3\%$. I observe an increased intensity as well as a yellow shift over the course of the 1170 hours of LID. The total intensity measured by PRC1 increased by +0.7% and the filtered short wavelength intensity measured by PRC2 decreased by -4.7%. For more details on the used LID set-up see Section 4.

For the complete duration of 1170 h of the LID the samples are not moved to a different set-up and only characterized in situ. Each solar cell is characterized every 24 minutes with an JV characteristic at the specific conditions the cell is degraded at. In between the JV measurements the samples are kept under the respective OC and SC conditions, where the connected cable ends are either open or connected to each other by a relay control. This means that the cable resistance determines the operating point for the cells kept in SC. The temperatures of the cells are tracked for every JV measurement with the encapsulated thermocouples. Due to the laboratory environment and the constant light source, the humidity can be considered very low during the LID. Furthermore, the samples are encapsulated ensuring no effects originating from water ingress.

4.4 Results

4.4.1 Performance before LID

Before the LID all 24 cells show a good and very similar performance. Table 4.1 summarizes the mean and standard deviation of the standard solar cell parameters (SSPs), i.e. V_{OC} , J_{SC} , V_{MPP} and J_{MPP} , measured under STC for the 24 used solar cells. For the open circuit voltage (V_{OC}) at STC I find a mean of 739.5 mV and a low standard deviation of 2.4 mV among the 24 cells. The maximum power point voltage (V_{MPP}) at STC shows a comparable scatter ($\sigma_{V_{OC}} = 2.5$ mV) with a mean of 596.7 mV. A larger scatter is observed for the short circuit and maximum power point current densities (J_{SC} and J_{MPP}) at STC with a mean of 33.8 mA/cm² respectively 30.6 mA/cm². This I attribute to the variations in the sample preparation described in Section 4.3.1. I assume constant dimensions given in Section 4.3.1, but deviations in the cell length and therefore in the cell area occur due to the cutting process. I find a standard deviation of 2.4% and 2.6% in J_{SC} and J_{MPP} respectively. Regarding the FF at STC I find very low scatter with a mean of 73.0% and a standard deviation of 0.6%. To compare the efficiency without the dominating effect of varying J_{SC} and J_{MPP} I correct the current density values of every individual STC measurement with the ratio of mean to measured J_{SC} . This way, I find an efficiency of $\eta_{corrected} = 17.83 \pm 0.14\%$ for the 24 solar cells .

Table 4.1: Standard solar cell parameters of the CIGS solar cells at STC before LID

V_{OC}	739.5	± 2.4	mV	V_{MPP}	596.7	± 2.5	mV
J_{SC}	33.8	± 0.8	mA/cm ²	J_{MPP}	30.6	± 0.8	mA/cm ²
FF	73.0	± 0.6	%	$\eta_{corrected}$	17.83	± 0.14	%

To further demonstrate the comparability in performance of the 24 solar cells Fig. 4.4.1 shows the JV characteristics of all 24 cells at STC conditions. Note that the current densities are normalized with J_{SC} to remove the effect of varying cell area. As described in Section 4.3.2 all 24 cells are characterized at the sun simulator at 11 further conditions. Regarding the scatter within the batch of

used solar cells for the performance at the other G and T conditions I find similar relative standard deviations (not shown) as for the case of the comparison shown for STC.

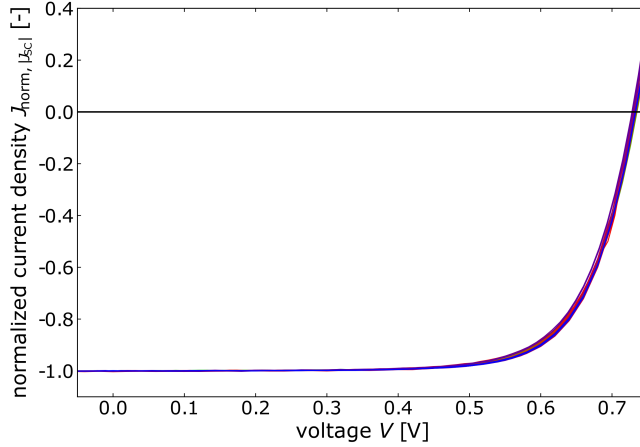


Figure 4.4.1: JV characteristics of the 24 used CIGS solar cells at STC. The current density is normalized to the individual short circuit current densities J_{SC} .

4.4.2 Degradation of the SSPs (in situ)

In situ the CIGS solar cells are characterized with an JV measurement every 24 minutes at the specific conditions the cell is degraded at like described in Section 4.3.3. First of all I observe no clear trend of decrease or increase in J_{SC} respectively J_{MPP} for all monitored cells¹³. Variations in J_{SC} respectively J_{MPP} occur mainly due to variations in the intensity of the light source and show a similar pattern for all monitored cells (not shown).

The development of the open circuit voltage V_{OC} and maximum power point voltage V_{MPP} exhibit a systematic evolution over time, depending on the bias conditions. For simplicity I concentrate on V_{OC} hereinafter. Figure 4.4.2 (a) shows ΔV_{OC} , the difference between the in situ measured V_{OC} and the first respective measured V_{OC} after the start of LID¹⁴ for the nine cells degraded at open circuit conditions and at irradiation intensities $G > 0$. I observe a clear trend of increasing V_{OC} for the first few hours of LID at open circuit conditions for all 9 cells. For the 3 cells degraded at $T_1 = 25^\circ\text{C}$ (dark blue) this increase is observed over the whole duration of LID of about 1170 hours. For the 3 cells degraded at $T_2 = 50^\circ\text{C}$ (yellow) the increase in V_{OC} saturates after approximately

¹³Note that 6 cells (SC and OC at 3 different degradation temperatures) are not illuminated and therefore the SSPs can not be analysed in situ.

¹⁴Note that I always compare V_{OC} to the first measurement of the respective cell. Since the cells are measured after each other in an interval of ≈ 1 min the first respective measurement takes place in between $0 \text{ min} < t < 24 \text{ min}$ after the start of LID. The order of measurements is oriented on the degradation temperature, where the cells degraded at $T_3 = 70^\circ\text{C}$ are measured first, followed by the cells degraded at $T_2 = 50^\circ\text{C}$ and the cells degraded at $T_1 = 25^\circ\text{C}$.

400 hours of degradation. For the 3 cells degraded at $T_3 = 70^\circ\text{C}$ (orange) the steep increase in V_{OC} within the first 40 hours of LID is followed by a less gradual decrease of V_{OC} . Furthermore, after approximately 78 hours of LID I observe an increased V_{OC} for around 35 hours, this coincides and is explained with a failure of the $T_3 = 70^\circ\text{C}$ water cooling/heating system and resulting slightly lower cell temperatures ($\Delta T = 4 - 8\text{K}$ depending on the sample). Comparing the development of ΔV_{OC} of cells degraded at the same degradation temperature and different irradiation intensities I observe a change in amplitude of the effects and trends described. Cells degraded at higher irradiation intensities show higher absolute changes in V_{OC} .

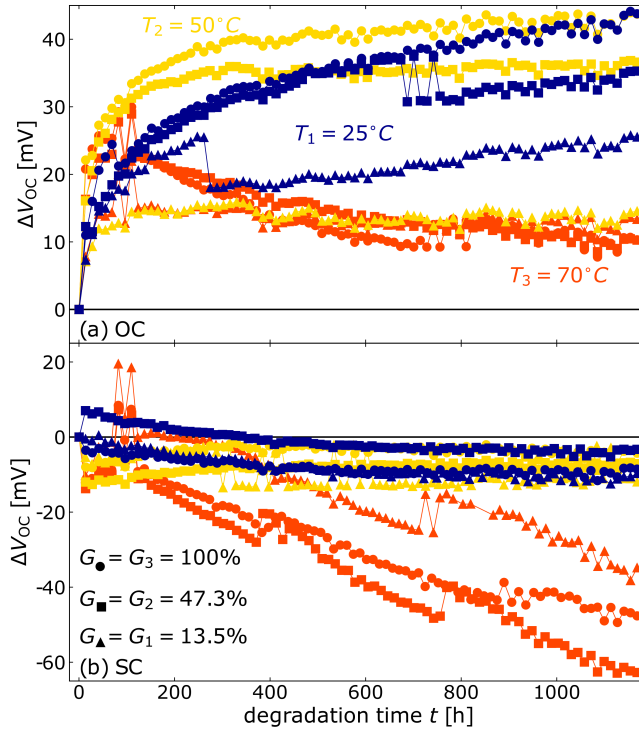


Figure 4.4.2: Absolute change in open circuit voltage ΔV_{OC} due to LID at (a) open circuit and (b) short circuit conditions. The color indicates the degradation temperatures, i.e. dark blue corresponds to $T = T_1 = 25^\circ\text{C}$, yellow to $T = T_2 = 50^\circ\text{C}$ and orange to $T = T_3 = 70^\circ\text{C}$. With the markers I differ for the irradiation levels of $G = G_1 = 13.5\%$ (triangles), $G = G_2 = 47.3\%$ (squares) and $G = G_3 = 100\%$ (circles) during degradation.

Figure 4.4.2 (b) shows the development of ΔV_{OC} for the nine cells degraded at short circuit conditions at irradiation intensities $G > 0$. Contrarily to the case of the samples degraded at OC I do not observe an increase in V_{OC} in the first few hours of LID. For the 3 cells degraded at $T_1 = 25^\circ\text{C}$ (dark blue) I observe a slow decrease of V_{OC} with a saturation towards the end of the 1170 hours of LID. The V_{OC} of the 3 cells degraded at $T_2 = 50^\circ\text{C}$ (yellow) is rather

stable during LID, with a slight decrease in V_{OC} at the start of LID. The biggest change in V_{OC} is observed for the 3 samples degraded at $T_3 = 70^\circ\text{C}$ (orange) at short circuit conditions, where ΔV_{OC} appears to be stable for the first few hours of LID followed by a steady decrease down to -40 mV ($G = G_1$), -45 mV ($G = G_3$) respectively -60 mV ($G = G_2$)¹⁵. Furthermore, for some samples I observe a sharp increase respectively a sharp decrease of varying amplitudes, e.g. in the case of the sample degraded at open circuit at $T = T_1$ and $G = G_1$ (dark blue triangles in Fig. 4.4.2 (a)) after approximately 300 hours of LID.

4.4.3 Degradation of the SSPs (ex situ)

As described in Section 4.3.2 all CIGS solar cells are characterized with JV characteristics at 12 conditions as well as with EL and PL images directly before (pre LID), directly after (after LID) and 2 weeks after the LID, where the cells are kept in dark (after dark storage, DS). Figure 4.4.3 (a) shows the change in V_{OC} at STC due to LID at open circuit conditions. All cells degraded at open circuit conditions show a low absolute change in V_{OC} after LID between $\pm 15\text{ mV}$. I observe a tendency for a positive ΔV_{OC} with increasing degradation temperatures. ΔV_{OC} is positive for seven out of 8 samples degraded at $T_2 = 50^\circ\text{C}$ (yellow) or $T_3 = 70^\circ\text{C}$ (orange), while ΔV_{OC} is negative for two out of four samples degraded at $T_1 = 25^\circ\text{C}$ (dark blue). Comparing these results to Fig. 4.4.2 (a) I find a lower (or even negative) ΔV_{OC} for the ex situ analysis in comparison to the last in situ measurements for the majority of the samples. I attribute this effect to a metastable increase in V_{OC} due to illumination during the LID (see also Section 4.2) as well as to a slightly different irradiation intensity and spectra in the two set-ups. Furthermore, the ΔV_{OC} shown in Fig. 4.4.2 are verified at the respective degradation conditions, impeding the comparison to the ΔV_{OC} verified ex situ at STC. After two weeks of dark storage (DS) I observe a slight enhancement of the V_{OC} for the majority of cells. Regarding the dependency from the illumination intensity during LID I do not observe a clear trend.

For the cells degraded at short circuit conditions I observe a very different degradation behaviour of V_{OC} as shown in Fig. 4.4.3 (b). ΔV_{OC} is negative for all 12 cells degraded at short circuit conditions. The absolute V_{OC} loss for the cells degraded under $G = G_0 = 0\%$ (stars) is small and comparable to the amplitude of V_{OC} change for the case of the cells kept in open circuit during LID (Fig. 4.4.3 (a)). For the 9 cells degraded at short circuit and light intensities $G > 0\%$ I observe a high loss in V_{OC} . The loss is highest for the cells degraded at $T_3 = 70^\circ\text{C}$ (orange) with a ΔV_{OC} between -70 and -100 mV . For the cells degraded at low or medium temperature ($T_1 = 25^\circ\text{C}$ in dark blue and $T_2 = 50^\circ\text{C}$ in yellow) I observe a V_{OC} loss in between -25 and -40 mV absolute. Again regarding differences due to the illumination intensity during LID I do not observe a clear trend. Comparing these results to Fig. 4.4.2 (b) I find similarly to the OC case a lower (more negative) ΔV_{OC} for the ex situ analysis in comparison to the last in situ measurements for all samples. As before, I attribute this to metastabilities as well as to the difference in the conditions

¹⁵Note that after 78 hours of LID I also observe an increased V_{OC} for around 35 hours due to failure of the $T_3 = 70^\circ\text{C}$ water cooling/heating system and resulting slightly lower cell temperatures ($\Delta T = 4 - 8\text{ K}$ depending on the sample).

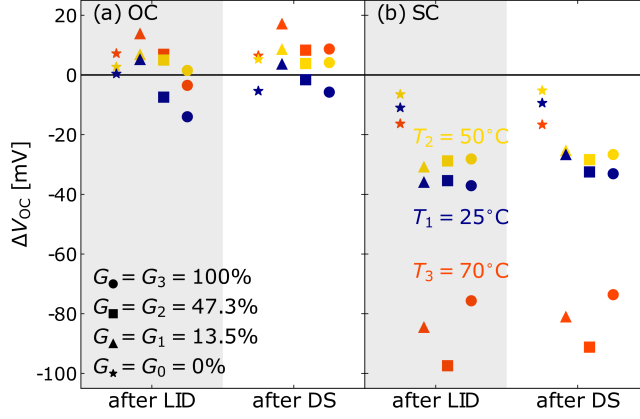


Figure 4.4.3: Absolute change in open circuit voltage ΔV_{OC} at STC due to LID at (a) open circuit and (b) short circuit conditions. The colour indicates the degradation temperatures, i.e. dark blue corresponds to $T = T_1 = 25^\circ\text{C}$, yellow to $T = T_2 = 50^\circ\text{C}$ and orange to $T = T_3 = 70^\circ\text{C}$. With the markers I differ for the irradiation levels of $G = G_0 = 0\%$ (stars), $G = G_1 = 13.5\%$ (triangles), $G = G_2 = 47.3\%$ (squares) and $G = G_3 = 100\%$ (circles) during degradation.

during the measurements. Analogous to the cells degraded at open circuit, for the cells degraded at short circuit I observe a slight enhancement of the V_{OC} for the majority of cells after two weeks of dark storage.

Figure 4.4.4 shows the absolute change in J_{SC} due to LID for all 24 degraded CIGS solar cells. With the exception of one cell (Degradation at $G = G_0$ and $T = T_3$) all cells degraded under short circuit conditions (Fig. 4.4.4 (b)) show nearly no change in short circuit current density after LID ($-0.1 \text{ mA/cm}^2 < \Delta J_{SC} < 0.15 \text{ mA/cm}^2$). For the cells degraded at open circuit (Fig. 4.4.4 (a)) I find a slightly reduced short circuit current density after LID (changes up to -0.65 mA/cm^2). I find no clear trend of degradation light intensity and or temperature. Note that even the highest absolute change in J_{SC} corresponds to a small relative change of less than 2% (compare with Tab. 4.1). After two weeks of DS I observe for the majority of cells a slight enhancement of J_{SC} w.r.t. the STC measurement directly after the LID.

The maximum power point voltage V_{MPP} at STC shows similar changes as V_{OC} at STC for all degraded cells, where cells degraded at short circuit and $T = T_3 = 70^\circ\text{C}$ exhibit the highest V_{MPP} loss and all cells degraded at open circuit show only slight changes in V_{MPP} (not shown). For the maximum power point current density J_{MPP} at STC I observe a slight loss (between -0.02 and -0.76 mA/cm^2) after LID for the cells degraded at open circuit (not shown) and by this similar absolute changes observed for J_{SC} . For cells degraded at short circuit I also observe only slight changes ($-0.35 \text{ mA/cm}^2 < \Delta J_{MPP} < 0.15 \text{ mA/cm}^2$, not shown), where a loss in J_{MPP} is correlated to the loss in V_{OC} (Fig. 4.4.3) and highest for the cells degraded at $T = T_3$.

Regarding the absolute change in efficiency at STC $\Delta\eta$ in Fig. 4.4.5 I observe, that all cells degraded at open circuit exhibit small changes due to LID ($-0.5\% <$

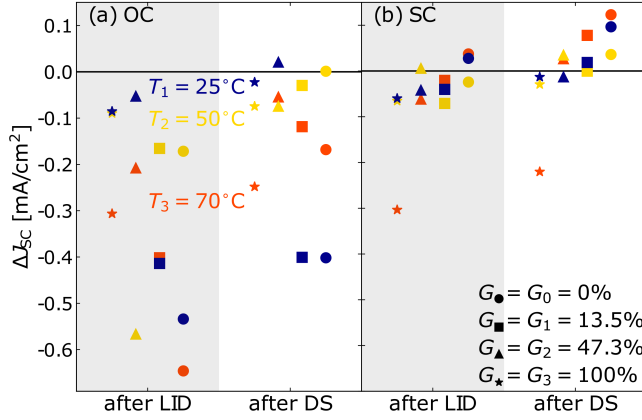


Figure 4.4.4: Absolute change in short circuit current density ΔJ_{SC} at STC due to LID at (a) open circuit and (b) short circuit conditions. To differ between all degradation conditions I use the same colour- and marker scheme used before in Fig. 4.4.2 and Fig. 4.4.3.

$\Delta\eta < 0.5\%$). For the cells degraded at short circuit I observe a loss in STC efficiency of up to -3.1% absolute. Comparing Fig 4.4.5 and Fig. 4.4.3 I observe, that the loss in efficiency is mainly connected with the loss in V_{OC} respectively V_{MPP} .

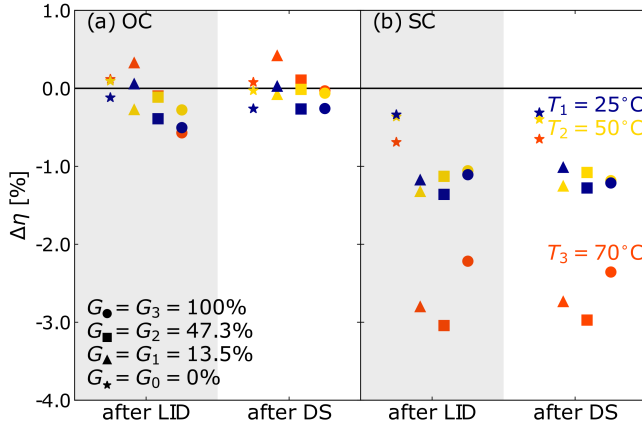


Figure 4.4.5: Absolute change in efficiency $\Delta\eta$ due to LID at (a) open circuit and (b) short circuit conditions. To differ between all degradation conditions I use the same colour- and marker scheme used before in Fig. 4.4.2 - 4.4.4.

4.5 Discussion

The results presented in Section 4.4 show that LID appears to be strongly dependent on the conditions CIGS solar cells are degraded at. The performance

of CIGS solar cells after 1170 hours of LID under open circuit conditions appears to be more or less stable with overall efficiency changes of less than $\pm 0.5\%$ absolute. For LID under short circuit conditions I find an efficiency drop of up to -3.1% absolute caused by a loss in V_{OC} respectively V_{MPP} . This drop in efficiency seems to be more or less independent of the degradation light intensity, while the degradation temperature appears to be crucial for the underlying mechanisms. The cells degraded at low ($T = T_1 = 25^\circ\text{C}$) respectively moderate ($T = T_2 = 50^\circ\text{C}$) temperatures degrade much less than cells degraded at high temperature ($T = T_3 = 70^\circ\text{C}$) conditions.

By now I only used the SSPs to evaluate changes due to LID. To further investigate the observed V_{OC} drop I use the open source PV-CRAZE library [31] described in detail in Section 2.3.3. PV-CRAZE provides a robust fitting algorithm for the one-diode model according to Shockley [107]. The current density voltage relation for a p-n junction reads:

$$J(V) = J_{ph} - J_0 \left[\exp \left(\frac{V + J(V)R_S A}{n_{id} k T / q} \right) - 1 \right] - \frac{V + J(V)R_S A}{R_{sh} A}, \quad (4.1)$$

where J_{ph} represents the photo generated current density, R_S respectively R_{sh} series respectively shunt resistance of the device, n_{id} the diode ideality factor, A the cell area, q the elementary charge, k the Boltzmann constant and

$$J_0 = J_{00} \left(\frac{T}{T_0} \right)^3 \exp \left(\frac{-E_g}{n_{id} k T} \right) \quad (4.2)$$

the dark saturation current density. Note that, I use 2.13 from Section 2.3 and account for area correction of the series and shunt resistance in 4.1¹⁶. As expected I find a strong correlation of J_{ph} and J_{SC} , i.e. only small changes in J_{ph} due to the LID are observed (not shown). Figure 4.5.1 shows the development of R_S for all 24 cells derived from the JV characteristics measured under STC directly before and after LID as well as after 2 weeks of DS. I find a clear trend of increasing series resistance after LID for the cells kept under short circuit conditions (see Fig. 4.5.1 (b)). For the series resistance of the cells degraded at open circuit conditions I find no clear trend (see Fig. 4.5.1 (a)). Overall the observed changes and absolute value of the series resistance is small for the whole batch with a maximum of 0.21Ω . Furthermore, due to the definition of the open circuit voltage ($J(V_{OC}) = 0$) an increased series resistance can not be the cause for the observed drop in V_{OC} .

For the comparison of the shunt resistance before and after LID as well as after dark storage I do not observe a clear trend regarding the degradation conditions (not shown). The shunt resistance varies within the batch of solar cells in a range of 200 to 1000Ω (outliers only occur to higher resistances) and to this end in a range, which has slight to no negative influence on the solar cells performance. Figure 4.5.2 shows the determined dark saturation current density J_0 for all cells before, after and after two weeks of dark storage after LID. Contrary to the observed V_{OC} loss in cells degraded at short circuit (see Fig. 4.4.3 (b)) J_0 tends to increase only slightly for the cells degraded at open circuit (Fig. 4.5.2

¹⁶This results in units for series and shunt resistance in Ω rather than Ωcm^2

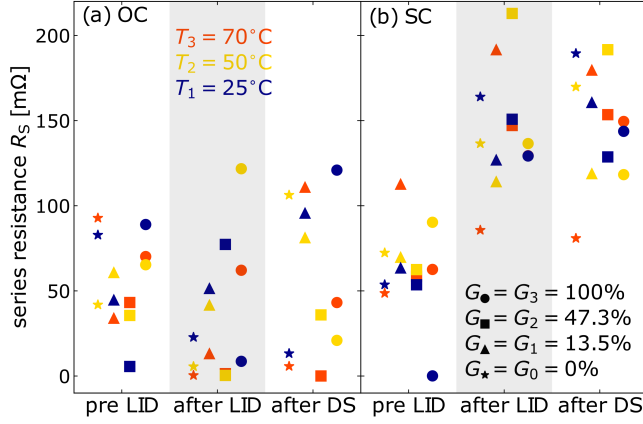


Figure 4.5.1: Series resistance R_S determined from a one-diode model fit to JV characteristics measured at STC conditions before and after LID as well as after two weeks of dark storage for the cells kept at (a) open circuit and (b) short circuit conditions during LID. The colour indicates the degradation temperatures, i.e. dark blue corresponds to $T = T_1 = 25^\circ\text{C}$, yellow to $T = T_2 = 50^\circ\text{C}$ and orange to $T = T_3 = 70^\circ\text{C}$. With the markers I differ for the irradiation levels of $G = G_0 = 0\%$ (stars), $G = G_1 = 13.5\%$ (triangles), $G = G_2 = 47.3\%$ (squares) and $G = G_3 = 100\%$ (circles) during degradation.

(a)), while a rather constant J_0 is observed for the cells degraded at short circuit (Fig. 4.5.2 (b)). Beside the dark saturation current density J_0 the diode ideality factor n_{id} is essential to describe recombination in solar cells according to the one-diode model. Figure 4.5.3 shows the determined diode ideality factor for all 24 cells. First of all I observe high ideality factors between 1.85 and 2.3 before LID. After LID the majority of cells degraded at open circuit show an enhanced ideality factor (Fig. 4.5.3 (a)), while the majority of cells degraded at short circuit exhibit a reduced ideality factor (Fig. 4.5.3 (b)).

To put the combined effects of the observed changes in J_0 and n_{id} into context I define a recombination current density J_x as

$$J_x = J_0 \exp \frac{qV_x}{n_{id}kT}, \quad (4.3)$$

which is the total recombination in the device as a fixed internal voltage V_x (i.e. voltage over the diode without series resistance). Figure 4.5.4 shows the recombination current density J_x at $V_x = 0.7\text{ V}$ and the from the one-diode model determined J_0 and n_{id} (Fig. 4.5.2 and 4.5.3). Note that I choose $V_x = 0.7\text{ V}$ to show the recombination current close to the initial V_{OC} (the mean V_{OC} across the used batch of CIGS solar cells before LID is 739.5 mV , see Tab. 4.1). I observe a strong correlation of J_x with the observed change in V_{OC} and efficiency η (see Fig. 4.4.3 and 4.4.5). I conclude, that LID of CIGS solar cells at short circuit leads to an enhanced recombination and associated open circuit voltage and efficiency loss.

It is interesting to note that the degradation on V_{OC} is associated with a reduction

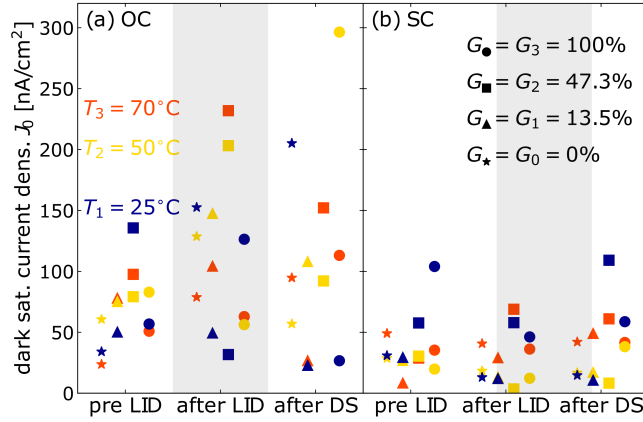


Figure 4.5.2: Dark saturation current density J_0 determined from a one-diode model fit to JV characteristics measured at STC conditions before and after LID at (a) open circuit and (b) short circuit conditions as well as after two weeks of dark storage (DS). To differ between all degradation conditions I use the same colour- and marker scheme used before in Fig. 4.4.2 - 4.5.1.

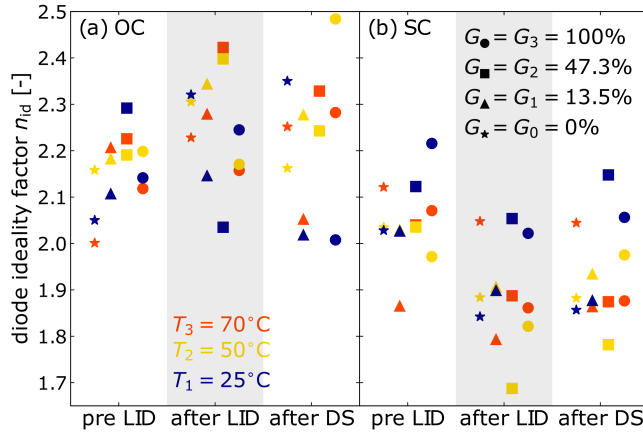


Figure 4.5.3: Diode ideality factor n_{id} determined from a one-diode model fit to JV characteristics measured at STC conditions before and after LID at (a) open circuit and (b) short circuit conditions as well as after two weeks of dark storage (DS). To differ between all degradation conditions I use the same colour- and marker scheme used before in Fig. 4.4.2 - 4.5.2.

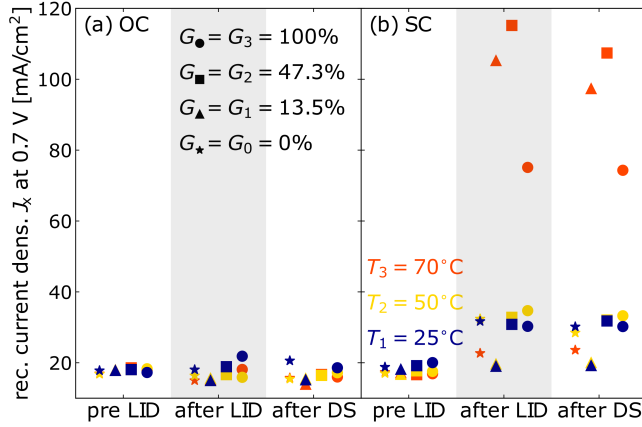


Figure 4.5.4: Recombination current density $J_x = J_0 \exp \frac{qV_x}{n_{id}kT}$ at $V_x = 0.7$ V determined from a one-diode model fit to JV characteristics measured at STC conditions before and after LID at (a) open circuit and (b) short circuit conditions as well as after two weeks of dark storage (DS). To differ between all degradation conditions I use the same colour- and marker scheme used before in Fig. 4.4.2 - 4.5.3.

in the ideality factor. This indicates that the increased recombination rate is due to a process with ideality factor 1. This in turn is an indication that additional recombination centres are located in quasi-neutral regions of the device (i.e. not within the depletion region) or are shallow traps^[108]¹⁷.

4.6 Summary and Outlook

The presented work describes a detailed study of temperature, bias and intensity dependent light induced degradation of industrially produced encapsulated CIGS solar cells. The results show, that the bias condition is crucial w.r.t. observed performance changes, where cells kept in OC appear to be stable under illumination and cells kept in SC exhibit a performance drop of up to 3% absolute (efficiency at STC) within 1170 h of light exposure. The performance drop is shown to be dependent on the temperature condition during LID, while being more or less independent on the irradiation intensity. Furthermore, it is observed that the overall recombination rate increases and, at the same time, the ideality factor is reduced. This indicates that the increased recombination rate is due to a process with ideality factor 1, consistent with recombination centres located in quasi-neutral regions of the device. Due to the observed big discrepancy between performance stability at OC and SC, and the fact that CIGS modules are operated somewhere between these extreme bias conditions (optimally at MPP), future work should concentrate on filling this gap, aiming

¹⁷The derivation in Section 2 of the paper from Pieters *et al.* focuses on the activation energy of recombination in p - i - n devices. However, the derivation is valid for Shockley-Read-Hall (SRH) recombination in general, and, the paper also discusses the spatial and energetic variation of the ideality factor of SRH recombination.

to depict a bias condition closer to real operation scenarios. Furthermore, from the results obtained it remains unclear, if the observed performance loss under SC conditions is permanent or if the underlying processes are reversible with e.g. LID under OC conditions (metastable behaviour). To this end future work should also investigate how the performance of CIGS solar cells develops due to LID under switched conditions and, if the processes are found to be reversible, on which time scales metastable effects influence the performance.

Chapter

5

Filtering of Outdoor data

5.1	Introduction	78
5.2	Mahalanobis distance filter	79
5.2.1	Concept	79
5.2.2	Mahalanobis filtering results	80
5.3	Gaussian Process Regression filter	85
5.3.1	Concept	85
5.3.2	GPR filtering results	87
5.4	Summary and Outlook	91

In this chapter, I will present two different approaches to filtering PV outdoor data. The first approach leverages correlations in the various dimensions of the data. By fitting appropriate physical models to utilize these correlations, implausible data, i.e., data where the different dimensions do not correlate according to the physical models, can be identified due to high deviations from expectations. The Mahalanobis distance (see Section 2.1.2) is used to quantify this deviation in terms of standard deviations.

The second approach focuses on filtering PV outdoor data where the measured in-plane irradiation intensity (G_{POA}) and the associated short circuit current (I_{SC}) do not fit, incorporating the temperature dependency of I_{SC} . The temperature and irradiation dependencies of I_{SC} are described using multiple Gaussian Process Regressions (GPRs) and a “mixture of experts” approach (see Section 2.1.3). The output of the multiple GPRs is overlaid, and data with high deviations between the actual measured I_{SC} and the predicted I_{SC} are iteratively filtered out.

5.1 Introduction

The first step in measurement data analysis should always be the validation and filtering of outliers, especially when measurements are performed under conditions that are difficult to control, as is generally the case for long-term outdoor experiments. Disruptions in normal operating conditions can significantly affect the results of an analysis. Since there is usually little information about the exact effects of outliers on the measurements, statistical analysis is often impeded. In general, the superposition of natural uncertainty on measured values with outlier deviations leads to a complex distribution.

For the evaluation of PV outdoor data, the PV community lacks standardizations for outdoor data filtering [32]. Common filtering approaches often rely on threshold filtering, such as imposing thresholds on values of plane of array irradiation (POA), nominal output power, ambient temperature, or module temperature. While thresholding has the advantage of simplicity, it generally leads to information loss as valid data is removed, and at the same time, invalid data points are not always reliably excluded. Furthermore, Jordan and Kurtz showed that different filtering approaches can lead to different results in PV degradation rate estimation [109]. Lindig *et al.* provide a comprehensive overview of common filtering approaches, highlighting their advantages and disadvantages [32].

An interesting approach is presented by Hansen, who proposes filtering by fitting a diode model to IV characteristics and removing data for which the diode model parameters are unreasonable [110]. This method, while potentially challenging due to the complexities of diode model parameterization, goes beyond simply imposing thresholds on individual measured values by filtering based on the shape of the complete IV characteristics, thus imposing restrictions on how the measured values correlate.

This chapter presents two different approaches, how to filter PV outdoor data. The first approach, published in [111], utilizes correlations in various dimensions of the data to identify and filter out implausible data points.

5.2 Mahalanobis distance filter

In the following, I will present a PV outdoor filtering approach that seeks to utilize the many correlations between the various measured dimensions in the data to evaluate each data point's plausibility. To this end, I simultaneously consider all Standard Solar cell Parameters (SSPs) — short circuit current (I_{SC}), open circuit voltage (V_{OC}), the maximum power point current (I_{MPP}), and the maximum power point voltage (V_{MPP}) - along with meteorological data. By using the Mahalanobis distance in combination with well-known correlations between these various dimensions, I integrate all this data into a plausibility measure. I demonstrate the presented method using publicly available data published by the National Renewable Energy Laboratory (NREL) [6].

5.2.1 Concept

An anomaly, according to Hawkins, is defined as any observation that deviates significantly from other observations, arousing suspicion that it was generated by a different mechanism [112]. In outdoor datasets, we have many strongly correlated measurements. Furthermore, a wide range of models describe these correlations. In a way, these models describe the mechanisms behind the data that we do *not* consider an anomaly. Thus, by comparing measured values with expected (modelled) values based on other measured quantities, we can assess to what degree a data point fits with the expected mechanisms. Formally, this is expressed as:

$$\begin{aligned}\Delta X_i &= X_i - X_{\text{expected},i} \\ dX_i &= \frac{\Delta X_i}{\sigma_{\Delta X}},\end{aligned}\tag{5.1}$$

where X_i is the i -th measurement of a particular parameter, $X_{\text{expected},i}$ is the corresponding expected value, and $\sigma_{\Delta X}$ is the standard deviation of ΔX over all measurements. From the NREL datasets introduced in Section 2.5.1, I used the time series of SSPs: I_{SC} , V_{OC} , I_{MPP} and V_{MPP} ; the module temperature, T_{Mod} , and the POA irradiation, G_{POA} , as well as GHI and DHI.

For the parameter G_{POA} , I use 5.1 and substitute $X = G_{\text{POA}}$. The expected G_{POA} (X_{expected}) is computed from the measured GHI and DHI using SSDP (see Section 2.5.3) and the given timestamp and coordinate (in Eugene and Cocoa). For the albedo, a value of 20% is assumed for the test locations in Eugene and Cocoa [113]. Note that occasionally there are missing values, and for simplicity, incomplete data points are removed.

In a similar fashion, I treat the SSPs. Here, I use the correction procedure 2 of

Table 5.1: Overview of dataset size and share of the filtered data for the 22 NREL datasets presented in Section 2.5.1.

	Eugene		Cocoa	
module type	dataset size	filtered	dataset size	filtered
mSi0166	42,908	12.8%	35,669	13.7%
mSi0188	42,773	13.8%	38,012	13.7%
mSi460A8	42,755	15.4%	37,864	15.3%
xSi12922	42,829	13.7%	37,905	14.1%
HIT05667	42,912	13.2%	37,313	12.8%
aSiMicro03036	42,969	15.7%	37,949	16.0%
aSiTandem72-46	42,905	16.1%	38,109	15.3%
aSiTriple28324	42,353	14.8%	37,407	14.2%
CdTe75638	41,959	11.7%	37,993	13.0%
CIGS8-001	42,791	13.3%	37,860	11.7%
CIGS39017	42,312	13.2%	33,791	12.0%
total	879,338	data points	13.9%	filtered

the IEC60891 standard (see Section 2.3.2) to correct each measured SSP to a set of standard conditions. In this work, the standard conditions are defined as $T_{\text{Mod},S} = 25^\circ\text{C}$ and $G_{\text{POA},S} = 500\text{Wm}^{-2}$. Note that I do not use the more commonly applied $G_{\text{POA},\text{STC}} = 1000\text{Wm}^{-2}$, as this high irradiation is somewhat atypical for actual operating conditions. Thus, I substitute a corrected SSP for X in Equation 5.1. The expected value, X_{expected} , is set to the mean of the corrected SSP. The correction procedure 2 of the IEC60891 standard requires six correction coefficients, which were obtained from regressions of the SSPs to the measured T_{Mod} and G_{POA} .

The Mahalanobis distances, $d_{M,i}$, are computed in five dimensions (dI_{SC} , dV_{OC} , dI_{MPP} , dV_{MPP} , dG_{POA}) using Equation 2.2. To filter the dataset, I use a threshold distance of $d_M = d_{M,\text{threshold}} \approx 3.884$, representing the distance where the CDF of a Mahalanobis distance distribution of a five-dimensional normally distributed vector reaches the quantile of $q = 99\%$.

5.2.2 Mahalanobis distance filtering results

The Mahalanobis distance filter is applied to the NREL datasets as described in Section 2.5.1 and, such, to in total $\approx 880,000$ data points divided in 22 datasets (11 modules at the two locations). Table 5.1 comprehensively shows the size of the used datasets (after measurement points with missing GHI and DHI are filtered out) and the share (percentage) of the data the Mahalanobis filter removes. In total $\approx 13.9\%$ of the data points are filtered out. For the individual datasets the share ranges from 11.7% to 16.1%.

To discuss the filtering results in more detail I hereinafter focus on the exemplary chosen Eugene mSi0166 dataset. The threshold distance of the Mahalanobis distance filter was set to filter 1% of the data in case the variables are normal distributed. Thus, the result of 12.8% filtered data for the Eugene mSi0166 dataset indicates a variable distribution with a significant deviation from a normal distribution.

Figure 5.2.1 shows the calculated Mahalanobis distance for the approximately 43000 data points (dark blue) as well as the theoretical Mahalanobis distance of a five dimensional (5D) normally distributed variable and the set threshold (orange). The Mahalanobis distance distribution appears to be heavily tailed, showing a strong deviation from the expected Mahalanobis distance for a normally distributed 5D variable. This indicates the presence of more systematic deviations from the applied physical models to utilize the correlations in the various dimensions of the data.

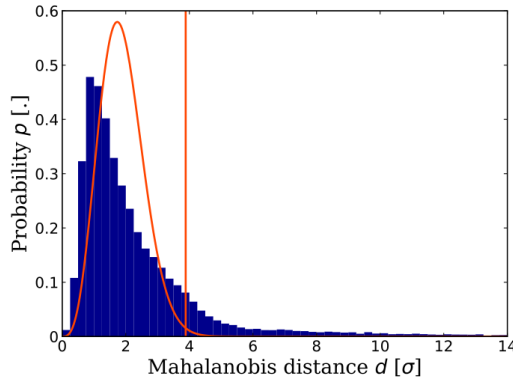


Figure 5.2.1: Histogram of the calculated Mahalanobis distance d for the Eugene mSi0166 dataset (dark blue) and theoretical Mahalanobis distance for a five dimensional normal distributed random variable (orange). The orange vertical line marks the used threshold to filter the dataset.

Figure 5.2.2 shows the scatter density plots for the modelled versus measured G_{POA} for the Eugene mSi0166 dataset. The colour of the points in the scatter plot represents the estimated local density of scatter points, with a logarithmic colour scale. From the unfiltered dataset in Figure 5.2.2, we see that the SSDP estimate of G_{POA} generally matches the measured data quite closely. The scatter points closely follow the identity line, with the density of the scatter points rapidly decreasing by more than an order of magnitude as they are located further from the identity line.

Comparing the unfiltered and filtered dataset in Fig. 5.2.2 (12.8% filtered out), one can see, that the Mahalanobis filter reliably removes measurements, where

the with SSDP modelled G_{POA} deviates from the measured G_{POA} .

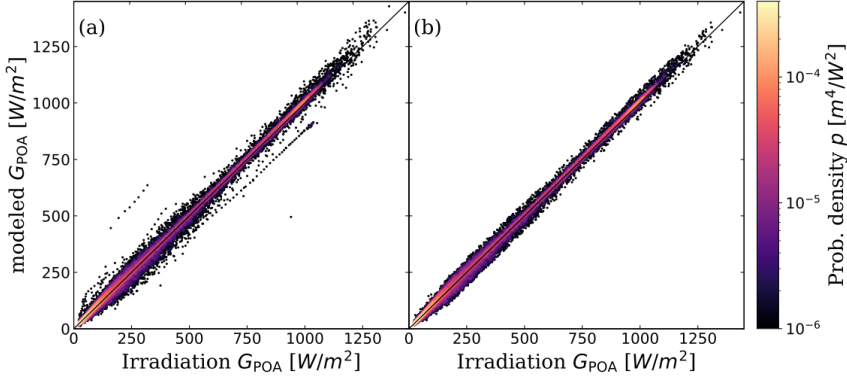


Figure 5.2.2: Scatter density plots of the modelled versus measured G_{POA} for the (a) unfiltered and (b) filtered Eugene mSi0166 dataset. The straight black line is the identity line, where the modelled and measured G_{POA} are equal.

Figure 5.2.3 shows scatter density plots for the corrected I_{SC} against measured G_{POA} for the unfiltered and filtered Eugene mSi0166 dataset. For irradiance values of $G_{\text{POA}} > 500 \text{ Wm}^{-2}$, the current correction is generally reliable. However, below this value, the corrected values of the unfiltered dataset in Figure 5.2.3 exhibit considerable scatter. The highest densities of scatter points are around the same corrected current values as for $G_{\text{POA}} > 500 \text{ Wm}^{-2}$, indicating that in most cases, the correction also works at low irradiation values. However, there is a clear structure in the scatter density plot where the corrected currents split into two branches, with one branch exhibiting considerably lower corrected currents. This branching appears to be the effect of partial shading of the module, occurring primarily in the mornings and evenings when the sun is low. Note that the irradiation sensors are positioned higher than the modules, according to the setup pictures in [6]. Thus, the modules may be partially shaded when the sensors are not. Partial shading appears to be the main cause of systematic deviations from the physical model in the dataset.

The corrected I_{SC} of the filtered dataset in Figure 5.2.3 shows that the large scatter for $G_{\text{POA}} < 500 \text{ Wm}^{-2}$ is trimmed off, thereby effectively filtering partial shading from the dataset. Finally, in Fig. 5.2.4, I show the scatter density plots for the corrected V_{OC} as a function of the measured T_{Mod} for the Eugene mSi0166 dataset. Here it is quite notable that for temperatures above 310 K ($\approx 37^\circ\text{C}$) there is considerably less scatter in the corrected V_{OC} values. Also here, it appears the partial shading events lead to an increased scatter. The filtered dataset exhibits considerably less scatter. In general, the Mahalanobis filter presented here is applicable when models are available for the mechanisms

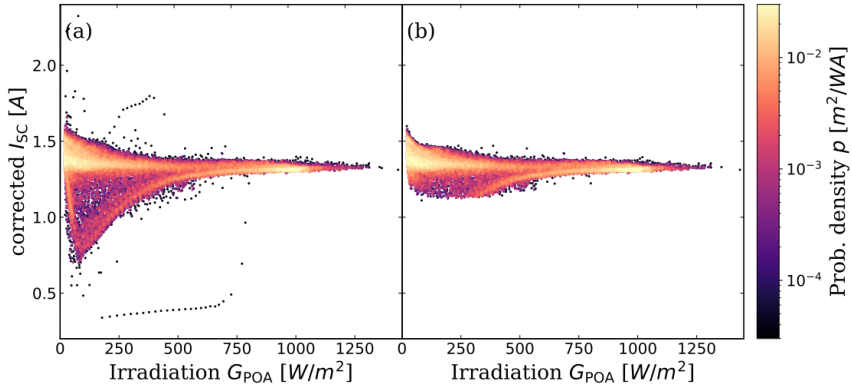


Figure 5.2.3: Scatter density plot of corrected I_{sc} versus measured G_{POA} for the (a) unfiltered and (b) filtered Eugene mSi0166 dataset.

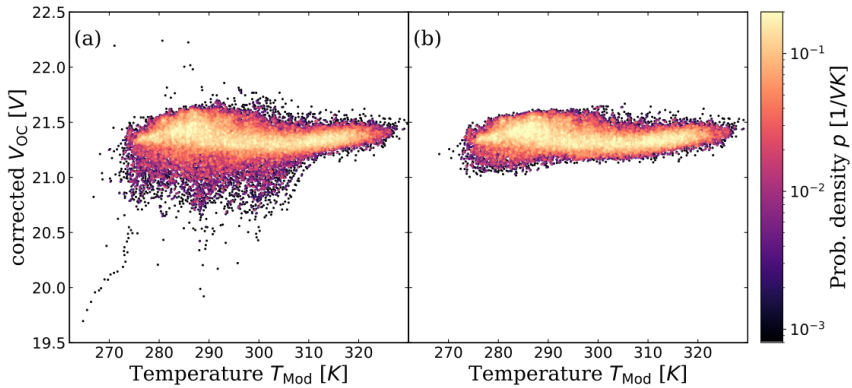


Figure 5.2.4: Scatter density plot of corrected V_{oc} versus T_{Mod} for the (a) unfiltered and (b) filtered Eugene mSi0166 dataset.

sought after. On one hand, this allows for a particularly targeted data filter. On the other hand, the method will filter out any data not fitting with the model, which are not necessarily measurement errors. In this case, I filtered out partial shading, which arguably is not a measurement error. Nonetheless, for many (most) performance analysis methods, the partial shading data is not desired.

5.3 Gaussian Process Regression filter

In the following, I will present a PV outdoor filtering approach aimed at identifying data points where the measured in-plane irradiation (G_{POA}) does not match the effective irradiation on the module. To achieve this, I consider the temperature and irradiation dependencies of the short circuit current (I_{SC}) using multiple Gaussian Process Regressions (GPRs) and a mixture of experts approach. Utilizing the inherent assumption in GPR of statistically distributed measurement uncertainties, I iteratively filter data points that show very large deviations from the expected values. Unlike the Mahalanobis filter presented in Section 5.3, this method does not assume a physical model. This concept is demonstrated on outdoor datasets provided by TÜV Rheinland, described in more detail in Section 2.5.1.

5.3.1 Concept

From the TÜV datasets, I used the time series of I_{SC} , T_{Mod} , and G_{POA} . The individual datasets consist of up to $N \approx 65,000$ data points. Assuming rather constant conditions of the PV module and neglecting second-order spectral effects, I_{SC} is dependent only on T_{Mod} and G_{POA} . To describe these dependencies of the short circuit current, I split the dataset into 20 random subsets and fit a GPR $f_{I_{\text{SC}},i} : \mathbb{R}^2 \rightarrow \mathbb{R}$ model (see Section 2.1.3) to each of the $i = 1, 2, \dots, 20$ subset. The splitting of the dataset is necessary as the computation time of GPRs scales with N^3 , i.e. I chose a “mixture of experts” approach [11] to reduce the computation time (see Section 2.1.3). For each subset, I fit the three hyperparameters of a two dimensional (2D) RBF kernel with an added white noise kernel. The kernel of each $f_{I_{\text{SC}},i}$ is defined as

$$K(X, \tilde{X}) = \exp \left(\frac{\|x_1 - \tilde{x}_1\|^2}{2l_1^2} + \frac{\|x_2 - \tilde{x}_2\|^2}{2l_2^2} \right) + \sigma^2 \mathbb{I}. \quad (5.2)$$

Where $X = [x_1, x_2] = [G_{\text{POA}}, T_{\text{Mod}}]$ represents the 2D input space of irradiation and temperature, \mathbb{I} is the identity matrix and l_1 , l_2 , and σ are the three hyperparameters to be optimized during training of each $f_{I_{\text{SC}},i}$. The hyperparameters are optimized using the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm implemented in SciPy [114].

Subsequently, I compute the predicted $I_{\text{SC},i,n}$ for each of the $n = 1, 2, \dots, N$ points of the dataset from each individual GPR and the, with the uncertainty $\sigma_{I_{\text{SC},i,n}}$ weighted, overlay of the $i = 1, 2, \dots, 20$ predictions $I_{\text{SC},n}$. I assume here statistical independence of the respective predictions to compute the uncertainty $\sigma_{I_{\text{SC},n}}$ neglecting the covariance between individual predictions. I justify this

assumption with the individual optimized hyperparameters and the individual subsets used for training¹⁸. The overlay of predictions has two advantages. First, an outlier only affects the prediction of a single $f_{I_{SC,i}}$ and thus only the i -th individual prediction $I_{SC,i,n}$ for each point n (5% of the predictions). Second, the predictions $I_{SC,i,n}$ of the i -th GPR that are trained with many outliers, as well as predictions $I_{SC,i,n}$ of the GPRs at points n with T_{Mod} and G_{POA} conditions far from the i -th training input, will result in $I_{SC,i,n}$ predictions with high associated uncertainty $\sigma_{I_{SC,i,n}}$. Overlaying multiple individual GPRs will suppress the influence of outliers and lack of information in individual GPRs¹⁹.

To filter the data, I use a z -score filter with a very high threshold deviation of the measured I_{SC} , set at 20 standard deviations $\sigma_{I_{SC,n}}$ from the predicted $I_{SC,n}$ ²⁰. This filter only excludes data points, which cannot be explained by statistical uncertainty in the measurement. In other words, the GPR predicts a probability density of measuring I_{SC} at the respective input conditions T_{Mod} and G_{POA} and the filter is designed to exclude data points, that show a very high deviation from this predicted probability distribution. Note, that this approach inherits the assumption, that all measurements exhibit ideally only statistical noise. This assumption is not valid in the presence of systematic errors. However, in case the data would only exhibit only statistical noise the filter presented would filter a share in the order of $10^{-87\%}$ (probability of a Normal distribution to find a measurement outside of 20 standard deviations) and such virtually no data. The high threshold is chosen to account for any possible changes over time of I_{SC} , which are not described within this filter. Effects of degradation and seasonality in the short-circuit current typically show low rates and exhibit low amplitudes over the approximately three-year span of operation in the datasets under consideration. The high threshold ensures that these effects are not filtered out.

Since this high threshold only excludes extreme outliers, the filtering process is applied iteratively. In each iteration, the most extreme outliers are detected and removed, which reduces the subsequent standard deviation (as the standard deviation, due to the quadratic weighting of deviations, is strongly affected by outliers). The iterative process can be continued until no extreme outliers are detected. This iterative method effectively rejects the tails of distributions, ensuring that the resulting distributions have no data points beyond 20 standard

¹⁸Note that the incorporation of the covariance would lead to slightly higher uncertainties. The used approximation consequently leads to an underestimation of the uncertainty.

¹⁹Each GPR only uses 5% of the available data as training input and might not cover the complete 2D input space of T_{Mod} and G_{POA} .

²⁰Note that since the incorporation of the covariance would lead to slightly higher uncertainties the used threshold of $20\sigma_{I_{SC,n}}$ reflects an effective threshold slightly lower than 20 standard deviations

deviations from the mean. Thus the distributions are filtered specifically to make them more Gaussian. In this work I limit the iterative process to a fixed number of iterations.

5.3.2 Gaussian Process Regression filtering results

The presented filtering concept is applied to all 45 datasets described in Section 2.5.1. The individual size of the datasets as well as the share of filtered data after four iterations is summarized in Tab. 5.2. The share of filtered data in the individual datasets varies between 1.6% and 4.4%, where in total 2.6% of the $\approx 2,200,000$ data points are filtered out. It is noticeable, that the share of filtered data does not vary much among one location. Since the modules operated at the same location are positioned close to each other, the probability of different effective irradiation of the PV module and the irradiation sensor leading to outliers is correlated. Furthermore, the same irradiation sensor is used for all datasets at the same location, resulting in (partial) shading of the irradiation sensor affecting all datasets of the location in the same way.

Figure 5.3.1 shows a scatter density plot of measured G_{POA} versus I_{SC} for the exemplary chosen unfiltered (a) and filtered (b) Italy CdTe1 dataset after one iteration of the applied GPR filter. As expected, the short circuit current is linear in the POA irradiation. It is observed, that the density of measured $[G_{\text{POA}}, I_{\text{SC}}]$ pairs decreases with increasing deviation from the linear trend, i.e. slightly different effective irradiation on the PV module and the irradiation sensor is more likely than large discrepancies. Furthermore, I find no clusters with a high scatter density deviating from the observed linear relationship. The first iteration of applying the GPR filter clearly removes only extreme outliers with a share of 1.0% of the dataset.

The second (scatter density shown in Fig.5.3.2 (a)) and third iteration of applying the GPR filter remove 1.1% and 0.8% of the exemplary chosen Italy CdTe1 dataset, respectively. The fourth and final iteration removes 0.4% of the Italy CdTe1 dataset. This trend of a decreasing share of removed data with more iterations of applying the GPR filter is observable across all 45 datasets. This occurs as the distributions becomes more normal and less tailed. As the distributions becomes more Gaussian, the share of data that can be accurately predicted using GPRs increases, resulting in fewer points being filtered out with each iteration. Finally, Fig. 5.3.2 (b) shows the scatter density plot of measured G_{POA} versus I_{SC} for the exemplary chosen filtered Italy CdTe1 dataset after all four iterations of the applied GPR filter. One can see, that the filter reliably removes outliers with high deviations from the expected relationship between

Table 5.2: Overview of dataset size and share of the filtered data for the 45 TÜV datasets described in Chapter 2.5.1. In total 2.6% of 2,205,946 data points are filtered out due to four iterations of the applied GPR filter.

	Pheonix (USA)		Thuwal (Saudi Arabia)		Cologne (Germany)		Ancona (Italy)		Chennai (India)	
module type	dataset size	filtered	dataset size	filtered	dataset size	filtered	dataset size	filtered	dataset size	filtered
c-Si	57,837	2.4%	31,857	2.2%	29,003	4.0%	57,091	3.6%	45,138	1.8%
ClGS1	65,158	1.9%	35,850	1.9%	47,500	4.0%	61,432	3.7%	50,204	2.1%
ClGS2	65,097	2.0%	35,862	1.9%	47,902	3.8%	61,208	3.8%	50,244	2.1%
ClGS3	64,351	2.3%	35,434	1.9%	48,561	4.4%	60,742	3.7%	49,601	1.9%
ClGS4	60,696	2.0%	32,539	2.0%	31,139	3.2%	58,060	3.7%	47,243	1.9%
CdTe1	64,754	1.9%	35,529	1.9%	49,037	3.9%	61,362	3.7%	49,974	1.6%
CdTe2	65,739	1.6%	36,132	1.9%	46,160	3.4%	61,141	3.5%	50,071	1.7%
CdTe3	61,282	1.7%	32,952	1.8%	30,098	3.6%	58,514	3.4%	47,271	1.6%
a-Si/ μ c-Si	58,703	1.9%	32,615	1.9%	30,097	2.9%	57,880	3.4%	46,886	1.9%
total	563,617	1.9%	308,770	1.9%	359,497	3.8%	537,430	3.6%	436,632	1.8%

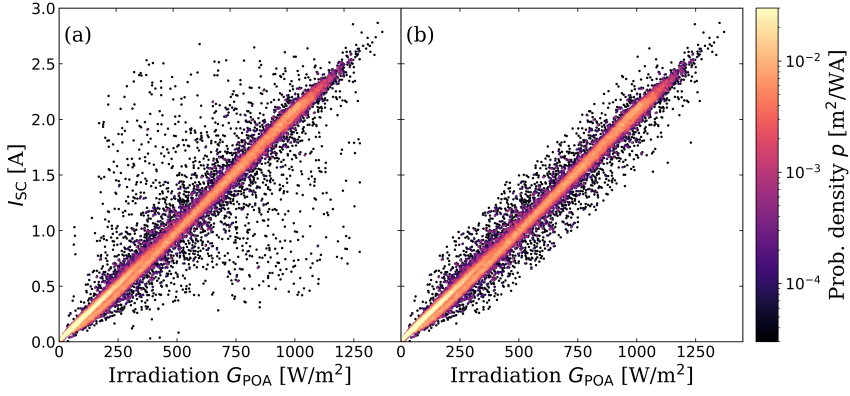


Figure 5.3.1: Scatter density plot of I_{SC} versus G_{POA} for the (a) unfiltered and (b) filtered Italy CdTe1 dataset after one iteration of applying the GPR filter.

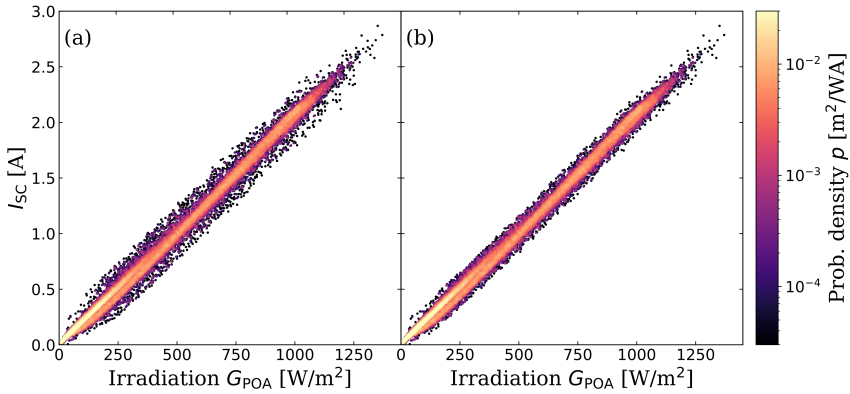


Figure 5.3.2: Scatter density plot of I_{SC} versus G_{POA} for the filtered Italy CdTe1 dataset after (a) two and (b) four iterations of applying the GPR filter.

I_{SC} and G_{POA} .

The applied GPR filter removes reliably data points where the measured G_{POA} , T_{Mod} and I_{SC} show a high deviation from the expected relationship. In general, the GPR filter presented here is applicable on any dataset, where G_{POA} , T_{Mod} and I_{SC} are available. *IV* characteristics, where the measured I_{SC} does not match (in a given uncertainty) the expectation based on the measured G_{POA} and T_{Mod} are hard to interpret, since an interpretation w.r.t. the PV modules performance or efficiency is only possible taking the conditions the *IV* is measured at into account. On one hand, the filter removes outliers that originate from different irradiation levels on the PV module and the irradiation sensor reliably. On the other hand, the filter only accounts for this particular cause of systematic errors in the dataset.

The major limitation of the filter is the inability to differentiate between a signal in the data and a systematic deviation, i.e. the GPRs do not incorporate the temporal development of the I_{SC} due to acclimatization, seasonality and degradation effects. These effects (the signal in the time series of I_{SC}) are overlaid with statistical noise and systematic deviations and leads to further broadening of the probability distribution measuring I_{SC} at conditions G_{POA} and T_{Mod} . A reduced threshold for the filter as well as each iteration of applying the filter increases the probability of removing valid data points. Due to this limitation the filter is utilized to remove only outliers with high deviations from the expectation, that cannot be explained with any mechanism relevant for the further analysis. Using scatter density plots I further demonstrate, that the data, which is filtered out does not show any clustering regarding high scatter densities at specific I_{SC} and G_{POA} combinations deviating from the linear relationship. Since the effects of acclimatization, degradation and seasonality exhibit, in general, low amplitudes, i.e. long lengthscales for changes in I_{SC} are observed, the signal in I_{SC} cannot explain unclustered deviations from the predicted probability density targeted and excluded with the applied filter.

5.4 Summary and Outlook

A variety of data filtering methods exist for the analysis of datasets for PV applications. However, most these filtering methods are based on simple thresholding. The Mahalanobis filter presents a plausibility filter concept, utilizing known correlations between the various measured quantities to define a measure of plausibility for each data point. To obtain such a measure of plausibility, the Mahalanobis distance combined with models to describe the correlations between several measured dimensions in a dataset are utilized. The models are fitted to the data in the dataset. This results in a very targeted filter, where data is filtered out if it does not conform to the selected models. I demonstrated this concept using the SSPs, T_{Mod} , G_{POA} , GHI and DHI. I show that a significant portion of the data is affected by partial shading. Furthermore, the method is effective in removing partial shading from the dataset, as partial shading does not conform to the selected models. We argue that this type of filter is, under circumstances, better at rejecting datapoints which underlie a specific mechanism (such as partial shading), which is not interesting for the further analysis than a purely statistical outlier metric. This advantage, however, is only an advantage if no interest exists in the different mechanisms, which cannot be described with the used physical models.

The Mahalanobis filtering concept may be extended to use more dimensions in the data, for example, provided wind speed, ambient temperature, G_{POA} , and T_{Mod} are available, the consistency between these dimensions may be validated with the Faiman model [115]. More generally, this approach may be used whenever a model is available which allows predicting one measured dimension from one or more other dimensions available in the data, and all data of interest more or less conform to the selected models. For my work I decided to *not* limit the data to only datapoints conforming to some selected physical models. For this reason I instead developed a filtering method based on purely statistical outlier metrics.

The results of the first presented filtering concept shows that partial shading is the main cause for systematic deviations from the applied physical models in the NREL datasets. In the second filtering concept, I do not assume any physical model but use a purely statistical outlier metric, where outliers, that cannot be described with Gaussian noise are filtered out. In this filter I only consider the measured POA irradiation G_{POA} and module temperature T_{Mod} as well as the short circuit current, I_{SC} . The underlying assumption is that the data can be described as Gaussian processes, the fundamental assumption of GPR. Thus, much like with the Mahalanobis filter, I compare measured and predicted I_{SC}

values. However, as I only consider G_{POA} , T_{Mod} and I_{SC} , I may use a simple z -score filter approach. As the distributions in our data are heavily tailed, I apply an iterative z -score filter.

The filter concentrates on data points, where the measured I_{SC} shows a high deviation from the predicted probability distribution based on the GPRs. I demonstrate the filter on several outdoor datasets, showing that in the filtered data I have little scatter in the G_{POA} and I_{SC} , indicating an effective filtering. The presented filter accounts for data points, where the measured performance (here I_{SC} as one specific performance measure) deviates from the expected performance of the PV module under the respective conditions. G_{POA} and T_{Mod} are in further analysis used to classify the measured PV module performance (i.e. the IV characteristic). The presented filter such excludes data points, where this classification based on the measured conditions would result in unreasonable conclusions about the module performance.

Limitations of the applicability of the GPR filter arise, since the effects of different irradiation levels is overlaid with other effects leading to deviations from the expected relationship between I_{SC} and G_{POA} and T_{Mod} . Variations in I_{SC} occur e.g. due to seasonality effects or degradation. The filter excludes only a small share of the data varying between 1.6% and 4.4% depending on the respective dataset. The scatter density plot further demonstrated, that the data, which is filtered out does not show any clustering, which indicates, that no signal resulting from acclimatization, seasonality and degradation is filtered out. In summary the applied metric excludes a small share of the dataset, which can be assigned to deviations between measured G_{POA} and T_{Mod} conditions and the effective conditions the PV module exhibits.

Analyzing PV Outdoor data

6.1	Introduction	94
6.2	A Gaussian process regression IV model	97
6.2.1	Filtering	97
6.2.2	ESPs	98
6.2.3	PCA	98
6.2.4	Mixture of Experts GPR	99
6.3	Results	101
6.3.1	Timeseries analysis	101
6.3.2	PR analysis	108
6.3.3	Physical analysis using the one-diode model	112
6.4	Validation	116
6.5	Summary and Outlook	123

In this chapter, I introduce a statistical model designed to analyse PV outdoor data. This model utilizes electrical performance measures, specifically the current-voltage (*IV*) characteristics and the extracted ESPs, alongside meteorological data, including plane of array irradiance (G_{POA}) and module temperature (T_{Mod}), as inputs over time.

First, a suitable filtering approach is applied to the data to ensure its quality and relevance. Following this, the ESPs undergo a principal component analysis (PCA), which reduces the dimensionality of the data while preserving its most significant features. Individual Gaussian process regressions (GPRs) are then trained on these principal components (PCs). Once the GPRs are trained, the model is capable of reproducing and predicting the complete *IV* characteristics at any given time t , for specified values of G_{POA} and T_{Mod} . This prediction includes an assessment of uncertainty, which is derived from data noise and the distance from the observations. This model serves as a versatile tool for various applications, such as analysing acclimatization effects, degradation trends, seasonal variations, and the performance ratio (PR) of PV modules or systems. Part of this work is submitted to “Progress in Photovoltaics: Research and Applications” and a preprint is published on Authorea [116].

6.1 Introduction

The reliability of PV modules is critically important for the amortization of PV systems. Long-term outdoor monitoring of PV modules is essential for analysing degradation, and, thus, also for assessing the quality and performance of PV products and technology. PV outdoor data generally includes electrical performance measures and additional meteorological and temperature measurements.

Electrical performance is commonly assessed through measurements of the complete current-voltage (IV) characteristics, which provide a reliable means to verify the status of individual modules over time[5, 117–119]. Meteorological and temperature measurements, such as module temperature (T_{Mod}) and plane-of-array irradiation (G_{POA}), are highly correlated with the measured performance, and thus of particular use to explain the measured electrical performance (see Section 2.5.1).

While many outdoor datasets provide the complete IV curve for analysis, the evaluation of outdoor PV data typically focuses on just a few key solar cell parameters: open-circuit voltage (V_{OC}), short-circuit current (I_{SC}), and the voltage and current at the maximum power point (V_{MPP} and I_{MPP}). This approach simplifies the analysis by reducing the complexity to only four parameters instead of the entire IV curve. However, this simplification can overlook potentially significant information. For example, relying solely on V_{OC} , I_{SC} , V_{MPP} , and I_{MPP} makes it impossible to distinguish between an IV curve with an "S"-shape (see for on overview of this phenomenon [120]) and one with high series resistance [27].

Another common way to simplify the information present in a measured IV characteristic is the extraction of the 5 one diode model parameters [110]. The one diode model is capable to extract more information then the SSPs as well as physical meaningful information, where its parameters provide a physical interpretation of the performance. While the one diode model is commonly used in the PV community, to obtain a robust and efficient one diode model parametrization method is a challenge itself. A more complete overview of proposed methods can be found in [24, 25, 121]. However, the one-diode model has inherent limitations: it cannot account for certain physical effects, such as IV curves with an S-shape, restricting its applicability as a feature vector for IV characteristics. Additionally, the complex non-linear optimization methods required to fit the one-diode model to measured IV curves pose further challenges. Arguably, the one diode model is beneficial for a physical interpretation of the IV characteristic under many circumstances. However, a more descriptive and comprehensive representation of the full IV curve shape can be achieved using

extended solar cell parameters (ESPs) [27]. The ESPs distill the information contained in the IV curve into ten key parameters (as detailed in Section 2.3.3), preserving much detail of the shape of the IV characteristic. By analyzing ten parameters instead of the conventional four or five, ESPs offer a more refined and accurate representation of the original IV curve. Moreover, these ten ESPs enable precise reconstruction of the original IV characteristic [27].

Analysing PV outdoor data remains challenging due to continuously changing conditions, such as variations in plane-of-array irradiance (G_{POA}) and module temperature (T_{Mod}). Individual IV characteristics, as well as their corresponding ESPs, cannot be directly compared without accounting for the influence of irradiance and temperature on the IV curve shape. Furthermore, additional factors such as (partial) shading and humidity affect data quality and introduce uncertainties. Another significant challenge in PV outdoor data analysis is the lack of standardized filtering methods within the PV community [32]. Different filtering approaches can lead to varying results, for instance, in PV degradation rate estimation [109].

In this chapter, I introduce a statistical model for analysing PV outdoor data. The model is designed to reproduce respectively predict the IV characteristic as well as its uncertainty at any given time and temperature and irradiation conditions. Such the model not only outputs the temperature- and irradiation-dependent temporal development of the IV characteristic, but also sets the output in context of a standard deviation arising through the uncertainty of the measurement itself and the lack of data, i.e. the measured IV characteristics are only available for distinct times at distinct conditions.

The presented model is depicted in Fig. 6.1.1. The arrows depict data flow between the processing steps, while the colour code and numbering reflects the structure of this chapter. At the root of the model stands PV outdoor raw data, acquired by TÜV Rheinland [5]. The data consists of up to three years of IV characteristic data as well as T_{Mod} , G_{POA} and the time t (up to approximately 60,000 data points each) for various commercial modules operated in different climate zones. Note that the modules were new at the time of installation, i.e. the data might be expected to show effects of acclimatization (a performance drop in the beginning of operation). The model builds upon a simple filtering concept (a brief overview is given in Section 6.2.1 and more details are discussed in Section 5.3.1), excluding incomplete data points as well as data where the effective irradiation on the PV module (i.e. I_{SC}) does not match the measured plane-of-array irradiation G_{POA} . The filtering approach ensures to keep a representative share of the available data, filtering only approximately 2 to 4% of the data. The time series of IV characteristics from the raw data is reduced to a 10-dimensional

time series of ESPs (see Section 6.2.2). After the ESP analysis and the filtering routine, a Principal Component Analysis (PCA) is applied to the ESP time series (see Section 6.2.3), resulting in a time series of 10 linearly uncorrelated principal components (PCs). These PCs, along with time t , plane-of-array irradiance (G_{POA}), and module temperature (T_{Mod}), serve as inputs for training multiple Gaussian Process Regressions (GPRs), where each PC is treated independently (see Section 6.2.4). To manage the challenges associated with large datasets ²¹, the PC time series are segmented into monthly subsets.

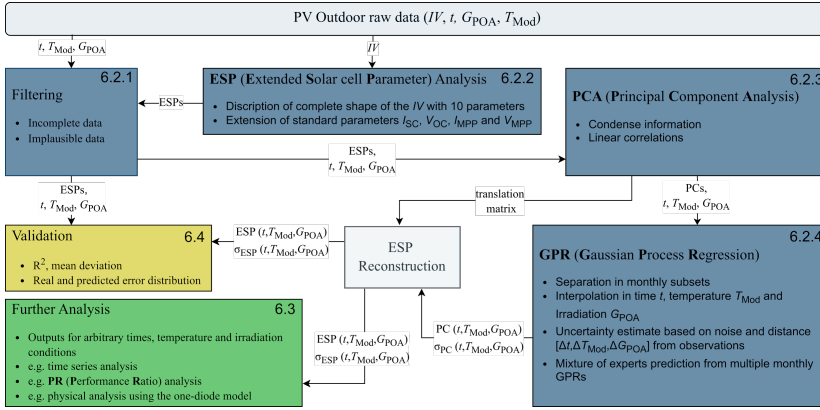


Figure 6.1.1: Illustration of the presented GPR IV model and its PV outdoor data processing steps. The arrows depict the data flow from one step to another, while the numbers and colour code give a reference to the structure of the chapter.

Once the individual GPRs are optimized and trained, their outputs can be combined to predict the PCs, and thereby reconstruct the ESPs and the complete shape of the IV characteristic for any given t , G_{POA} , and T_{Mod} . Note that the model is not capable respectively designed to extrapolate (e.g. forecasting). Additionally, the GPRs provide insights into the uncertainty of the predictions, which arises from both data noise and gaps in the input data. Specifically, predictions for outputs far from observed data points will have higher uncertainty compared to those close to observations.

The result of this approach is a compact yet comprehensive description of the temporal, temperature, and irradiance dependencies of the modules ESPs, and thus IV characteristics. The applicability of this concept is demonstrated using the TÜV datasets, which consist of up to three years of outdoor monitoring data for various modules operating in different climate zones, as thoroughly

²¹In particular the computational complexity, as GPR optimization scales with N^3 , where N is the number of data points.

presented in Section 2.5.1²². Furthermore, the output of the GPRs can be used for validation (see Section 6.4) using the $\text{ESP}(t, T_{\text{Mod}}, G_{\text{POA}})$ predictions at the measured input conditions. Such, the from the *IV* characteristics determined ESPs are reconstructed, while test and training data is temporally separated with a common training to test ratio of 80:20.

In Section 6.2, I provide a detailed explanation of the model concept, highlighting how it integrates the data filtering approach introduced in Section 5.3.1, ESPs, PCA, and multiple GPRs. Following this, the concept is applied to an exemplary dataset from the TÜV outdoor data in Section 6.3, where one possible model output is compared with a traditional performance ratio (PR) analysis approach. Finally, in Section 6.4, I evaluate the model's accuracy before summarizing the major findings in Section 6.5.

6.2 A Gaussian process regression IV model

6.2.1 Filtering

An essential step in all PV outdoor data analyses is the proper filtering of data to eliminate erroneous measurements (e.g., missing data, physically unreasonable *IV* shapes, or shading-induced errors). To achieve this, ESPs (see Section 6.2.2) are extracted from an *IV* dataset using the PV-CRAZE library [31]. During ESP extraction, PV-CRAZE automatically flags *IV* data with unusual properties. Specifically, PV-CRAZE flags *IV* characteristics that do not allow for reliable ESP extraction due to issues such as excessive noise or insufficient data. Additionally, non-monotonic *IV* curves are identified and removed, including cases where the slope at V_{OC} is negative or where the maximum power point current I_{MPP} exceeds the short-circuit current I_{SC} . Furthermore, incomplete data points (those missing *IV* characteristics, G_{POA} , or T_{Mod}) are also discarded. This initial filtering step effectively removes the most obvious errors from the dataset, ensuring higher data reliability for further analysis.

Beyond filtering for obvious errors, we also apply the more nuanced filtering approach presented in the Section 5.3.1. This method involves filtering out data points where there is a mismatch in the effective irradiation levels experienced by the PV module and the irradiation sensor, addressing the most common source of systematic errors in the datasets under consideration. The short-circuit current (I_{SC}) is modelled as a function of G_{POA} and T_{Mod} using multiple GPRs,

²²Note that the presented model is exemplary applied to PV outdoor data consisting of complete *IV* characteristics alongside G_{POA} and T_{Mod} , but is in principal applicable to any kind of PV outdoor data consisting of an electrical performance measure (e.g. AC or DC power) and data capturing the conditions, the performance is measured at.

randomly splitting the complete dataset into 20 subsets, similar to a bootstrap aggregating approach. We compare the overlaid prediction of the 20 GPRs with the measured I_{SC} . A high filter threshold is set so that only significant discrepancies between the expected and actual I_{SC} are filtered out. Furthermore, we use an iterative approach, resulting in an effective filter excluding roughly 2 to 4% of the data points. Details to the filtering approach can be found in the Section 5.3.1.

6.2.2 ESPs

For completeness and readability, I begin with a brief introduction to the extended solar cell parameters (ESPs); a more detailed discussion can be found in Section 2.3.3. The ESPs comprise a set of 10 parameters that describe the IV characteristics of a solar cell. As the name suggests, the ESPs extend beyond the standard solar cell parameters (SSPs) of V_{OC} , I_{SC} , V_{MPP} , and I_{MPP} . The ESPs extend the SSPs with the slopes of the IV at short and open circuit (G_{SC} and R_{OC}). Savitzki-Golay filters are used to obtain noise robust estimators for the slopes at short and open circuit [28]. Furthermore, the ESPs add two additional key points ('upper quasi maximum power point' [I_{qmp+} , V_{qmp+}] and 'lower quasi maximum power point' [I_{qmp-} , V_{qmp-}]) to comprehensively describe the complete shape of the IV characteristic.

The lower and upper quasi-maximum power points incorporate information on the shape of the IV characteristic between short circuit (SC) and maximum power point (MPP) and between MPP and open circuit (OC), respectively. These points are defined using the power curve ($P(V)$) associated with the IV . V_{qmp+} is defined as the voltage between MPP and OC, where the power curve exhibits the largest difference from a straight line drawn between the two coordinates [V_{MPP}, P_{MPP}] and [$V_{OC}, 0$]. I_{qmp+} is then the associated current of the IV at $V = V_{qmp+}$. Analogously, V_{qmp-} is defined as the voltage, where the power curve exhibits the largest difference from a straight line drawn between the two coordinates [$0, I_{SC}$] and [V_{MPP}, P_{MPP}] and I_{qmp-} is the associated current at $V = V_{qmp-}$. Thus, the ESPs provide a general-purpose parameterization of IV characteristics using ten parameters. Compared to other parameterization methods, ESPs result in less information loss, as they capture more subtle features of the IV curve.

6.2.3 PCA

After fitting the ESPs and applying the filters described in the previous sections, I obtain a time series of 10 ESPs that vary along with G_{POA} and T_{Mod} . Since

these time series are generally correlated, I further apply a principal component analysis (PCA) to extract possible linear correlations. To this end, the ESPs are normalized by their respective mean values. As described in Section 2.1.1, a PCA involves a change in the basis of the coordinate system used to describe a set of n -dimensional data points. Illustratively, this basis change aligns the first principal component (PC) with the direction that best represents the linear fit to the m -dimensional point cloud of the dataset. Subsequent PCs represent directions that best fit the data, subject to the constraint of being orthogonal to all previous PCs.

Formally, the j -th PC can be expressed as:

$$\text{PC}j = \sum_{k=1}^{10} p_{j,k} \frac{\text{ESP}k - \overline{\text{ESP}k}}{\overline{\text{ESP}k}}, \quad (6.1)$$

where the matrix elements $p_{j,k}$ of the transformation matrix \mathbb{P} serve as the weights that translate the mean-normalized 10 ESPs into the 10 PCs. The primary advantage of using PCs to describe a dataset is the elimination of redundant information in the form of linear correlations. Additionally, a PCA can reduce the dimensionality of the dataset by revealing whether it can be adequately described with $d < m$ principal components. However, it is important to note that in this context, the PCA is not necessarily applied to reduce the data's dimensionality but rather to concentrate redundant information that manifests as linear correlations among the ESPs.

6.2.4 Mixture of Experts GPR

After applying the PCA, I obtain 10 PC time series, which are treated separately in the subsequent analysis. To describe the temporal, G_{POA} , and T_{Mod} dependencies of each $\text{PC}j$, I use multiple Gaussian Process Regressions (GPRs). As described in Section 2.1.3, a GPR is a probabilistic method that allows for predictions of unsampled inputs, assuming a Gaussian distribution of measurements. Note that this assumption refers to a single measurement being Gaussian distributed in case it is repeated under the exact same conditions and time, i.e. I do not assume that the modelled time series is Gaussian distributed. I utilize GPRs to interpolate desired outputs, e.g. a time series of IV characteristics under constant G_{POA} and T_{Mod} conditions, and to estimate the associated uncertainty based on the available discrete data. I.e. the model is designed to predict the IV characteristic under given inputs of t , G_{POA} and T_{Mod} within the parameter space spanned by the observations (not for extrapolation respectively forecasting).

Each PC j time series contains of up to $N \approx 65,000$ data points. To mitigate the computational burden, which scales with N^3 for a single GPR (see Section 2.1.3), I adopt a "mixture of experts" approach [11]. I split each PC j time series into n monthly subsets, indexed by i , and train individual GPRs $f_{\text{PC}j_i} : \mathbb{R}^3 \rightarrow \mathbb{R}$. Each $f_{\text{PC}j_i}$ is optimized using five hyperparameters within a 3D radial basis function (RBF) kernel, which is multiplied by a constant kernel and supplemented by a white noise kernel. The kernel of each $f_{\text{PC}j_i}$ is defined as:

$$K(X, \tilde{X}) = c \exp \left(\frac{\|x_1 - \tilde{x}_1\|^2}{2l_1^2} + \frac{\|x_2 - \tilde{x}_2\|^2}{2l_2^2} + \frac{\|x_3 - \tilde{x}_3\|^2}{2l_3^2} \right) + \sigma^2 \mathbb{I}, \quad (6.2)$$

where $X = [x_1, x_2, x_3] = [G_{\text{POA}}, T_{\text{Mod}}, t]$ represents the 3D input space of irradiation, temperature, and time, \mathbb{I} is the identity matrix, and c , l_1 , l_2 , l_3 , and σ are the five hyperparameters optimized during training. The use of an RBF kernel assumes that each PC j varies smoothly with time, temperature, and irradiation. The most influential hyperparameters for the accuracy of GPR predictions are the lengthscales parameters l_1 , l_2 , and l_3 , which define the correlation length in the directions of time t , irradiation G_{POA} , and temperature T_{Mod} . By training separate GPRs for each month, the model allows for temporal variation in these parameters, thereby adjusting the sensitivity to small changes over time. The hyperparameters are optimized using the limited-memory BFGS algorithm with parameter constraints (L-BFGS-B) implemented in SciPy [114]. For a more reliable result the optimization routine is applied 5 times with random start parameters. The parameter constraints ensure a limitation of the parameter space to meaningful variation of the hyperparameters.²³

After training the individual $10n$ GPRs $f_{\text{PC}j_i} : \mathbb{R}^3 \rightarrow \mathbb{R}$, the models can predict $\text{PC}j_i(G_{\text{POA}}, T_{\text{Mod}}, t)$ and the associated uncertainty $\sigma_{\text{PC}j_i}(G_{\text{POA}}, T_{\text{Mod}}, t)$ ²⁴. In addition to providing predictions and uncertainties, the model outputs can be interpreted as probability densities, with each GPR prediction corresponding to a Gaussian distribution $\mathcal{N}(\text{PC}j_i(G_{\text{POA}}, T_{\text{Mod}}, t), \sigma_{\text{PC}j_i}(G_{\text{POA}}, T_{\text{Mod}}, t))$.

The ESPs can be derived from the PC predictions using the matrix elements $\hat{p}_{k,j}$ of the inverse transformation matrix \mathbb{P}^{-1} as follows:

$$\text{ESP}k_i(G_{\text{POA}}, T_{\text{Mod}}, t) = \overline{\text{ESP}k} \left(1 + \sum_{j=1}^{10} \hat{p}_{k,j} \text{PC}j_i(G_{\text{POA}}, T_{\text{Mod}}, t) \right). \quad (6.3)$$

This allows the model to predict the complete IV characteristic at any given

²³E.g. if a lengthscales parameter becomes very large w.r.t the variation of the dataset in the respective direction, a GPR outcome is visually not affected by further increasing the lengthscales parameter.

²⁴Note that the uncertainty arises from both data noise and the distance from observed data points.

t , G_{POA} , and T_{Mod} . The uncertainty of the ESP predictions can be calculated from $\sigma_{\text{PC}j_i}(G_{\text{POA}}, T_{\text{Mod}}, t)$ using the following equation:

$$\sigma_{\text{ESP}k_i}(G_{\text{POA}}, T_{\text{Mod}}, t) = \overline{\text{ESP}k_i} \sqrt{\sum_{j=1}^{10} \hat{p}_{j,k}^2 \sigma_{\text{PC}j_i}^2(G_{\text{POA}}, T_{\text{Mod}}, t)}, \quad (6.4)$$

where I assume that the $j = 1, 2, \dots, 10$ predictions $\text{PC}j_i(G_{\text{POA}}, T_{\text{Mod}}, t)$ are statistically independent, i.e., the covariance matrix elements $\sigma_{\text{PC}j_i, \text{PC}l_i}$ are zero for $j \neq l$. Given that the PCs are linearly uncorrelated (by definition of the PCA) and that each GPR uses individual hyperparameters for predictions, this assumption is a reasonable approximation. Consequently, the ESP predictions can also be interpreted as Gaussian probability densities $\mathcal{N}(\text{ESP}k_i(G_{\text{POA}}, T_{\text{Mod}}, t), \sigma_{\text{ESP}k_i}(G_{\text{POA}}, T_{\text{Mod}}, t))$, since the linear combination of statistically independent Gaussian variables is itself Gaussian.

The model's accuracy is highest for test points $[G_{\text{POA}}, T_{\text{Mod}}, t]$ close to observed data, and the model is not intended for extrapolation. Within these limitations, the model is capable of extracting, for example, temperature or irradiation dependencies at a specific time, or the temporal development of the *IV* characteristic under constant conditions. By splitting the dataset into $i = 1, 2, \dots, n$ monthly subsets, each $f_{\text{PC}j_i}$ becomes an "expert" for the respective month. As the temporal distance from observations increases, the prediction $\text{PC}j_i(G_{\text{POA}}, T_{\text{Mod}}, t)$ and consequently $\text{ESP}k_i(G_{\text{POA}}, T_{\text{Mod}}, t)$ exhibit increased uncertainty.

6.3 Results

The presented concept is applied to all 45 datasets provided by TÜV Rheinland. For simplicity, I use an exemplary dataset of a CIGS module (CIGS4 in Tab. 5.2) operated for approximately 2.8 years, beginning on the 1st of November 2013, in Ancona, Italy to provide a compact overview of possible applications of the presented GPR *IV* model concept. I provide in the following three examples of how to utilize the presented concept to analyse PV outdoor data, concentrating on the temporal development of the shape of the complete *IV* characteristic, the determination of the performance ratio (PR) in comparison to a classical temperature-corrected performance ratio PR_T and utilizing the models output for a physical analysis using the one-diode model.

6.3.1 Timeseries analysis

Figure 6.3.1 shows one possible output of the GPR prediction, a timeseries of the 10 PCs over the complete operation time of approximately 2.8 years at

constant $T_{\text{Mod}} = 40^\circ\text{C}$ and $G_{\text{POA}} = 750 \text{ Wm}^{-2}$ for the exemplary chosen Italy CIGS4 dataset. Note that the constant conditions are chosen to represent a medium high irradiation with a realistic module temperature at such irradiation levels. Compare e.g. to typical nominal operating cell temperature (NOCT) at $G_{\text{POA}} = 800 \text{ Wm}^{-2}$ and $T_{\text{amb}} = 20^\circ\text{C}$ being in the range of 40 to 50°C [122].

The output shown in Fig. 6.3.1 is produced with a five-month sliding window approach, i.e. I use the output $\text{PC}j_i(t)$ (constant T_{Mod} and G_{POA}) of $f_{\text{PC}j_i}$ for $i = m - 2, m - 1, m, m + 1, m + 2$ to predict $\text{PC}j(t)$ for times t in month m using

$$\text{PC}j(t) = \sigma_{\text{PC}j}^2(t) \sum_{i=m-2}^{m+2} \frac{\text{PC}j_i(t)}{\sigma_{\text{PC}j_i}^2(t)}, \quad (6.5)$$

where the standard deviation of the prediction $\sigma_{\text{PC}j}(t)$ is determined from the uncertainty outputs of the five GPR predictions according to

$$\sigma_{\text{PC}j}(t) = \frac{1}{\sqrt{\sum_{i=m-2}^{m+2} \frac{1}{\sigma_{\text{PC}j_i}^2(t)}}} \quad (6.6)$$

and the prediction of the five GPRs are weighted with the respective uncertainty output for the predictions. The window size of five months is chosen to suppress the influence of observations with a temporal distance of more than three months, i.e. three months are used as a threshold for the temporal correlation. The 95% confidence interval, defined as $[\text{PC}j(t) - 1.96\sigma_{\text{PC}j}(t), \text{PC}j(t) + 1.96\sigma_{\text{PC}j}(t)]$, is denoted with the blue area.

From Fig. 6.3.1 one can see, that the amplitude of the PCs gets lower with increasing number. This is expected with the nature of a PCA, reflecting the direction of most information (highest variance) in the first PC and decreasing degree of information with higher number of the respective PC. The first (Fig. 6.3.1 (a)) and second (Fig. 6.3.1 (b)) PC show a clear seasonality in the temporal development with a period of one year. Note, that all PCs reflect a linear combination of the mean-normalized ESPs $\frac{\Delta \text{ESP}}{\text{ESP}}$. To visualize this linear combination Fig. 6.3.2 (a) and (b) exemplary show the direction of PC1 and PC2, respectively. PC1 (negative in Fig. 6.3.1 (a)) reflects a direction towards negative deviations from the mean currents and mean G_{SC} and (with a lower amplitude) in the direction of negative deviations from the mean voltages as well as in the direction of positive deviation from the mean R_{OC} . PC2 (positive in Fig. 6.3.1 (b)) reflects a direction towards positive deviations from the mean currents and mean G_{SC} and (as well with a lower amplitude) in the direction of negative deviations from the mean voltages as well as in the direction of positive deviation from the mean R_{OC} . First of all, the directions of PC1 and PC2 show, that R_{OC} ,

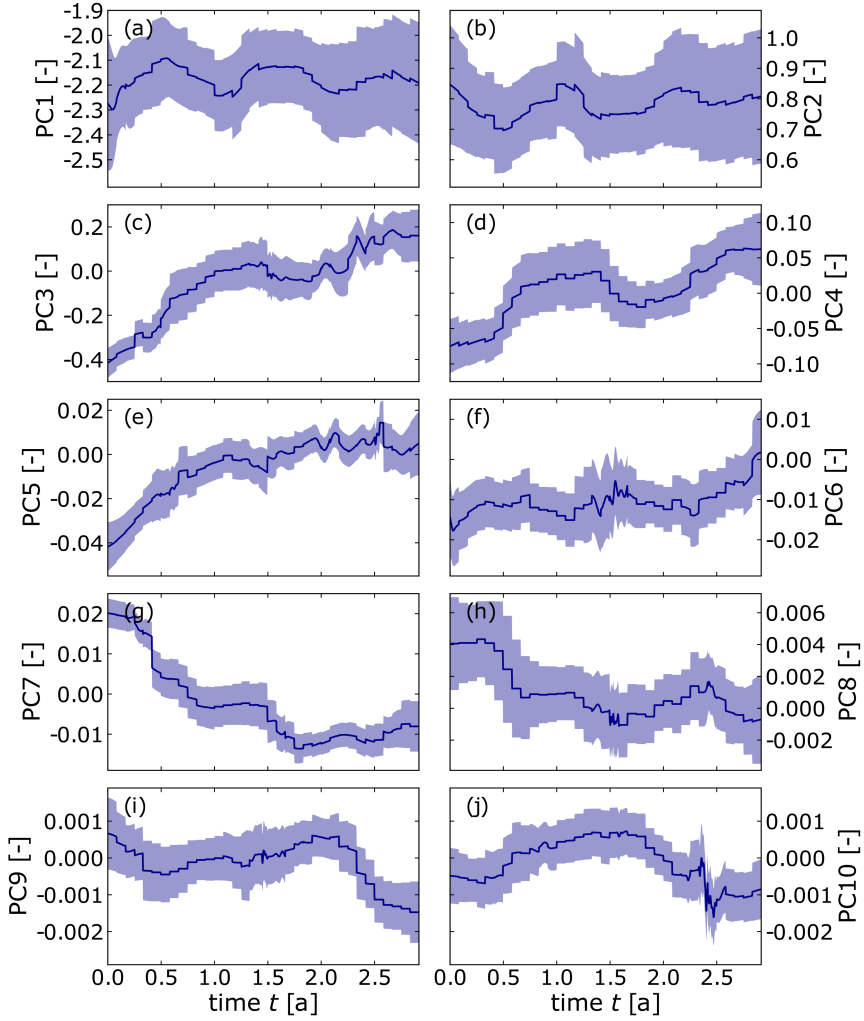


Figure 6.3.1: Output of the GPR prediction with a 95% confidence interval for a timeseries of the 10 PCs over the complete operation time of approximately 2.8 years at constant $T_{\text{Mod}} = 40^\circ\text{C}$ and $G_{\text{POA}} = 750 \text{ Wm}^{-2}$ for the exemplary chosen Italy CIGS4 dataset. For the prediction a five-month sliding window approach is chosen, where $\text{PC}_j(t)$ and $\sigma_{\text{PC}_j(t)}$ are determined for times t in the m -th month according to 6.5 and 6.6, respectively.

the current parameters and G_{SC} exhibit a higher (with the mean normalized) variance throughout the dataset w.r.t. the voltage parameters. Note, that R_{OC} and G_{SC} reflect slopes of the IV characteristics, where a slight change in the condition may lead to big discrepancy between the two slopes at OC respectively SC. The higher variance of the current parameters w.r.t. the variance of the voltage parameters also reflects their sensitivity to condition changes. While the current parameters are expected to change e.g. linear with irradiation, the voltage parameters are expected to exhibit a logarithmic dependency on the irradiation. Beside the overall variance of the ESPs, the directions of PC1 and PC2 set their temporal development in Fig. 6.3.1 into context of describing a strong seasonality in the current parameters and G_{SC} , while the seasonality of the voltage parameters and R_{OC} is less pronounced, where the seasonality in PC1 is partially compensated by the seasonality of PC2.

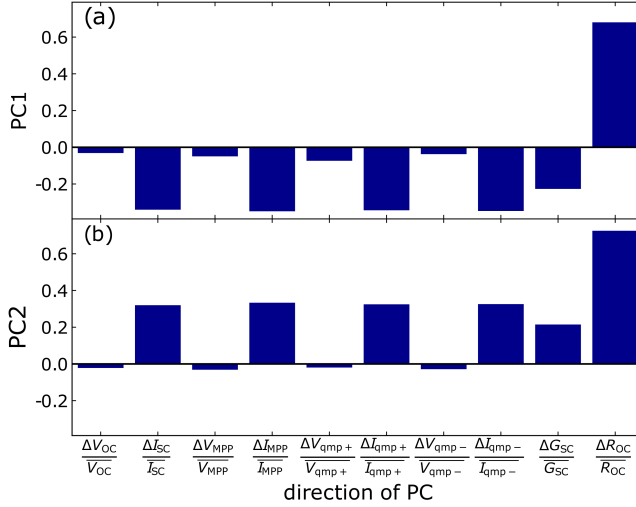


Figure 6.3.2: Direction of (a) PC1 and (b) PC2 for the Italy CIGS4 dataset. PC1 reflects a direction towards negative deviations from the mean currents and mean G_{SC} and in the direction of negative deviations from the mean voltages as well as in the direction of positive deviation from the mean R_{OC} . PC2 reflects a direction towards positive deviations from the mean currents and mean G_{SC} and in the direction of negative deviations from the mean voltages as well as in the direction of positive deviation from the mean R_{OC} .

The example of the direction of PC1 and PC2 show how linear correlations of the ESPs are reflected in the PCs. Regarding the uncertainty of PC1 and PC2, I find rather high uncertainty estimates w.r.t. the amplitude of PC1 and PC2. PC1 and PC2 clearly describe a seasonal behaviour, typically induced by metastabilities in CIGS modules. Considering the typical time scales of

metastabilities ranging from days to years, the observed high uncertainty might be a result of lower time scales, which are not tracked accordingly due to the GPRs temporal lengthscale parameter or just due to a lack of data points needed to fully resolve shorter time scales.

The directions of PC3 till PC10 are less intuitive to interpret, but also reflect a linear combination of the mean-normalized ESPs $\frac{\Delta \text{ESP}}{\text{ESP}}$. The temporal development of PC3, PC4 and PC5 (Fig. 6.3.1 (c), (d) and (e)) all show an increase of the respective PC over the operation time, where the increase of PC3 can be considered more or less linear, the increase in PC4 is superimposed with a periodic variation and the increase in PC5 saturates towards the end of the operation time considered. The sixth PC (Fig. 6.3.1 (f)) appears to be constant over most of the time period with a slight increase at the beginning and towards the end of the time period. PC7 (Fig. 6.3.1 (g)) depicts a step-wise decrease after half a year and one and a half years of operation. The uncertainty of the prediction of PC3 till PC7 is comparably low w.r.t. the uncertainty of PC1 and PC2. Considering, that the temporal development of PC3 till PC7 all exhibit a (stepwise) increase or decrease, one might argue, that these PC are somehow connected to effects of degradation and acclimatization. The lower uncertainty such reflects a higher confidence of present degradation and acclimatization effects w.r.t. seasonality effects. PC8, PC9 and PC10 (Fig. 6.3.1 (h), (i) and (j)) all show very low amplitudes 3 orders of magnitude lower than the first PC. The temporal development clearly shows a signal, which implies, that these PCs carry information present in the dataset. The amplitude of this information, however, is at a level, where it becomes negligible w.r.t. the amplitudes of PC1 till PC7 once the ESPs are reconstructed from the PCs.

While the representation of the PCs may be beneficial to decompose several effects visible in the performance development of the operated CIGS module, the reconstruction of the ESPs allow for a more meaningful interpretation regarding PV module performance changes during outdoor operation. Figure 6.3.3 shows the output of the GPR prediction for the 10 reconstructed ESP timeseries including the 95% confidence interval at the same constant $T_{\text{Mod}} = 40^\circ\text{C}$ and $G_{\text{POA}} = 750 \text{ Wm}^{-2}$ condition for the exemplary chosen Italy CIGS4 dataset. $\text{ESP}_k(t)$ and $\sigma_{\text{ESP}_k}(t)$ are determined from $\text{PC}_j(t)$ and $\sigma_{\text{PC}_j}(t)$ (see Fig. 6.3.1) according to 6.3 and 6.4. A clear seasonality superimposed with a slight linear decrease is visible for the 4 current parameters I_{SC} , I_{MPP} , $I_{\text{qmp+}}$ and $I_{\text{qmp-}}$ (Fig. 6.3.3 (a), (c), (e) and (g)). For the 4 voltage parameters V_{OC} , V_{MPP} , $V_{\text{qmp+}}$ and $V_{\text{qmp-}}$ (Fig. 6.3.3 (b), (d), (f) and (h)) the temporal development shows a more or less linear decrease, where the slope of $V_{\text{qmp+}}$ is observed to flatten out. The slope of the IV characteristic at open circuit is described by the parameter R_{OC}

(Fig. 6.3.3 (i)), which is observed to be constant over the operation period with a high uncertainty of the prediction ²⁵. For the slope of the *IV* characteristic at short circuit, depicted by the parameter G_{SC} (Fig. 6.3.3 (j)), one can observe a steady increase superimposed with a periodic change over the course of the operation time. Considering the temporal development of PC1 and PC2 in Fig. 6.3.1 as well as their directions in terms of the mean-normalized ESPs $\frac{\Delta ESP}{ESP}$ (see Fig. 6.3.2), I find the expected seasonality in the current parameters and G_{SC} , while the seasonality contribution of PC1 and PC2 are mostly compensated by each other in the temporal development of the voltage parameters and R_{OC} .

The representation of the complete shape of the *IV* characteristic in terms of the ESPs enables to quantify the temporal development of the PV modules performance. First of all it is clear, that a strong seasonality affects the current level of the *IV* characteristic, while a slight degradation is visible in both the current and voltage parameters. The difference in the development of I_{SC} and I_{MPP} already indicates a loss in the FF over time, since the degradation appears to be more pronounced in I_{MPP} . Considering the temporal development of G_{SC} as well as of V_{qmp+} one can follow, that the slope at SC as well as the complete shape of the *IV* between SC and MPP changes, while the shape of the *IV* between MPP and OC remains quantitatively the same (rather constant R_{OC} as well as similar temporal development of V_{OC} , V_{MPP} and V_{qmp-}).

Regarding the uncertainty estimate of the ESPs I find the highest confidence (lowest relative uncertainty) for the temporal development of the parameter G_{SC} (Fig. 6.3.3 (j)). The uncertainty on the temporal development of the 4 current parameters (Fig. 6.3.3 (a), (c), (e) and (g)) as well as on the temporal development of V_{qmp+} (Fig. 6.3.3 (f)) is comparably higher. Nonetheless, one can clearly observe a high confidence for the qualitative temporal development, i.e. the amplitude of seasonality and linear degradation is comparable to the amplitude of the 95% confidence interval. For the 3 voltage parameters V_{OC} , V_{MPP} and V_{qmp-} (Fig. 6.3.3 (b), (d) and (h)) and especially for the parameter R_{OC} (Fig. 6.3.3 (i)), I find a comparably high uncertainty on the temporal development, where the amplitude of the signal (temporal variation of the prediction) is lower than the respective confidence interval.

²⁵Note that negative R_{OC} values do not occur in *IV* characteristics measured under steady conditions. The uncertainty estimate for R_{OC} thus primarily arises from deviations toward higher R_{OC} values, resulting in a one-sided tailed distribution. One could argue that the graph and the lower confidence interval in Fig. 6.3.3 (i) should be constrained to zero as a physically meaningful boundary for R_{OC} . However, since Gaussian process regression (GPR) predictions follow a Gaussian distribution, the uncertainty output is also Gaussian. As a result, the representation in Fig. 6.3.3 (i) is mathematically more accurate.

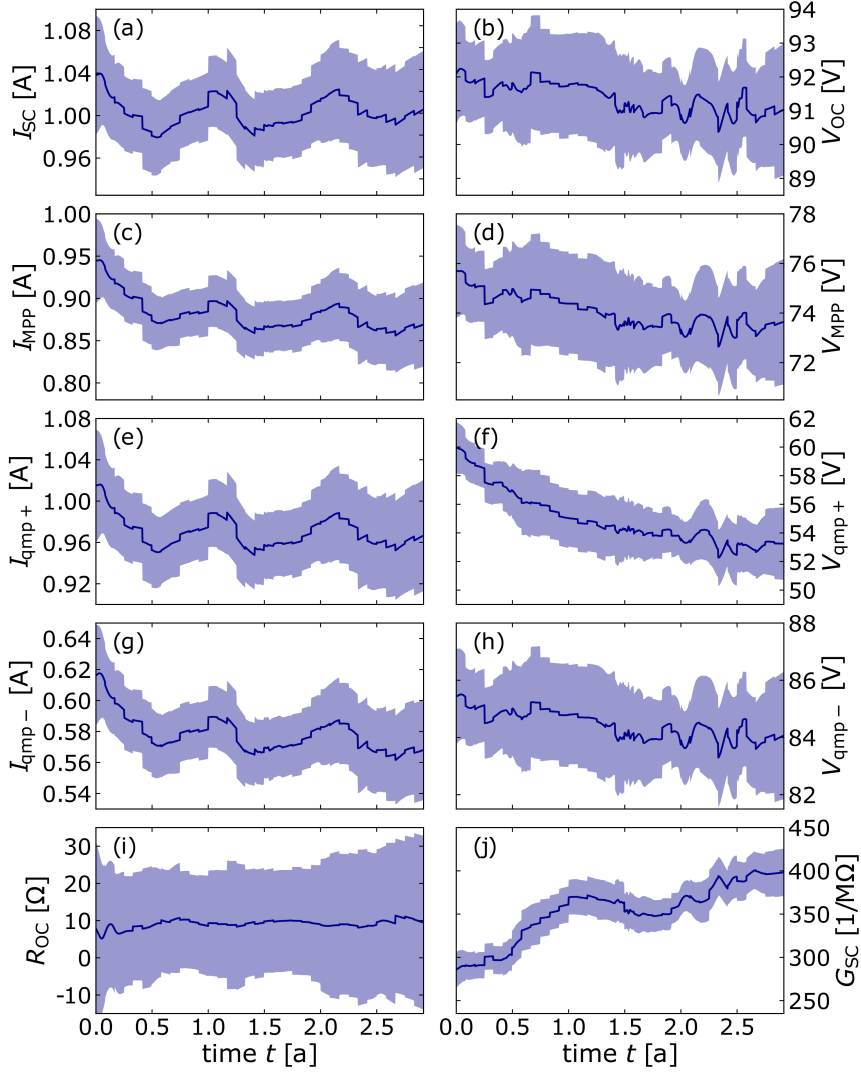


Figure 6.3.3: Output of the GPR prediction with a 95% confidence interval for a timeseries of the 10 reconstructed ESPs over the complete operation time of approximately 2.8 years at constant $T_{Mod} = 40^\circ\text{C}$ and $G_{POA} = 750 \text{ Wm}^{-2}$ for the exemplary chosen Italy CIGS4 dataset. For the prediction a five-month sliding window approach is chosen, where $ESP_k(t)$ and $\sigma_{ESP_k}(t)$ are determined for times t in the m -th month according to 6.3 and 6.4 from $PC_j(t)$ and $\sigma_{PC_j}(t)$.

6.3.2 PR analysis

As stated in Section 2.5.2 a common way to verify a PV modules performance over time is the performance ratio PR respectively the temperature corrected performance ratio PR_T . Usually both PR and PR_T are referenced to STC conditions. Since STC conditions are uncommon in real operation scenarios and GPRs are not able to extrapolate to inputs far away from the discrete inputs they are trained on, a models output like shown in Fig. 6.3.3 is not suitable to evaluate on classical STC referenced PR analysis. However, the PR for distinct constant conditions can be computed, comparing the MPP power output $P_{MPP,GPR}(t) = V_{MPP,GPR}(t)I_{MPP,GPR}(t)$ to the initial $P_{MPP,GPR}(t=0)$. Since there is no producer information on the nominal output power at the respective conditions available and taking only one single value as reference ($P_{ref} = P_{MPP,GPR}(t=0)$) might result in high deviations when $P_{MPP,GPR}(t=0)$ is not accurate, I define the initial $P_{MPP,GPR}$ as the mean of $P_{MPP,GPR}(t)$ in the first 7 days of operation $\overline{P_{MPP,GPR}}|_{t \leq 7d}$ ²⁶. Thus, I define a generalized PR from the GPR IV model at constant G_{POA} and T_{Mod} conditions as

$$PR_{GPR}(t) = \frac{P_{MPP,GPR}(t)}{\overline{P_{MPP,GPR}}|_{t \leq 7d}} = \frac{V_{MPP,GPR}(t)I_{MPP,GPR}(t)}{\overline{V_{MPP,GPR}}|_{t \leq 7d} \overline{I_{MPP,GPR}}|_{t \leq 7d}}. \quad (6.7)$$

Regarding the uncertainty estimate of the such defined PR_{GPR} , one has to consider the uncertainty of $V_{MPP,GPR}$ and $I_{MPP,GPR}$. As already discussed in Section 6.2.4 the predictions of the ESPs can be interpreted as a Gaussian probability density $\mathcal{N}(ESP_k(G_{POA}, T_{Mod}, t), \sigma_{ESP_k}(G_{POA}, T_{Mod}, t))$. For the uncertainty prediction of $PR_{GPR}(t)$ the correlation between the two probability densities of $V_{MPP,GPR}(t)$ and $I_{MPP,GPR}(t)$ needs to be considered²⁷. Since all ESPs are a linear combination of the PCs, I can generalize 6.4 to determine the covariance between two ESPs via the matrix elements $\hat{p}_{k,j}$ of the inverse translation matrix \mathbb{P}^{-1} and the matrix elements $\sigma_{PC_j}^2$ of the diagonal covariance matrix of the PCs.

²⁶Note, that the choice of the reference output power is not arbitrary and might affect the absolute PR.

²⁷Arguably, the uncertainty on the reference power output $\overline{P_{MPP,GPR}}|_{t \leq 7d}$ might need to be considered as well. Here, I neglect the influence of the uncertainty on the reference power output for two reasons. First, this reference value would be ideally substituted by available manufacturer data on the modules performance, making the task of the uncertainty estimate of the reference power output obsolete. Second, a possible deviation of the actual initial power output and the chosen reference affects all computed PR estimates (and the later computed uncertainties) by the same factor. This way the uncertainty originating from the reference cannot be directly compared to the GPR uncertainty prediction resulting from noise and missing data and would need to be treated separately.

Formally I compute the covariance between $\text{ESP}k(t)$ and $\text{ESPl}(t)$ given by

$$\sigma_{\text{ESP}k, \text{ESPl}}(t) = \overline{\text{ESP}k \text{ESPl}} \sum_{j=1}^{10} \hat{p}_{k,j} \sigma_{\text{PC}_j}^2(t) \hat{p}_{l,j}, \quad (6.8)$$

where $\hat{p}_{l,j} = \hat{p}_{j,l}^\top$ are the matrix elements of the transposed matrix $\mathbb{P}^{-1\top}$. With the covariance of $V_{\text{MPP,GPR}}(t)$ and $I_{\text{MPP,GPR}}(t)$ the standard deviation $\sigma_{\text{PR}_{\text{GPR}}}(t)$ can be computed to

$$\begin{aligned} \sigma_{\text{PR}_{\text{GPR}}}(t) = & \left[\left(\frac{\sigma_{V_{\text{MPP,GPR}}}(t) I_{\text{MPP,GPR}}(t)}{\overline{V_{\text{MPP,GPR}}|_{t \leq 7d}} \overline{I_{\text{MPP,GPR}}|_{t \leq 7d}}} \right)^2 + \left(\frac{V_{\text{MPP,GPR}}(t) \sigma_{I_{\text{MPP,GPR}}}(t)}{\overline{V_{\text{MPP,GPR}}|_{t \leq 7d}} \overline{I_{\text{MPP,GPR}}|_{t \leq 7d}}} \right)^2 \right. \\ & \left. + \frac{2V_{\text{MPP,GPR}}(t) I_{\text{MPP,GPR}}(t) \sigma_{I_{\text{MPP,GPR}}, V_{\text{MPP,GPR}}}(t)}{(\overline{V_{\text{MPP,GPR}}|_{t \leq 7d}} \overline{I_{\text{MPP,GPR}}|_{t \leq 7d}})^2} \right]^{\frac{1}{2}}. \quad (6.9) \end{aligned}$$

While the standard deviation of the probability density for $\text{PR}_{\text{GPR}}(t)$ can be computed, defining a confidence interval similar to the confidence interval given in Fig. 6.3.1 and Fig. 6.3.3 directly from $\sigma_{\text{PR}_{\text{GPR}}}(t)$ might not be accurate, since the probability density of a product of two correlated Gaussian distributed random variables is in general not Gaussian. While there is an expression for an exact solution of the probability density (compare e.g. with [123]), the computation of the exact solution involves an infinite sum of modified Bessel functions and is not applicable without an approximation (e.g. setting a limit for the infinite sum). To compute a confidence interval for $\text{PR}_{\text{GPR}}(t)$, I use an alternative approach approximating the probability density with a numerical solution. I generate $n_r = 10^6$ random distributed variables according to the 2D multivariate Gaussian distributed probability density ($\mathcal{N}(I_{\text{MPP,GPR}}(t), \sigma_{I_{\text{MPP,GPR}}}(t)), \mathcal{N}(V_{\text{MPP,GPR}}(t), \sigma_{V_{\text{MPP,GPR}}}(t))$) with covariance $\sigma_{I_{\text{MPP,GPR}}, V_{\text{MPP,GPR}}}(t)$ and compute the probability density of $\text{PR}_{\text{GPR}}(t)$ for each t^{28} .

Figure 6.3.4 shows the temporal development of the PR_{GPR} (blue, including 95% confidence interval) for the constant $G_{\text{POA}} = 750 \text{ Wm}^{-2}$ and $T_{\text{Mod}} = 40^\circ\text{C}$. In Fig. 6.3.4 the PR_{GPR} is further compared to a classical determined monthly temperature corrected performance ratio PR_T (red, including 95% confidence) referenced to $G_{\text{POA,ref}} = 750 \text{ Wm}^{-2}$ and $T_{\text{Mod,ref}} = 40^\circ\text{C}$. For the determination of the classical PR_T I use 2.27 and substitute the STC with the reference conditions, insert $P_{\text{DC}} = P_{\text{MPP,meas}}$ and further use $\overline{P_{\text{MPP}}|_{t \leq 7d}}$ from the GPR prediction as reference for the initial output power at the reference conditions.

²⁸Note that the probability density computed this way approaches the exact solution with $n_r \rightarrow \infty$.

This results in the expression

$$\text{PR}_T = \frac{750 \text{ Wm}^{-2}}{G_{\text{POA,meas}}} \frac{P_{\text{MPP,meas}}}{\overline{P_{\text{MPP,GPR}}|_{t \leq 7d}} (1 + \gamma(T_{\text{Mod,meas}} - 40^\circ\text{C}))}, \quad (6.10)$$

where the temperature correction coefficient γ is determined by a linear fit to the filtered dataset. The shown temperature corrected PR is a commonly used monthly PR_T , where all measurements in a narrow irradiation band $550 \text{ Wm}^{-2} < G_{\text{POA}} < 950 \text{ Wm}^{-2}$ around the reference irradiation $G_{\text{POA}} = 750 \text{ Wm}^{-2}$ are used for the determination of the average $\overline{\text{PR}_T}$ and its 95% confidence interval²⁹. Thus, the 95% confidence interval is defined as the interval between the boundaries set by the 2.5% and 97.5% quantile of the distribution of determined PR_T in the respective month.

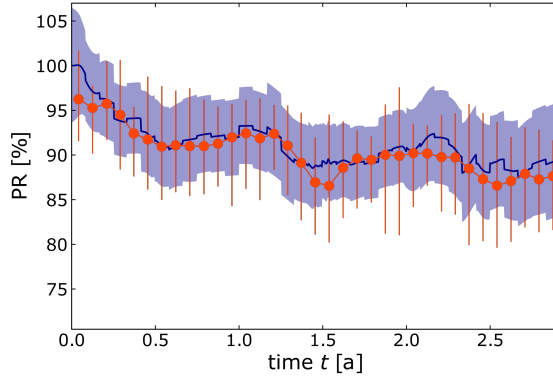


Figure 6.3.4: GPR performance ratio $\text{PR}_{\text{GPR}}(t)$ (blue) prediction for the constant $G_{\text{POA}} = 750 \text{ Wm}^{-2}$ and $T_{\text{Mod}} = 40^\circ\text{C}$ and classical PR_T estimate (red) using a narrow irradiation band filter of 400 Wm^{-2} referenced to $G_{\text{POA}} = 750 \text{ Wm}^{-2}$ and $T_{\text{Mod}} = 40^\circ\text{C}$. The blue area and the red errorbars depict a 95% confidence determined from the numerical approximation to the theoretical probability distribution of $\text{PR}_{\text{GPR}}(t)$ and the distribution of the determined PR_T , respectively.

One can observe a strong acclimatization effect within the first few months of operation overlaid with a seasonality and a long-term degradation for both PR estimates. The PR_T , determined with the classical monthly approach matches qualitatively the PR_{GPR} , while a slightly higher uncertainty can be observed for the classical approach. Furthermore, I observe for some months PR_T values several percent below the PR_{GPR} , remaining within the 95% confidence interval of the GPR prediction. While the GPR prediction uses all available information

²⁹Note that for the determination of PR_T the dataset is also filtered with the GPR filter as described in Section 5.3.1, before applying the additional narrow band irradiation filter. Such the presented PR_T is already corrected from possible influences due to outliers resulting from shading of the module respectively the irradiation sensor.

of the performance measures at different conditions within one month ³⁰, the narrow irradiation band filter reduces the dataset to $\approx 22\%$ and such to a share of the dataset, that can not be expected to be representative for the complete dataset.

Despite reducing the dataset to a small share of itself, these narrow irradiation band filters are a common choice while determining the PR respectively PR_T of PV modules respectively systems. Beside effects of low-light performance of PV modules gaining more influence on the result, for the determination of the PR a broader irradiation band filter would further increase the influence of the operating temperature, highly correlated with the G_{POA} . Such a PR estimate would experience an increased uncertainty. For the PR_T the effect of the operating temperature is reduced due to the linear correction coefficient, but can still affect the result in the presence of non-linear temperature effects. Figure 6.3.5 underlines the increasing uncertainty of the PR_T with increasing width of the irradiation band filter. In comparison to the PR_T shown in Fig. 6.3.4, the PR_T shown in Fig. 6.3.5 is determined with a doubled width of the irradiation band filter, averaging P_{MPP} for all measurements for G_{POA} with $350 \text{ Wm}^{-2} < G_{POA} < 1150 \text{ Wm}^{-2}$. Note that despite the choice of a broad irradiation band filter, the dataset is reduced to $\approx 42\%$ in this case.

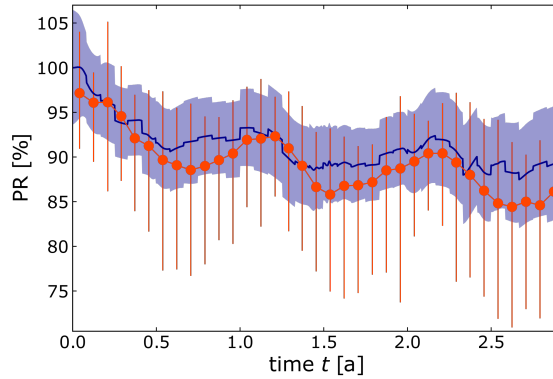


Figure 6.3.5: GPR performance ratio $PR_{GPR}(t)$ (blue) prediction for the constant $G_{POA} = 750 \text{ Wm}^{-2}$ and $T_{Mod} = 40^\circ\text{C}$ and classical PR_T estimate (red) using a broad irradiation band filter of 800 Wm^{-2} referenced to $G_{POA} = 750 \text{ Wm}^{-2}$ and $T_{Mod} = 40^\circ\text{C}$. The blue area and the red errorbars depict a 95% confidence determined from the numerical approximation to the theoretical probability distribution of $PR_{GPR}(t)$ and the distribution of the determined PR_T , respectively.

From the comparison of Fig. 6.3.4 and Fig. 6.3.5 I find as expected an increased uncertainty for the PR_T estimate with broader irradiation band filter. Further-

³⁰The contributions of each measurement are weighted differently.

more, one can observe clear deviations of $PR_{GPR}(t)$ and PR_T , where PR_T is found to be reduced in the summer months³¹ and this way in operation at higher ambient and resulting higher operation temperatures. This effect of temperature in the time series of PR_T shows one major downfall of the classical PR approach. Equation 2.27 only accounts for temperature and irradiation dependencies in terms of linear dependencies using a single correction coefficient. In comparison, the IEC60891 norm (see Section 2.3.2) accounts for 6 correction coefficients. Furthermore, while observing degradation mechanisms in PV modules it is expected, that correction coefficients might also change over time leading to increased inaccuracy of the classical PR_T method.

The PR analysis using GPR prediction on the others side allows for arbitrary temperature and irradiation dependencies in the data, allowing the dependencies further to change over time. Furthermore, the GPR PR analysis does not require to exclude large part of the data, since the influence of each data point is weighted according to the respective distance in temperature, irradiation and time. Therefore, I argue the GPR PR analysis is a more accurate alternative to classical PR analysis techniques.

6.3.3 Physical analysis using the one-diode model

In Fig. 6.3.3, I show one possible output of the GPR IV model, where the IV characteristic respectively the ESPs are shown over time for one constant set of irradiation and temperature conditions. However, from the trained GPRs one can reproduce the ESPs at arbitrary conditions, enabling multiple applications³². One possible application is to fit a physical model to a given produced output. To cover the range of irradiation intensities and module temperatures occurring in operation, I compute the timeseries prediction of the ESPs for the exemplary chosen Italy CIGS4 dataset for multiple combinations of the two conditions, where I use $T_{Mod} = 10, 20, 30, 40, 50, 60^\circ\text{C}$ and $G_{POA} = 100, 250, 500, 750, 1000 \text{ Wm}^{-2}$. With the output of in total 30 timeseries of ESPs at constant conditions, the temperature and irradiation dependent one diode model is fitted to the predicted IV data.

Rewriting 2.13 for the current, substituting R_P by R_{Sh} , accounting for N number of cells connected in series and assuming a temperature activated recombination

³¹Note that the operation of the CIGS4 module operated in Italy started on the 1st of November 2013.

³²Note that for the predictions there are limitations arising due to the parameter space of $[G_{POA}, T_{Mod}, t]$ covered by the observations and associated high uncertainties for predictions that need to be extrapolated.

current results in

$$I(V) = I_{00} \left(\frac{T}{T_{\text{STC}}} \right)^3 \exp \left(\frac{(T - T_{\text{STC}})E_a}{kT_{\text{STC}}T} \right) \left[\exp \left(\frac{e(T_{\text{STC}} - T)(V - I(V)R_s)}{Nn_{\text{id}}kT_{\text{STC}}T} \right) - 1 \right] + \frac{V - I(V)R_s}{R_{\text{Sh}}} - I_{\text{Ph}}(G, T), \quad (6.11)$$

where E_a is an activation energy and

$$I_{\text{Ph}}(G, T) = \frac{G}{G_{\text{STC}}} (1 + \alpha(T - T_{\text{STC}})) I_{\text{Ph,STC}} \quad (6.12)$$

similar to the IEC60891 temperature dependency of the currents in 2.18 substituting $I_1 = I_{\text{Ph,STC}}$. With the fit of 6.11 and 6.12 to the timeseries of ESPs at the distinct conditions using the PV-CRAZE library [31] the result is a timeseries of the parameters R_s , R_{Sh} , $I_{\text{Ph},0}$, Nn_{id} , I_{00} , α and E_a . I find a better fit is obtained with PV-CRAZE, if the fitting is repeated several times and the best optimum is selected (some fitting algorithms used by PV-CRAZE are probabilistic). In this work the fit was repeated 23 times in a trade off between computation time and fit error. Figure 6.3.6 shows the development of the extracted parameters R_s , R_{Sh} , $I_{\text{Ph},0}$, Nn_{id} , I_0 , α and E_a over time, enabling to physically interpret the temporal development of the performance of the operated CIGS module. Furthermore, the root mean squared error (RMSE) of the fit is shown.

First of all, I observe a strong acclimatization effect most prominent in the shunt resistance R_{Sh} (Fig. 6.3.6 (e)), the activation energy E_a (Fig. 6.3.6 (g)), as well as in I_{00} , Nn_{id} and R_s (Fig. 6.3.6 (a), (b) and (d)). In the first few months of operation the R_{Sh} is reduced from over 5 k Ω to less than 2 k Ω and the activation energy drops from 0.7 to 0.5 eV. Comparing the temporal development of I_{00} and Nn_{id} (Fig. 6.3.6 (a) and (b)), I find a strong correlation of the two parameters. The observed increase in dark saturation current I_{00} is associated with an enhanced ideality factor (Nn_{id} increases and N is constant). This suggests, that there is a change in the most prominent recombination mechanism, i.e. either a recombination process with low ideality is reduced or a recombination process with high ideality enhanced. Furthermore, the series resistance shows an acclimatization effect, where the diode model fit finds two solutions, one close to the initial R_s between 5 and 6 Ω and one solution between 1 and 2 Ω . This effect where the solution alternates between two values is also visible in the RMSE, indicating the solution alternates between two local minima with distinct R_s values. Inspection shows the slightly smaller errors are associated with the more constant R_s solution between 5 and 6 Ω .

Beside the described acclimatization effects a strong seasonality is visible for the

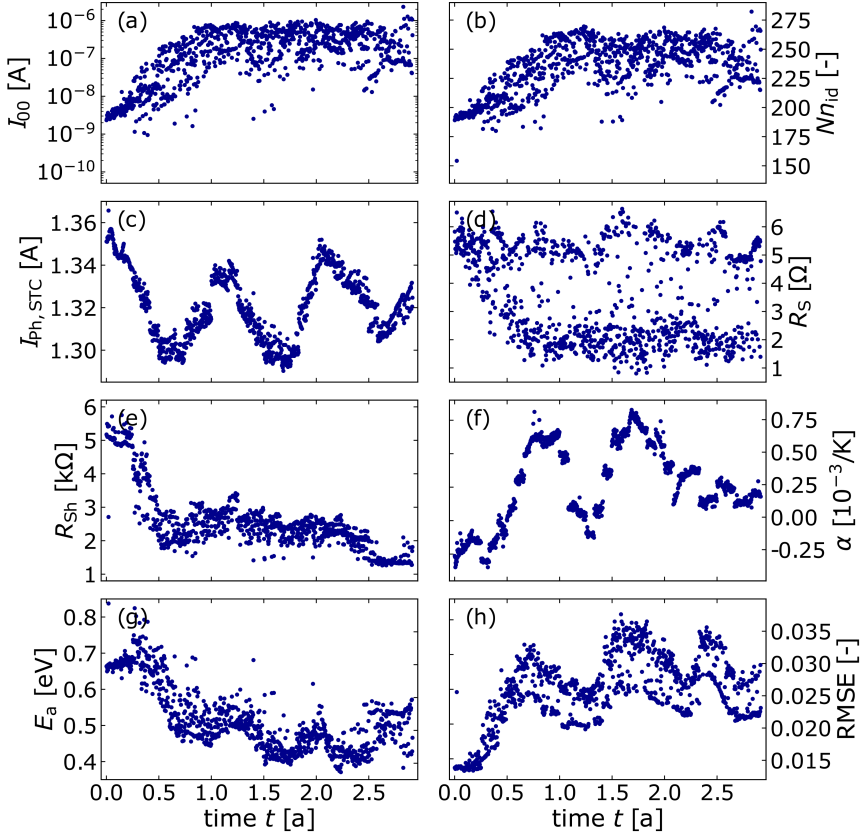


Figure 6.3.6: One diode model parameters timeseries prediction fitted to a GPR predicted timeseries matrix at different irradiation and temperature conditions for the Italy CIGS4 dataset.

photo-generation current $I_{\text{Ph,STC}}$ (Fig. 6.3.6 (c)). With the metastable nature of CIGS solar cells and modules in mind this clear difference in performance in winter and summer is not surprising. Furthermore, a seasonality contribution is visible in E_a (Fig. 6.3.6 (g)), the temperature coefficient α (Fig. 6.3.6 (f)) as well as in the RMSE (Fig. 6.3.6 (h)), which implies a slightly better fit accuracy in winter.

I find contributions of a linear degradation for the parameter R_{Sh} and for α (Fig. 6.3.6 (e) and (f)). A linear degradation over the complete operation period is further observed for the activation energy E_a , where a slight linear reduction is overlaid with the seasonality and the acclimatization. Assuming e.g. SRH recombination, the activation energy is expected to be dependent on the ideality factor, where the activation energy can be replaced by the quotient of the band gap energy and the ideality factor ($E_a \approx \frac{E_g}{n_{\text{id}}}$). To this end, I correct the activation energy for the ideality factor influence. Figure 6.3.7 shows the temporal development of the product $Nn_{\text{id}}E_a$. I find a clear linear degradation overlaid with a seasonal variation³³. This further indicates a more complex change in the prominent recombination mechanism over time.

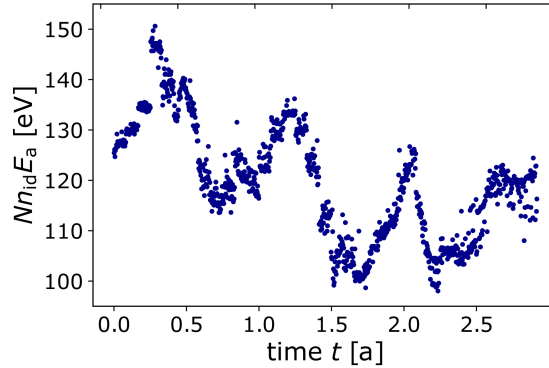


Figure 6.3.7: Product of the one diode model parameters timeseries prediction $Nn_{\text{id}}E_a$ fitted to a GPR predicted timeseries matrix at different irradiation and temperature conditions for the Italy CIGS4 dataset.

Comparing the qualitative development of the diode parameters with the PR_{GPR} in Fig. 6.3.4 or 6.3.5, I attribute the acclimatization effect to a change in the prominent recombination mechanism overlaid with a reduced shunt resistance. The relative amplitude of the seasonality matches well with the seasonality observed for the photo current and the long term degradation is influenced

³³Note that the seasonality appears to have a slightly lower period than one year. I attribute this discrepancy to the overlay of effects in the product of Nn_{id} (showing strong acclimatization) and E_a (showing an overlay of acclimatization, seasonality and degradation).

by multiple factors, such as the further reduced activation energy and shunt resistance.

Without further knowledge on the exact production and design of the CIGS module chosen as an exemplary outdoor study, it is hard to directly link the observed trends to the accelerated light induced degradation experiments on CIGS cells from a different producer presented in Chapter 4. Furthermore, the results of the outdoor study presented reflect many more influences affecting a PV modules performance over time than the isolated influence of irradiation at constant temperature and bias conditions. Nonetheless, I find only a slight reduction of the open circuit voltage (see Fig. 6.3.3) and argue, that the observed outdoor degradation is better reflected by LID at OC conditions. Furthermore the increasing dark saturation current I_{00} and ideality factor n_{id} match the tendency of a slight increase in the respective parameters due to LID at OC (compare Fig. 4.5.2 and 4.5.3).

6.4 Validation

With the three examples shown, I have introduced to some possible outputs of the GPR IV model without addressing to its accuracy. To this end, I use the trained GPRs f_{PCj_i} to validate their performance regarding recreating the training data (observations). Since each GPR is inherited to recreate its own input with an accuracy of the estimated noise level, validating a GPR on its own training input has little to no information on the performance. For validation there needs to be a clear separation of training and test data, where no test data is used for training the model. To further ensure a temporal separation of training and test data, I compare the input of the $PCj_m(t)$ ($ESPj_m(t)$), i.e. the observations in the m -th month, with the overlaid output of f_{PCj_i} for $i = m - 2$, $i = m - 1$, $i = m + 1$ and $i = m + 2$, i.e. the predictions of the four GPRs trained with the data of the two months before and after the month of consideration. Formally, I compute

$$PCj_{val,m}(G_{POA}, T_{Mod}, t) = \sigma_{PCj}^2(G_{POA}, T_{Mod}, t) \sum_{\substack{i=m-2, m-1, \\ m+1, m+2}} \frac{PCj_i(G_{POA}, T_{Mod}, t)}{\sigma_{PCj_i}^2(G_{POA}, T_{Mod}, t)}, \quad (6.13)$$

for all distinct observations in $[G_{POA}, T_{Mod}, t]$ for t in month m and determine $ESPk_{val,m}(G_{POA}, T_{Mod}, t)$ analogously to 6.3. This way, I use four instead of five months (w.r.t. the predictions shown in Section 6.3) for training and the fifth month as test data repeating the procedure for each month of the complete

dataset, resulting in a training-to-test ratio of approximately 80:20³⁴. Due to the reduced input for validating the prediction for the first and last month (67:33 ratio) as well as for the second and second to last month (75:25 ratio) of each dataset, the actual training to test ratio is a bit lower than the commonly used 80:20 ratio.

I compute the difference of prediction output ($\text{ESP}_{k_{\text{val},m}}(G_{\text{POA}}, T_{\text{Mod}}, t)$) and the test observations ($\text{ESP}_{k_{\text{test}}}(G_{\text{POA}}, T_{\text{Mod}}, t)$) for each month m of the dataset and set the performance of the models prediction in comparison to the difference of observations to their mean using

$$R_{\text{ESP}k}^2 = 1 - \frac{\sum (\text{ESP}_{k_{\text{test}}} - \text{ESP}_{k_{\text{val}}})^2}{\sum (\text{ESP}_{k_{\text{test}}} - \overline{\text{ESP}_{k_{\text{test}}}})^2}. \quad (6.14)$$

Illustratively, a $R^2 = 0$ translates to the prediction having the same RMSE of a model that would predict a constant ESP for all measurements independent of G_{POA} , T_{Mod} and t . The maximum possible value for R^2 is one (100%), translating to a perfect prediction with no deviation for all test data points.

For the ten ESPs, I find median R^2 values between 91.7% and 99.4% among all 45 datasets (five locations, nine modules). Figure 6.4.1 shows a boxplot of the R^2 values of each ESP, reconstructed from 10 PCs. The orange line denotes the median R^2 , the box denotes the range of the first to third quartile, the whiskers extend from the box to the farthest data point lying within 1.5 times the inter-quartile range (IQR, the distance between first and third quartile) and the single points denote individual R^2 values outside the described ranges. Note, that the logarithmic scale of the y axis is chosen such, that it captures the relevant range of R^2 values with increasing resolution towards most often observed high R^2 values close to 100%. For the four current parameters, I find high R^2 values among all 45 datasets with only a few outliers below 98% and the majority above 99%. The three voltage parameters V_{OC} , V_{MPP} and $V_{\text{qmp+}}$ are predicted with median R^2 values around 96%. The median accuracy of $V_{\text{qmp-}}$, R_{OC} and G_{SC} is between 91.5 and 94.5%. In total I find only very few outliers of predictions with R^2 values below 70%. Note that the R^2 values set the models performance in comparison to the difference of observations to their mean. Since the measured voltage parameters are better represented by their mean w.r.t. representing the current parameters by their mean, the observation of significant higher R^2 values for the current parameters is not surprising.

Due to the decreasing amplitude of the PCs, the validation can further be used

³⁴Note, that due to the sliding window approach, the validation of the model does not result in a high additional computation time, since the already trained GPRs respectively optimized hyperparameters are used for the predictions.

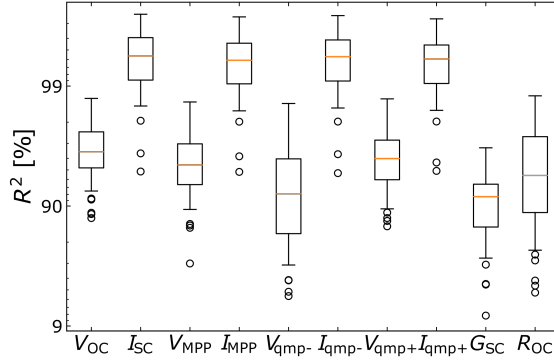


Figure 6.4.1: GPR performance validation in terms of the R^2 values of each ESP prediction, reconstructed from 10 PCs. The PC prediction is computed according to 6.13.

to evaluate, if the PCA reduces the data to a lower dimensional representation. I find no improvement regarding the R^2 values using more than seven and eight PCs, respectively, depending on the dataset. Thus, the ESPs can be reduced to a seven to eight-dimensional representation, and the computation time can be reduced to 70 – 80%. To visualize this high performance already observed using seven PCs, Fig. 6.4.2 shows the boxplot of the R^2 values for the ESPs reconstructed from seven PCs (computed with limits of the sum in 6.3 of [1,7]). Comparing Fig. 6.4.1 and 6.4.2, I find only slight deviations for single parameters in some datasets.

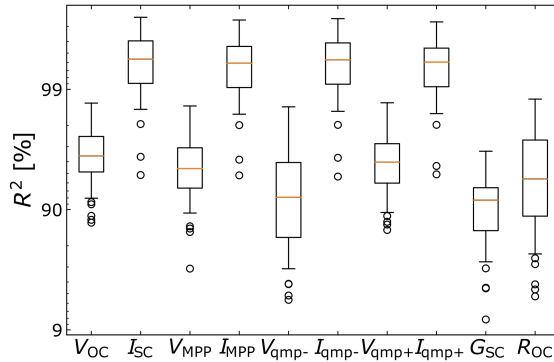


Figure 6.4.2: GPR performance validation in terms of the R^2 values of each ESP prediction, reconstructed from 7 PCs. The PC prediction is computed according to 6.13.

R^2 values give a reasonable indication how good a model performs setting the performance in comparison to the deviation of the respective data from its mean. For PV outdoor data analysis it might be from greater interest to evaluate on

(mean) absolute deviations of the prediction, since they determine the accuracy regarding a PV modules performance (and yield). Figure 6.4.3 shows the histogram (dark blue) of the absolute deviations $\Delta \text{ESP}k = \text{ESP}k_{\text{GPR}} - \text{ESP}k_{\text{test}}$ of the from seven GPR predicted PCs reconstructed ESPs ($\text{ESP}k_{\text{GPR}}$) from the real test data of ESPs ($\text{ESP}k_{\text{test}}$) for the exemplary chosen Italy CIGS4 dataset. The standard deviation of the histogram σ_{hist} is given in each subplot in dark blue.

Furthermore, Fig. 6.4.3 depicts the from the GPRs predicted probability density of deviations (red), i.e. the prediction of the histogram according to the GPR output. Each GPR prediction has an individual predicted standard deviation (compare with 6.6), i.e. every prediction $\text{PC}j(G_{\text{POA}}, T_{\text{Mod}}, t)$ of the GPRs $f_{\text{PC}j_i}$ can be interpreted as a predicted Gaussian probability density $\mathcal{N}(\text{PC}j(G_{\text{POA}}, T_{\text{Mod}}, t), \sigma_{\text{PC}j}(G_{\text{POA}}, T_{\text{Mod}}, t))$. The prediction $\text{ESP}k(G_{\text{POA}}, T_{\text{Mod}}, t)$ is a linear combination of the predictions $\text{PC}j(G_{\text{POA}}, T_{\text{Mod}}, t)$ (compare 6.3). In the approximation, that the 10 predictions $\text{PC}j(G_{\text{POA}}, T_{\text{Mod}}, t)$ are statistically independent for each input of $[G_{\text{POA}}, T_{\text{Mod}}, t]$ (see Section 6.2.4), the prediction $\text{ESP}k(G_{\text{POA}}, T_{\text{Mod}}, t)$ can be interpreted as a Gaussian probability density $\mathcal{N}(\text{ESP}k(G_{\text{POA}}, T_{\text{Mod}}, t), \sigma_{\text{ESP}k}(G_{\text{POA}}, T_{\text{Mod}}, t))$, where $\sigma_{\text{ESP}k}(G_{\text{POA}}, T_{\text{Mod}}, t)$ is determined analogously to 6.4 from the predictions $\text{PC}j(G_{\text{POA}}, T_{\text{Mod}}, t)$ and standard deviations $\sigma_{\text{PC}j}(G_{\text{POA}}, T_{\text{Mod}}, t)$. Considering deviations from the prediction sets the mean of each normal distribution to zero resulting in the prediction of the probability of deviations to $\mathcal{N}(0, \sigma_{\text{ESP}k}(G_{\text{POA}}, T_{\text{Mod}}, t))$ for each coordinates $[G_{\text{POA}}, T_{\text{Mod}}, t]$. The set of N test data points and the respective $n = 1, 2, \dots, N$ input vectors $[G_{\text{POA}}, T_{\text{Mod}}, t]_n$ can essentially be interpreted as a random choice for a subset of coordinates $[G_{\text{POA}}, T_{\text{Mod}}, t]$. Exploiting the law of total probability and substituting n for $[G_{\text{POA}}, T_{\text{Mod}}, t]_n$ the probability densities for the deviation of each test point n ($p_n(\Delta \text{ESP}k) = \mathcal{N}(0, \sigma_{\text{ESP}k}(n))$) can be summed up to receive the predicted joint probability of deviations from the model prediction

$$P(\Delta \text{ESP}k) = \frac{1}{N} \sum_{n=1}^N p_n(\Delta \text{ESP}k) \quad (6.15)$$

where the sum of the predicted deviations is divided by the number of tests data points N for normalization. In general, this sum of multiple Gaussian probability densities reflects a Gaussian mixture model and might be Gaussian in specific scenarios, e.g. in case all individual standard deviations are identical. However, regardless of this, the resulting standard deviation can be computed

and is given by

$$\begin{aligned}\sigma_{\text{GPR}} &= \sqrt{\int_{-\infty}^{\infty} (\Delta \text{ESP} k)^2 P(\Delta \text{ESP} k) d\Delta \text{ESP} k} \\ &= \sqrt{\int_{-\infty}^{\infty} (\Delta \text{ESP} k)^2 \frac{1}{N} \sum_{n=1}^N p_n(\Delta \text{ESP} k) d\Delta \text{ESP} k.}\end{aligned}\quad (6.16)$$

Since $p_n(\Delta \text{ESP} k)$ is a probability density, it holds $p_n(\Delta \text{ESP} k) \geq 0$ and according to the Fubini-Tonelli theorem [124] it follows

$$\begin{aligned}\sigma_{\text{GPR}} &= \sqrt{\frac{1}{N} \sum_{n=1}^N \int_{-\infty}^{\infty} (\Delta \text{ESP} k)^2 p_n(\Delta \text{ESP} k) d\Delta \text{ESP} k} \\ &= \sqrt{\frac{1}{N} \sum_{n=1}^N \sigma_{\text{ESP} k}^2(n)},\end{aligned}\quad (6.17)$$

where I use, that the standard deviation of $p_n(\Delta \text{ESP} k)$ is given by $\sigma_{\text{ESP} k}(n)$. In every subplot in Fig. 6.4.3 σ_{GPR} is given in orange for reference.

I find standard deviations of the deviation histograms for the current parameters I_{SC} , I_{MPP} , $I_{\text{qmp+}}$ and $I_{\text{qmp-}}$ in the range of 14.8 to 23.8 mA (compare σ_{hist} , dark blue). If I compare these absolute deviations to the mean of the time series data in Fig. 6.3.3 I find relative deviations in the range between 2.3 and 2.6%. All absolute and relative deviations of the prediction are summarized in Tab. 6.1. For the voltage parameters V_{OC} , V_{MPP} , $V_{\text{qmp+}}$ and $V_{\text{qmp-}}$ I find absolute deviation in the range of 1.17 to 1.36 V (σ_{hist} , dark blue) and compared to the time series at $G_{\text{POA}} = 750 \text{ Wm}^{-2}$ and $T_{\text{Mod}} = 40^\circ\text{C}$, relative deviations in the range of 1.3 to 2.4%.

Table 6.1: Absolute and relative deviation of the GPR ESP prediction

parameter	σ_{hist}		σ_{GPR}	
	abs.	rel.	abs.	rel.
I_{SC}	23.8 mA	2.3%	24.7 mA	2.4%
I_{MPP}	22.1 mA	2.5%	22.3 mA	2.5%
$I_{\text{qmp+}}$	23.5 mA	2.4%	24.1 mA	2.5%
$I_{\text{qmp-}}$	14.8 mA	2.6%	14.5 mA	2.5%
V_{OC}	1.17 V	1.3%	0.84 V	0.9%
V_{MPP}	1.36 V	1.8%	1.10 V	1.5%
$V_{\text{qmp+}}$	1.32 V	2.4%	1.12 V	2.0%
$V_{\text{qmp-}}$	1.26 V	1.5%	0.96 V	1.1%
G_{SC}	$24.4 \frac{1}{\text{M}\Omega}$	6.9%	$12.4 \frac{1}{\text{M}\Omega}$	3.5%
R_{OC}	59.0Ω	648%	10.5Ω	115%

Comparing the histogram with the GPR predicted probability density (orange)

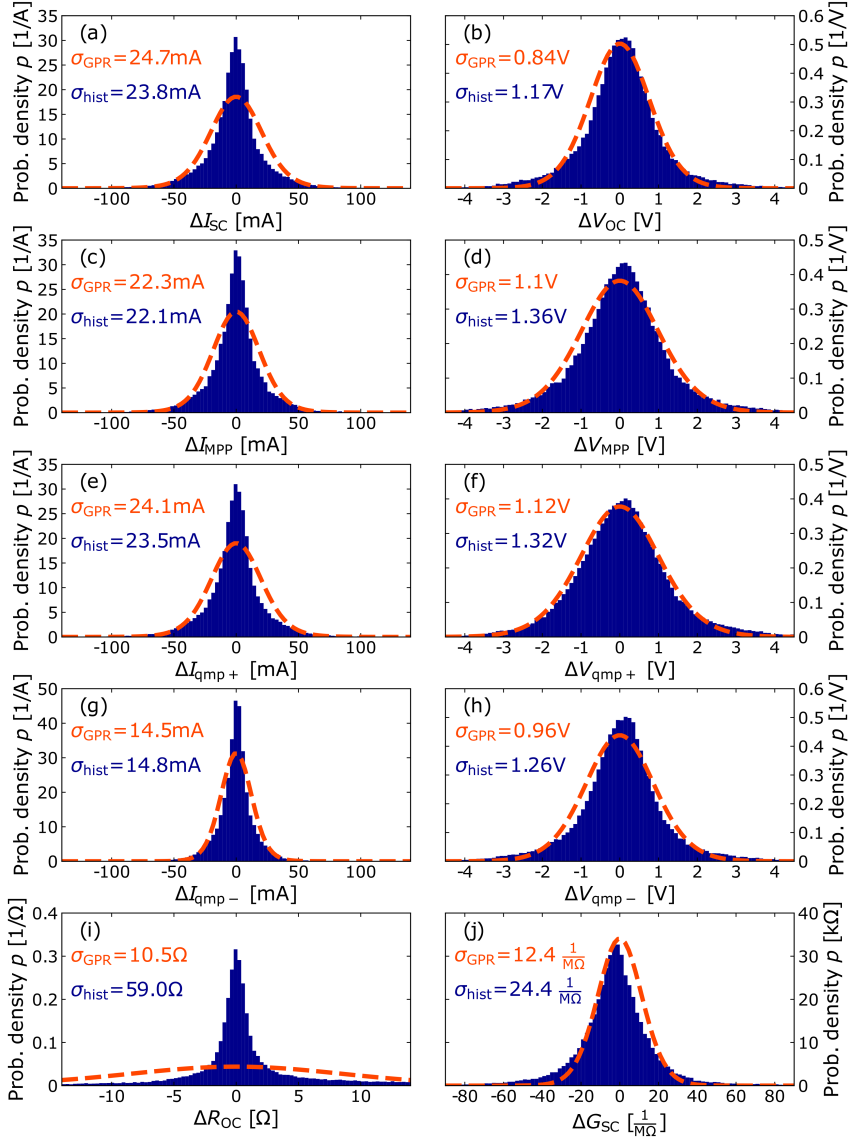


Figure 6.4.3: Histogram (dark blue) of the absolute deviations $\Delta \text{ESP} = \text{ESP}_{\text{GPR}} - \text{ESP}_{\text{test}}$ of the from seven GPR predicted PCs reconstructed ESPs from the real test data of ESPs for the exemplary chosen Italy CIGS4 dataset. The histograms standard deviation σ_{hist} is given in dark blue. The GPR predicted probability density of deviations (with the standard deviation of σ_{GPR}) is depicted in orange.

for the current and voltage parameter, I find, that all eight deviation distributions are heavy tailed compared to the GPR predicted probability density, where the effect is more pronounced for the four current parameters. Especially for the four voltage parameters, the GPR predicted deviation qualitatively matches the real distribution of deviations, while the standard deviation of the real distribution of deviations σ_{hist} matches the predicted GPR uncertainty σ_{GPR} for the four current parameters. Table 6.1 further summarizes the GPR predicted uncertainty σ_{GPR} for all parameters. For the 8 current and voltage parameters the GPR predicted uncertainty is in the range of 0.9 to 2.5% relative to the mean of the time series at $G_{\text{POA}} = 750 \text{ Wm}^{-2}$ and $T_{\text{Mod}} = 40^\circ\text{C}$.

For the parameter G_{SC} I find similar to the eight current and voltage parameters a heavy tailed distribution, which is qualitatively well described by the GPR predicted probability density. Furthermore, one can see, that the tail of the distribution is more pronounced for negative ΔG_{SC} . The relative deviation of ΔG_{SC} , however, is much higher. I find an absolute deviation of G_{SC} of $24.4 \frac{1}{\text{M}\Omega}$ and a relative deviation of 6.9% compared to the mean of G_{SC} in Fig. 6.3.3. This comparably high relative deviation can be explained by the definition of G_{SC} as the slope at short circuit. In comparison to the current or voltage parameters G_{SC} is not directly dependent on irradiation and temperature, thus more influenced by more subtle factors, like noise and such expected to have a higher relative variance. Furthermore, the definition of G_{SC} as a slope can be used the observed one-sided heavily tailed distribution.

For the parameter R_{OC} I find a high discrepancy between the shown histogram (dark blue) and the GPR predicted probability density (orange). This high discrepancy can be explained with (systematic) outliers to high R_{OC} values in the dataset, which are not captured with the GPR prediction, since they cannot be described adequately with the assumption of Gaussian noise. The uncertainty prediction of the GPR is, however, clearly affected by the blatant outliers predicting an unreasonable high uncertainty of the prediction, e.g. the GPR predicted probability distribution (orange) does not match the histogram (dark blue). Neglecting these outliers would result in a standard deviation of the histogram σ_{hist} more than one order of magnitude lower, depending on the cut-off deviation, after which outliers are neglected. Since also R_{OC} describes a slope (at OC), a similar argument as for G_{SC} applies, where a higher relative variance is already expected.

6.5 Summary and Outlook

This Chapter has introduced a statistical model how to analyse PV outdoor data. First the general concept of the model has been discussed. Starting with the compact description of *IV* characteristics using the ESPs, a suitable filtering approach is applied. Afterwards, the ESPs are processed with a principal component analysis (PCA). The main step is then temporally splitting the PC timeseries and training individual Gaussian process regressions (GPRs) using G_{POA} and T_{Mod} and the time t as input. Once the GPRs are trained the model can reproduce respectively predict the complete *IV* characteristic at any given time t , G_{POA} and T_{Mod} , incorporating an uncertainty of the ESP prediction based on the data noise and distance from the observations.

To underline its usability, the presented concept is applied to represent the *IV* characteristic in terms of the 10 ESPs (and PCs) clearly showing superimposed acclimatization, degradation and seasonality effects in the respective parameters. Furthermore, the model is utilized to determine a PV modules performance ratio over time and compared to a classical monthly temperature corrected performance ratio approach. Especially regarding the uncertainty estimate resulting from the GPR predicted uncertainty I find an improvement of the PR_{GPR} estimate over the classical PR_T approach. A set of predicted ESP timeseries is further used as input for a one diode model fit, allowing for a more physical interpretation of the presented performance variation and changes over time. I find a strong seasonality reflected mostly in the photo current and a more complex acclimatization and long-term degradation reflected in multiple parameters. Further I connect the outdoor analysis to the LID studies discussed in Chapter 4. I argue, that LID at OC conditions is reflects realistic operation conditions better.

Finally, the presented concept is validated using a common 80:20 training to test data ratio, while keeping training and test data temporally separated. The overview of the models performance measure R^2 over all 10 ESPs and all 45 datasets shows a consistently good accuracy of the model for a wide range of technologies operated in different climates. For one exemplary subset, the accuracy of the model is discussed in more detail, showing low relative deviations in the prediction of the majority of ESPs, while being consistent with the models prediction for its uncertainty.

The presented model widens the possibilities in the field of PV reliability monitoring and PV degradation analysis. The manageability of big datasets is ensured due to the compact description in terms of the ESPs and a temporal split of the data. The accuracy of the model is validated for different locations (climates) on

different module types. Furthermore, only the possibility to measure in field IV characteristics, module temperature and irradiation is presupposed. As a result, that provides for the concepts presented many applications. Especially the possibility to set the results into context of their uncertainty enables to quantify the confidence of the results. Further, if only parts of the *IV* is available, e.g. the SSPs or only AC or DC output voltage, current or power, the concept of training monthly GPRs is still applicable.

In the presence of rising accessibility to high computational power and the possibility to utilize Graphics Processing Units (GPUs) for matrix multiplications, GPRs become more and more applicable to big datasets [125]. This way, the presented approach could be updated in future to bigger subsets or by a single GPR on each timeseries. Furthermore, the model can be utilized to compare different filtering approaches, since the applicability of GPRs to data remains on Gaussian distributed errors in the data and can help to detect the presence of systematic errors.

Summary

The rapid extension of the renewable energy sector is essential for the compliance with climate goals especially considering the rising global energy demand. Increasing the energy production from PV technologies is one key aspect to achieve a clean and reliable energy supply system. Beside the need for an overall growth in production and capacity, improvements regarding the outdoor performance and sustainability of PV modules are inalienable for an economy driven extension of renewable energy production.

This thesis has addressed multiple topics related to PV module outdoor performance and sustainability. Chapter 3 has introduced a model to extrapolate CIGS module performance from cell performance as well as to extrapolate module yield in different climates from module performance. The in the PV community well implemented IEC60891 norm for irradiation and temperature corrections has been utilized to gain a compact description of a solar cells respectively PV modules irradiation and temperature dependent performance. Afterwards, standard reference climates defined in the IEC61853 norm have been used to determine an expected yield in different climates and such quantified the potential of a PV module operated in various locations around the world. Furthermore, the effect of potential improvements on cell level on the yield have been discussed. This way the presented model represents a possibility how to interpret cell efficiencies and performance in terms of energy production and yield in different climates and translates fundamental cell improvements to the resulting energy production potential.

Beside the (potential) outdoor performance of PV modules and systems, their reliability and performance stability is crucial for the energy yield during operation over many years. Monitoring the PV outdoor performance is ultimately the best method to analyse the performance stability considering all effects on the PV modules performance occurring in a real operation scenario. However, in

PV outdoor data multiple influences affecting the performance are overlaid. To separate these multiple influences laboratory experiments can be used, where the conditions of degradation can be tuned and are well defined. One aspect of degradation and performance variations in operation arises through illumination of the semiconductor layer stack. In Chapter 4 laboratory experiments were presented, studying light induced degradation of CIGS solar cells as well as the dependency on temperature, irradiation intensity and applied bias condition. The results remarkably showed, that the operation bias is crucial for the effects triggered due to illumination of CIGS solar cells. Cells kept under OC appear stable under illumination while cells kept in SC exhibit a significant performance drop. This performance drop is shown to be more or less independent on the intensity, but shown to be affected by the operation temperature. High operation temperatures lead to the highest amplitude of up to -3% absolute loss in efficiency. The efficiency loss is caused by an increase of the overall recombination rate, while the ideality factor is reduced. This indicates an enhanced probability of a recombination process with ideality factor 1, consistent with e.g. recombination centres located in quasi-neutral regions of the device.

Accelerated testing procedures like the presented LID experiment cannot replace actual long-term studies. Monitoring the PV outdoor performance over years of operation is ultimately the best method to test and validate the results of accelerated ageing tests, tracking performance stability considering all effects on the PV modules performance occurring in a real operation scenario. The evaluation of the PV performance requires information on the conditions the performance is measured at, e.g. meteorological data as well as a suitable approach to validate the data quality. Chapter 5 has introduced two different approaches for PV outdoor data quality assessment and filtering. The first approach is based on utilizing the many correlations present in the electrical performance and meteorological data. Physical models are used to compare individual measured parameters with the expectation of these parameters based on other information available from the data profile. Further, the Mahalanobis distance is used to weight correlated deviations from the expectation accordingly. The second filtering approach presented is designed to selectively filter out data points where the effective irradiation on the PV module does not match the measured in plane irradiation. For that I utilize the stochastic nature of GPRs to find data points that cannot be described adequately with Gaussian deviations from the expectation and can such only originate from a systematic deviation from the expected dependency of the short circuit current I_{SC} on irradiation G_{POA} and temperature T_{Mod} . An iterative approach is hereby chosen, to suppress the influence of outliers on the prediction.

Once the data quality is assessed and low quality data is filtered, analysing PV outdoor data still needs to overcome challenges arising through the data quantity and correlations in the different dimensions impeding to separate multiple effects influencing the PV module performance. In Chapter 6 a model for statistically analysing PV outdoor data has been presented. The GPR *IV* model is based on dimensionality reduction through the use of the ESPs and a subsequent PCA. The timeseries of linear uncorrelated PCs is then processed using multiple GPRs and a mixture of experts approach overlaying the predictions. After the model is trained, it can predict the PCs, this way the ESPs and such the complete *IV* characteristic at any time, irradiation and temperature condition. Furthermore the model provides information about the uncertainty of the prediction resulting from the noise level of the data and the distance (in time, irradiation and temperature) of the prediction from observations used for training. The model has been shown to be an alternative to analyse PV outdoor data and gain insights on e.g. acclimatization, seasonality and long-term degradation. Furthermore, the model has been shown to outperform a classical PR estimate, gaining results with lower uncertainty, while being not dependent on the choice of an additional filter, commonly used in the classical approach.

7.1 Outlook

Wafer-based crystalline silicon (c-Si) solar cells are the dominant PV technology for more then 6 decades [126]. Consequently, the reliability and performance stability of the established glass-backsheet and glass-glass encapsulated c-Si PV modules has been tested and verified sufficiently. Arguably, reliability and performance stability under accelerated tests or outdoor operation become less and less important in the PV sector. Considering the future of the PV sector, however, reliability and performance stability analysis and by this the topics of extrapolation of module performance, outdoor data analysis and accelerated ageing tests are expected to gain importance:

- Currently, the most promising (new) PV technologies are based on perovskites, where high efficiencies have been demonstrated on cell level as well as in tandem configurations [2]. Poor long-term stability is the major challenge hindering the commercialization of perovskite PV technologies [127]. Long-term outdoor studies and especially accelerated ageing tests are needed to validate the long-term performance of perovskite and perovskite tandem modules. Furthermore, insights gained from such studies might help addressing the long-term performance issue and such play a crucial role for the future of perovskites in the PV sector.

- The PV sector is currently an example for a classical linear economy, where PV modules are mostly not recycled at the end of life (EoL) and end up in landfill [128]. The transition to a circular economy, where ideally all resources needed for production of new devices are recovered from EoL devices is a huge challenge. While the share of recycling of EoL PV modules increases politically driven, methods of recycling of commercial glass-backsheet and glass-glass encapsulated c-Si PV modules ³⁵ lack on recovery rates or purity and are in general not profitable, due to the amount of energy needed to separate glass, metals and semiconductor materials from the encapsulant and each other. Considering an expectation of cumulative 78 million tons of PV waste by 2050 [129], module concepts, where the recyclability at the EoL is part of the design (design for recycling, DFR), are a necessity for a circular economy. For the successful implementation of new DFR PV module concepts in the market, long-term outdoor studies and accelerated ageing tests are needed.

Beside the overall prospect of the topic of PV module performance stability and reliability I further want to give a short term outlook on the topics covered in this thesis:

- Chapter 3 has introduced a model how to extrapolate from CIGS cell performance to module yield. While part of the model is transferable to other PV technologies, future work could work on a generalization of the presented model to e.g. c-Si solar cells or tandem applications.
- Chapter 4 has presented the results of LID accelerated ageing tests on industrially produced CIGS solar cells. While the difference in stability of the solar cells at OC and SC is shown to be remarkable, the question arises how a bias condition between SC and OC would affect the observed performance changes in CIGS solar cells under illumination. An experiment applying a bias, which is common in operation, i.e. a bias close to MPP, during the LID might gain valuable insights on how the effects observed translate to real operation scenarios.
- Chapter 5 has introduced two possibilities of filtering PV outdoor data. The two approaches address different systematic errors typically present in PV outdoor data sets, but do not take into account all possible influences affecting the data quality. Furthermore the question remains, how different filtering approaches affect further analysis, e.g. PR estimates or degradation studies.

³⁵more than 90% market share over the last 40 years [128]

- Chapter 6 has presented a model how to statistically analyze PV outdoor data. While the general applicability of the model has been demonstrated, a physical model how to describe the observed trends of acclimatization, seasonality and degradation as well as the extrapolation of the observed temporal trends is beyond the scope of this work. The overlaid nature of many effects affecting a PV modules performance as well as the accuracy of the measurements itself impede further analysis. The classification of the data in terms of an uncertainty estimate from a given time, temperature and irradiation as well as the possibility to produce arbitrary outputs with the presented model, however, might enable the application of further timeseries analysis methods. One promising approach (especially for extrapolation and such for predicting future performance) is the autoregressive integrated moving average (ARIMA) approach. An ARIMA timeseries analysis requires a regularly spaced timeseries, which is in general not the case for PV outdoor data, but can be produced with the presented model.
- Combining the topics addressed, it will be a question for future work, how to combine the prediction of outdoor performance in various climates with a degradation model based on the insights gained through accelerated ageing tests and long-term outdoor studies. The insights gained through outdoor and laboratory experiments might be used as input for a more realistic yield prediction, where performance variations occurring in operation are included in the yield calculations in Chapter 3.

Bibliography

- [1] N. M. Haegel, P. Verlinden, M. Victoria, P. Altermatt, H. Atwater, T. Barnes, C. Breyer, C. Case, S. De Wolf, C. Deline, *et al.*, *Photovoltaics at multi-terawatt scale: Waiting is not an option*, Science **380**, 39 (2023).
- [2] NREL, Best research-cell efficiency chart (2024).
- [3] *Photovoltaic (PV) module performance testing and energy rating - Part 4: Standard reference climatic profiles*, Standard IEC 61853-4 (International Electrotechnical Commission, Geneva, CH, 2018).
- [4] *Photovoltaic (PV) module performance testing and energy rating - Part 3: Energy rating of PV modules*, Standard IEC 61853-3 (International Electrotechnical Commission, Geneva, CH, 2018).
- [5] M. Schweiger, J. Bonilla, W. Herrmann, A. Gerber, and U. Rau, *Performance stability of photovoltaic modules in different climates*, Progress in Photovoltaics: research and applications **25**, 968 (2017).
- [6] W. Marion, A. Anderberg, C. Deline, S. Glick, M. Muller, G. Perrin, J. Rodriguez, S. Rummel, K. Terwilliger, and T. Silverman, *User's manual for data for validating models for PV module performance*, Tech. Rep. (National Renewable Energy Lab.(NREL), Golden, CO (United States), 2014).
- [7] K. Pearson, *Liii. on lines and planes of closest fit to systems of points in space*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **2**, 559 (1901).
- [8] P. C. Mahalanobis, *On the generalized distance in statistics*, Sankhyā: The Indian Journal of Statistics, Series A (2008-) **80**, S1 (2018).

- [9] A. C. Rencher, *Multivariate statistical inference and applications*, Vol. 338 (Wiley New York, 1998).
- [10] C. K. Williams and C. E. Rasmussen, *Gaussian processes for machine learning*, Vol. 2 (MIT press Cambridge, MA, 2006).
- [11] H. Liu, Y.-S. Ong, X. Shen, and J. Cai, *When gaussian process meets big data: A review of scalable gps*, IEEE transactions on neural networks and learning systems **31**, 4405 (2020).
- [12] A. Vladimirescu, A. R. Newton, and D. O. Pederson, *SPICE Version 2G.1 user's Guide* (University of California Berkeley, Ca, 1980).
- [13] U. Rau, A. Jasenek, H. Schock, F. Engelhardt, and T. Meyer, *Electronic loss mechanisms in chalcopyrite based heterojunction solar cells*, Thin Solid Films **361**, 298 (2000).
- [14] F. Padovani and R. Stratton, *Field and thermionic-field emission in schottky barriers*, Solid-State Electronics **9**, 695 (1966).
- [15] A. Riben and D. Feucht, *Electrical transport in nge-pgaas heterojunctions*, International Journal of Electronics **20**, 583 (1966).
- [16] J. H. Werner and H. H. Güttler, *Barrier inhomogeneities at schottky contacts*, Journal of applied physics **69**, 1522 (1991).
- [17] Y. Yu, X. Bai, S. Li, J. Shi, L. Wang, F. Xi, W. Ma, and R. Deng, *Review of silicon recovery in the photovoltaic industry*, Current Opinion in Green and Sustainable Chemistry , 100870 (2023).
- [18] A. Singh, *Laser processing for interdigitated back-contacted silicon heterojunction solar cells*, Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen, Aachen (2021), veröffentlicht auf dem Publikationsserver der RWTH Aachen University 2022; Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen, 2021.
- [19] B. Misić, *Analysis and Simulation of Macroscopic Defects in Cu(In,Ga)Se₂ Photovoltaic Thin Film Modules*, Ph.D. thesis, Forschungszentrum Jülich GmbH, Institute of Energy and Climate Research, IEK-5 Photovoltaics (2017).
- [20] T. Kirchartz, D. Abou-Ras, and U. Rau, *Introduction to thin-film photovoltaics*, Advanced characterization techniques for thin film solar cells **1**, 1 (2016).

- [21] TC 82 - Solar photovoltaic energy systems, *Photovoltaic Devices—Procedure for Temperature and Irradiance Corrections to Measured IV Characteristics*, Standard IEC 60891 (International Electrotechnical Commission, Geneva, CH, 2009).
- [22] *Photovoltaic devices – Procedures for temperature and irradiance corrections to measured I-V characteristics*, Standard IEC 60891 (International Electrotechnical Commission, Geneva, CH, 2021).
- [23] B. Li, D. Diallo, A. Migan-Dubois, and C. Delpha, *Performance evaluation of iec 60891: 2021 procedures for correcting i-v curves of photovoltaic modules under healthy and faulty conditions*, Progress in Photovoltaics: Research and Applications **31**, 474 (2023).
- [24] D. Cotfas, P. Cotfas, and S. Kaplanis, *Methods to determine the dc parameters of solar cells: A critical review*, Renewable and Sustainable Energy Reviews **28**, 588 (2013).
- [25] A. Ortiz-Conde, F. J. García-Sánchez, J. Muci, and A. Sucre-González, *A review of diode and solar cell equivalent circuit model lumped parameter extraction procedures*, Facta Universitatis, Series: Electronics and Energetics **27**, 57 (2014).
- [26] R. Venkateswari and N. Rajasekar, *Review on parameter estimation techniques of solar photovoltaic systems*, International Transactions on Electrical Energy Systems **31**, e13113 (2021).
- [27] B. E. Pieters, *Extended solar cell parameters - general purpose descriptive i/v parameters for solar cells*, Authorea (2023).
- [28] A. Savitzky and M. J. E. Golay, *Smoothing and differentiation of data by simplified least squares procedures*, Analytical Chemistry **36**, 1627 (1964).
- [29] S. Karmalkar and S. Haneefa, *A physically based explicit j-v model of a solar cell for simple design calculations*, IEEE electron device letters **29**, 449 (2008).
- [30] L. Shao and H. Zhou, *Curve fitting with bezier cubics*, Graphical models and image processing **58**, 223 (1996).
- [31] B. Pieters, Pv-craze, <https://github.com/PVCRAZE/> (2024).
- [32] S. Lindig, A. Louwen, D. Moser, and M. Topic, *Outdoor pv system monitoring—input data quality, data imputation and filtering approaches*, Energies **13**, 5099 (2020).

- [33] I. Høiaas, K. Grujic, A. G. Imenes, I. Burud, E. Olsen, and N. Belbachir, *Inspection and condition monitoring of large-scale photovoltaic power plants: A review of imaging technologies*, Renewable and Sustainable Energy Reviews **161**, 112353 (2022).
- [34] *Photovoltaic System Performance-Part 1: Monitoring*, Standard IEC 61724-1 (International Electrotechnical Commission, Geneva, CH, 2017).
- [35] B. Litzenburger, S. Pingel, S. Janke, S. Held, and R. Alam, in *Conf Pap* (2014).
- [36] G. Mannino, G. M. Tina, M. Cacciato, L. Todaro, F. Bizzarri, and A. Canino, *A photovoltaic degradation evaluation method applied to bifacial modules*, Solar Energy **251**, 39 (2023).
- [37] G. Belluardo, P. Ingenhoven, W. Sparber, J. Wagner, P. Weihs, and D. Moser, *Novel method for the improvement in the evaluation of outdoor performance loss rate in different pv technologies and comparison with two other methods*, Solar Energy **117**, 139 (2015).
- [38] SSDP: Simple Sky Dome Projector, github.com/IEK-5/SSDP, accessed: Oct, 2022.
- [39] R. Perez, R. Seals, and J. Michalsky, *All-weather model for sky luminance distribution—preliminary configuration and validation*, Solar energy **50**, 235 (1993).
- [40] E. Sovetkin, J. Noll, N. Patel, A. Gerber, and B. E. Pieters, *Vehicle-integrated photovoltaics irradiation modeling using aerial-based lidar data and validation with trip measurements*, Solar RRL , 2200593 (2022).
- [41] D. Faïman, *Assessing the outdoor operating temperature of photovoltaic modules*, Progress in Photovoltaics: Research and Applications **16**, 307 (2008).
- [42] E. Barykina and A. Hammer, *Modeling of photovoltaic module temperature using Faïman model: Sensitivity analysis for different climates*, Solar Energy **146**, 401 (2017).
- [43] T. M. Friedlmeier, P. Jackson, A. Bauer, D. Hariskos, O. Kiowski, R. Menner, R. Wuerz, and M. Powalla, *High-efficiency Cu(In,Ga)Se₂ solar cells*, Thin Solid Films **633**, 13 (2017), e-MRS 2016 Spring Meeting, Symposium V, Thin-Film Chalcogenide Photovoltaic Materials.

- [44] M. Zinßer, T. Helder, A. Bauer, T. M. Friedlmeier, J. Zillner, J.-P. Becker, and M. Powalla, *Optical and Electrical Loss Analysis of Thin-Film Solar Cells Combining the Methods of Transfer Matrix and Finite Elements*, IEEE Journal of Photovoltaics **12**, 1154 (2022).
- [45] N. E. I. Boukortt and S. Patané, *Single junction-based thin-film CIGS solar cells optimization with efficiencies approaching 24.5%*, Optik **218**, 165240 (2020).
- [46] S. Dabbabi, T. B. Nasr, and N. Kamoun-Turki, *Parameters optimization of CIGS solar cell using 2D physical modeling*, Results in Physics **7**, 4020 (2017).
- [47] W. De Soto, S. A. Klein, and W. A. Beckman, *Improvement and validation of a model for photovoltaic array performance*, Solar energy **80**, 78 (2006).
- [48] B. Marion, B. Kroposki, K. Emery, J. Del Cueto, D. Myers, and C. Osterwald, *Validation of a photovoltaic module energy ratings procedure at NREL*, Tech. Rep. (National Renewable Energy Lab.(NREL), Golden, CO (United States), 1999).
- [49] T. Ma, H. Yang, and L. Lu, *Solar photovoltaic system modeling and performance prediction*, Renewable and Sustainable Energy Reviews **36**, 304 (2014).
- [50] M. Zaimi, H. El Achouby, A. Ibral, and E. Assaid, *Determining combined effects of solar radiation and panel junction temperature on all model-parameters to forecast peak power and photovoltaic yield of solar panel under non-standard conditions*, Solar Energy **191**, 341 (2019).
- [51] M. Zinßer, B. Braun, T. Helder, T. Magorian Friedlmeier, B. Pieters, A. Heinlein, M. Denk, D. Göddeke, and M. Powalla, *Irradiation-dependent topology optimization of metallization grid patterns and variation of contact layer thickness used for latitude-based yield gain of thin-film solar modules*, MRS Advances **7**, 706 (2022).
- [52] H. Yousuf, M. A. Zahid, M. Q. Khokhar, J. Park, M. Ju, D. Lim, Y. Kim, E.-C. Cho, and J. Yi, *Cell-to-Module Simulation Analysis for Optimizing the Efficiency and Power of the Photovoltaic Module*, Energies **15**, 1176 (2022).
- [53] V. Bermudez and A. Perez-Rodriguez, *Understanding the cell-to-module efficiency gap in Cu (In, Ga)(S, Se) 2 photovoltaics scale-up*, Nature Energy **3**, 466 (2018).

- [54] P. Gečys, E. Markauskas, S. Nishiwaki, S. Buecheler, R. De Loor, A. Burn, V. Romano, and G. Račiukaitis, *CIGS thin-film solar module processing: case of high-speed laser scribing*, Scientific reports **7**, 40502 (2017).
- [55] J. Johansson, U. Zimmermann, and M. Edoff, in *22nd EU PVSEC, 22nd European Photovoltaic Solar Energy Conference, in Milan, Italy, 3-7 September 2007* (WIP-Renewable Energies, 2007) pp. 1922–1925.
- [56] T. Vaas, B. Pieters, T. M. Friedlmeier, A. Gerber, and U. Rau, *Extrapolating cigs module performance from laboratory cell performance using iec61853 standard reference climatic profiles*, IEEE Journal of Photovoltaics **13**, 882 (2023).
- [57] A. Kanevce, S. Paetel, D. Hariskos, and T. M. Friedlmeier, *Impact of RbF-PDT on Cu (In, Ga) Se₂ solar cells with CdS and Zn (O, S) buffer layers*, EPJ Photovoltaics **11**, 8 (2020).
- [58] Y. Gupta, H. Liers, S. Woods, S. Young, R. DeBlasio, and L. Mrig, in *16th photovoltaic specialists conference* (1982) pp. 1092–1101.
- [59] J.-H. Yoon, J.-K. Park, W. M. Kim, J. Lee, H. Pak, and J.-h. Jeong, *Characterization of efficiency-limiting resistance losses in monolithically integrated cu (in, ga) se₂ solar modules*, Scientific Reports **5**, 7690 (2015).
- [60] M. Theelen, A. Liakopoulou, V. Hans, F. Daume, H. Steijvers, N. Barreau, Z. Vroon, and M. Zeman, *Determination of the temperature dependency of the electrical parameters of CIGS solar cells*, Journal of Renewable and Sustainable Energy **9**, 021205 (2017), https://pubs.aip.org/aip/jrse/article-pdf/doi/10.1063/1.4979963/15698944/021205_1_online.pdf .
- [61] M. Theelen and F. Daume, *Stability of Cu(In,Ga)Se₂ solar cells: A literature review*, Solar Energy **133**, 586 (2016).
- [62] J. Kettle, M. Aghaei, S. Ahmad, A. Fairbrother, S. Irvine, J. J. Jacobson, S. Kazim, V. Kazukauskas, D. Lamb, K. Lobato, *et al.*, *Review of technology specific degradation in crystalline silicon, cadmium telluride, copper indium gallium selenide, dye sensitised, organic and perovskite solar cells in photovoltaic modules: Understanding how reliability improvements in mature technologies can enhance emerging technologies*, Progress in Photovoltaics: Research and Applications **30**, 1365 (2022).
- [63] G. Makrides, B. Zinsser, M. Schubert, and G. E. Georghiou, *Performance loss rate of twelve photovoltaic technologies under field conditions using statistical techniques*, Solar Energy **103**, 28 (2014).

- [64] M. N. Ruberto and A. Rothwarf, *Time-dependent open-circuit voltage in CuInSe₂/CdS solar cells: Theory and experiment*, Journal of Applied Physics **61**, 4662 (1987).
- [65] D. Willett and S. Kuriyagawa, in *Conference Record of the Twenty Third IEEE Photovoltaic Specialists Conference-1993 (Cat. No. 93CH3283-9)* (IEEE, 1993) pp. 495–500.
- [66] R. Sasala and J. Sites, in *Conference Record of the Twenty Third IEEE Photovoltaic Specialists Conference-1993 (Cat. No. 93CH3283-9)* (IEEE, 1993) pp. 543–548.
- [67] F. Engelhardt, M. Schmidt, T. Meyer, O. Seifert, J. Parisi, and U. Rau, *Metastable electrical transport in Cu(In,Ga)Se₂ thin films and ZnO/CdS/Cu(In,Ga)Se₂ heterostructures*, Physics Letters A **245**, 489 (1998).
- [68] M. Igalson, A. Kubiacyk, and P. Zabierowski, *Deep centers and fill factor losses in the CIGS devices*, MRS Online Proceedings Library (OPL) **668**, H9 (2001).
- [69] C. Walkons, M. Jahandardoost, T. M. Friedlmeier, W. Hempel, S. Paetel, M. Nardone, B. Ursprung, E. S. Barnard, K. E. Kweon, V. Lordi, *et al.*, *Behavior of Na and RbF-treated CdS/Cu(In,Ga)Se₂ solar cells with stress testing under heat, light, and junction bias*, physica status solidi (RRL)–Rapid Research Letters **15**, 2000530 (2021).
- [70] T. Vaas, B. Pieters, D. Roosen-Melsen, M. van den Nieuwenhof, A. Kingma, M. Theelen, C. Zahren, A. Gerber, and U. Rau, *Light induced degradation of cigs solar cells*, Solar Energy Materials and Solar Cells **275**, 113036 (2024).
- [71] M. Sodhi, L. Banaszek, C. Magee, and M. Rivero-Hudec, *Economic lifetimes of solar panels*, Procedia CIRP **105**, 782 (2022).
- [72] P.-O. Westin, P. Neretnieks, and M. Edoff, in *21st European Photovoltaic Solar Energy Conference* (2006) pp. 2470–2473.
- [73] D.-W. Lee, O.-Y. Kwon, J.-K. Song, C.-H. Park, K.-E. Park, S.-M. Nam, and Y.-N. Kim, *Effects of ZnO:Al films on CIGS PV modules degraded under accelerated damp heat*, Solar Energy Materials and Solar Cells **105**, 15 (2012).

- [74] D.-w. Lee, W.-j. Cho, C.-i. Jang, J.-k. Song, C.-h. Park, K.-e. Park, J.-s. Ryu, H. Lee, and Y.-n. Kim, *Damp heat and thermal cycling-induced degradation mechanism of AZO and CIGS films in Cu(In,Ga)Se₂ photovoltaic modules*, Current Applied Physics **15**, 285 (2015).
- [75] M. Theelen, R. Hendrikx, N. Barreau, H. Steijvers, and A. Böttger, *The effect of damp heat-illumination exposure on CIGS solar cells: A combined XRD and electrical characterization study*, Solar Energy Materials and Solar Cells **157**, 943 (2016).
- [76] M. Theelen, K. Beyeler, H. Steijvers, and N. Barreau, *Stability of CIGS solar cells under illumination with damp heat and dry heat: A comparison*, Solar Energy Materials and Solar Cells **166**, 262 (2017).
- [77] J. Malmström, J. Wennerberg, and L. Stolt, *A study of the influence of the Ga content on the long-term stability of Cu(In,Ga)Se₂ thin film solar cells*, Thin Solid Films **431**, 436 (2003).
- [78] M. Theelen, K. Polman, M. Tomassini, N. Barreau, H. Steijvers, J. van Berkum, Z. Vroon, and M. Zeman, *Influence of deposition pressure and selenisation on damp heat degradation of the Cu(In,Ga)Se₂ back contact molybdenum*, Surface and Coatings Technology **252**, 157 (2014).
- [79] M. Theelen, S. Harel, M. Verschuren, M. Tomassini, A. Hovestad, N. Barreau, J. van Berkum, Z. Vroon, and M. Zeman, *Influence of Mo/MoSe₂ microstructure on the damp heat stability of the Cu(In,Ga)Se₂ back contact molybdenum*, Thin Solid Films **612**, 381 (2016).
- [80] V. Fjällström, P. M. P. Salomé, A. Hultqvist, M. Edoff, T. Jarmar, B. G. Aitken, K. Zhang, K. Fuller, and C. K. Williams, *Potential-induced degradation of CuIn_{1-x}Ga_xSe₂ thin film solar cells*, IEEE Journal of Photovoltaics **3**, 1090 (2013).
- [81] P. Yilmaz, J. de Wild, R. Aninat, T. Weber, B. Vermang, J. Schmitz, and M. Theelen, *In-depth analysis of potential-induced degradation in a commercial CIGS PV module*, Progress in Photovoltaics: Research and Applications (2023).
- [82] S. Lee, S. Bae, S. J. Park, J. Gwak, J. Yun, Y. Kang, D. Kim, Y.-J. Eo, and H.-S. Lee, *Characterization of potential-induced degradation and recovery in CIGS solar cells*, Energies **14**, 4628 (2021).

- [83] S. P. Harvey, H. Guthrey, C. P. Muzzillo, G. Teeter, L. Mansfield, P. Hacke, S. Johnston, and M. Al-Jassim, *Investigating PID shunting in polycrystalline CIGS devices via multi-scale, multi-technique characterization*, IEEE Journal of Photovoltaics **9**, 559 (2019).
- [84] S. Yamaguchi, S. Jonai, K. Hara, H. Komaki, Y. Shimizu-Kamikawa, H. Shibata, S. Niki, Y. Kawakami, and A. Masuda, *Potential-induced degradation of Cu(In,Ga)Se₂ photovoltaic modules*, Japanese Journal of Applied Physics **54**, 08KC13 (2015).
- [85] O. Salomon, W. Hempel, O. Kiowski, E. Lotter, W. Witte, A. Ferati, A. Schneikart, G. Kaune, R. Schäffler, M. Becker, *et al.*, *Influence of molybdenum back contact on the PID effect for Cu(In,Ga)Se₂ solar cells*, Coatings **9**, 794 (2019).
- [86] C. P. Muzzillo, K. Terwilliger, P. Hacke, H. R. Moutinho, M. R. Young, S. Glynn, B. Stevens, I. L. Repins, and L. M. Mansfield, *Potential-induced degradation of cu(in,ga)se₂ can occur by shunting the front i-zno and by damaging the p-n junction*, Solar Energy **232**, 298 (2022).
- [87] V. Fjällström, P. Szaniawski, B. Vermang, P. M. P. Salomé, F. Rostvall, U. Zimmermann, and M. Edoff, *Recovery after potential-induced degradation of CuIn_{1-x}Ga_xSe₂ solar cells with CdS and Zn(O,S) buffer layers*, IEEE Journal of Photovoltaics **5**, 664 (2015).
- [88] U. Rau, M. Schmitt, J. Parisi, W. Riedl, and F. Karg, *Persistent photoconductivity in Cu(In,Ga)Se₂ heterojunctions and thin films prepared by sequential deposition*, Applied Physics Letters **73**, 223 (1998).
- [89] T. Meyer, M. Schmidt, F. Engelhardt, J. Parisi, and U. Rau, *A model for the open circuit voltage relaxation in Cu(In,Ga)Se₂ heterojunction solar cells*, The European Physical Journal Applied Physics **8**, 43 (1999).
- [90] J. Heath, J. Cohen, and W. Shafarman, *Distinguishing metastable changes in bulk CIGS defect densities from interface effects*, Thin Solid Films **431**, 426 (2003).
- [91] I. Eisgruber, J. Granata, J. Sites, J. Hou, and J. Kessler, *Blue-photon modification of nonstandard diode barrier in CuInSe₂ solar cells*, Solar Energy Materials and Solar Cells **53**, 367 (1998).
- [92] M. Igalson and H. Schock, *The metastable changes of the trap spectra of CuInSe₂-based photovoltaic devices*, Journal of Applied Physics **80**, 5765 (1996).

- [93] V. Nádaždy, M. Yakushev, E. Djebbar, A. Hill, and R. Tomlinson, *Switching of deep levels in CuInSe₂ due to electric field-induced Cu ion migration*, Journal of applied physics **84**, 4322 (1998).
- [94] M. Igalson, M. Cwil, and M. Edoff, *Metastabilities in the electrical characteristics of CIGS devices: Experimental results vs theoretical predictions*, Thin Solid Films **515**, 6142 (2007).
- [95] I. Repins, S. Glynn, T. J. Silverman, R. Garriss, K. Bowers, B. Stevens, and L. Mansfield, *Large metastability in Cu(In,Ga)Se₂ devices: The importance of buffer properties*, Progress in Photovoltaics: Research and Applications **27**, 749 (2019).
- [96] I. Repins, S. Glynn, K. Bowers, B. Stevens, C. L. Perkins, and L. Mansfield, *Using hole injection layers for decreased metastability and higher performance in Cu(In,Ga)Se₂ devices*, Solar Energy Materials and Solar Cells **215**, 110597 (2020).
- [97] S. Lany and A. Zunger, *Light-and bias-induced metastabilities in Cu(In,Ga)Se₂ based solar cells caused by the (V_{Se} - V_{Cu}) vacancy complex*, Journal of Applied Physics **100** (2006).
- [98] M. Maciaszek and P. Zabierowski, in *2015 IEEE 42nd Photovoltaic Specialist Conference (PVSC)* (IEEE, 2015) pp. 1–3.
- [99] K. Macielak, M. Maciaszek, M. Igalson, P. Zabierowski, and N. Barreau, *Persistent photoconductivity in polycrystalline Cu(In,Ga)Se₂ thin films: Experiment versus theoretical predictions*, IEEE Journal of Photovoltaics **5**, 1206 (2015).
- [100] F. Obereigner, N. Barreau, W. Witte, and R. Scheer, *Open-circuit and doping transients of Cu(In,Ga)Se₂ solar cells with varying Ga content*, Journal of Applied Physics **117** (2015).
- [101] J.-F. Guillemoles, L. Kronik, D. Cahen, U. Rau, A. Jasenek, and H.-W. Schock, *Stability issues of Cu(In,Ga)Se₂-based solar cells*, The Journal of Physical Chemistry B **104**, 4849 (2000).
- [102] S. Ishizuka, N. Taguchi, and P. J. Fons, *Similarities and critical differences in heavy alkali-metal rubidium and cesium effects on chalcopyrite Cu(In,Ga)Se₂ thin-film solar cells*, The Journal of Physical Chemistry C **123**, 17757 (2019).

- [103] M. Gostein and L. Dunn, in *2011 37th IEEE Photovoltaic Specialists Conference* (IEEE, 2011) pp. 003126–003131.
- [104] A. Kanevce, S. Paetel, D. Hariskos, and T. M. Friedlmeier, *Impact of RbF-PDT on Cu(In,Ga)Se₂ solar cells with CdS and Zn(O,S) buffer layers*, EPJ Photovoltaics **11**, 8 (2020).
- [105] C. Zhao, S. Yu, W. Tang, X. Yuan, H. Zhou, T. Qi, X. Zheng, D. Ning, M. Ma, J. Zhu, *et al.*, *Advances in CIGS thin film solar cells with emphasis on the alkali element post-deposition treatment*, Materials Reports: Energy , 100214 (2023).
- [106] MiaSolé Hi-Tech Corp, Flexible solar cell datasheet, Brochure (2022), accessed on November 5, 2023.
- [107] W. Shockley, *The theory of p-n junctions in semiconductors and p-n junction transistors*, Bell system technical journal **28**, 435 (1949).
- [108] B. Pieters, H. Stiebig, M. Zeman, and R. Van Swaaij, *Determination of the mobility gap of intrinsic μ c-Si: H in pin solar cells*, Journal of Applied Physics **105** (2009).
- [109] D. C. Jordan and S. R. Kurtz, *The dark horse of evaluating long-term field performance—data filtering*, IEEE Journal of Photovoltaics **4**, 317 (2013).
- [110] C. Hansen, *Parameter estimation for single diode models of photovoltaic modules*, Tech. Rep. (Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2015).
- [111] T. S. Vaas, J. Körtgen, E. Sovetkin, U. Rau, and B. E. Pieters, in *2023 IEEE 50th Photovoltaic Specialists Conference (PVSC)* (2023) pp. 1–4.
- [112] D. Hawkins, *Identification of outliers*, Monographs on applied probability and statistics (Chapman and Hall, London [u.a.], 1980).
- [113] B. Marion, M. G. Deceglie, and T. J. Silverman, *Analysis of measured photovoltaic module performance for florida, oregon, and colorado locations*, Solar energy **110**, 736 (2014).
- [114] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen,

- E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, Nature Methods **17**, 261 (2020).
- [115] D. Faiman, *Assessing the outdoor operating temperature of photovoltaic modules*, Progress in Photovoltaics: Research and Applications **16**, 307 (2008).
- [116] T. S. Vaas, B. Pieters, E. Sovetkin, A. Gerber, and U. Rau, *A gaussian process regression iv model for pv outdoor data*, Authorea (2024).
- [117] A. Ndiaye, C. M. Kébé, A. Charki, P. A. Ndiaye, V. Sambou, and A. Kobi, *Degradation evaluation of crystalline-silicon photovoltaic modules after a few operation years in a tropical environment*, Solar Energy **103**, 70 (2014).
- [118] O. S. Sastry, S. Saurabh, S. Shil, P. Pant, R. Kumar, A. Kumar, and B. Bandopadhyay, *Performance analysis of field exposed single crystalline silicon modules*, Solar energy materials and solar cells **94**, 1463 (2010).
- [119] A. Phinikarides, G. Makrides, B. Zinsser, M. Schubert, and G. E. Georghiou, *Analysis of photovoltaic system performance time series: Seasonality and performance loss*, Renewable Energy **77**, 51 (2015).
- [120] R. Saive, *S-shaped current–voltage characteristics in solar cells: A review*, IEEE Journal of Photovoltaics **9**, 1477 (2019).
- [121] A. R. Jordehi, *Parameter estimation of solar photovoltaic (pv) cells: A review*, Renewable and Sustainable Energy Reviews **61**, 354 (2016).
- [122] A. Dolara, S. Leva, and G. Manzolini, *Comparison of different physical models for pv power output prediction*, Solar energy **119**, 83 (2015).
- [123] G. Cui, X. Yu, S. Iommelli, and L. Kong, *Exact distribution for the product of two correlated gaussian random variables*, IEEE Signal Processing Letters **23**, 1662 (2016).
- [124] G. Fubini and L. Tonelli, *Sulla derivata seconda mista di un integrale doppio*, Rendiconti del Circolo Matematico di Palermo (1884-1940) **40**, 295 (1915).
- [125] K. Wang, G. Pleiss, J. Gardner, S. Tyree, K. Q. Weinberger, and A. G. Wilson, *Exact gaussian processes on a million data points*, Advances in neural information processing systems **32** (2019).

- [126] Z. Li, *Prospects of photovoltaic technology*, Engineering **21**, 28 (2023).
- [127] T. A. Chowdhury, M. A. B. Zafar, M. S.-U. Islam, M. Shahinuzzaman, M. A. Islam, and M. U. Khandaker, *Stability of perovskite solar cells: issues and prospects*, RSC advances **13**, 1787 (2023).
- [128] C. Farrell, A. Osman, R. Doherty, M. Saad, X. Zhang, A. Murphy, J. Harrison, A. Vennard, V. Kumaravel, A. Al-Muhtaseb, *et al.*, *Technical challenges and opportunities in realising a circular economy for waste photovoltaic modules*, Renewable and Sustainable Energy Reviews **128**, 109911 (2020).
- [129] A. Mulazzani, P. Eleftheriadis, and S. Leva, *Recycling c-si pv modules: A review, a proposed energy model and a manufacturing comparison*, Energies **15**, 8419 (2022).

Curriculum Vitae

10.10.2025	Doctoral Examination for Dr. rer. nat.
2020–2024	Research Assistant and Doctoral Candidate Institute of Energy Materials and Devices, Photovoltaics (IMD-3) RWTH Aachen University, Germany
2014–2019	Studies of Physics (Bachelor and Master) RWTH Aachen University, Germany
December 2019	Master of Science RWTH Aachen University
September 2017	Bachelor of Science
2006–2014	General Higher Education Entrance Qualification (Abitur) Gymnasium Marienschule Euskirchen, Euskirchen, Germany
1996	born in Euskirchen, Germany

Acknowledgements

First of all, I would like to thank my parents Uwe and Adelheid. You have been nothing but supportive throughout my time as bachelor, master and PhD student. My appreciation cannot fully be expressed in words so I keep it simple: I love you and I am grateful for all you have done to pave my way.

I would like to thank Dr. Bart Pieters. There is no doubt, that he had the biggest impact on my work, providing me with feedback and ideas whenever needed. I do not only appreciate his expertise on physics and electrical engineering, but most I appreciate, that despite the discrepancy of my and his knowledge he always valued my ideas and opinion. I truly think I could not have any better supervisor, always encouraging me and helping me to improve on so many aspects, that you need as a scientist. Further, I am grateful for the (I guess meanwhile) more than a hundred meetings, which not only once ended up in completely unrelated discussions. I could spent hours just discussing with Bart about anything and that pretty much should show, how much I also value him apart being an excellent scientist.

Further, I would like to thank Prof. Uwe Rau for his expertise, his input on my research and his intransigence regarding improvements on my publications. Prof. Rau always encouraged me to go the extra mile and to not be pleased by "good enough".

I would like to thank Prof. Dr. Marko Topič for his feedback he gave and time he spent as second examiner.

I am also grateful to Dr. Andreas Gerber. I could always rely on his advice not only regarding my scientific work, but also on a personal level.

Furthermore, the whole "Messtechnik"-group deserves a big thank you. Thanks to Christoph Zahren and Niklas Bongartz for your help with the laboratory experiments. Thanks to Sabine Kasper for the help at all set-ups I used. Thanks to Thomas Birrenbach for the help on the encapsulator. Thanks to Michael Gordon, Jan Körtgen, and especially to Evgenii Sovetkin for your advice and feedback to my work, that involved programming in any way. Furthermore,

thank you to Jonas Noll, Dirk Erdweg, Wilfried Reetz, Willi Hilgers, Jürgen Radde, Resi Oligschleger, Suheir Nofal, Armin Matindoust, Jonathan Werner and Vladimir Smirnov. Whatever problem or question I had, someone in the group was always there to help out.

I would like to thank Markus Hülsbeck, Daniel Weigand, Prof. Maurice Nuys and Prof. Stefan Haas. Even though they were not directly involved into my topics, they all were willing to help me with their experience.

Thanks to everyone, that was involved in the CIGS TheoMax project. I got very much valuable feedback from Dr. Ana Kanevce, Mario Zinßer, Dr. Theresa Magorian-Friedlmeier, Tim Helder, Dr. Andreas Bauer, Seçil Güler, Prof. Thomas Kirchartz and of course also from Dr. Bart Pieters and Prof. Uwe Rau. Thank you as well to Dr. Dorrit Roosen-Melsen, Dr. Aldo Kingma, Dr. Monique van den Nieuwenhof and Dr. Mirjam Theelen from TNO, partner in Solliance, from where I got the samples for the laboratory experiments. I enjoyed my work travels to Eindhoven and appreciate all the input I got from planning the experiment to finalizing the publication.

I further would like to thank all the people at IEK-5/IMD-3. The institute is not only full of great scientists, but also full of excellent people. I am very grateful to have colleagues like you. Thank you for making my phd time so much better to Daniel, Markus, Toby, Sergey, Seçil, Thérèse, Joachim, Maurice, Yanxin, Mohammed, Ifeanyi, Susanne, Yannick, Gina, Janine, Christian, Alain, Pascal, Lukas, Lennard, Armin, Sandheep, Johanna and so many more. A special thank you goes to Suheir and Niklas for tolerating me in the shared office. I guess our office will be for some more time the only office in the institute, which has had its own christmas tree including self-made cookies and (fake) presents. Thanks a lot to Basita, who has become such a good friend I can always rely on. Thanks as well to Alex and Benedikt for the help and the good time in Puerto Rico. Thank you also to all BSG members I enjoyed so many hours on the football pitch with.

I would like to thank my family and in particular Gerrit, Maike, Eva, Frank, Karl, Helena, Hector and Lore. There is nothing more important in life than family.

Last but not least I would like to thank all my friends. Tristan, Tim, Philipp, Gerrit, Kübra, Olivia, Elsa, David, Flori, Floflo, Basti, Thomas, Knese, Lars, Chris, Tim, Lea, Martin, Mehmet, Alex and Alex all helped me a lot to balance all the stress a PhD thesis comes with.

Band / Volume 671

Improving Energy Efficiency of Public Buildings by Influencing Occupant Behaviour using Dashboards and Gamification

E. Ubachukwu (2025), xxi, 191 pp

ISBN: 978-3-95806-840-7

Band / Volume 672

Exploring Plant Responses to Changing Environments: Integrating Phenotyping and Modeling Across Scales

F. M. Bauer (2025), xxix, 188 pp

ISBN: 978-3-95806-845-2

Band / Volume 673

A constitutive theory to represent non-idealities in contacting of SOC interconnect contacts

R. M. Pinto (2025), xii, 139 pp

ISBN: 978-3-95806-846-9

Band / Volume 674

Strontium titanate based materials for use as oxygen transport membranes in membrane reactors

Y. Tang (2025), XIV, 132 pp

ISBN: 978-3-95806-849-0

Band / Volume 675

Scaling Methods for the Production of Tungsten Fiber-Reinforced Composites via Chemical Vapor Deposition

A. Lau (2025), untersch. Pag.

ISBN: 978-3-95806-851-3

Band / Volume 676

Nanoscale analysis of high-temperature oxidation mechanisms of Cr₂AlC MAX phase and W-Cr-Y self-passivating tungsten alloy

A.J. S. Reuban (2025), ix, 142 pp

ISBN: 978-3-95806-855-1

Band / Volume 677

First principles simulations of high-entropy materials for energy storage

Y. Ting (2025), xviii, 169 pp

ISBN: 978-3-95806-858-2

Band / Volume 678

Deployment of Fuel Cell Vehicles in Road Transport and the Expansion of the Hydrogen Refueling Station Network

T. Grube, M. Sander (2025), iv, 61 pp

ISBN: 978-3-95806-859-9

Band / Volume 679

Entwicklung von nickelbasierten katalysatorbeschichteten Diaphragmen für die alkalische Wasserelektrolyse

C. B. Karacan (2025), 146 pp

ISBN: 978-3-95806-860-5

Band / Volume 680

Bewertung lokaler Eigenspannungsverteilungen bei der lokalen Bauteilreparatur durch Kaltgasspritzen

J.-C. Schmitt (2025), 154, xxvii pp

ISBN: 978-3-95806-861-2

Band / Volume 681

First principles study of the effect of substitution\doping on the performance of layered oxide cathode materials for secondary batteries

N. Yacoob (2025), iii, 126 pp

ISBN: 978-3-95806-864-3

Band / Volume 682

Field assisted sintering technology/spark plasma sintering in the direct recycling of hot-deformed Nd-Fe-B scrap and PM T15 steel swarf

M. T. M. Keszler (2025), viii, 173 pp

ISBN: 978-3-95806-866-7

Band / Volume 683

Assessment of erosion in recessed areas of fusion devices using multi-scale computer simulations

S. Rode (2025), viii, 196 pp

ISBN: 978-3-95806-867-4

Band / Volume 684

Europäische Energiewende – Deutschland im Herzen Europas

T. Klütz, P. Dunkel, T. Busch, J. Linssen, D. Stolten (2025), IV, 56 pp

ISBN: 978-3-95806-870-4

Band / Volume 685

Performance and stability of solar cells and modules: From laboratory characterization to field data analysis

T. S. Vaas (2025), xvii, 146 pp

ISBN: 978-3-95806-871-1

Energie & Umwelt / Energy & Environment
Band / Volume 685
ISBN 978-3-95806-871-1