

AALBORG UNIVERSITY

MASTER THESIS

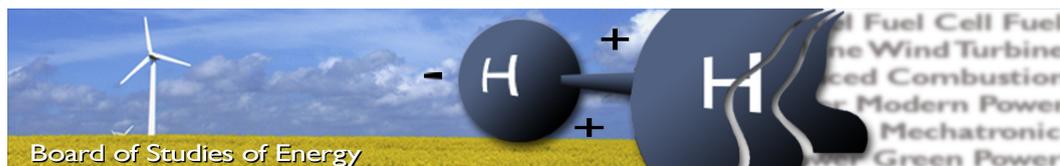
Performance Modelling and Monitoring of Thin-film Photovoltaic Systems

PED4-1048

Department of Energy Technology

Anamaria Florina GAVRILUȚĂ

May 2016



Title: Performance Modelling and Monitoring of Thin-film Photovoltaic Systems
Semester: 10
Semester theme: Short Master Thesis
Project period: 15.02.2016 to 1.06.2016
ECTS: 30
Supervisor: Sergiu Spătaru, Dezso Sera
Project group: PED4-1048

Anamaria Gavriluta

SYNOPSIS:

Denmark has reached 790 MW in late 2015 with the vast majority of this capacity installed in just four years. The goal for the Danish Photovoltaic energy sector is to have 1000 MW by 2020 and 3400 MW by 2030. This thesis is part of the the research project Fully Automated Service Execution platform for Photovoltaic power plants (FASE) financed by Innovation Fund Denmark and developed in collaboration with Danica Pension, Econ GmbH, GreenGo Energy A/S and DTU Compute Technical University of Denmark. The projects vision is to significantly improve the profitability of PV Plants by integrating state-of-the-art PV fault diagnosis with an automatic service execution platform. The main challenges addressed in this master thesis is developing the necessary PV performance models for fault diagnosis in thin-film based PV plants.

Keywords: solar power, performance modelling, performance monitoring, FASE, PV models

Copies: 1
 Pages, total: 75
 Appendix: 1
 Supplements: 0

By signing this document, each member of the group confirms that all group members have participated in the project work, and thereby all members are collectively liable for the contents of the report. Furthermore, all group members confirm that the report does not include plagiarism.

Preface

The present report represents the final project of the Power Electronics and Drives master program from Aalborg University. This master thesis project, entitled *Performance Modelling and Monitoring of Thin-film Photovoltaic Systems* was realised to support the research project Fully Automated Service Execution platform for Photovoltaic power plants (FASE) financed by Innovation Fund Denmark and developed in collaboration with Danica Pension, Econ GmbH, GreenGo Energy A/S and DTU Compute Technical University of Denmark. The projects vision is to significantly improve the profitability of PV Plants by integrating state-of-the-art PV fault diagnosis with an automatic service execution platform. The main challenges addressed in this master thesis is developing the necessary PV performance models for fault diagnosis in thin-film based PV plants.

The project is documented in a main report and appendices. The main report contains theory background and analysis of PV panel field performance measurements, while the appendices contain more information regarding the implementation of the performance modelling. The chapters in this project are consecutive numbered whereas the appendixes are labelled with letters. The figures, equations and tables are numbered in succession within the chapters.

Acknowledgement

I would like to thank Sergiu and Prof. Sera for their advice and guidance throughout this project and to the National Renewable Energy Laboratory for providing me with access to their data.

Contents

Preface	i
Contents	ii
List of Figures	iv
List of Tables	vii
Abbreviations	viii
Symbols	ix
1 Introduction	1
1.1 Background	1
1.2 Problem Formulation	2
1.3 Objectives	6
1.4 Project limitations	7
2 Performance characterization of Photovoltaic Technologies	8
2.1 Analysis of metastability for Thin-Film technologies	8
2.1.1 Seasonal effects	11
2.1.2 Diurnal effects	19
2.2 Conclusions	23
3 Performance Modeling	24
3.1 Performance models	24
3.2 Irradiance profile for Denmark	27
3.3 Method for model identification	29
3.4 PVWatts vs. SAPM	31
3.4.1 SAPM	31
3.4.2 PVWatts	33
4 Results	36
4.1 Definition of Test Cases	36
4.2 Comparison of SAPM vs. PVWatts	38
4.3 Analysis of dataset size	38
4.4 Analysis of dataset distribution	40
4.5 Analysis of measurement procedure	45
4.6 Conclusions	48

5	Conclusions	50
5.1	Future work	51
A	Appendix	53
A.1	Matlab Code	53
A.2	PVWatts vs SAPM	62
	Bibliography	63

List of Figures

1.1	Solar PV Total Global Capacity, 2004–2015	1
1.2	Cumulated PV capacity in megawatt (MW) in Denmark since 2004 . . .	2
1.3	Diagram of a simplified PV system and its performance model structure with practical challenges highlighted with red dotted line and identified with letters	5
2.1	Spectral response of pyranometer and different solar cells devices (CIGS, CdTe, a-Si, c-Si)	9
2.2	Modeling the short-circuit current dependence on AM-units for a-Si (blue circles) and c-Si (red squares) PV modules. Dark solid lines are linear fits to the modeled data	10
2.3	Seasonal Variation of the Operating Efficiency of Devices	11
2.4	Normalized performance measurements from modules in Florida for a year. Blue - Irradiance, Orange - Module Temperature, Green - Current, Light Blue - Voltage	12
2.5	Performance Ratio of c-Si technologie at Cocoa beach for an year using NREL measurements showing seasonal effect for $G = [950 : 1050]W/m^2$ and $T_{mod} = [48 : 52]^{\circ}C$	13
2.6	Performance Ratio of c-Si technologies at Cocoa beach for an year using NREL measurements showing seasonal effect for different irradiance and module temperature values	14
2.7	Performance Ratio of a-Si technology at Cocoa beach for an year using NREL measurements showing seasonal effect for $G = [950 : 1050]W/m^2$ and $T_{mod} = [48 : 52]^{\circ}C$	15
2.8	Performance Ratio of a-Si technology at Cocoa beach for an year using NREL measurements showing seasonal effect for different irradiance and module temperature values	15
2.9	Performance Ratio of CdTe technology at Cocoa beach for an year using NREL measurements showing seasonal effect for $G = [950 : 1050]W/m^2$ and $T_{mod} = [48 : 52]^{\circ}C$	16
2.10	Performance Ratio of CdTe technology at Cocoa beach for an year using NREL measurements showing seasonal effect for different irradiance and module temperature values	16
2.11	Performance Ratio for thin-film CIGS system over the period April 2011- March 2016 in Golden Colorado for $G = [950 : 1050]W/m^2$ and $T_{mod} =$ $[48 : 52]^{\circ}C$	17
2.12	Performance Ratio for thin-film CIGS system over the period April 2011- March 2016 in Golden Colorado for different irradiance and module tem- perature values	17

2.13	Performance ratio for thin-film CIGS system over the period April 2011 - March 2016 in Cocoa Florida for $G = [950 : 1050]W/m^2$ and $T_{mod} = [48 : 52]^{\circ}C$	18
2.14	Performance ratio for thin-film CIGS system over the period April 2011 - March 2016 in Cocoa Florida for different irradiance and module temperature values	18
2.15	Module temperature effect on the PR of the module in spring period for a c-Si at Cocoa beach Florida	19
2.16	PR for a clear sky day in spring for Florida Cocoa beach comparing the four technologies	20
2.17	PR vs. Irradiance for a day in spring season showing the difference when there is no temperature correction and when the temperature is corrected at $25^{\circ}C$	21
2.18	PR for a clear sky day in summer for Florida Cocoa beach comparing the four technologies	21
2.19	PR vs. Irradiance for a day in summer season showing the difference when there is no temperature correction and when the temperature is corrected at $25^{\circ}C$	22
2.20	PR for a clear sky day in autumn for Florida Cocoa beach comparing the four technologies	22
2.21	PR vs. Irradiance for a day in autumn season showing the difference when there is no temperature correction and when the temperature is corrected at $25^{\circ}C$	23
3.1	Example of sunny and cloudy day	26
3.2	Distribution of measurement samples vs. irradiance for a sunny and a cloudy sky day	27
3.3	Monthly average of irradiance and module temperature together with the number of sunny days/month for three different geographic locations, Cocoa-Florida, Golden-Colorado and Aalborg-Denmark	28
3.4	Number of sunny days/month for three different geographic locations, Cocoa-Florida, Golden-Colorado and Aalborg-Denmark	29
3.5	Diagram of training and test phase showing how the PV model parameters are obtained and validated	30
4.1	Measured power vs. predicted power using a varying training dataset from winter and tested on a fixed size test dataset number from summer and winter	40
4.2	Distribution of irradiance vs. number of samples for the four training datasets	42
4.3	Training dataset made of 4 sunny days and test dataset of 2 sunny and 2 cloudy days ($4s - 2s2c$)	43
4.4	Training dataset made of 4 cloudy days and test dataset of 2 sunny and 2 cloudy days ($4c - 2s2c$)	43
4.5	Training dataset made of 2 sunny and 2 cloudy days and test dataset of 2 sunny and 2 cloudy days ($2s2c - 2s2c$)	44
4.6	Training dataset made of 1 sunny and 9 cloudy days and test dataset of 2 sunny and 2 cloudy days ($1s9c - 2s2c$)	44

4.7	Training dataset made of 4 sunny days and test dataset of 2 sunny and 2 cloudy days ($4s - 2s2c$) using effective irradiance	46
4.8	Training dataset made of 4 cloudy days and test dataset of 2 sunny and 2 cloudy days ($4c - 2s2c$) using effective irradiance	47
4.9	Training dataset made of 2 sunny and 2 cloudy days and test dataset of 2 sunny and 2 cloudy days ($2s2c - 2s2c$) using effective irradiance	47
4.10	Training dataset made of 1 sunny and 9 cloudy days and test dataset of 2 sunny and 2 cloudy days ($1s9c - 2s2c$) using effective irradiance	48
A.1	Modeling error statistics for P_m using SAPM and PVWatts models on a CIGS module considering the four different cases	62

List of Tables

2.1	Temperature coefficients used for temperature correction of the power . .	13
4.1	Defined TC by training dataset and test dataset	37
4.2	RMSE and MAE statistics for P_{mp} by model and case	38
4.3	RMSE and MAE statistics for P_m using PVWatts model for four cases in order to establish the training dataset size	39
4.4	RMSE and MAE statistics for P_m using PVWatts model for all test cases studying the irradiance distribution impact at Cocoa Beach, Florida . . .	41
4.5	RMSE and MAE statistics for P_m using PVWatts model for all test cases studying the irradiance distribution impact at Golden, Colorado	41
4.6	RMSE and MAE statistics for P_m using PVWatts model for all test cases studying the effective irradiance distribution impact at Cocoa Beach, Florida	45

Abbreviations

PV	Photovoltaic
GW	Giga Watts
MW	Mega Watts
STC	Standard Testing Conditions (1000 W/m^2 , 25°C , 1.5 AM)
AM	Air Mass
c-Si	Crystalline Silicon
x-Si	Mono-crystalline Silicon
m-Si	Multi-crystalline Silicon
a-Si	Amorphous Silicon
CIS	Copper Indium Selenide
CIGS	Copper Indium Gallium Selenide
CdTe	Cadmium Telluride
FASE	Fully Automated Service Execution
SAPM	Sandia Array Performance Model
SWE	Staebler Wronski Effect
NREL	National Renewable Energy Laboratory
PR	Performance Ratio
MPP	Maximum Power Point
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
TC	Test Cases

Symbols

V	Voltage	[V]
I	Current	[A]
P	Power	[W]
R	Resistance	[Ω]
T	Temperature	[$^{\circ}C$]
E	Irradiance	[W/m^2]
λ	Wavelength	[nm]
PR	Performance Ratio	[%]
T	Temperature	[$^{\circ}C$]
$\alpha_{I_{mp}}$	Temperature coefficient for current	[$1/^{\circ}C$]
$\alpha_{V_{mp}}$	Temperature coefficient for voltage	[$V/^{\circ}C$]
k	Boltzmann constant	[J/K]
q	Electron charge	[coulomb]
$RMSE$	Root mean square error	[%]
MAE	Mean absolute error	[%]

Chapter 1

Introduction

1.1 Background

Renewable energy technologies have matured considerably in the last decades, and are now becoming competitive with traditional fossil based energy generation in many countries. The photovoltaic (PV) energy industry has been steadily growing for the last decades at a very fast rate, with around 30% annual increase in cumulative installed capacity worldwide. At the end of 2015, the capacity reached a global total of about 227.1 GW as shown in Figure 1.1 [1].

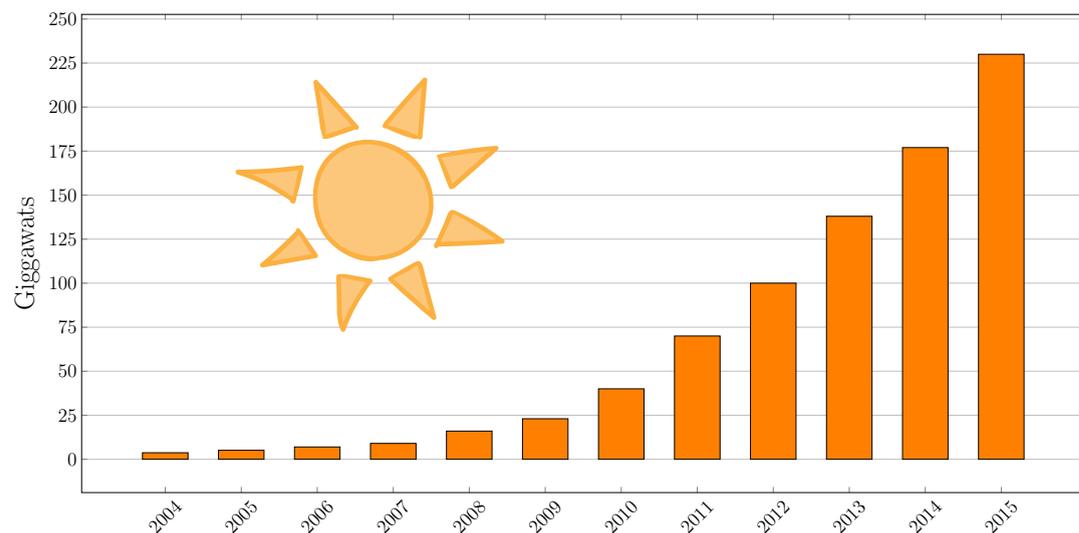


FIGURE 1.1: Solar PV Total Global Capacity, 2004–2015 [2]

Denmark has reached 790 MW in late 2015 with the vast majority of this capacity installed in just four years [2]. The rise in capacity during the past 11 years can be seen in Figure 1.2. The goal for the Danish Photovoltaic energy sector is to have 1000 MW by 2020 and 3400 MW by 2030¹.

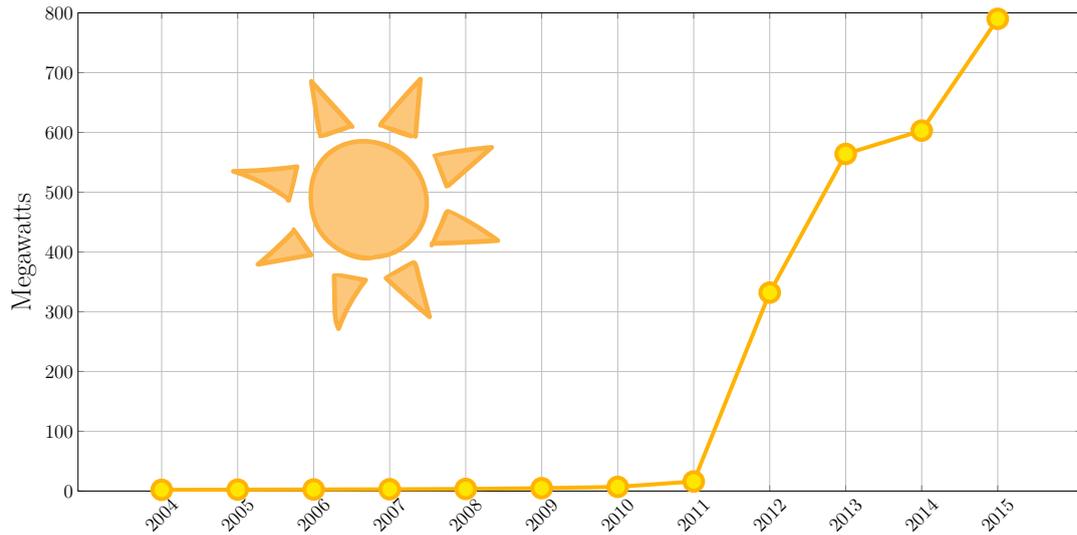


FIGURE 1.2: Cumulated PV capacity in megawatt (MW) in Denmark since 2004

With the growing industry, the initial cost of PV systems has been showing a fast decline, with over 75% drop over the last 10 years [3]. In parallel and as a result, the operation and maintenance costs have become a bigger part of the total lifetime cost of PV plants.

This motivates the development of technologies to reduce the operation and maintenance costs for PV plants over their lifetime.

1.2 Problem Formulation

Nowadays, the most widespread PV technology is based on c-Si, representing around 90% of the annual solar cell production [4]. This is due to the stability of the technology, its efficiency and the fact that it has proven over the years to be a reliable choice. As a drawback, c-Si cells and modules have a relatively high manufacturing cost. The other major solar cell technology in the market, the Thin-Film (TF) solar cells are typically less expensive, however they have lower efficiency and stability as well. This technology is less mature than the c-Si, however it has significantly improved over the last years [4]

¹<http://um.dk/en/news/newsdisplaypage/?newsID=25147B44-3DCE-4647-8788-AD9243C22DF2>

and it has the long term potential to become the main solar cell technology. A short overview of the main PV cell types is given below:

- Crystalline silicon (c-Si)^{1,2}: There are two types of crystalline silicon solar cells used in PV: mono-crystalline silicon (x-Si) and multi-crystalline silicon(m-Si). x-Si is very expensive due to the purity of the silicon from which is made, and has the biggest efficiency, 27.6% while m-Si has an efficiency of 21.3%.
- Cadmium telluride (CdTe)^{1,2}: CdTe is a PV technology that is based on the use of cadmium and telluride. Due to the cadmium presence, it provides an environmental issue when recycled or disposed. CdTe has a maximum efficiency in the laboratory of 22.1%.
- Amorphous silicon (a-Si)^{1,2}: Is the non-crystalline form of silicon used for solar cells. a-Si cells featured low efficiency therefore was used for small scale applications, but due to the progress made over the last two decades in improving their performance, they became more attractive for larger scale applications. a-si has a maximum efficiency in the laboratory of 13.6%.
- Copper Indium Selenide (CIS)^{1,2}: Like a-Si and CdTE, CIGS layers are thin enough to be flexible, allowing them to be deposited on flexible substrates. Due to the use of high-temperature deposition techniques, the best performance is achieved when deposited on glass. CIS based solar cells are becoming one of the leading technologies for solar energy generators, having the highest efficiency among thin-film devices, around 22.3%.

A key element in having a good performance of the PV systems is implementing a performance monitoring system that will predict the power output of a PV system. The inputs for the model will vary with meteorological data and location, but the model will have to take into account the characteristics of the applied solar cell technology. However, some TF technologies are still in an early stage of deployment and ongoing development therefore, there is not as much data on their field performance and reliability yet.

Several models have been developed and implemented during the years for c-Si technologies, but just a few for TF. The most known and used models are PVWatts[5] and Sandia

¹http://www.nrel.gov/ncpv/images/efficiency_chart.jpg

²<http://energyinformative.org/best-solar-panel-monocrystalline-polycrystalline-thin-film/>

Array Performance Model (SAPM) [6] that provided good results for c-Si technologies but also for TF. These models are used to evaluate system performance over time and will determine if performance suddenly decreases and troubleshooting is necessary.

The focus of this master thesis is performance modelling and monitoring of PV plants in order to detect power loss and represents a part of the Fully Automated Service Execution (FASE) project. FASE aims to significantly improve the profitability of PV power plants, through a platform, implemented as a cloud based IT solution. This project is done by the PV Systems department of Aalborg University in collaboration with Danica Pension, Econ GmbH, GreenGo Energy A/S and DTU Compute Technical University of Denmark. Taking as an example GreenGo, they sell products with guaranteed performance, monitoring and service, therefore, an accurate, real time detection and localization of the power loss in the plant is crucial. The PV plants are based on CIS modules for which there is not a lot of experience in performance modelling and monitoring. Also, the performance of CIS modules is not well evaluated in Danish climate. Therefore, a study of the metastable behaviour of thin-film modules, and modelling of performance with focus on CIS in comparison with c-Si technology is carried out throughout the next chapters of this project.

There are several practical issues that have to be considered when developing an accurate PV performance model for practical thin-film PV systems. Some of these issues are specific to thin-film technology, whereas others need to be addressed in order to develop an accurate and efficient PV monitoring system. The most important aspects that need to be addressed are summarized in Figure 1.3 and described below.

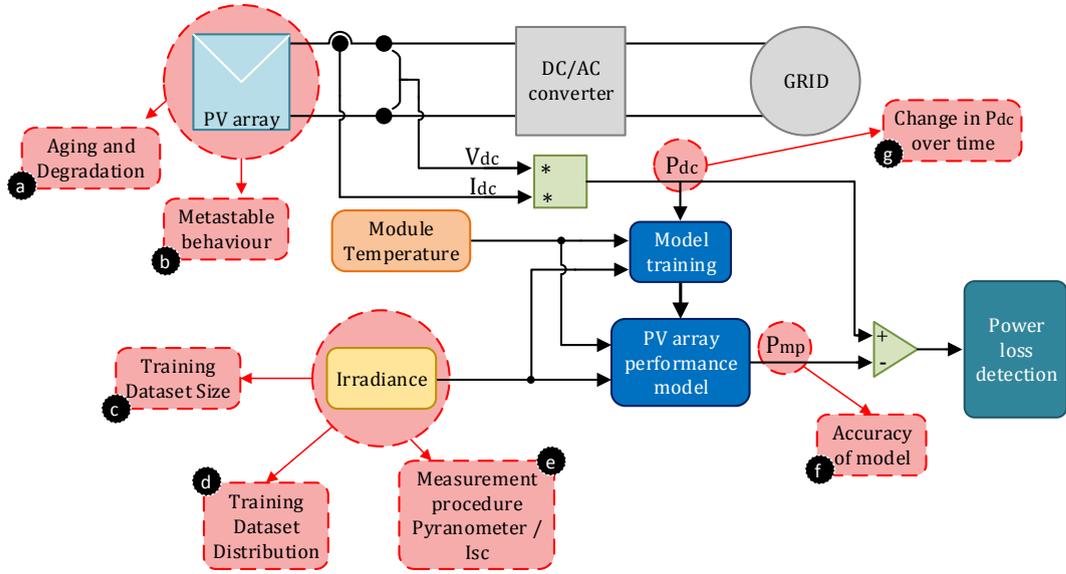


FIGURE 1.3: Diagram of a simplified PV system and its performance model structure with practical challenges highlighted with red dotted line and identified with letters

The main practical challenges highlighted in Figure 1.3 above, that the performance model needs to take into account, are described in the following:

- (a) Ageing and Degradation: The natural degradation expected for thin-film is 1.5% /year if the panels are installed prior to the year 2000, and 1% /year if they are installed after the year 2000 [7]. However, degradation is influenced by climate, such as very low temperatures, heavy wind, snow loads or exposure to high levels of UV. If the performance model does not take this into account, it will increasingly overestimate the available power as the system ages.
- (b) Metastable behaviour: The widely varying spectral responses, temperature coefficients and metastable behaviours of different thin-film technologies lead to special challenges in outdoor performance analysis and modelling. Modules must be stabilised before performance measurement to avoid the influence of metastability on the results. This is normally done by light exposure or “light soaking” [8]. Furthermore, the model must account for this behaviour to avoid seasonal errors in the performance estimation.
- (c) Training Dataset size: The data used for the model training is the measured irradiance, the measured module temperature and the output DC power. In order to

generate accurate models the training needs to be provided with sufficient measurement data. In practice, it is not convenient to wait a long time on the data collection. Therefore, we need to know the minimum amount of time needed in order to achieve an accurate performance model of the PV system. On the other hand, providing too much time/data will cause overfitting and may not lead to significant model accuracy improvements.

- (d) Training Dataset irradiance distribution: For similar reasons as the size of the dataset, the irradiance distribution of the location where the PV panels will operate is also important such that the resulting PV performance models are not biased for certain irradiance values.
- (e) Measurement procedure: The solar irradiance in PV plants is usually measured with pyranometers and/or reference cells. The pyranometer is able to accurately measure the all available solar radiation (global irradiation) under all conditions. Reference cells measure only that part of solar radiation that can be used by cells (effective irradiance) of identical material, and are available only for c-Si technologies [9]. Due to long wavelength response (above 1200 nm) the pyranometer behaves different compared to PV technologies. Spectral effects are causing the pyranometers to deviate from the irradiance perceived by the PV module, reaching a monthly deviation of over 3%. C-Si reference devices also show significant mismatch when used on thin-film module [9].

1.3 Objectives

The main goal of this project is to build an accurate PV performance model for thin-film PV technologies, that is able to address the challenges outlined in section 1.2. The models should be suitable for implementation in performance monitoring platforms of thin film PV Plants. In order to do this, several key tasks were set for this project:

- Analyse the metastable behaviour of thin-film technologies.
- Find a solution to account for ageing and degradation of PV panels in the performance model.
- Find a suitable PV performance model for thin-film applications.

- Determine the optimum size for the training dataset.
- Determine the optimum distribution of irradiance for the training dataset.
- Quantify the improvement in accuracy when using effective irradiance (from I_{sc}) instead of global irradiance (from pyranometer).

1.4 Project limitations

The project was subjected to certain limitations that are listed here:

- While it is shown that all TF technologies exhibit metastable behaviour to various extent, the performance models in this project focus only on the CIGS technology (and its comparison to x-Si).
- While there are many models available in literature, only two mainstream models have been considered for analysis and comparison.
- The assessment of performance model is based on outdoor test facility data and not PV plant data. Consequently, the effects of inverters, MPPT etc are not considered.
- The measurement data was from Florida and Colorado and was only for one year period.

Chapter 2

Performance characterization of Photovoltaic Technologies

2.1 Analysis of metastability for Thin-Film technologies

One important challenge of modelling the performance of TF is their metastable behaviour, which is not an issue for c-Si. Their metastability can cause changes in performance which affect the accuracy of the model, therefore a prior analysis is requested in order to predict and adjust to the outcome. The performance of PV modules are influenced by several factors like: [8]

- spectrum of incoming light
- intensity of incoming light
- temperature of the module

The output power is rated at standard testing conditions (STC) that indicate an irradiance of 1000 W/m^2 , a module temperature of 25°C and a solar spectrum corresponding to an air mass (AM)1.5. Real outdoor conditions rarely correspond to this STC values. Due to variations of the solar altitude from season to season or during a day, the AM changes. The performance of c-Si is not significantly affected by the spectral variations. It is not the same case for thin-film devices, especially a-Si where the effect of spectral variations on their performance is far more pronounced [8]. In Figure 2.1 the spectral

response for different thin-film solar cells in comparison with several types of *c-Si* solar cells are shown. The most stable and well known technology is *c-Si*, therefore, the behaviour of the thin-film technologies used in this project will be done in comparison with *c-Si*.

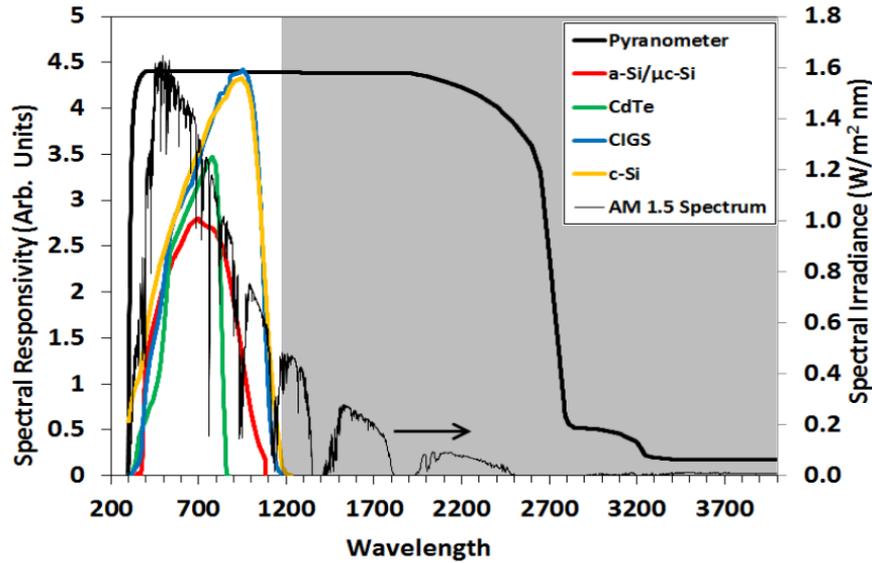


FIGURE 2.1: Spectral response of pyranometer and different solar cells devices (CIGS, CdTe, a-Si, c-Si) [10]

The spectral response of a-Si is more limited (350 - 800 nm) than for c-Si (300 - 1200 nm). As a consequence, this technology is more sensitive to blue components of the spectrum than to red ones. During the course of a day, the spectrum changes with a shift towards red light in the morning and evening. Also, during the summer, the distance passed by the light from the sun through the atmosphere is shorter than in the winter, due to the higher elevation of the sun. Therefore, the blue component becomes larger than *AM1.5* and will cause a blue-shift. Opposite, in winter, the distance will be larger and will cause a red-shift. This trend is shown in Figure 2.2 and leads to a higher module power in the summer and a lower one in the winter for a-Si and a reverse change in module power of c-Si [11].

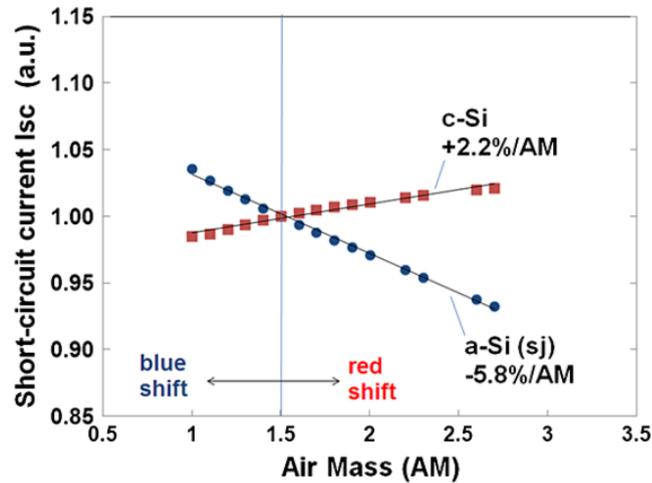


FIGURE 2.2: Modelling the short-circuit current dependence on AM-units for a-Si (blue circles) and c-Si (red squares) PV modules. Dark solid lines are linear fits to the modelled data [11]

CdTe spectral response is not as limited compared to a-Si but it is more narrow when compared to c-Si. It is expected that this devices will vary from +4% to -6% around the annual average when the seasons are changing [8]. For CIGS devices the spectral response is even broader compared to c-Si devices. The outcome of this broad response is less effect on device performance caused by spectral variations.

The variation of the intensity of incoming light from the STC towards lower values, will cause a reduction in the efficiency of the cell. How much the efficiency is reduced is dependent on the shunt resistance R_{shunt} of the cell which depends on the fabrication process. Usually a-Si technologies have a lower value of R_{shunt} than c-Si and has a good performance at low light intensities [12].

During the first months of operation the efficiency of an a-Si cell suffers a drop of about 10 - 30 %. This is known as the Staebler-Wronski effect (SWE) and represents a loss in electrical output caused by changes in photo-conductivity and dark conductivity induced by prolonged exposure to sunlight [8]. C-Si solar cells do not exhibit this effect, but this degradation is reversible upon annealing the a-Si cells at or above $150^{\circ}C$ [13].

Module temperature related losses in performance for a-Si devices are much smaller compared to c-Si cells . The temperature coefficients of the maximum power are around $-0.2\%/^{\circ}C$ for a-Si and $-0.36\%/^{\circ}C$ to $-0.42\%/^{\circ}C$ for CIGS. The values for CIGS approach

the values usually observed for c-Si, around $-0.45/\text{°C}\%$. Increased temperature leads to decreased performance, which means that the temperature effects are opposite compared to the spectral effect. The temperature coefficients for CdTe are larger than the a-Si coefficient but the effect is still smaller than c-Si with coefficient values of $-0.25/\text{°C}$.

2.1.1 Seasonal effects

The seasonal effects for a-Si devices are caused by temperature and spectrum variation. The performance increases due to thermal annealing and decreases due to negative temperature coefficient. As previously mentioned, the spectrum variation will lead to maximum performance during summer and minimum during winter. This trend is shown in Figure 2.3 for different types of a-Si junctions. [14]

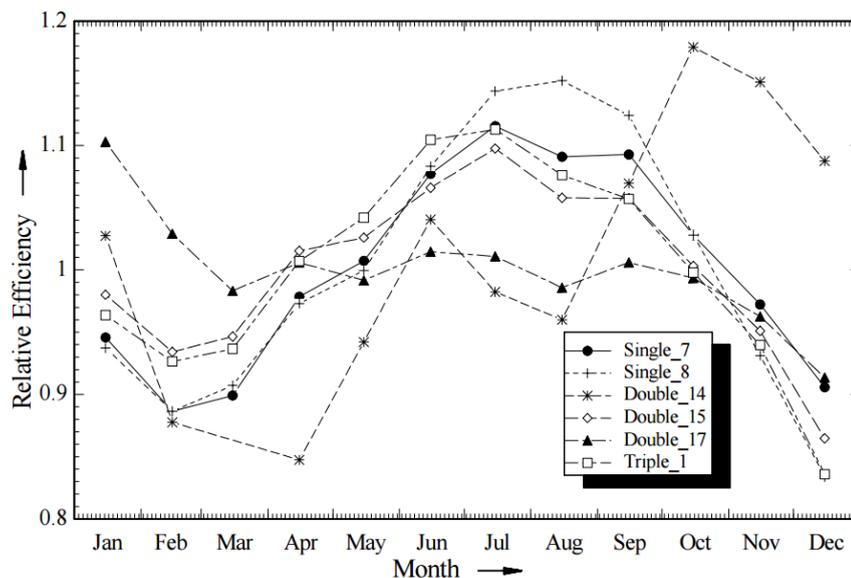


FIGURE 2.3: Seasonal Variation of the Operating Efficiency of Devices [14]

The seasonal performance changes are less pronounced for CIGS and CdTe compared to a-Si. CIGS devices are less affected by the spectral variations of the incoming light, therefore will exhibit a behaviour similar to c-Si during the change of seasons [8]. The high dependence of their performance on operating temperature, will cause a drop in efficiency during the summer. CdTe behaves similar in respect with seasonal variations, but the changes are less obvious due to the low temperature dependency [8].

In this chapter, four PV technologies were analysed and compared in order to establish the presence of metastability in the technologies. The technologies are: x-Si, a-Si,

CdTe, and two types of CIGS(CIGS Type1 and CIGS Type2). For CIGS Type 2, there are 5 years of available data that are used to establish the loss in performance due to ageing and degradation. The rest of the data is for 1 year period, from February 2011 to February 2012. The data was provided by the National Renewable Energy Laboratory (NREL) and for CIGS Type 2, from the NREL's PVDAQ public photovoltaic performance database¹. Figure 2.4 shows typical NREL measured data of irradiance, module temperature and module current and voltage for the modules situated at Cocoa beach in Florida. The values are down-sampled and normalized, and are displaying the quality and the available data. This data is used further to assess the metastability of the technologies just mentioned and to create and test the models implemented in chapter 3 for this location.

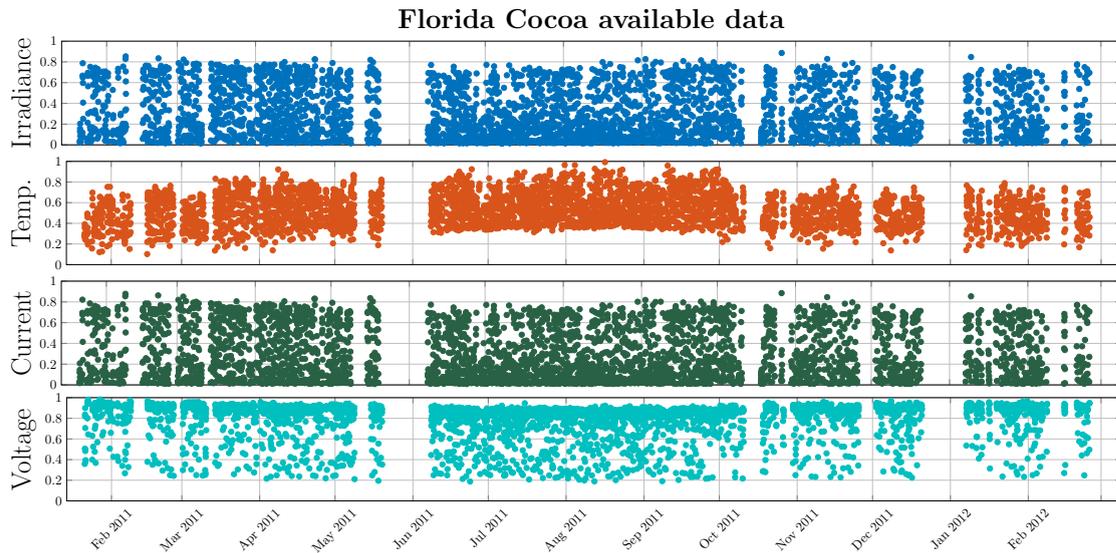


FIGURE 2.4: Normalized performance measurements from modules in Florida for a year. Blue - Irradiance, Orange - Module Temperature, Green - Current, Light Blue - Voltage

In order to investigate the seasonal effects, the performance ratio (PR) of the modules was calculated. The PR refers to the relationship between actual production and targeted production. For the project's purpose, the PR was calculated for three different ranges of irradiance and temperature. For high ($1000 \pm 50 \text{ W/m}^2$, $50 \pm 2 \text{ }^\circ\text{C}$), medium, case present in Denmark, ($700 \pm 50 \text{ W/m}^2$, $42 \pm 2 \text{ }^\circ\text{C}$) and low irradiances ($200 \pm 30 \text{ W/m}^2$, $30 \pm 2 \text{ }^\circ\text{C}$) conditions. The formulas used to calculate this PRs are presented in Equation 2.1, Equation 2.2 and Equation 2.3 where P_{mp} is the maximum DC power, G is the irradiance, $\delta_{P_{mp}}$ is the temperature coefficient used for correction, whose value can be found in Table 2.1, T_{mod} is the module temperature, and the reference values

¹<http://maps.nrel.gov/pvdaq>

for temperature and irradiance are: $T_{ref1000} = 50^\circ C$, $T_{ref700} = 42^\circ C$, $T_{ref200} = 30^\circ C$, $G_{ref1000} = 1000W/m^2$, $G_{ref700} = 700W/m^2$, $G_{ref200} = 200W/m^2$. The PR values are normalized, and displayed for a year along with their monthly average, except the PR values for CIGS Type 2, which are averaged for each 3 months.

TABLE 2.1: Temperature coefficients used for temperature correction of the power

Module	$\delta P_{mp}[\%/^\circ C]$
x-Si	-0.31
CIGS	-0.51
CdTe	-0.21
a-Si	-0.22

$$PR_{1000} = \frac{P_{mp}}{1 + \delta P_{mp}(T_{mod} - T_{ref1000})} \frac{G_{ref1000}}{G} [W] \quad (2.1)$$

$$PR_{700} = \frac{P_{mp}}{1 + \delta P_{mp}(T_{mod} - T_{ref700})} \frac{G_{ref700}}{G} [W] \quad (2.2)$$

$$PR_{200} = \frac{P_{mp}}{1 + \delta P_{mp}(T_{mod} - T_{ref200})} \frac{G_{ref200}}{G} [W] \quad (2.3)$$

In Figure 2.5 PR_{1000} for c-Si can be seen. This device varies with 2% around the annual average of 98.30% when the seasons are changing for this conditions.

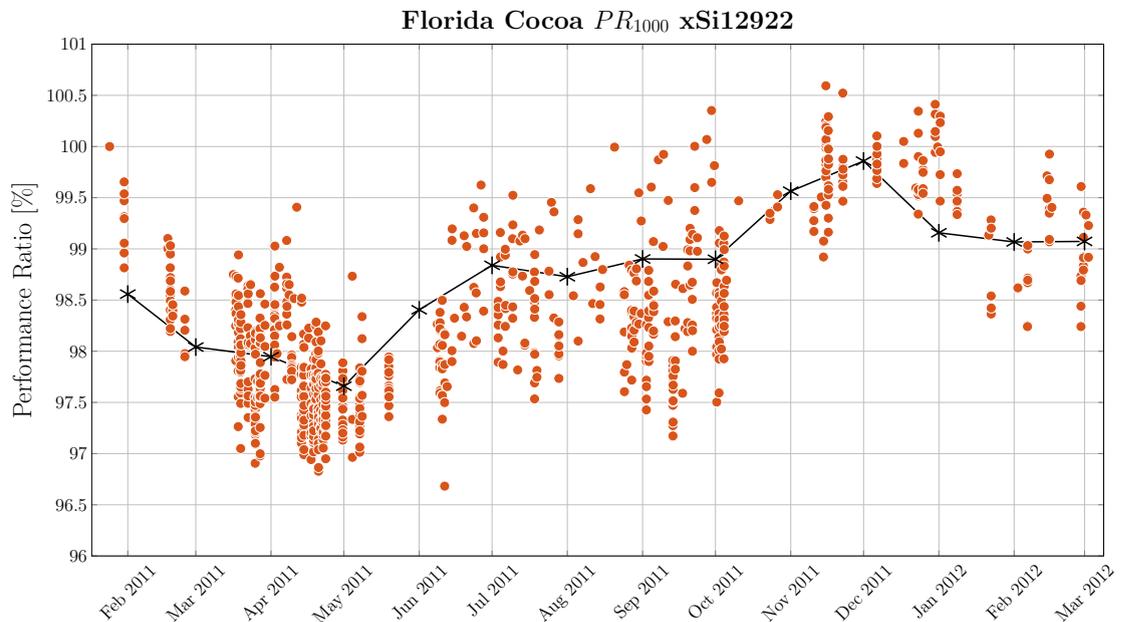


FIGURE 2.5: Performance Ratio of c-Si technologie at Cocoa beach for an year using NREL measurements showing seasonal effect for $G = [950 : 1050]W/m^2$ and $T_{mod} = [48 : 52]^\circ C$

It can be seen that this technology behaves as expected, having performance maxima during winter and minima in the summer due to the high operating temperature. In Figure 2.6 the performance ratio for lower irradiances is displayed. PR_{700} exhibits the same behaviour as PR_{1000} concluding that at medium irradiances, x-Si still has a good performance and does not vary that much with the change in seasons. The same cannot be said about PR_{200} , where it can be seen a difference in performance up to 40% when the seasons are changing.

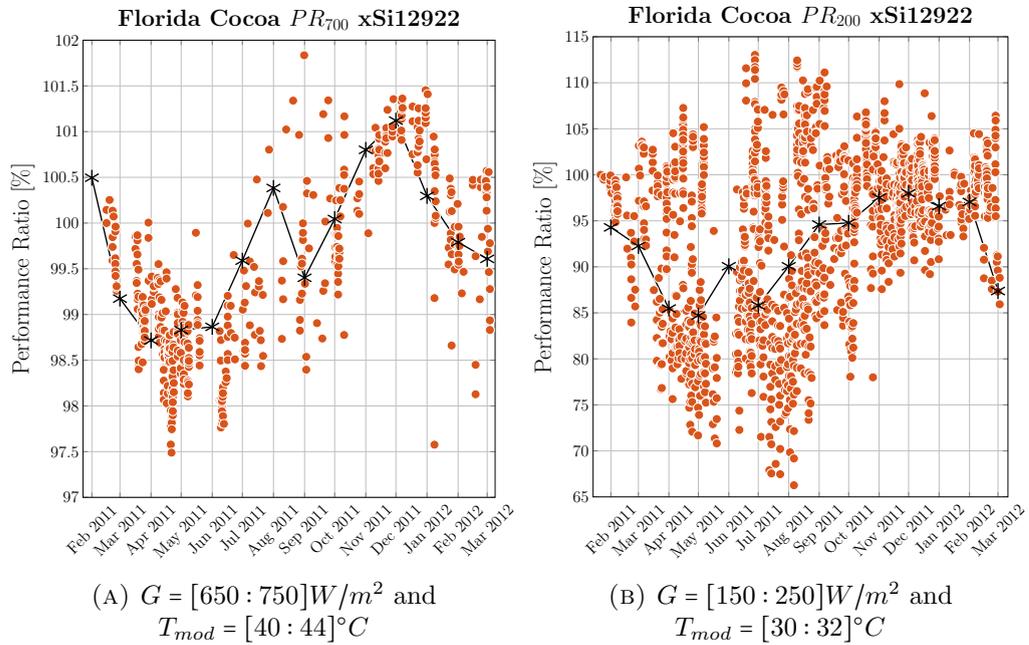


FIGURE 2.6: Performance Ratio of c-Si technologies at Cocoa beach for an year using NREL measurements showing seasonal effect for different irradiance and module temperature values

Figure 2.7 shown the PR_{1000} for a-Si, and can be seen exhibiting a variation of 6% around the annual average of 99.35% when the seasons are changing. This technology, as presented before, behaves opposite compared with c-Si, and has a maximum performance in the summer and a minimum in the winter. From Figure 2.8 it can be seen that PR_{700} has a similar trend as for high irradiances. At low irradiances, alongside the PR drop, the trend present in the other two cases is not visible here any more.

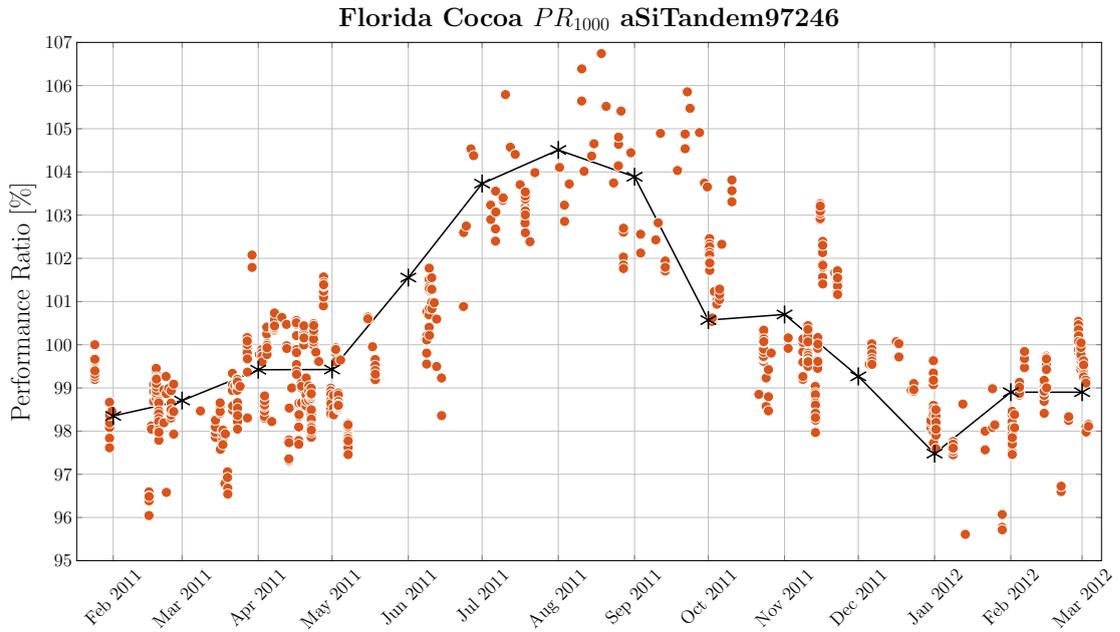


FIGURE 2.7: Performance Ratio of a-Si technology at Cocoa beach for an year using NREL measurements showing seasonal effect for $G = [950 : 1050] W/m^2$ and $T_{mod} = [48 : 52] °C$

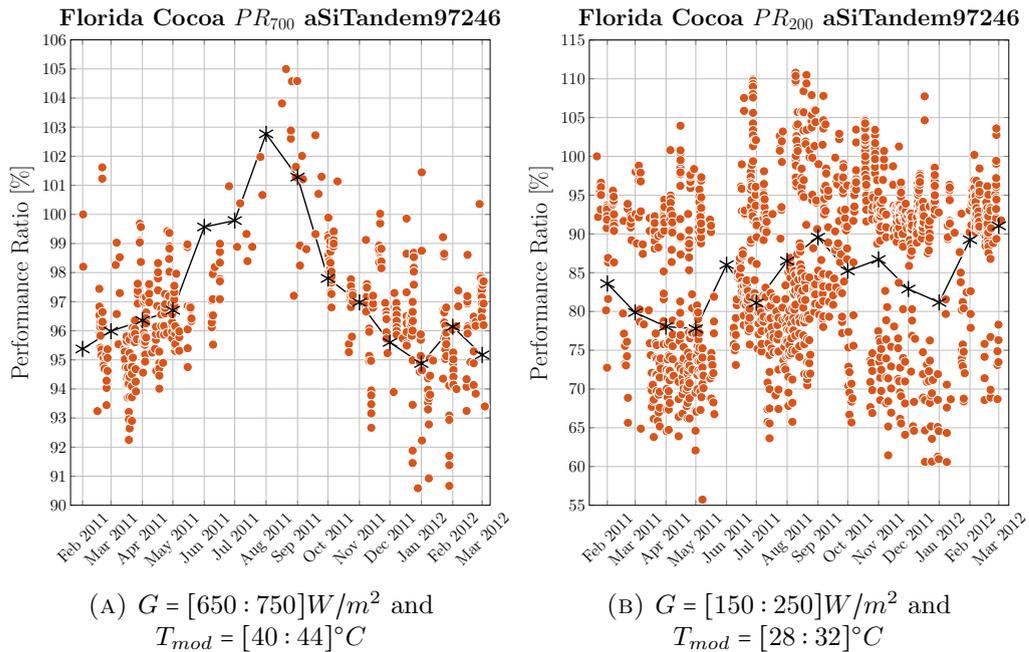


FIGURE 2.8: Performance Ratio of a-Si technology at Cocoa beach for an year using NREL measurements showing seasonal effect for different irradiance and module temperature values

The performance of CdTe increases and decreases with spectrum wavelength. Similar to a-Si, CdTe has a maximum performance in the summer and a minimum in the winter. When the seasons are changing, PR_{1000} varies with 4% around the annual average of

100.04%. The spectral effect of this technology can be seen in Figure 2.9. Similar to a-Si, at medium irradiance, CdTe still has a good performance but at low irradiances, the trend visible with the changes in seasons cannot be seen any more and the performance variation is very big. This can be seen in Figure 2.10.

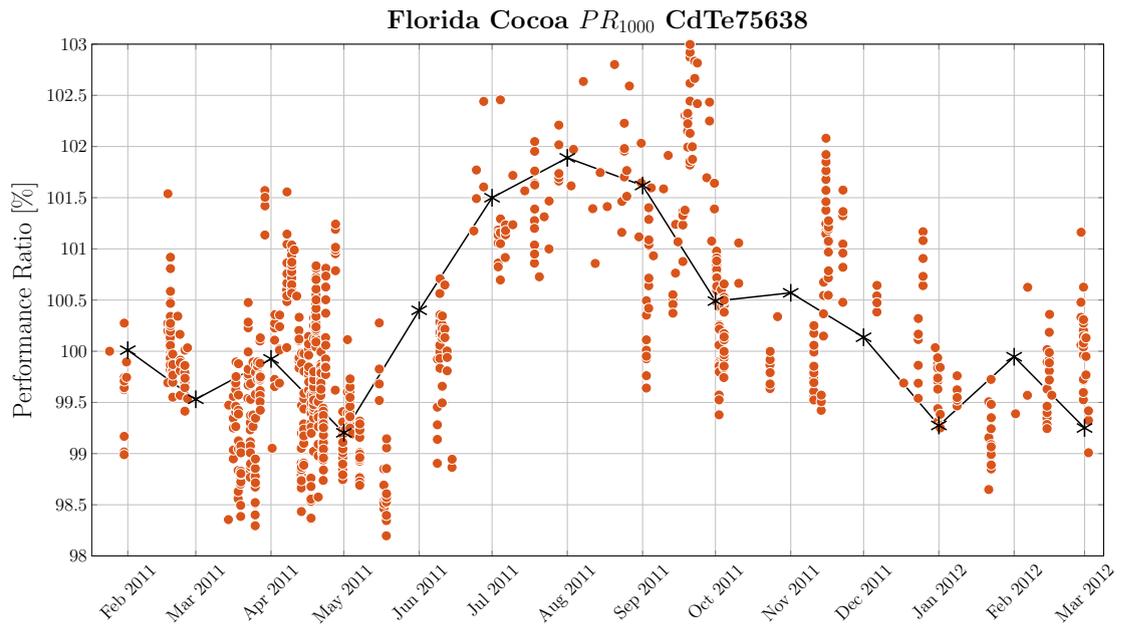


FIGURE 2.9: Performance Ratio of CdTe technology at Cocoa beach for an year using NREL measurements showing seasonal effect for $G = [950 : 1050] W/m^2$ and $T_{mod} = [48 : 52] ^\circ C$

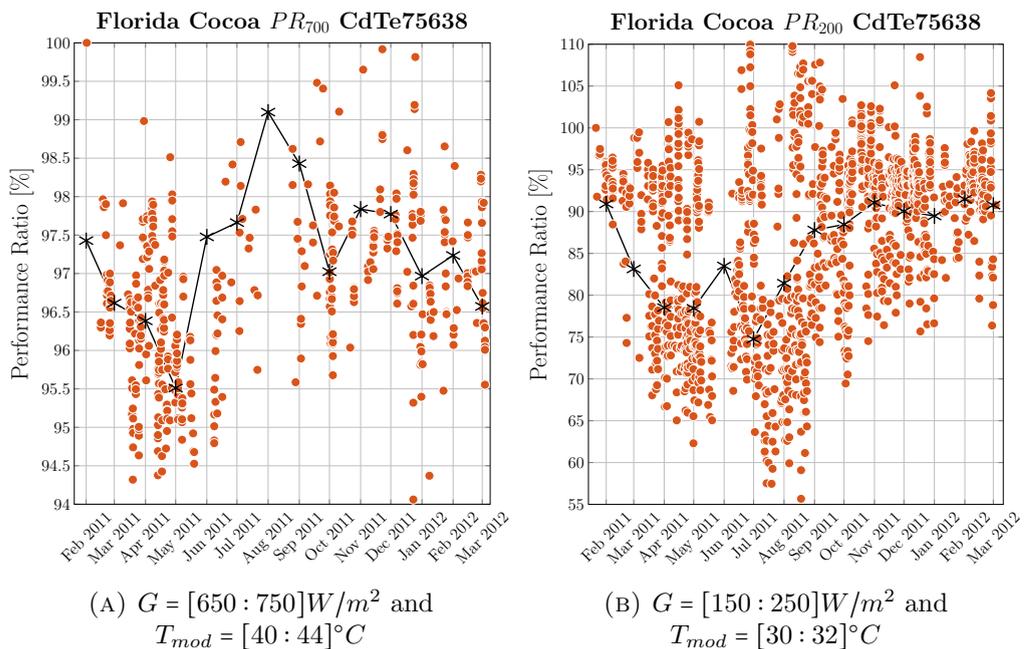


FIGURE 2.10: Performance Ratio of CdTe technology at Cocoa beach for an year using NREL measurements showing seasonal effect for different irradiance and module temperature values

PR_{1000} for CIGS can be seen in Figure 2.11. In this conditions, his device varies with 4.12% around the annual average of 96.99% when the seasons are changing. This technology behaves similar to c-Si. In February 2012, the PR has a 4% decrease in value from the value seen in February 2011. This can be associated with degradation of the module.

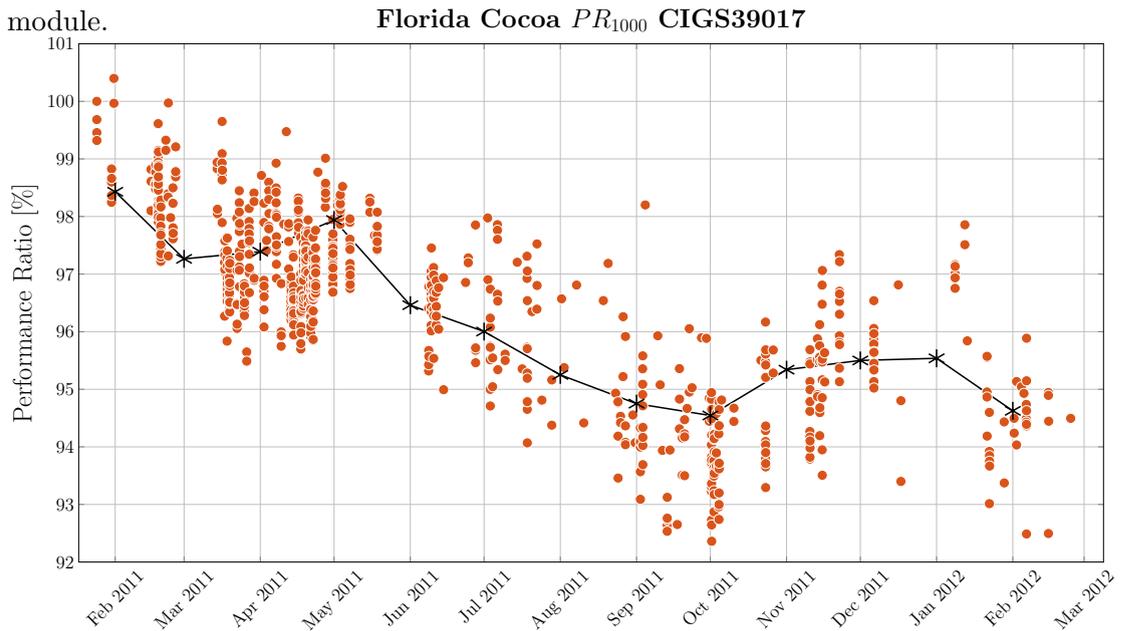


FIGURE 2.11: Performance Ratio for thin-film CIGS system over the period April 2011-March 2016 in Golden Colorado for $G = [950 : 1050]W/m^2$ and $T_{mod} = [48 : 52]^{\circ}C$

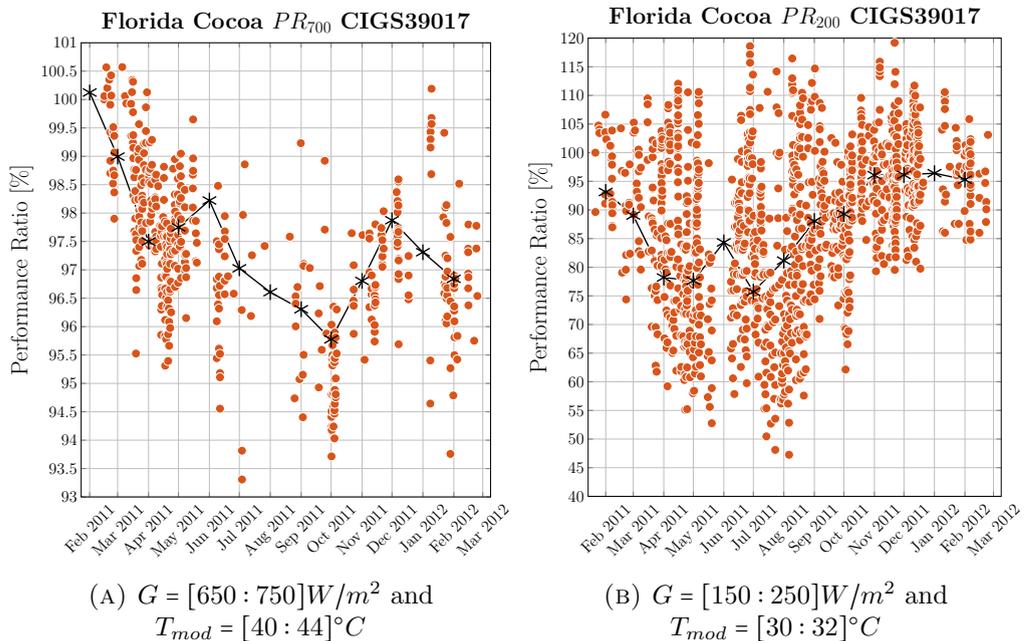


FIGURE 2.12: Performance Ratio for thin-film CIGS system over the period April 2011-March 2016 in Golden Colorado for different irradiance and module temperature values

For CIGS Type 2 a drop in PR_{1000} is noticeable, the module starts in 2011 with a average PR of 89% and until 2015 the PR drops approximately 11% reaching an average PR of 78%. In the literature, the degradation for this device, as exposed in the introduction is expected to be 1.5% /year if the panels are installed prior to the year 2000, and 1% /year if they are installed after the year 2000.

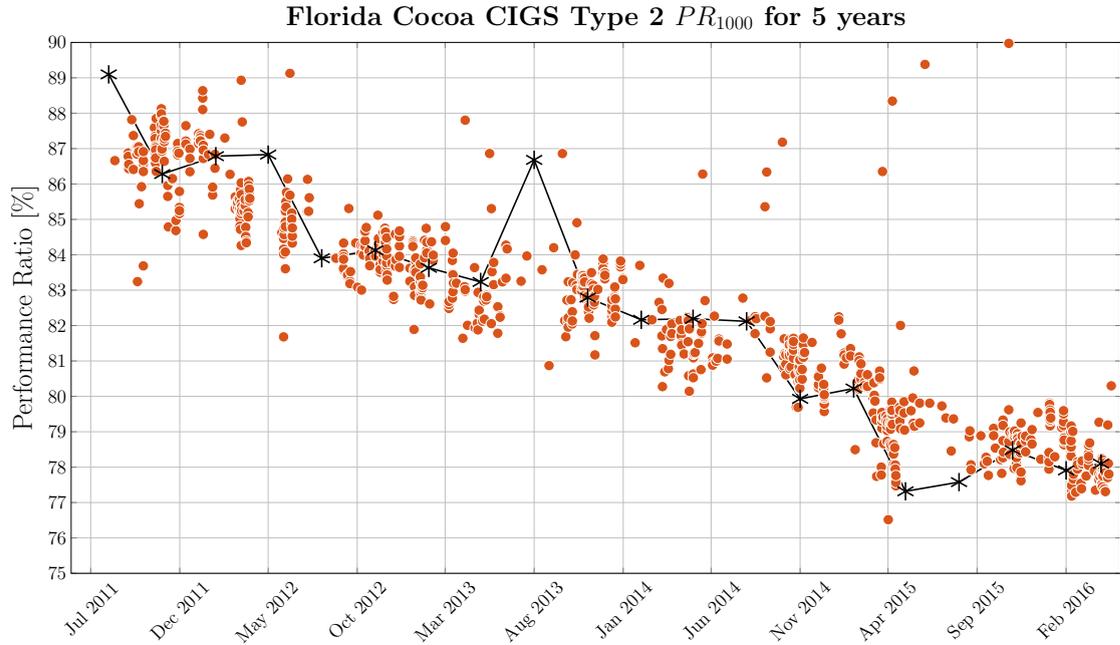


FIGURE 2.13: Performance ratio for thin-film CIGS system over the period April 2011 - March 2016 in Cocoa Florida for $G = [950 : 1050]W/m^2$ and $T_{mod} = [48 : 52]^{\circ}C$

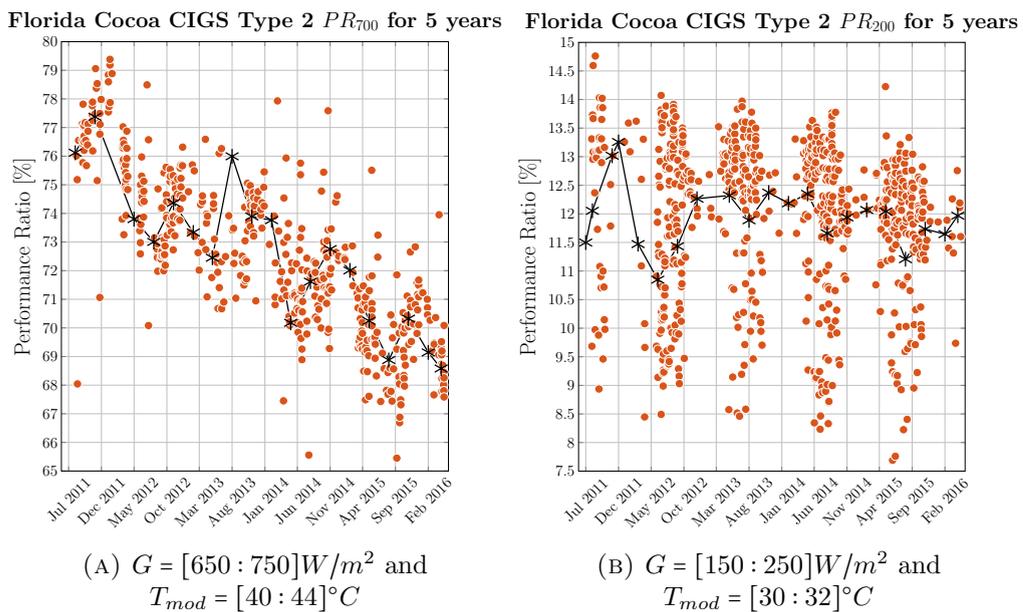


FIGURE 2.14: Performance ratio for thin-film CIGS system over the period April 2011 - March 2016 in Cocoa Florida for different irradiance and module temperature values

The average performance at medium irradiances PR_{700} decreased with 7%. The degradation is visible also at low irradiances but in 2015, the average PR_{200} is just with 2% lower than in 2011. The response for medium and low irradiance can be seen in Figure 2.14

2.1.2 Diurnal effects

The PR of a-Si technologies is decreasing with light exposure, meaning that we should see a bigger PR in the morning than for the rest of the day. The PR formula used for this study is given in Equation 2.4. Due to the effect of the module temperature, that can be seen in Figure 2.15, the PR was temperature corrected and its formula can be found in Equation 2.5, where the temperature reference, in this case is $T_{ref} = 25^{\circ}C$ and the irradiance reference is $G_{ref} = 1000W/m^2$.

$$PR = P_{mp} \frac{G}{G_{ref}} \quad (2.4)$$

$$PR_{corrected} = \frac{P_{mp}}{1 + \delta_{P_{mp}}(T_{mod} - T_{ref})} \frac{G_{ref}}{G} \quad (2.5)$$

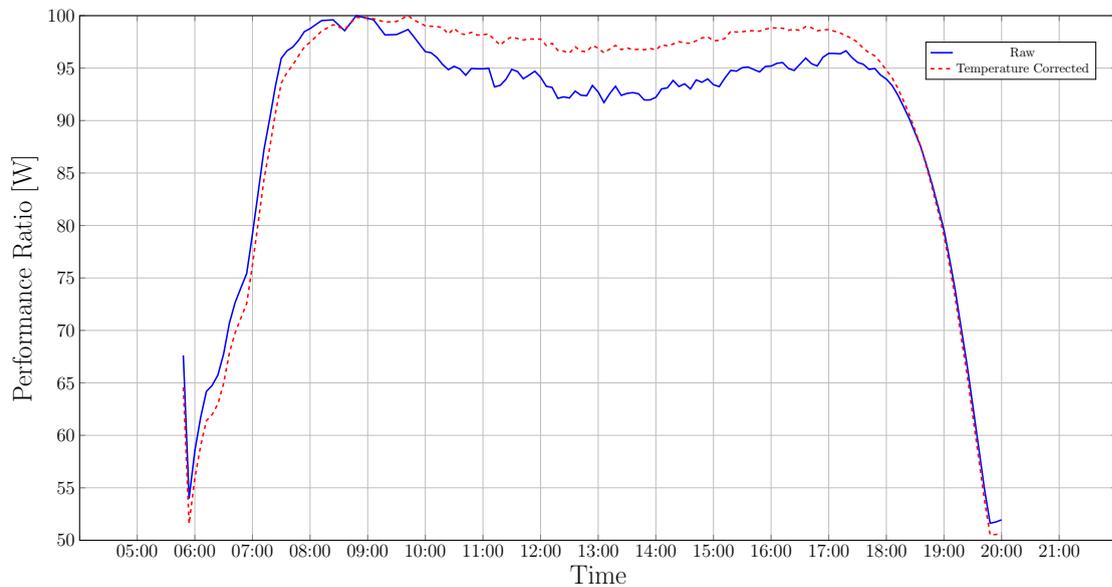


FIGURE 2.15: Module temperature effect on the PR of the module in spring period for a c-Si at Cocoa beach Florida

In general the PR of thin-film technologies, corrected for temperature, decreases at increasing wavelength of the solar spectrum. This effect is more evident in a-Si, where it has to be pointed out that the Staebler-Wronski effect plays a certain role.

In Figure 2.16, a clear sky day from spring showing the PR of the PV technologies studied in this chapters is displayed. To further check the presence of diurnal effect, the PR as a function of irradiance, with and without temperature correction, was plotted in Figure 2.17 for c-Si and a-Si technologies.

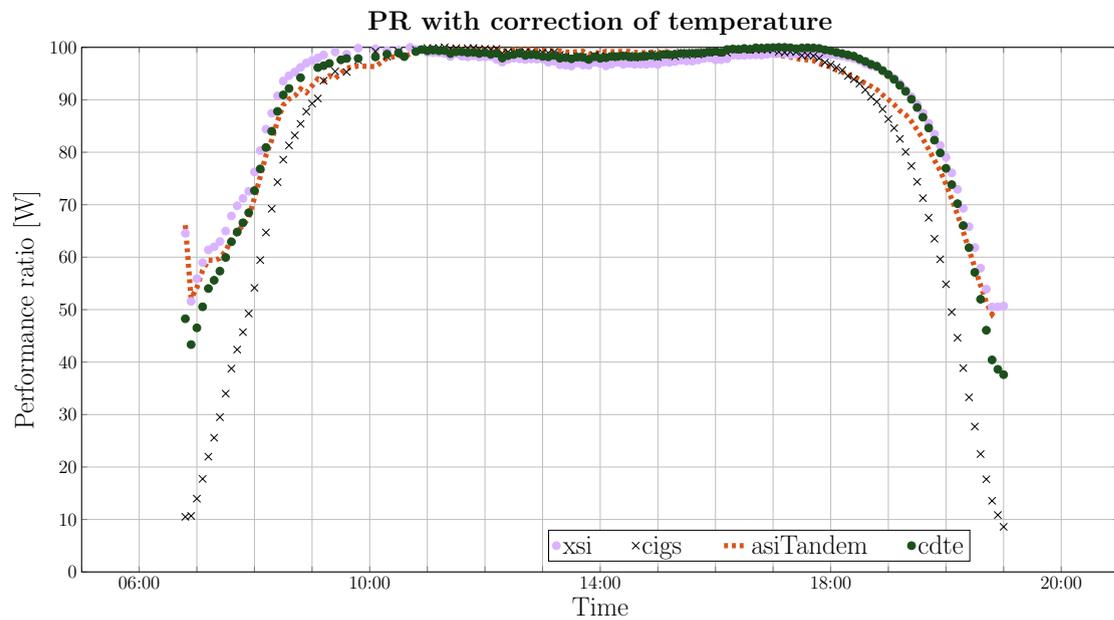


FIGURE 2.16: PR for a clear sky day in spring for Florida Cocoa beach comparing the four technologies

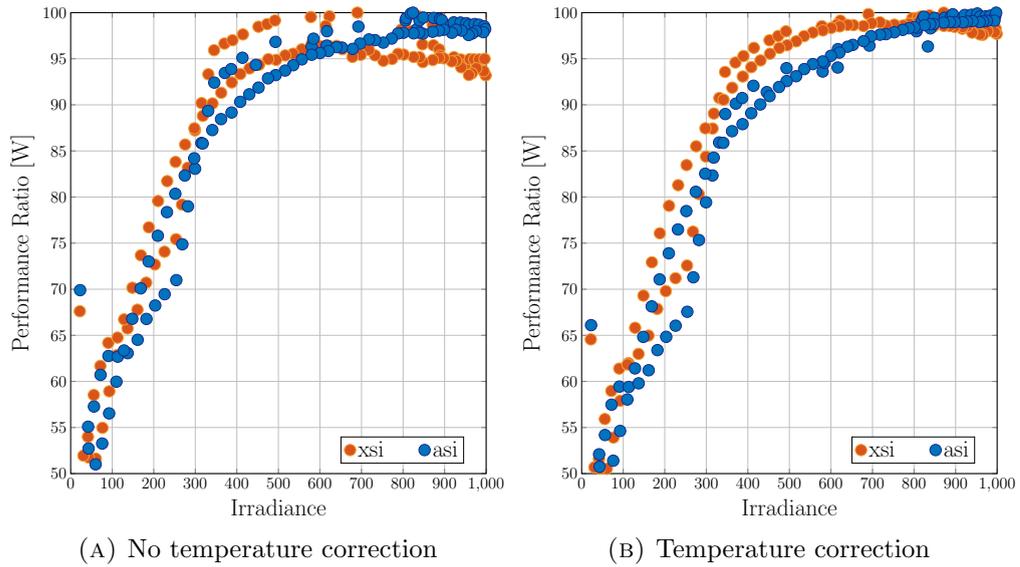


FIGURE 2.17: PR vs. Irradiance for a day in spring season showing the difference when there is no temperature correction and when the temperature is corrected at 25°C

A clear sky day from the summer is plotted in Figure 2.18 showing the PR of the technologies. In Figure 2.19, the PR as a function of irradiance for a-Si and c-Si is shown. No diurnal effect can be seen in this case.

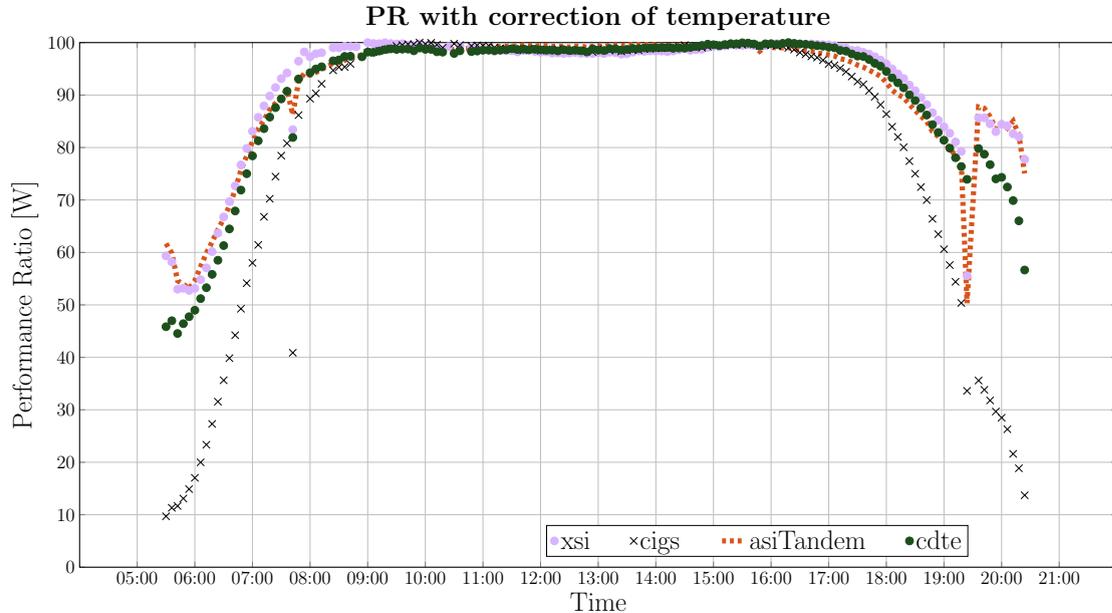


FIGURE 2.18: PR for a clear sky day in summer for Florida Cocoa beach comparing the four technologies

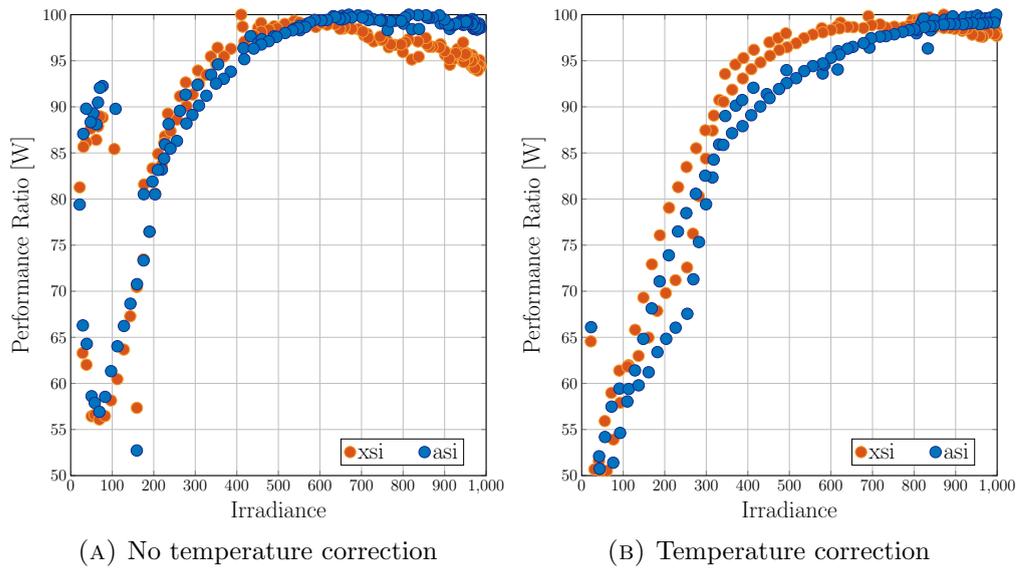


FIGURE 2.19: PR vs. Irradiance for a day in summer season showing the difference when there is no temperature correction and when the temperature is corrected at 25°C

In autumn, the temperatures are lower, therefore a drop in PR can be seen in Figure 2.20 for a-Si. This technology still does not present any signs of diurnal effect. The PR as a function of irradiance can be seen in Figure 2.21.

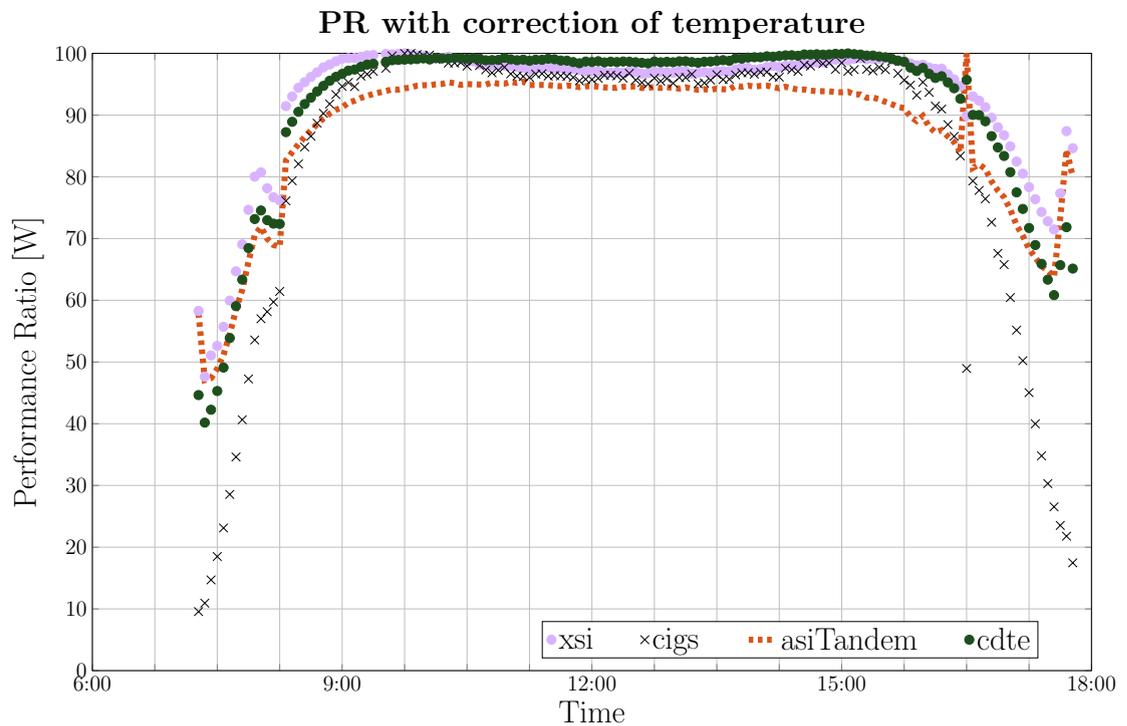


FIGURE 2.20: PR for a clear sky day in autumn for Florida Cocoa beach comparing the four technologies

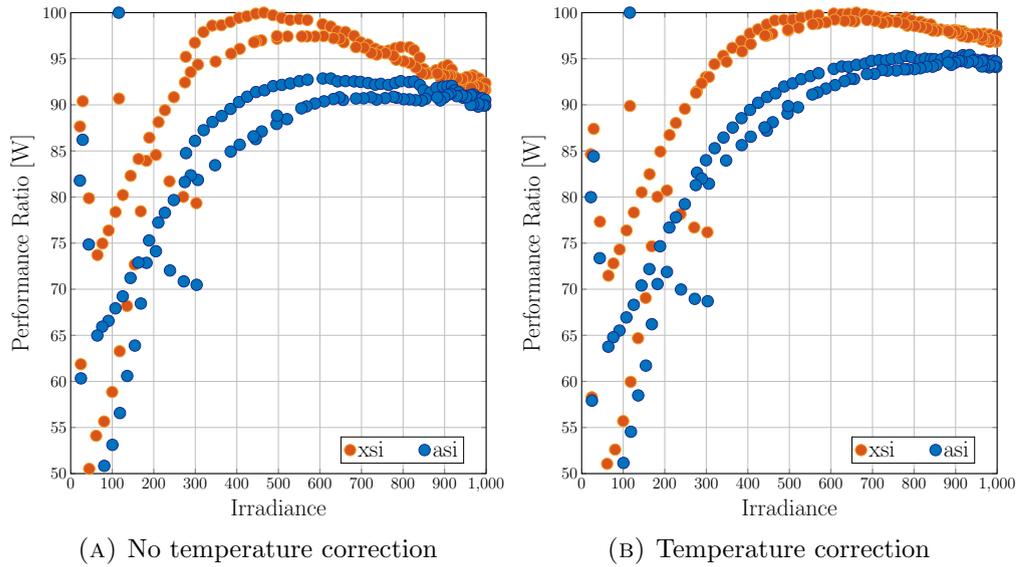


FIGURE 2.21: PR vs. Irradiance for a day in autumn season showing the difference when there is no temperature correction and when the temperature is corrected at 25°C

2.2 Conclusions

During this analysis, several conclusions were drawn assessing the metastability of this four PV modules. a-Si module displays the biggest seasonal effect but does not present any diurnal effect.

At medium irradiances, the modules still have a good PR, therefore, working with thin-film technologies in Denmark should not have any downside coming from the irradiance values.

Due to the PR variation caused by the change of season and the degradation of the module each year, it is recommended to retrain the model data twice per year. Once at the beginning of the year and once in the summer period in order to keep the accuracy of the model high as possible.

Chapter 3

Performance Modeling

3.1 Performance models

PV performance models are used to estimate how much power a PV system will produce at a given location, with characteristic weather conditions, as a function of irradiance and module temperature. Many computer models and algorithms for determining the performance of the PV systems were developed in the last couple of years. The report in [15] presents the most important PV performance models developed. The models were developed for different applications such as PV plant planning, yield assessment or performance monitoring, and thus take different modeling approaches, however they all predict the energy as a function of irradiance and cell temperature. While some models represent the full IV curve, others estimate only the maximum power point (MPP), the short circuit current and open circuit voltage [16]. In this project we focus on point-value models, that estimate the MPP from local irradiance and module temperature measurements.

From all the technologies available, only a CIGS and a xSi were chosen for this part of the study. The focus is on CIGS from the thin-film technologies due to the fact that, in the FASE project the PV plants are based on CIGS modules.

Four modules are selected, two CIGS type 1 and two xSi. A pair of CIGS-xSi were situated at Cocoa beach in Florida and the other pair was located in Golden, Colorado. It is important to see how different geographical locations having different meteorological characteristics affect the performance of the modules and the data set requirement for

the model identification. Data from each module was collected over the course of a year starting in Cocoa in January 2011 and in Golden in August 2012. In order to compare the performance of the technologies and to deal with the challenges that appear, IV data was used to estimate the parameters of the chosen model, using regression methods that are explained later in this chapter. In the problem formulation all the challenges were presented and in this chapter the following are analysed :

- Training dataset size
- Training dataset irradiance distribution
- Measurement procedure of irradiance

As explained in the problem formulation, in practice we need to know the minimum amount of time/data needed in order not to cause underfitting/overfitting of the performance model. This analysis is done using measured data from a CIGS type 1 module. Two different test datasets were selected having the same size from two different seasons, summer and winter. The size of the training dataset varies and data was used also from different seasons. Statistical errors are calculated for each training-test dataset combination. The results of this analysis are presented in section 4.3.

The resulting PV performance models should not be biased for certain irradiance values. Therefore the training dataset becomes dependent on a good distribution of irradiance data. Due to the fact that the accuracy of the model is influenced by the geographical location and weather conditions, for this analysis, data from a CIGS type 1 and a xSi module from Cocoa and another CIGS type 1 and xSi module from Golden, were used to create training and test datasets. The datasets have different irradiance distribution. Each training dataset is tested on different test datasets to establish what combination of training-test dataset irradiance distribution is the best. Statistical errors are calculated for every case. The results of this analysis are presented in section 4.4.

Another analysis done in this project is the measurement procedure and the difference that appears in the accuracy of the model when using effective irradiance instead of global irradiance as an input. The irradiance data for the modules used in this project are measured using a pyranometer, therefore we used the global irradiance for all the analysis done so far. Using effective irradiance E_e removes the losses caused by reflectance and the

solar spectrum variation effects[5]. The influence of irradiance and module temperature becomes more easily studied in the performance of the model. The formula for E_e is given in Equation 3.1 as a function of I_{sc} [6]. The same analysis as for the training dataset distribution was done, with the difference that now, as an input to the model, we use E_e instead of global irradiance. Statistical errors are calculated for every case. The results of this analysis are presented in section 4.5.

$$E_e = \frac{I_{sc}}{I_{sc0}[\alpha_{I_{sc}}(T_c - T_0)]} \quad (3.1)$$

Where $\alpha_{I_{sc}}[1/^\circ C]$ is the normalized temperature coefficient for I_{sc} , I_{sc0} is the short-circuit current at STC, T_c is the module temperature and T_0 is the reference module temperature, $T_0 = 25^\circ C$.

The analysis is conducted based on periods of clear sky, cloudy sky days and a combination between this two. A clear sky day, or a sunny day is considered, when the irradiance does not present significant variations and has a smooth evolution/profile throughout the day, like in Figure 3.1a. A cloudy day presents big variations between the irradiance points caused by the presence of clouds. An example of cloudy day can be seen in Figure 3.1b.

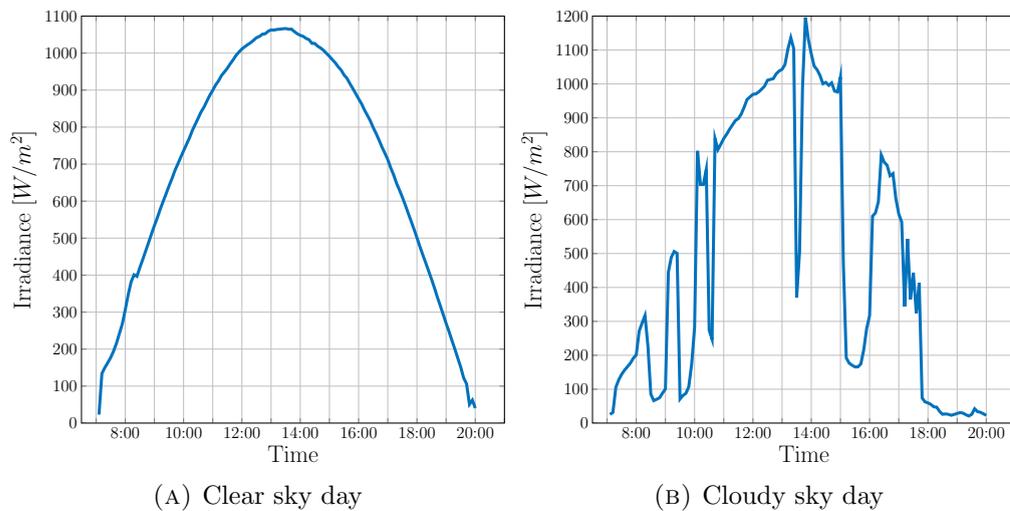


FIGURE 3.1: Example of sunny and cloudy day

Looking at the irradiance distribution, a sunny day will always have the more samples at high irradiance, whilst a cloudy day will have more at low irradiances. In Figure 3.2 the distribution of irradiance from the days shown above is displayed.

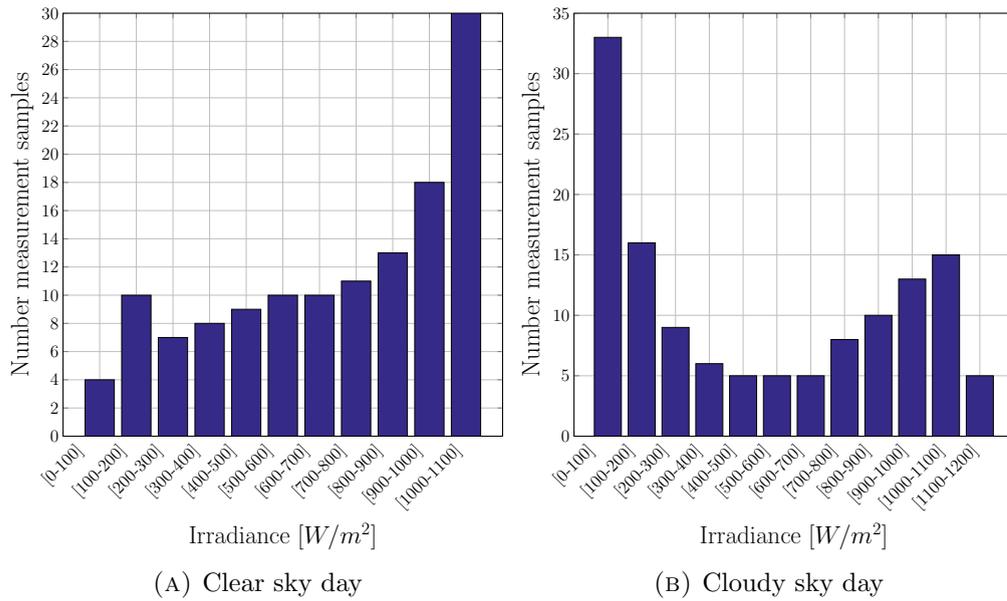


FIGURE 3.2: Distribution of measurement samples vs. irradiance for a sunny and a cloudy sky day

3.2 Irradiance profile for Denmark

In Figure 3.3 the monthly average irradiance distribution and average module temperature are shown, together with the number of sunny days available each month in three different geographical locations. It can be seen that the highest average irradiance are at Cocoa, followed by Golden. In these two locations, the seasonal change does not have a big impact on the values of the irradiance and temperature. When looking at Aalborg, besides the big differences between seasons, the average irradiance and temperature is smaller in comparison with the other two locations.

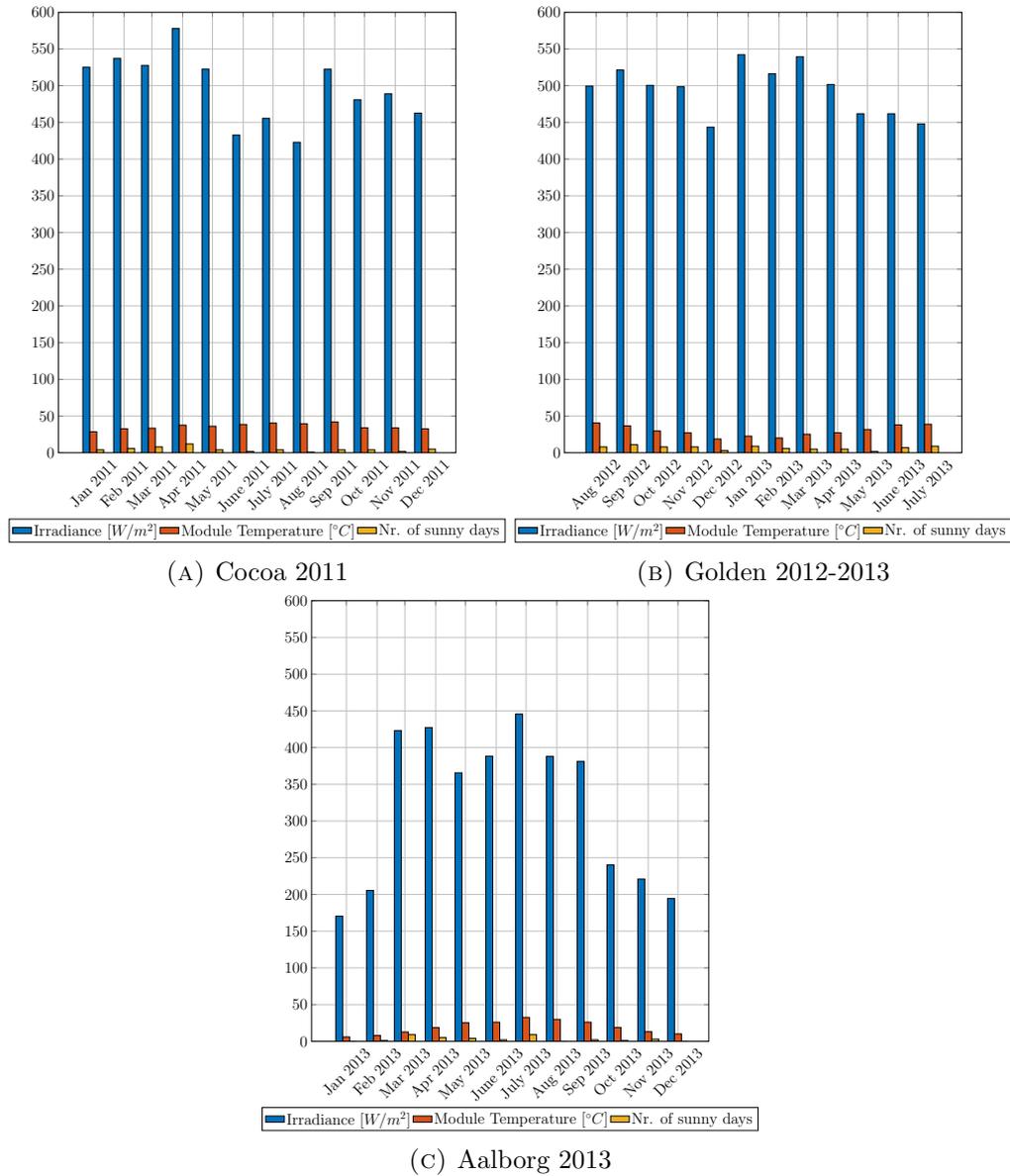


FIGURE 3.3: Monthly average of irradiance and module temperature together with the number of sunny days/month for three different geographic locations, Cocoa-Florida, Golden-Colorado and Aalborg-Denmark

Figure 3.4 shows only the number of sunny days/month provided at each location for the available period. In the year 2013, Aalborg had only 36 days of clear sky. Less than the other two locations. In can be said that in Aalborg, for year 2013, there was 10% of clear sky days. This will help create one of the test cases that will be used to analyze how the dataset irradiance distribution, the dataset size and the method of irradiance measurements affects the accuracy of the model implemented.

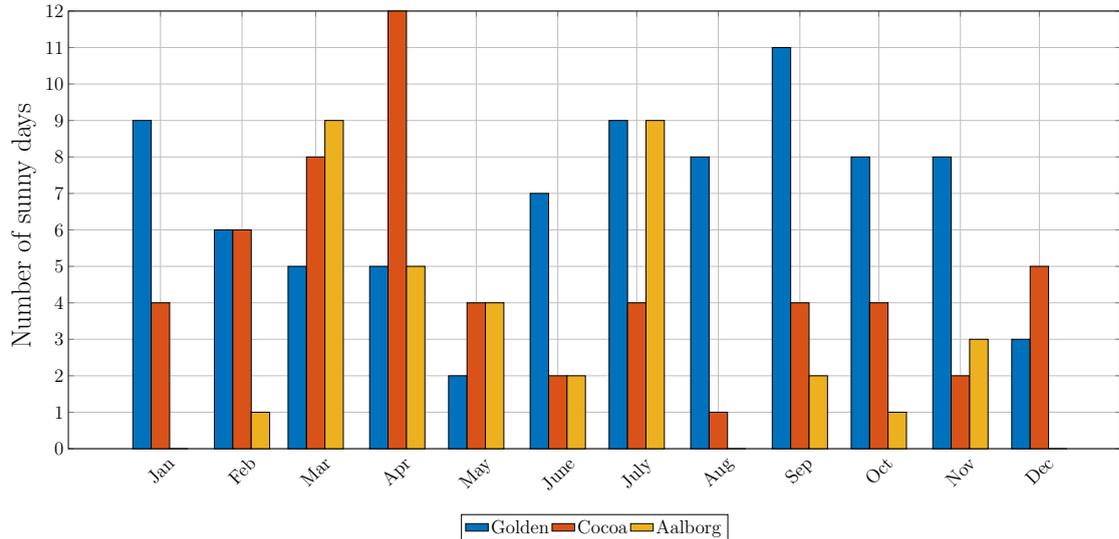


FIGURE 3.4: Number of sunny days/month for three different geographic locations, Cocoa-Florida, Golden-Colorado and Aalborg-Denmark

3.3 Method for model identification

In this project, linear regression is used in order to parametrize the PV performance model. The performance parameters of the PV panels change in time due to degradation and ageing, and will differ from the initial datasheet values, thus we need to identify them periodically. Each model has a different number of parameters that need to be identified. The formula for the generic multiple linear regression model is shown in Equation 3.2.[17]

$$\hat{P}_{mp} = \beta_0 + \sum_{i=1}^k \beta_i x_i + \epsilon = y \quad (3.2)$$

Where, y is the response variable (in this case P_{mp}) that is related to k regressors variables, β_i are the regression coefficients that are to be calculated and ϵ is an error variable that adds noise to the linear relationship between the regression coefficients and the predictor values [17].

Figure 3.5 shows the parametrization and validation of the PV model parameters. This process consists in two big phases, the training phase and the testing phase.

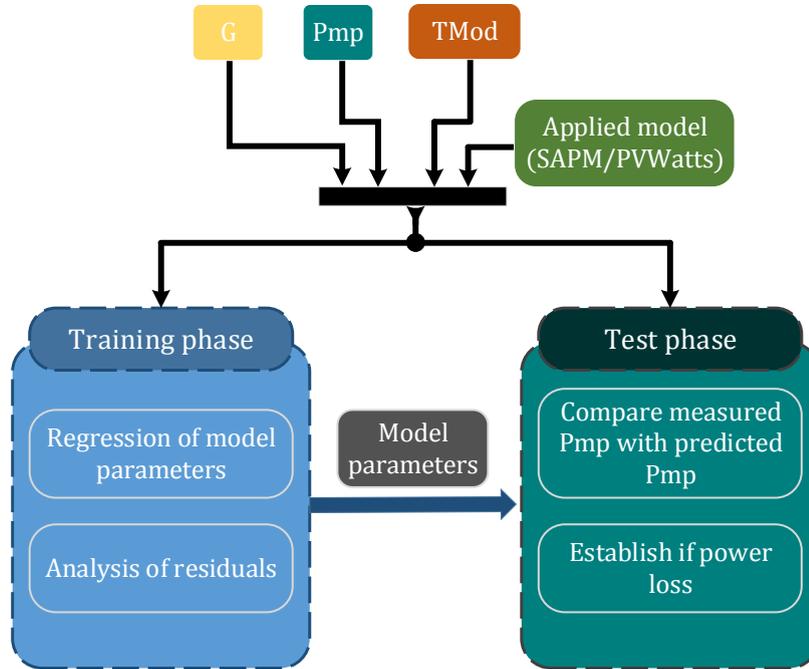


FIGURE 3.5: Diagram of training and test phase showing how the PV model parameters are obtained and validated

The training phase consists in gathering measurement data (power, irradiance and module temperature) from the PV system functioning normal and without faults. In order to improve the prediction result, low quality data is filtered out, by removing measurements at low irradiances (below 50 W/m^2). Once the training dataset has a sufficient size, a PV performance model is specified. Having all this information available, the training phase is parametrizing the model using regression modelling. The parameters are estimated using the method of least squares, which is a typical method used to estimate the regression coefficients in a multiple linear regression model. The method optimizes the regression coefficients (β) such as the sum of the squares of the errors is minimized. The normal equations are expressed in matrix notation for it to be easier to solve. Therefore, Equation 3.2 may be written in matrix notation as Equation 3.3 [17].

$$y = X\beta + \epsilon \quad (3.3)$$

$$\text{Where: } y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

After calculations and simplifications that can be found explained more detailed in [17], the least squares estimator of β is given in Equation 3.4, providing the final values for the model parameters.

$$\hat{\beta} = (X' \cdot X)^{-1} \cdot X' \cdot y \quad (3.4)$$

Besides the power, irradiance, module temperature, and the model type, the test phase receives as an input from the training phase the model parameters. This phase predicts the power of the PV module. In order to determine the accuracy of the models, a prediction error is calculated for each point using Equation 3.5.

$$Error_{prediction} = \frac{Y - Y_{predicted}}{Y} \cdot 100 \quad (3.5)$$

Where, Y is the measured value and $Y_{predicted}$ is the estimated value. Root-mean-square-error and mean-absolute-error are calculated to have an overview of the accuracy on the entire test period. Formulas and explanation of this statistical errors are given below.

3.4 PVWatts vs. SAPM

3.4.1 SAPM

SAPM is a photovoltaic array performance model developed at Sandia National Laboratories and is empirically based. Individual equations used in the model are derived from individual solar cell characteristic therefore the model holds a good accuracy. Extensive outdoor module testing and several comparison studies validated the approach of this model [6]. On the official Sandia website¹ a database of module performance parameters is maintained containing results from outdoor performance tests on the available commercial modules .

¹<http://www.sandia.gov/pv>

Testing and modeling of PV module performance is challenging due to the influence of the environment and the physics of the solar cell. An effective performance model should be able to separate and establish the effect of all influencing factors.

The equations defining the model are expressed below and are describing the electrical performance for individual photovoltaic modules, and can be adapted to any number of series or parallel modules in an array.

The equations used in this project are given in Equation 3.6 through Equation 3.8 and they are used to calculate the expected power, having local ambient information available, like irradiance and module temperature. The parameters highlighted in red are found using the regression modeling method explained previously in section 3.3.

$$I_{mp} = I_{mp0} [C_0 E + C_1 E^2] [1 + \alpha_{I_{mp}} (T_c - T_0)] \quad (3.6)$$

$$V_{mp} = V_{mp0} + C_2 N_s \delta(T_c) \ln(E) + C_3 N_s \delta(T_c) \ln(E)^2 + \beta_{V_{mp}}(E)(T_c - T_0) \quad (3.7)$$

$$P_{mp} = I_{mp} V_{mp} \quad (3.8)$$

Parameter definition:

- E = Solar irradiance. [W/m^2]
- I_{mp} = Current at the maximum power point [A]
- I_{sc} = Short-circuit current [A]
- I_{sc0} = Short-circuit current at STC [A]
- V_{mp} = Voltage at maximum power point [V]
- I_{mp0} = Maximum current at STC [A]
- V_{mp0} = Maximum voltage at STC [V]
- P_{mp} = Power at maximum power point [W]
- C_0, C_1 = Empirically determined coefficients relating I_{mp}
- C_2, C_3 = Empirically determined coefficients relating V_{mp} ($C_3[1/V]$)

- n = Diode factor
- T_c = Module temperature [$^{\circ}C$]
- T_0 = Reference cell temperature, typically $25^{\circ}C$
- E_0 = Reference solar irradiance, typically $1000W/m^2$
- $\alpha_{I_{mp}}$ = Normalized temperature coefficient for $I_{mp}[1/^{\circ}C]$
- $\beta_{V_{mp}}$ = Normalized temperature coefficient for $V_{mp}[V/^{\circ}C]$
- $\delta(T_c) = \frac{n \cdot k(T_c + 273.15)}{q}$ - ‘Thermal voltage’ per cell at temperature T_c .
- $k = 1.3806503 \cdot 10^{-23}$ - Boltzmann constant [J/K]
- $q = 1.6022 \cdot 10^{-19}$ - Electron charge [coulomb]
- N_s = Number of cells in series in a module’s cell-string

A generic model is derived, \hat{P}_{mpSAPM} where the non-linearities of this model are separated in linear terms standing for predictor variables. The predictor variables are defined below and the the linear regression model defined in Equation 3.2 is applied [17].

$$\begin{aligned} x_1 &= \frac{N_s \cdot k \cdot \delta(T_c) \cdot \ln(E)}{q}; & x_2 &= \frac{N_s \cdot k \cdot \delta(T_c) \cdot \ln(E)^2}{q}; & x_3 &= T_c - T_0; & x_4 &= E; \\ x_5 &= E^2; & x_6 &= E(T_c - T_0); & x_7 &= E^2(T_c - T_0); \end{aligned}$$

3.4.2 PVWatts

PVWatts is a simpler PV performance model, alternative for SAPM that can be found in [5]. A comparison between measured and modeled maximum power P_{mp} for PV modules that rely on different technologies was conducted. After the evaluation of the three models presented in this paper, an improved power temperature coefficient model with an irradiance non-linearity correction was developed, known as the PVWatts model. This model is necessary when at lower irradiances, the PV module efficiency starts to show a strong nonlinearity with the irradiance, and thus a linear approximation is no longer possible this causing the model to overestimate performance. PVWatts is taking into account two irradiance regions and inserts an irradiance correction factor. The irradiance regions are $\leq 200W/m^2$ and $> 200W/m^2$. The equation for the irradiance

correction factor can be found in Equation 3.9, while Equation 3.10 and Equation 3.11 are providing the formulas representing the model. Again, the parameters highlighted in red are found using regression modelling methods.

$$k = \frac{P_m(E_L, T) - P_{meas}(E_L, T)}{P_{m_0}} \quad (3.9)$$

$$P_m = P_{m_0} \left[\frac{E}{E_0} [1 + \gamma(T - T_0)] - k \frac{E_0 - E}{E_0 - 200} \right], \quad \text{for } E > 200W/m^2 \quad (3.10)$$

$$P_m = P_{m_0} \left[\frac{E}{E_0} [1 + \gamma(T - T_0)] - k \left[1 - \left(1 - \frac{E}{200} \right)^4 \right] \right], \quad \text{for } E \leq 200W/m^2 \quad (3.11)$$

Parameter definition:

- E = Solar irradiance. $[W/m^2]$
- k = Irradiance correction factor
- E_L = Effective low irradiance, $200W/m^2$
- E_0 = Reference solar irradiance, typically $1000W/m^2$
- T = Module temperature $[^\circ C]$
- T_0 = Reference cell temperature, typically $25^\circ C$
- $P_m(E_L, T)$ = The value of P_m for E_L and T without the irradiance correction factor, k
- $P_{meas}(E_L, T)$ = measured P_m for E_L and T
- γ = maximum power correction factor for temperature $[1/^\circ C]$

The generic model derived from PVWatts is $\hat{P}_{mpPVWatts}$, where the non-linearities of this model are separated in linear terms standing for predictor variables. The predictor variables are defined below for the two different ranges of irradiance.

$$\text{For } E \leq 200W/m^2: \quad x_1 = \frac{E}{1000}; \quad x_2 = \frac{E}{1000} \cdot (T_c - T_0); \quad x_3 = \left(1 - \frac{E}{200} \right)^4;$$

$$\text{For } E > 200W/m^2: \quad x_1 = \frac{E}{1000}; \quad x_2 = \frac{E}{1000} \cdot (T_c - T_0);$$

The model's estimated power and measured power were compared using root-mean-square-error (RMSE) and mean-absolute-error (MAE) statistics. RMSE shows the difference between the estimated values and the measured values, while MAE provides the average absolute deviation of the modeled values from the measured values [5]. Both have positive values and their formulas are shown in Equation 3.12 and Equation 3.13. The errors are calculated as a percentage of the measured value.

$$RMSE = 100\% \left[\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2 \right]^{\frac{1}{2}} \div \left[\frac{1}{n} \sum_{i=1}^n x_i \right] \quad (3.12)$$

$$MAE = 100\% \left[\frac{1}{n} \sum_{i=1}^n |y_i - x_i| \right] \div \left[\frac{1}{n} \sum_{i=1}^n x_i \right] \quad (3.13)$$

Where y_i is the i^{th} estimated value, x_i is the i^{th} measured value and n is the number of measured and estimated values.

Chapter 4

Results

4.1 Definition of Test Cases

In order to analyse the influence of the dataset distribution on the PV model accuracy, different sets of conditions are created, under which the accuracy of the model is tested based on the irradiance distribution of the training and test dataset. Many cases were found to be relevant, but it was not possible to implement and test them all. We found it important to see how a dataset distribution having only sunny or cloudy days would affect the accuracy of the performance models. Therefore, four sunny and four cloudy days were used to create the first two datasets. A distribution created by equally combining sunny and cloudy days was also considered. Two sunny and two cloudy days formed the third dataset. Due to the irradiance profile in Denmark, a distribution where 10% of the irradiance data came from sunny days was studied. One sunny day and nine cloudy days formed the last dataset considered in this analysis. These four test cases(TC) are proposed and tested as shown in Table 4.1.

TABLE 4.1: Defined TC by training dataset and test dataset

Training dataset	Test dataset	Code
4 sunny days	4 sunny days	4s-4s
	4 cloudy days	4s-4c
	2 sunny and 2 cloudy days	4s-2s2c
	1 sunny and 9 cloudy days	4s-1s9c
4 cloudy days	4 sunny days	4c-4s
	4 cloudy days	4c-4c
	2 sunny and 2 cloudy days	4c-2s2c
	1 sunny and 9 cloudy days	4c-1s9c
2 sunny and 2 cloudy days	4 sunny days	2s2c-4s
	4 cloudy days	2s2c-4c
	2 sunny and 2 cloudy days	2s2c-2s2c
	1 sunny and 9 cloudy days	2s2c-1s9c
1 sunny and 9 cloudy days	4 sunny days	1s9c-4s
	4 cloudy days	1s9c-4c
	2 sunny and 2 cloudy days	1s9c-2s2c
	1 sunny and 9 cloudy days	1s9c-1s9c

Each train-test case was given a code to simplify further explanations and display of results. The numbers in the code represent how many days were used to form the dataset and the letter represents the type of days(sunny/cloudy) used in the dataset. The purpose of testing these TC is to observe how the accuracy of the model changes depending on the irradiance distribution.

All the errors statistics are expressed in tables for a specific applied model and different train-test datasets. The errors are calculated for three irradiance regions to establish which region is more problematic for the accuracy of the model:

- L - low irradiances, $< 400W/m^2$
- M - medium irradiances, $400W/m^2 - 800W/m^2$
- H - high irradiances, $> 800W/m^2$

4.2 Comparison of SAPM vs. PVWatts

Table 4.2 gives modeling error statistics for P_{mp} using SAPM and PVWatts models on a CIGS module considering four different cases.

TABLE 4.2: RMSE and MAE statistics for P_{mp} by model and case

Code	PVWatts						SAPM					
	RMSE [%]			MAE[%]			RMSE[%]			MAE[%]		
	L	M	H	L	M	H	L	M	H	L	M	H
4s-4s	7.34	0.94	0.35	3.80	0.80	0.29	7.77	0.98	0.35	5.44	0.84	0.29
4s-4c	17.20	2.14	0.93	13.04	1.84	0.74	18.58	2.23	0.97	15.24	1.87	0.79
4c-4s	11.84	1.48	0.94	8.53	1.22	0.89	12.59	1.44	0.94	9.59	1.12	0.86
4c-4c	12.24	1.12	0.54	8.85	1.01	0.41	12.85	1.26	0.56	9.26	1.12	0.45

In Table 4.2 the comparison between models illustrate that for all modules, both models can predict maximum power point with similar level of accuracy, but at low irradiance, the PVWatts model is slightly better. At medium and high irradiances, both models provide a good accuracy of approximately 1-2 %. Due to the good accuracy and low complexity of the model, the PVWatts model was applied further in the analysis conducted in the project. These results are visually represented in section A.2.

4.3 Analysis of dataset size

Two different test datasets were built having a 10 days of data measured each 5 minutes from two different seasons, summer and winter. The size of the training varies (10 days, 20 days, 40 days) and are collected also once from summer and once from winter. In Table 4.3 statistical errors are calculated for each train-test dataset combination. To be specified that the days used to build the datasets are consecutive and the irradiance used as an input to the model is the global irradiance. The model applied is the PVWatts model. The statistical errors are calculated for three different irradiance ranges.

Taking each situation separately, we have:

TABLE 4.3: RMSE and MAE statistics for P_m using PVWatts model for four cases in order to establish the training dataset size

Test	Training	Nr.days of training	PVWatts					
			RMSE [%]			MAE [%]		
			L	M	H	L	M	H
10 days of Summer	Summer	10	11.22	1.65	1.39	7.67	1.25	0.88
		20	11.5	2.47	1.51	8.25	2.04	0.87
		40	12.40	1.65	1.3	8.38	1.25	0.79
	Winter	10	13.15	2.1	0.85	9.88	1.17	0.68
		20	10.48	3.33	1.17	7.66	2.75	1.04
		40	10.46	2.07	0.71	7.65	1.14	0.56
10 days of winter	Summer	10	16.84	2.38	1.75	12.08	1.94	1.44
		20	16.21	2.06	1.65	11.63	1.61	1.3
		40	15.92	1.96	1.57	11.4	1.99	1.18
	Winter	10	8.6	1.92	0.6	5.74	1.02	0.47
		20	8.55	1.91	0.6	5.52	1.01	0.47
		40	9.26	1.91	0.57	5.54	1.05	0.45

- Case 1: If the training dataset is collected during the summer and tested in the summer, the differences in accuracy provided by the size of the training dataset are very small.
- Case 2: If the training dataset is collected during the winter and tested in the summer, it seems to provide better accuracy a training dataset having a bigger size.
- Case 3: If the training dataset is collected during the summer and tested in the winter, the errors increased comparing them to Case 1, and increasing the size of the training dataset does not have a big influence.
- Case 4: If the training dataset is collected during the winter and tested in the winter the errors are smaller than compared with Case 2, due to the similar irradiance distribution of the training dataset and test dataset . Increasing the size of the training dataset does not have a big influence.

In Figure 4.1 the measured power vs. the predicted power is plotted to highlight the difference of the fitting when the training dataset varies and is formed in winter and tested on a fixed test dataset in summer and winter.

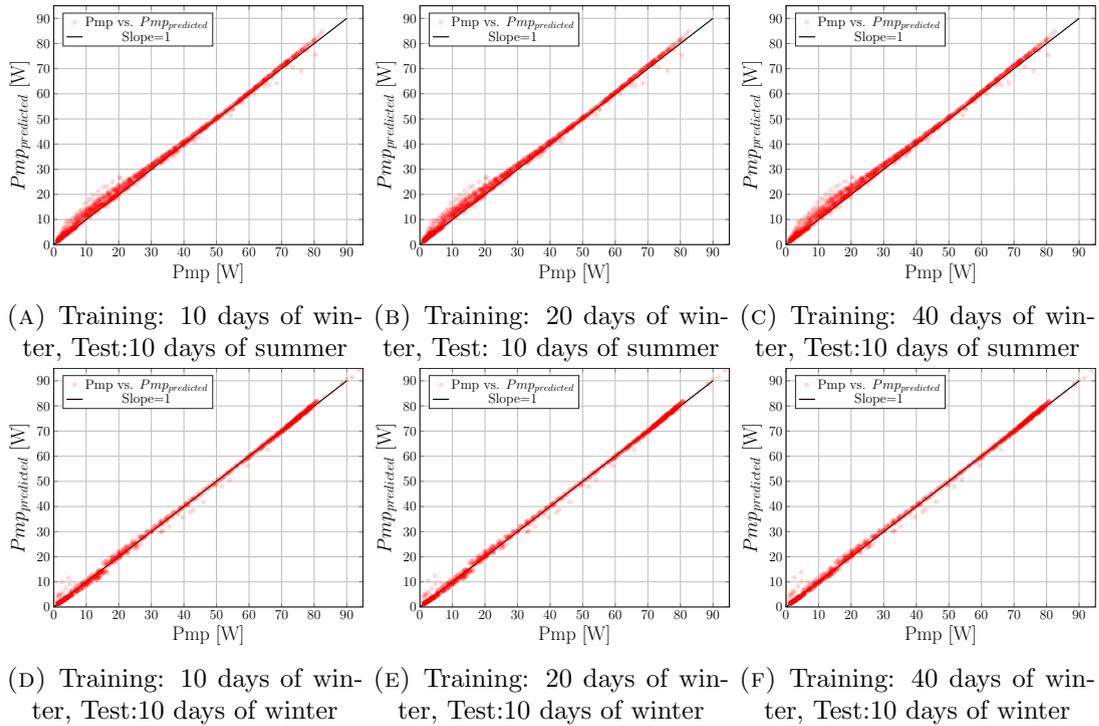


FIGURE 4.1: Measured power vs. predicted power using a varying training dataset from winter and tested on a fixed size test dataset number from summer and winter

4.4 Analysis of dataset distribution

All TC named by code discussed in section 4.1 are implemented and tested in order to see how the distribution of irradiance affects the accuracy of the model. Considering the different periods available from the two locations and the different weather conditions, the days used for the TC do not correspond, but they are from the same periods. For example, if for the 4s dataset in Cocoa, the 4 days are from 19 -22 September, in Golden, those 4 days will not be the same, but they will still be 4 sunny days from September or the beginning of October. Also, the number of samples from each location is different, at Cocoa, the measurements are done each 5 minutes, while at Golden they are done each 15 minutes. The TC are compared by using statistical errors and in Table 4.4 and Table 4.5 the results of the analysis conducted for the data from Cocoa and Golden are presented. To be noted that the irradiance used for this analysis is the global irradiance. The best and worse scenario are highlighted using green and red colours.

In Figure 4.2 the irradiance distribution for the four training datasets is displayed.

TABLE 4.4: RMSE and MAE statistics for P_m using PVWatts model for all test cases studying the irradiance distribution impact at Cocoa Beach, Florida

Code	Location: Cocoa , Model: PVWatts											
	Module: CIGS						Module: xSi					
	RMSE [%]			MAE [%]			RMSE [%]			MAE [%]		
	L	M	H	L	M	H	L	M	H	L	M	H
4s-4s	7.35	0.94	0.36	3.81	0.80	0.30	8.71	1.16	0.75	4.79	1.03	0.64
4s-4c	17.20	2.14	0.93	13.04	1.84	0.74	15.65	1.87	1.01	11.48	1.44	0.86
4s-2s2c	22.91	4.55	2.62	18.05	3.94	2.58	15.07	2.4	0.9	12.43	1.46	0.75
4s-1s9c	17.01	2.91	0.98	14.3	2.51	0.84	13.06	2.56	1.68	11.36	1.91	1.32
4c-4s	11.84	1.48	0.95	8.54	1.22	0.89	12.47	0.75	0.73	8.66	0.64	0.59
4c-4c	12.25	1.12	0.55	8.85	1.01	0.41	13.22	1.39	0.64	7.88	1.11	0.52
4c-2s2c	14.35	2.64	1.7	11.93	2.08	1.63	11.16	1.83	0.69	9.32	1.27	0.55
4c-1s9c	8.51	1.59	0.88	6.94	1.24	0.67	11.28	1.71	0.88	9.15	1.37	0.71
2s2c-4s	22.02	3.99	2.56	19.65	3.81	2.48	14.77	0.92	0.80	11.68	0.72	0.60
2s2c-4c	13.49	3.04	1.49	8.84	2.50	1.37	10.84	1.62	0.85	7.04	1.24	0.60
2s2c-2s2c	12.39	2.42	0.55	9.53	1.98	0.37	9.85	1.80	0.53	7.35	1.35	0.42
2s2c-1s9c	8.01	1.54	0.83	6.13	1.24	0.69	6.13	3.56	2.49	4.83	3.35	2.39
1s9c-4s	16.74	2.22	0.83	14.57	1.83	0.72	13.38	0.82	0.79	10.72	0.66	0.63
1s9c-4c	10.88	2.10	0.82	6.98	1.67	0.61	9.21	1.75	0.75	5.87	1.43	0.57
1s9c-2s2c	13.53	2.79	2.10	11.10	2.25	2.04	10.19	1.81	0.57	7.33	1.35	0.43
1s9c-1s9c	8.06	1.73	0.85	6.11	1.29	0.68	6.76	1.44	0.66	5.14	1.17	0.52

TABLE 4.5: RMSE and MAE statistics for P_m using PVWatts model for all test cases studying the irradiance distribution impact at Golden, Colorado

Code	Location: Golden , Model: PVWatts											
	Module: CIGS						Module: mSi					
	RMSE [%]			MAE [%]			RMSE [%]			MAE [%]		
	L	M	H	L	M	H	L	M	H	L	M	H
4s-4s	2.65	0.93	0.45	1.75	0.84	0.38	2.41	0.82	0.46	1.67	0.72	0.40
4s-4c	11.87	1.22	0.83	7.24	1.05	0.76	13.08	1.83	1.20	8.13	1.69	0.94
4s-2s2c	10.53	2.89	2.66	6.24	2.56	2.49	10.81	1.52	0.85	7.18	1.28	0.72
4s-1s9c	23	4.71	2.87	20.05	4.23	2.82	21.8	3.51	0.66	19.63	2.55	0.52
4c-4s	7.05	1.28	0.68	6.27	1.10	0.59	8.17	1.16	1	7.36	0.94	0.66
4c-4c	8.25	0.94	0.46	7.05	0.80	0.39	7.85	1.12	0.58	6.61	0.92	0.52
4c-2s2c	8.64	1.44	0.89	7.65	1.25	0.81	9.78	1.15	1.08	8.48	0.74	0.86
4c-1s9c	18.02	3.12	0.84	16.37	2.55	0.63	18.48	2.56	1.69	16.78	1.59	1.49
2s2c-4s	5.91	1.18	0.49	4.06	1.09	0.40	5.18	0.69	1.35	4.00	0.58	1.22
2s2c-4c	30.41	3.68	1.01	27.87	3.45	0.77	26.83	1.47	1.09	24.54	1.37	0.98
2s2c-2s2c	5.36	1.37	0.75	3.96	1.22	0.65	4.80	1.20	0.55	3.51	1.08	0.48
2s2c-1s9c	18.56	3.31	0.83	15.72	2.69	0.71	20.5	3.71	0.99	17.38	2.87	0.85
1s9c-4s	17.51	1.83	0.91	16.54	1.39	0.87	17.37	2.33	0.59	16.54	1.99	0.54
1s9c-4c	9.49	2.69	2.50	7.49	2.34	2.47	8.45	0.92	0.27	6.64	0.65	0.20
1s9c-2s2c	21.59	2.43	0.62	20.29	1.70	0.48	19.32	2.32	0.38	18.16	1.76	0.28
1s9c-1s9c	9.30	2.21	0.64	7.08	1.75	0.50	8.58	1.88	0.46	6.61	1.50	0.37

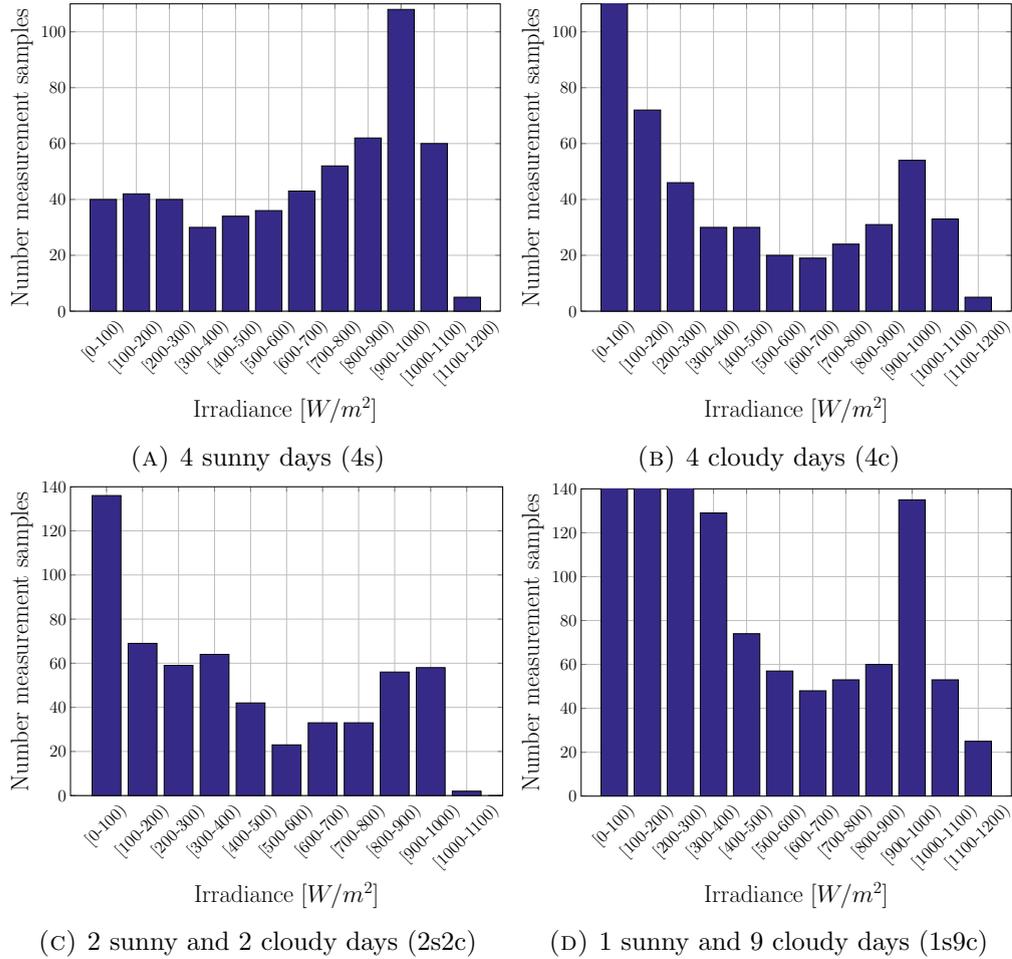


FIGURE 4.2: Distribution of irradiance vs. number of samples for the four training datasets

There are no big differences between the accuracy of the model applied on the two modules even though they are different technologies. At low irradiances, the errors are large for both modules. At medium irradiance, the errors for all the TC applied for CIGS module are below 5%, while for xSi, are slightly smaller, below 4%. At high irradiances, both modules have errors below 1% for the majority of the cases. The values are different for the two locations which proves that the geographical location and weather conditions will change the behaviour of the model. It can be seen that having a similar distribution for the training dataset as for the test dataset will provide better results. From Figure 4.3 to Figure 4.6 it is shown how different training dataset distributions affects the accuracy of the model by presenting the measured and predicted power and the prediction error for the entire test period, when the test dataset remains the same.

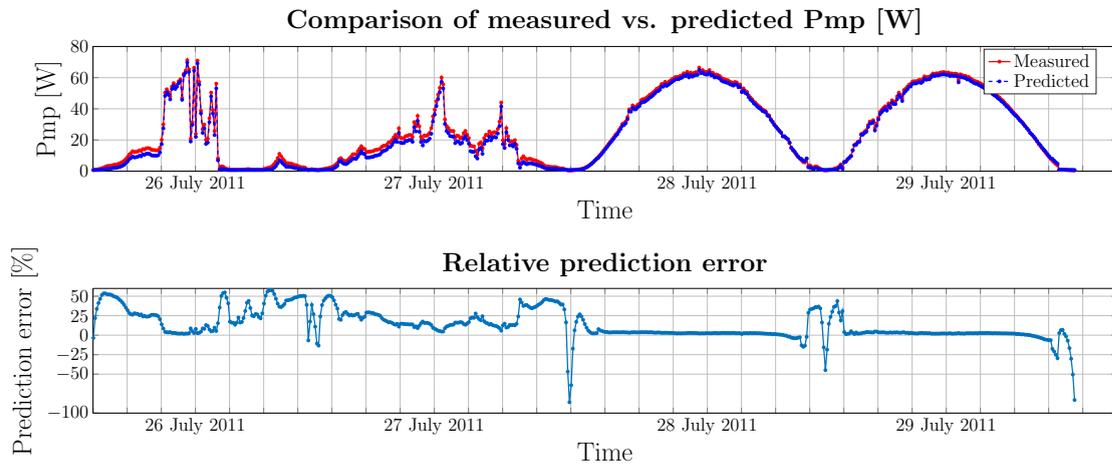


FIGURE 4.3: Training dataset made of 4 sunny days and test dataset of 2 sunny and 2 cloudy days ($4s - 2s2c$)

For the $4s - 2s2c$ case, the irradiance distribution of the training dataset is different from the test dataset, having more sample points at high irradiances. It can be seen in Figure 4.3 that during the sunny days, the model is accurate, and the prediction error is quite small, except for the points at low irradiance. For the cloudy days, the difference between the predicted and measured power becomes larger and therefore the prediction error is increased, reaching in some points over 50%.

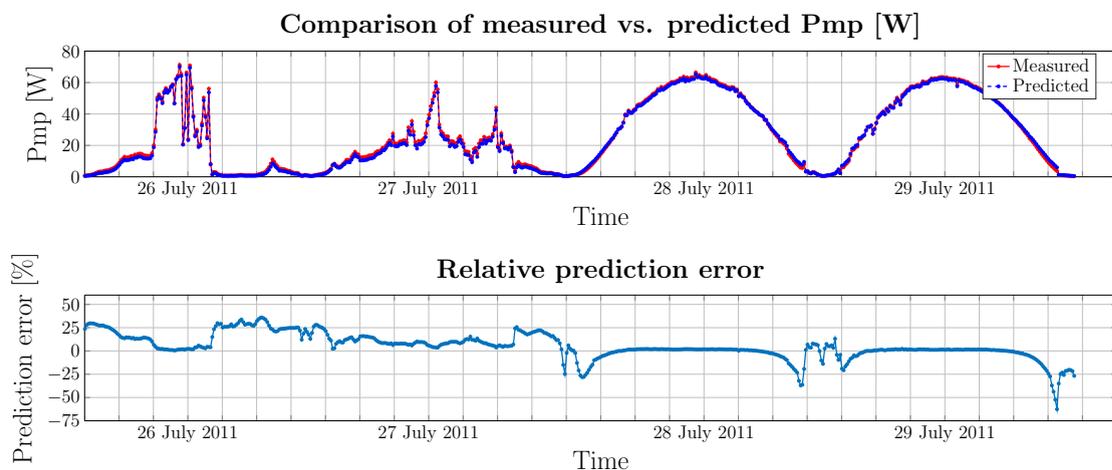


FIGURE 4.4: Training dataset made of 4 cloudy days and test dataset of 2 sunny and 2 cloudy days ($4c - 2s2c$)

For the $4c - 2s2c$ case, the irradiance distribution of the training dataset is similar to the test dataset. It can be seen in Figure 4.4 that the the estimated power fits better the measured power for cloudy days and does not seem to have a decrease of accuracy when applied on the sunny days. The prediction error is below 50% for all points and

from a RMSE of 22.91% at low irradiances, reaches a RMSE of 14.35%. Also for the other two irradiance ranges, the errors drop to half.

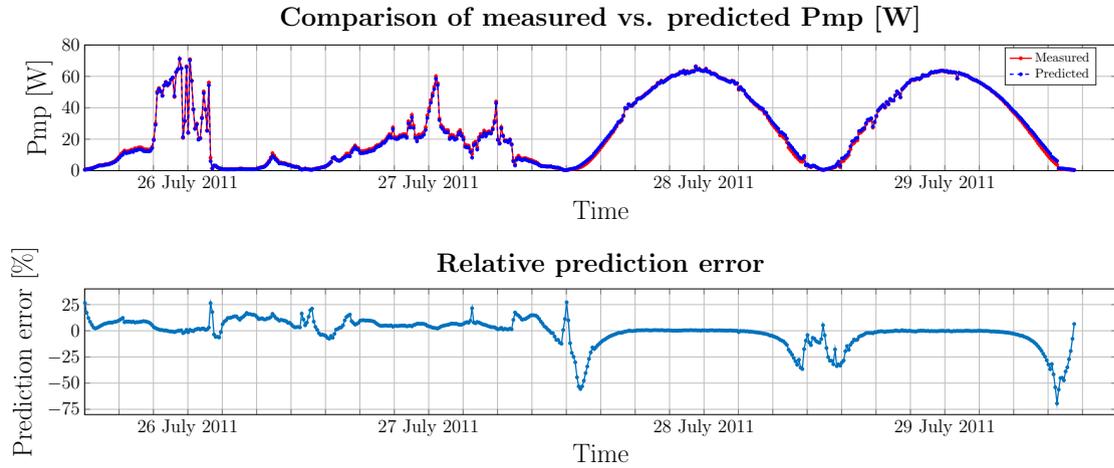


FIGURE 4.5: Training dataset made of 2 sunny and 2 cloudy days and test dataset of 2 sunny and 2 cloudy days ($2s2c - 2s2c$)

The case where the distribution of the training dataset is the same as the test dataset is the best scenario. Therefore, the $2s2c - 2s2c$ case can be seen in Figure 4.5. The prediction error dropped below 25% for all points and the RMSE and MAE values are the best so far, with a RMSE at high irradiances below 1%.

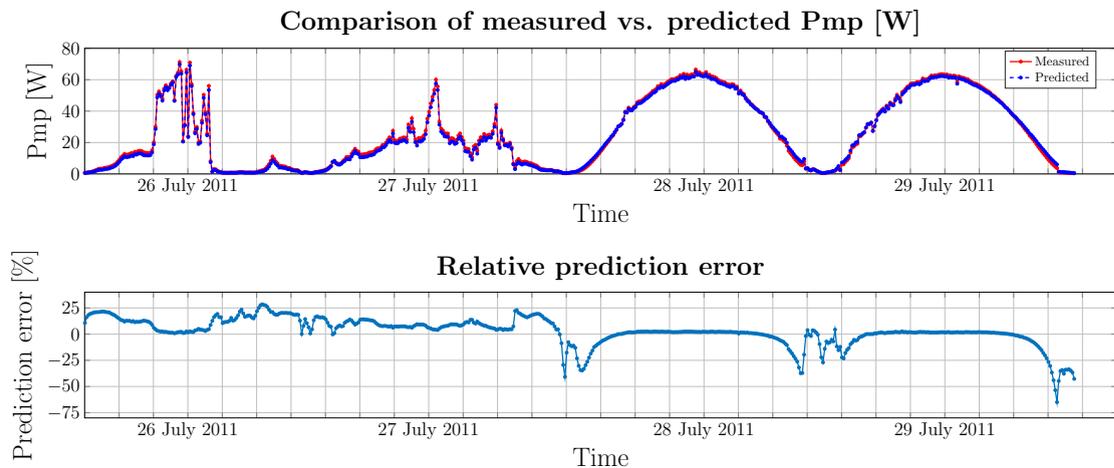


FIGURE 4.6: Training dataset made of 1 sunny and 9 cloudy days and test dataset of 2 sunny and 2 cloudy days ($1s9c - 2s2c$)

The $1s9c - 2s2c$ case has a training dataset distribution similar with the test dataset at medium and low irradiances, but has extra sample points at high irradiance. This does not affect the accuracy of the model due to the fact that the test dataset does not include many samples at high irradiance. Therefore, in Figure 4.6 the prediction error

is still below 25%, but the RMSE values for the three different irradiance ranges are bigger compared with the previous case.

4.5 Analysis of measurement procedure

Instead of global irradiance, in this study, the effective irradiance was used. The TC are kept the same and also the model used. In Table 4.6 the statistical errors are presented. It can be seen that at low irradiances, the model has a better accuracy when using effective irradiance for both modules. But, for xSi, the error for all ranges of irradiance is below 2%. The best and worse scenario is highlighted using green and red colours.

TABLE 4.6: RMSE and MAE statistics for P_m using PVWatts model for all test cases studying the effective irradiance distribution impact at Cocoa Beach, Florida

Code	Location: Cocoa , Model: PVWatts											
	Module: CIGS						Module: xSi					
	RMSE [%]			MAE [%]			RMSE [%]			MAE [%]		
	L	M	H	L	M	H	L	M	H	L	M	H
4s-4s	3.73	1.05	0.34	2.39	0.98	0.27	0.84	0.50	0.40	0.53	0.32	0.32
4s-4c	5.21	1.72	0.69	3.63	1.38	0.56	1.08	0.89	1.21	0.69	0.64	0.88
4s-2s2c	9.07	3.81	3.11	7.59	3.61	3.08	0.61	0.51	0.61	0.40	0.39	0.51
4s-1s9c	4.15	1.71	0.85	2.97	1.36	0.67	1.35	0.58	1.15	0.83	0.44	0.86
4c-4s	4.04	1.24	0.57	3.17	1.05	0.50	1.63	0.75	0.33	1.14	0.58	0.28
4c-4c	4.57	1.61	0.43	2.52	1.47	0.32	1.41	0.55	0.53	0.75	0.40	0.41
4c-2s2c	9.03	3.19	2.47	7	3.03	2.44	1.24	0.66	0.34	0.81	0.53	0.28
4c-1s9c	3.95	1.31	0.68	2.71	1.08	0.57	1.13	0.65	0.61	0.73	0.54	0.45
2s2c-4s	7.53	3.95	2.96	5.94	3.78	2.89	1.31	0.59	0.70	0.77	0.43	0.54
2s2c-4c	8.96	4.20	2.30	6.82	3.96	2.23	0.89	0.63	0.92	0.62	0.49	0.63
2s2c-2s2c	3.00	1.08	0.36	2.47	0.92	0.29	0.68	0.53	0.42	0.46	0.40	0.35
2s2c-1s9c	6.13	3.56	2.49	4.83	3.35	2.39	0.97	0.56	0.86	0.65	0.44	0.65
1s9c-4s	3.97	1.37	0.74	3.12	1.15	0.63	1.39	0.80	0.53	0.91	0.67	0.40
1s9c-4c	4.04	1.66	0.63	3.07	1.53	0.42	1.10	0.60	0.63	0.65	0.48	0.43
1s9c-2s2c	6.95	2.84	2.48	5.38	2.61	2.44	1.12	0.79	0.35	0.69	0.68	0.28
1s9c-1s9c	3.15	1.33	0.66	2.30	1.17	0.54	1.03	0.78	0.55	0.67	0.68	0.41

It was expected that the accuracy of the model to be better when using effective irradiance instead of global irradiance due to the fact that the reference cells behaves similar to PV panels. Even though they are only available for the c-Si technologies, therefore the better results provided by the xSi module, the reference cells removes the losses caused by reflectance and the solar spectrum variation effects. From Figure 4.7 to Figure 4.10 it is shown how using effective irradiance improves the accuracy of the models at low irradiances, by presenting the measured and predicted power and the prediction error for the entire test period in comparison with the prediction error given by the model when using global irradiance. The test dataset size and distribution does not change.

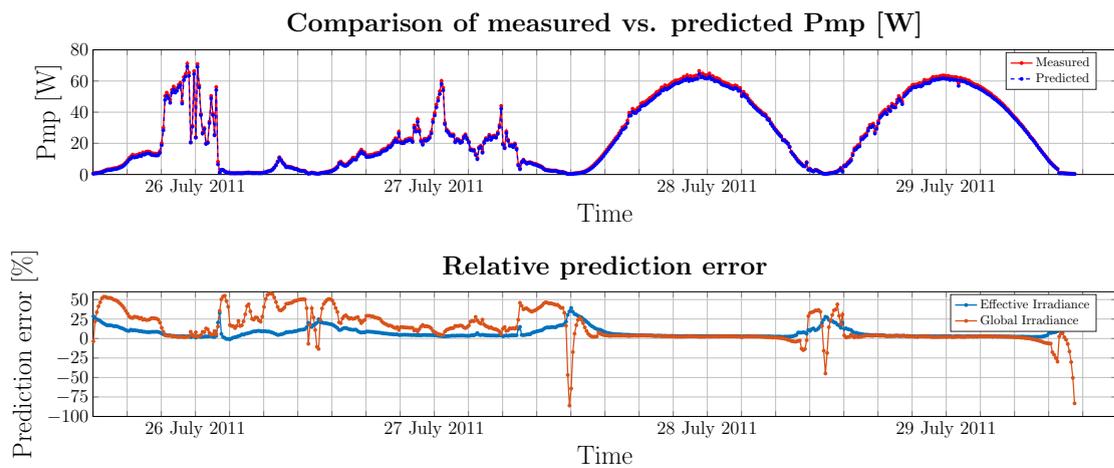


FIGURE 4.7: Training dataset made of 4 sunny days and test dataset of 2 sunny and 2 cloudy days ($4s - 2s2c$) using effective irradiance

For the $4s - 2s2c$ case, the fitting improved for low irradiances and cloudy days compared to the same case but when global irradiance was used. The RMSE dropped with 13% percent at low irradiance, but having a different distribution of datasets creates a bigger mismatch between the predicted and measured power.

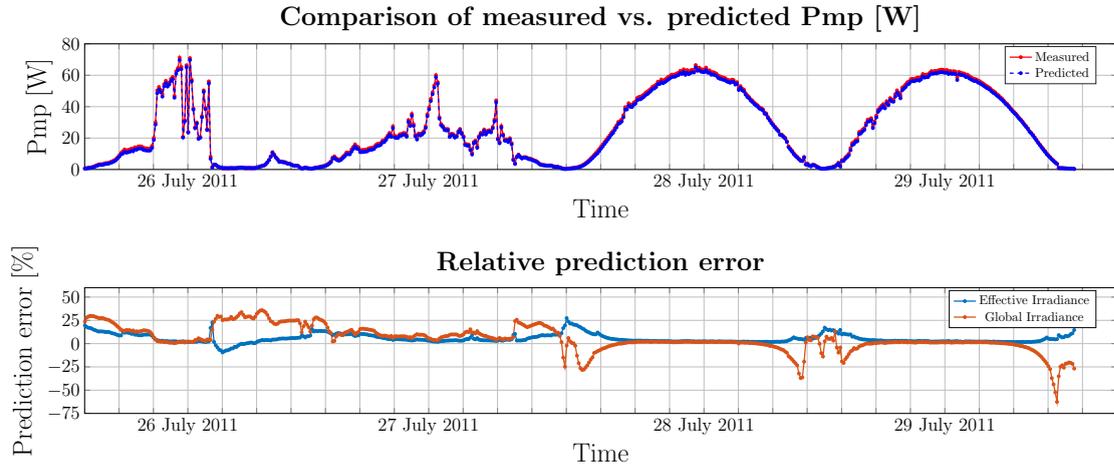


FIGURE 4.8: Training dataset made of 4 cloudy days and test dataset of 2 sunny and 2 cloudy days ($4c - 2s2c$) using effective irradiance

In the $4c - 2s2c$ case, the RMSE dropped only with 5% percent at low irradiance. This case responds slightly better at high irradiance than the $4s - 2s2c$ case due to the similar distribution of the training and test data.

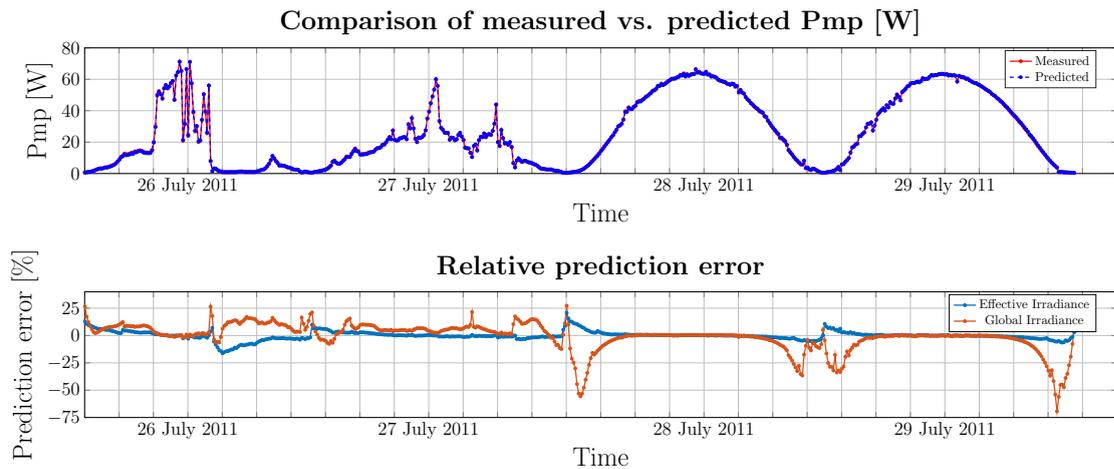


FIGURE 4.9: Training dataset made of 2 sunny and 2 cloudy days and test dataset of 2 sunny and 2 cloudy days ($2s2c - 2s2c$) using effective irradiance

Having a distribution of the training dataset identical with the test dataset and using effective irradiance is decreasing the RMSE at low irradiance until 3%. Also at medium and high irradiances, the RMSE is smaller than when using global irradiance.

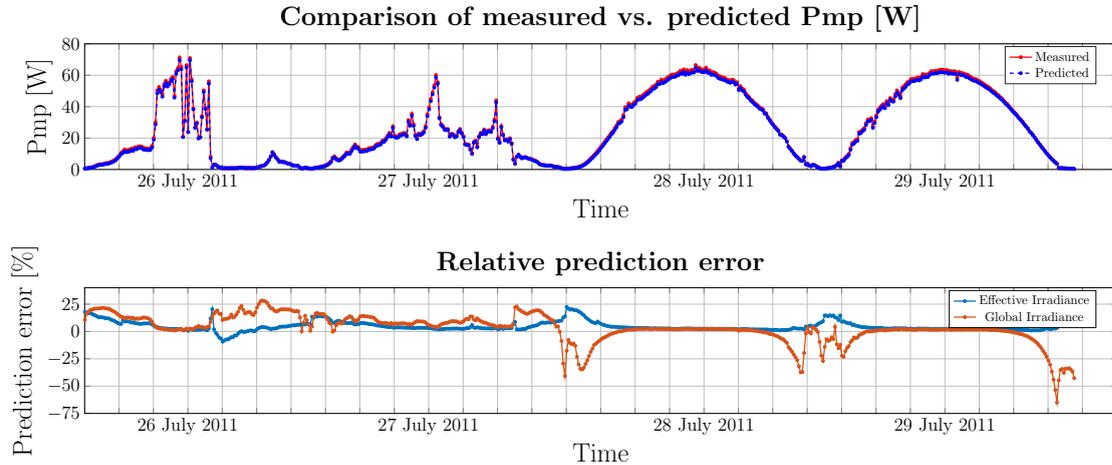


FIGURE 4.10: Training dataset made of 1 sunny and 9 cloudy days and test dataset of 2 sunny and 2 cloudy days ($1s9c - 2s2c$) using effective irradiance

The similar distribution in the datasets provided by the $1s9c - 2s2c$ and a better fit at low irradiances gives for this case a RMSE of 7% at low irradiance while still keeping the RMSE at medium and high irradiances below 3%.

4.6 Conclusions

Both proposed and tested models provided a good accuracy, but PVWatts was chosen due to better response at low irradiances and lower complexity of the model.

The training dataset is not that important if the data used for training is gathered in the summer. As proved in this chapter, training data from winter and testing it in the summer will double the size of the errors compared with a test dataset from the same season. It is recommended to have a training dataset for each season.

Having the same irradiance distribution in the training dataset as in the test dataset provides the best result. It is recommended to have as many training datasets with different distribution pattern so that can fit better the distribution required by the tested period.

There is no significant difference between the accuracy of the model when implemented on thin film compared with crystalline using global irradiance as the input.

Different locations and weather conditions affect the accuracy of the model.

Using the effective irradiance as an input to the model provides better results at low irradiances for both technologies, but is significantly better for the crystalline technology.

Chapter 5

Conclusions

A PV array performance modelling method for CIS models is proposed based on the PVWatts model which is parametrized using regression methods that can detect power loss.

From the analysis of the metastable behaviour we saw that from all the four thin-film modules studied, a-Si module displays the biggest seasonal effect and presents no sign of diurnal effect.

For a Danish climate (medium irradiances), the modules still provides a good PR, therefore working with this type of modules in Denmark should not be a problem.

Due to the variation of the PR caused by the change in season and/or degradation of the module each year, retraining the model data twice/year will be a good solution in keeping a high accuracy.

PVWatts and SAPM both provide a good accuracy but implementing the PVWatts model proved to be less complicated.

As long as the data for the training is collected in the summer, the size of the training dataset is not that important, but collecting it in another season significantly affects the accuracy of the model. It is recommended to have a training dataset for each season to keep a high performance.

No significant difference was seen between the accuracy of the model when implemented on CIGS compared to x-Si using global irradiance as an input. But when using effective irradiance, the performance of x-Si is significantly better.

Different locations and weather conditions affect the accuracy of the model.

The worst accuracy of the model is present at low irradiance.

5.1 Future work

As future work, it is recommended to:

- Add more TC in the analysis: It is important to check the accuracy of the model in the presence of more combinations between sunny days and cloudy days. Analysing more TC might establish an optimum distribution of irradiance that will fit with all types of test dataset.
- Implement other models and assess their accuracy: Many other performance models were developed and proved to have a good accuracy when applied on c-Si. Testing them on thin-film technologies should be considered.
- Analyse data from Denmark: Danish climate is very different than the climate at Golden or Cocoa. Having access to measurements data from Denmark will add more validity to this study.
- Evaluate prediction for periods longer than 1 year: It is important to evaluate the accuracy of the model after several years to quantify how much the degradation and ageing of the module affected the performance.
- Implement the performance modelling method in a performance monitoring system: This method of performance modelling should be implemented in a performance monitoring system and the performance should be automatically checked. If power loss is detected, a warning should be sent and the source of the loss will be investigated and fixed as soon as possible.
- Develop fault type detection: In order to have an even more efficient performance monitoring system, a fault type detection should be developed. As soon as the

power loss is detected, an identification of the type of fault is automatically done by the monitoring system. Having this feature, will avoid too much production loss.

- Analyse in depth the behaviour of performance models at low irradiance: Low irradiance seem to be problematic for the accuracy of performance models. A more in depth investigation should be conducted in how to lower the effect of low irradiances on the precision of the models.

Appendix A

Appendix

A.1 Matlab Code

Implementation of PVWatts model:

```
%Pmpp_PVWatts_v3 Implementation of the (version 3) PV Watts model
%   G: Irradiance [W/m^2]
%   T: Module temperature [Celsius]

function [ y, X] = Pmpp_PVWatts_v3( G, TMod, theta )
[n, m] = size(G);

% input data arrays are row vectors
if m > 1
    G = G';
    TMod = TMod';
end
dT = TMod - 25;

if max(G(1)) < 200
    x1 = G/1000;
    x2 = x1.*dT;
    x3 = (1-G/200).^4;
    X = [x1 x2 x3];
else
    G = G/1000;
```

```

    x1 = G;
    x2 = G.*dT;
    X = [x1 x2];
end

m = length(G);
X = [ones(m, 1) X];

if nargin == 2
    y = [];
else
    y = X * theta;
end
end

```

Implementation of SAPM model:

```

%PMPP_SAPM_FUNCTION_V1 Summary of this function goes here
function [ y, X] = Pmpp_SAPM_v1( G, T, theta )

[n, m] = size(G);

% input data arrays are row vectors
if m > 1
    G = G';
    T = T';
end

k = 1.3806503e-23; %Boltzmann constant
q = 1.6022e-19;    %Electron charge
ns = 60;
T = T+273.15;
dT = T - 297;

G = G/1000;

x1 = ns * k * T.* log(G) / q;
x2 = ns * k * (T.* log(G)).^2 / q;
x3 = dT;
x4 = (1-G).*dT;

```

```
x5 = G;
x6 = G.^2;
x7 = G.*dT;
x8 = G.^2 .* dT;
x9 = 1./G;

X = [x1 x3 x5 x7 x9];

m = length(G);
X = [ones(m, 1) X];

if nargin == 2
    y = [];
else
    y = X * theta;
end
end
```

Implementation of the training and test phase using PWVatts model:

```
clear
close all
clc

addpath('...\models\PVWatts');
addpath('...\models');
addpath('...\utils');

%% Train model
load('...\TestCase1\testdata\Cocoa.CIGS00408_12Sept_15Sept_sunny.mat')

Vmp = VdcA;
Imp = IdcA;
n = length(Imp);
Imp = reshape(Imp, n,1);
Isc = reshape(Isc, n,1);
Vmp = reshape(Vmp, n,1);
Timestamp = reshape(Timestamp, n,1);
G = reshape(G, n,1);
TMod = reshape(TMod, n,1);
```

```
ISC0 = 2.6;
ALFA = 0.01;
T0 = 25;
Pmp = PdcA;

%% using Effective Irradiance
%
Ee = Isc./(ISC0*(1+ALFA*(T0-TMod)))*1000;
G=Ee;

%% apply model
idx_highG = G > 200;
model_function = @Pmpp_PVWatts_v3;

model_params_highG = Train.Performance.Model(model_function, G(idx_highG), ...
TMod(idx_highG), Pmp(idx_highG), 'Pmp');
model_params_lowG = Train.Performance.Model(model_function, G(~idx_highG), ...
TMod(~idx_highG), Pmp(~idx_highG), 'Pmp');

%% test model
clearvars -except model_params_highG model_params_lowG model_function
load('...\TestCase1\testdata\Cocoa-CIGS00408_12Sept_15Sept_sunny.mat');

Vmp = VdcA;
Imp = IdcA;
n = length(Imp);
Imp = reshape(Imp, n,1);
Isc = reshape(Isc, n,1);
Vmp = reshape(Vmp, n,1);
Timestamp = reshape(Timestamp, n,1);
G = reshape(G, n,1);
TMod = reshape(TMod, n,1);
ISC0 = 2.6;
ALFA = 0.01;
T0 = 25;
Pmp = PdcA;

%% using Effective Irradiance
Ee = Isc./(ISC0*(1+ALFA*(T0-TMod)))*1000;
G=Ee;
```

```

%% apply parameters and test
idx_highG = G > 200;
Pmp_predicted_highG = Test_Performance_Model(model_function, ...
model_params_highG, G(idx_highG), TMod(idx_highG));
Pmp_predicted_lowG = Test_Performance_Model(model_function, ...
model_params_lowG, G(~idx_highG), TMod(~idx_highG));

Pmp_predicted = nan(size(Pmp));
Pmp_predicted(idx_highG) = Pmp_predicted_highG;
Pmp_predicted(~idx_highG) = Pmp_predicted_lowG;

%% calculate errors for 3 different irradiance intervals
idx_small = G < 400;
idx_medium = G >= 400 & G <= 800;
idx_high = G > 800;
RMSE_0_400 = sqrt(mean((Pmp_predicted(idx_small) - Pmp(idx_small)).^2))...
./mean(Pmp(idx_small))*100
RMSE_400_800 = sqrt(mean((Pmp_predicted(idx_medium) - Pmp(idx_medium)).^2))...
./mean(Pmp(idx_medium))*100
RMSE_800_max = sqrt(mean((Pmp_predicted(idx_high) - Pmp(idx_high)).^2))...
./mean(Pmp(idx_high))*100
MAE_0_400 = mean(abs(Pmp_predicted(idx_small) - Pmp(idx_small)))...
./mean(Pmp(idx_small))*100
MAE_400_800 = mean(abs(Pmp_predicted(idx_medium) - Pmp(idx_medium)))...
./mean(Pmp(idx_medium))*100
MAE_800_max = mean(abs(Pmp_predicted(idx_high) - Pmp(idx_high)))...
./mean(Pmp(idx_high))*100

%% plot estimated power versus measured day
evaluatePrediction(datetime(2011, 9, 12, 0, 0, 0), ...
datetime(2011, 9, 15, 20, 0, 0), Timestamp, G, Pmp, Pmp_predicted, 'Pmp [W]')
% matlab2tikz

```

TrainPerformanceModel function:

```

%TRAIN_LINEAR_MODEL Summary of this function goes here
function [ model_params ] = Train_Performance_Model(model_function, G, TMod, ...
y_measured, y_name)

[n, m] = size(G);

```

```
% input data arrays are row vectors
if m > 1
    G = G';d
    TMod = TMod';
    y_measured = y_measured';
end

[~, X] = model_function(G, TMod);

X_lm = X(:, 2:end);

initial_model_params = LinearModel.fit(X_lm, y_measured);

disp('++++++Initial model parameters+++++')
disp(initial_model_params)

[y_pred ~] = predict(initial_model_params, X_lm, 'Prediction', 'observation');

%% compare measured vs predicted values
scatterPlot(G, y_measured, TMod, 'G', y_name, 'TMod');
hold on;
plot(G, y_pred, 'rx');
xlabel('G [W/m^2]');
ylabel('y_name')

scatterPlot(y_measured, y_pred, TMod, ['Measured ' y_name], ...
['Predicted ' y_name], 'TMod');
title('Initial regression model');

%% plot residuals
figure
plotResiduals(initial_model_params, 'histogram');
figure
plotResiduals(initial_model_params, 'fitted');
figure
plotResiduals(initial_model_params, 'caseorder');
figure
plotResiduals(initial_model_params, 'probability');

model_params = initial_model_params;
```

```
end
```

TestPerformanceModel function:

```
%TEST_PERFORMANCE_MODEL Summary of this function goes here
function [ y_pred ] = TestPerformanceModel(model_function, model_params, G, TMod)
%   Detailed explanation goes here

[n, m] = size(G);

% input data arrays are row vectors
if m > 1
G = G';
TMod = TMod';
end

[~, X] = model_function(G, TMod);

X_lm = X(:, 2:end);

[y_pred, ~] = predict(model_params, X_lm, 'Prediction', 'observation');

end
```

evaluatePrediction function:

```
%PLOT PMP PREDICTION DAY Summary of this function goes here
function [ output_args ] = evaluatePrediction(startDate, stopDate, ...
Timestamp, G, y, y_predicted, y_name)

idx = Timestamp > startDate & Timestamp < stopDate;
y = y(idx);
y_predicted = y_predicted(idx);
G = G(idx);
Timestamp = Timestamp(idx);
y_error = ((y-y_predicted)./y*100);

figureFullScreen
```

```
subplot(2, 2, 1)
plot(Timestamp, y, 'r.-', 'Displayname', 'Measured')
hold on
plot(Timestamp, y_predicted, 'b.--', 'Displayname', 'Predicted')
grid on;
xlabel('Time')
ylabel(y_name)
legend('show', 'location', 'best')
title(['Comparison of measured vs. predicted ' y_name]);
%
subplot(2, 2, 3)
[hAx,hLine1,hLine2] = plotyy( Timestamp,y_error,Timestamp, G);
set([hLine1,hLine2],'Marker', '.');
grid on;
title('Relative prediction error')
xlabel('Time')
ylabel(hAx(1), 'Prediction error [%]') % left y-axis
ylabel(hAx(2), 'Irradiance [W/m^2]') % right y-axis

GLevels = 0:100:1200;
n = length(GLevels)-1;
m = 0;
for i=1:n

    nn(i) = i;
    x_labels(i) = {sprintf('%d-%d', GLevels(i), GLevels(i+1))};
    idxG = G >= GLevels(i) & G < GLevels(i+1);
    y_error_G(i) = {y_error(idxG)};
    y_error_G_count(i) = sum(idxG);
    if sum(idxG) > m
        m = sum(idxG);
    end
end

ydata = nan(m, n);

for i=1:n
    m = length(y_error_G{i});
    if m > 0
        ydata(1:m, i) = y_error_G{i};
    end
end
```

```
end

subplot(2, 2, 2)
bar(nn, [y_error_G_count])
xlabel('Irradiance [W/m^2]');
ylabel('Number measurement samples');
grid on
set(gca, 'xticklabel', x_labels);
xlim([1 11])
title('Distribution of measurement samples vs. irradiance')

subplot(2, 2, 4)
boxplot(ydata)
xlabel('Irradiance [W/m^2]');
ylabel('Prediction error [%]');
grid on
xlim([1 11])
set(gca, 'xticklabel', x_labels);
title('Distribution of error vs. irradiance')

end
```

A.2 PVWatts vs SAPM

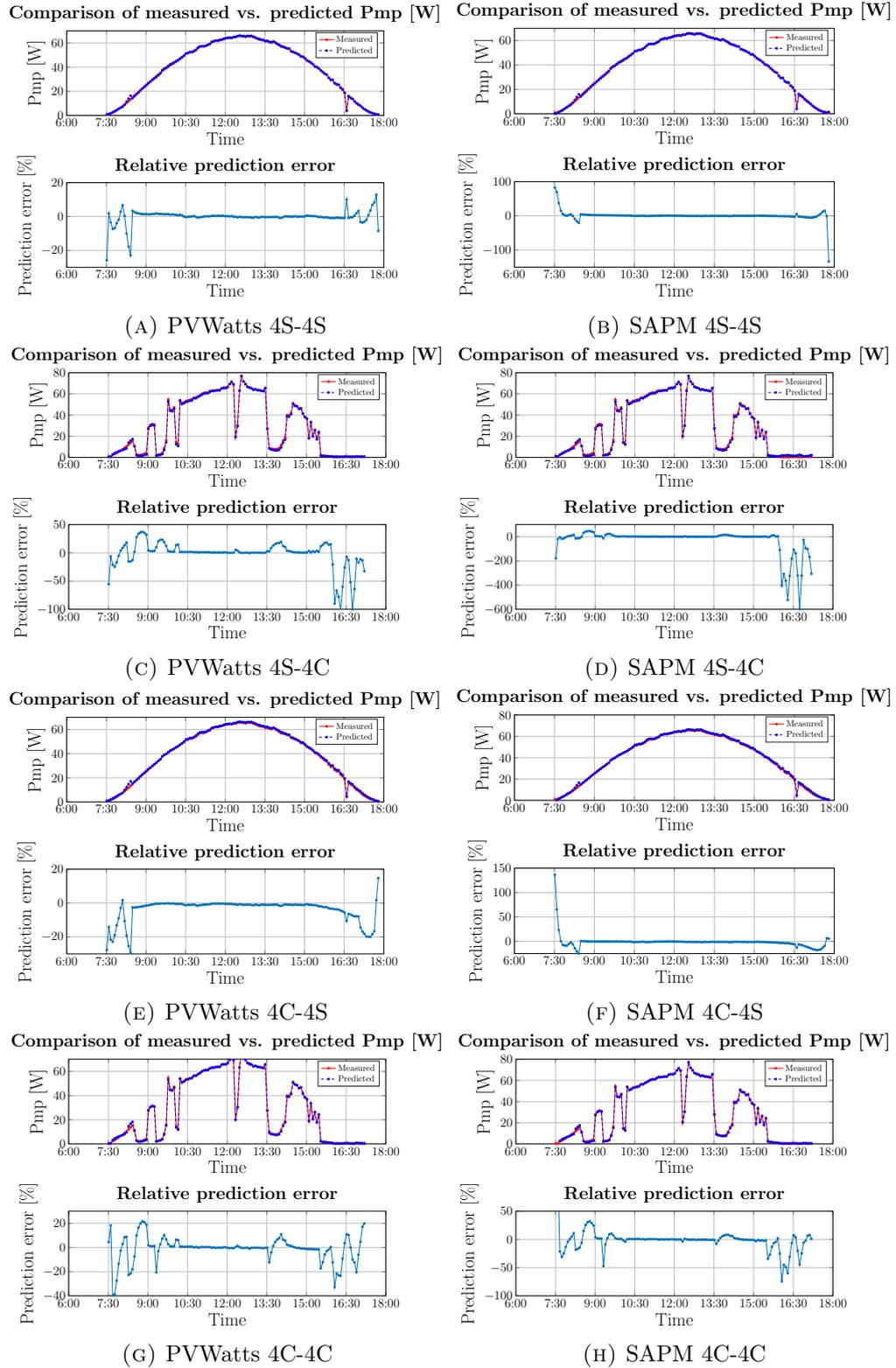


FIGURE A.1: Modeling error statistics for P_m using SAPM and PVWatts models on a CIGS module considering the four different cases

Bibliography

- [1] International Energy Agency Photovoltaic Power System. *A Snapshot of Global Photovoltaic Markets*. International Energy Agency Photovoltaic Power System Programme, 2016.
- [2] K. Petrichenko W. Rickerson J. L. Sawin K. Seyboth J. Skeen B. Sovacool F. Sverrisson Laura E. Williamson C. Lins, E. Musolino. *Global Status Report*. Ren21 Secretariat, 2015.
- [3] M. Reking G. Masson, S. Orlandi. *Global Market Outlook For Photovoltaics 2014-2018*. European Photovoltaic Industry Association (EPIA), 2014.
- [4] C. Kost S. Nold S. Philipps R. Preu T. Schlegl G. Stryi-Hipp G. Willek H. Wirth I. Brucker A. Häberle W. Warmuth B. Burger, K. Kiefer. *Photovoltaics Report*. Fraunhofer Institute for Solar Energy, Systems, ISE with support of PSE AG Freiburg,, 2016.
- [5] B. Marion. *Comparison of Predictive Models for Photovoltaic Module Performance*. Photovoltaic Specialists Conference, 2008. PVSC '08. 33rd IEEE, 2008.
- [6] Jay A. Kratochvil David L. King, William E. Boyson. *Photovoltaic Array Performance Model*. Sandia National Laboratories, 2004.
- [7] D. C. Jordan and S. R. Kurtz. *Photovoltaic Degradation Rates — An Analytical Review*. National Renewable Energy Laboratory (NREL), Golden, CO, USA, June 2012.

-
- [8] G. Friesen M. Pravettoni S. ISAAC M. Apolloni A Louwen WGJHM van Sark M. Schweiger G. Belluardo J. Wagner A. Tetzlaff P. Ingenhoven D. Moser T. J Silverman, U. Jahn. *Characterisation of Performance of Thin-film Photovoltaic Technologies*. International Energy Agency Photovoltaic Power System Programme, May 2014.
- [9] Hukseflux Thermal Sensors. *Outdoor PV performance monitoring: Pyranometers versus reference cells*. Hukseflux Thermal Sensors, 2011.
- [10] M. Gostein L. Dunn. *Spectral Effects in Performance Ratio Measurement Comparing PV Reference Devices and Pyranometers*. IEEE 39th Photovoltaic Specialists Conference (PVSC), 2012.
- [11] L. Fanni A. Virtuani. *Seasonal power fluctuations of amorphous silicon thin-film solar modules: distinguishing between different contributions*. 2012.
- [12] Luis Castaner Augustin McEvoy, Tom Markvart. *Practical Handbook of Photovoltaic, Fundamentals and Applications Second Eddition*. Elsevier B.V, 2012.
- [13] Luis Castaner Tom Markvart. *Practical Handbook of Photovoltaic, Fundamentals and Applications, pp 282-308*. Elsevier B.V, 2003.
- [14] T. Betts R. Gottschalg and S. Williams. *A critical appraisal of the factors affecting energy production from amorphous silicon photovoltaic arrays in a maritime climate*. Dec 2004).
- [15] G. T. Klise and J. S. Stein. *Models used to asses the Performance of Photovoltaic Systems*. Sandia National Laboratories, Albuquerque, NM, 2009.
- [16] J. Sutterlueti J. S. Stein. *Outdoor PV Performance Evaluation of Three Different Models: Single-Diode, SAPM and Loss Factor Model*. Sandia National Laboratories , Albuquerque,United States, 1 September 2013.
- [17] Douglas C. Montgomery. *Design and Analysis of Experiments,8th eddition, Chapter 10*. John Wiley & Sons, Inc., 2009.