

**Faculty of Science and Engineering  
School of Electrical Engineering, Computing, and Mathematical  
Sciences**

**Random Finite Sets Based Very Short-Term Solar Power  
Forecasting Through Cloud Tracking**

**Florian Benjamin Eric Barbieri**

**This thesis is presented for the Degree of  
Doctor of Philosophy  
of  
Curtin University**

**June 2019**

## **DECLARATION AND SIGNATURE**

---

To the best of my knowledge and belief, this thesis contains no material previously published by any other person except where due acknowledgment has been made. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature: .....

Date: .....13/06/2019.....

# LIST OF PUBLICATIONS

---

## **Peer-to-peer reviewed journal articles:**

“Very short-term photovoltaic power forecasting with cloud modeling: A review” by Florian Barbieri, Sumedha Rajakaruna, and Arindam Ghosh, published in Renewable and Sustainable Energy Reviews, Elsevier in August 2017 (DOI.[10.1016/j.rser.2016.10.068](https://doi.org/10.1016/j.rser.2016.10.068))

This paper endeavours to review exhaustively the state of the art and advances in terms of very short-term photovoltaic power forecast in literature.

“Intra-Hour Cloud Tracking Based on Probability Hypothesis Density Filtering” by Florian Barbieri, Corentin Riffart, Ba-Tuong Vo, Sumedha Rajakaruna, and Arindam Ghosh, published in the peer-review journal IEEE Transactions on Sustainable Energy in January 2018, Volume 9(1) pp. 340-349 (DOI.[10.1109/TSTE.2017.2733258](https://doi.org/10.1109/TSTE.2017.2733258))

This paper presents the results of using a probability hypothesis density filter, cardinalised probability hypothesis filter method to track clouds with an original bounding box and centroid method and compares those results with those obtained by using Lukas-Kanade optical flow. An original method is also presented to select the most accurate pixel discrimination criterion for a given sky camera.

## **Conference articles:**

“A Comparative Study of Clear-Sky Irradiance Models for Western Australia” by Florian Barbieri, Corentin Riffart, Ba-Tuong Vo, Sumedha Rajakaruna, and Arindam Ghosh, published at the occasion of the IEEE Power and Energy Society General Meeting conference in Boston in July 17 2016 (DOI.[10.1109/TSTE.2017.2733258](https://doi.org/10.1109/TSTE.2017.2733258))

This article compares the most common fast computing empirical clear-sky irradiance models with some collected data to see which ones are the most accurate for Western Australia.

# TABLE OF CONTENTS

---

Declaration and Signature.....	ii
List of Publications .....	iii
Table of Contents.....	iv
List of Figures.....	ix
List of Tables .....	xix
Acknowledgements.....	xxii
Abstract.....	xxiii
Statement of Attribution by Others.....	xxiv
Nomenclature.....	xxv
1 Introduction and Problematic.....	27
1.1 Main Objectives of this Study.....	27
1.2 Contributions of This Study .....	27
1.3 Advantages of PV Generation.....	28
1.4 Fast Increasing Popularity of PV Generation.....	29
1.5 Challenges of PV Generation.....	29
1.6 Sunlight Fluctuations and Their Effect on PV Power .....	30
1.7 The Dynamics of PV Generation.....	31
1.8 The Case Study of PV Generation in Western Australia .....	33
1.9 Solar Power Forecasting as a Solution and Similarities with Wind Forecasting .....	34
1.10 Outline of This Thesis.....	34
2 State of the Art of Short-Term Solar Power Forecasting.....	37
2.1 Relevant Methods used in the Reference Field of Wind Power Forecasting.....	37
2.1.1 Analysis of the Reference Field of Wind Power Forecasting .....	37
2.1.2 Conclusions on What Can Be Learned From the Field of Wind Power Forecasting....	41
2.2 Complementary Information on Statistical Methods for PV Forecasting.....	42
2.2.1 Review of the Statistical Methods for PV Forecasting .....	42
2.2.2 Conclusions of the Review of Statistical Methods Used for PV Forecasting .....	44
2.3 Comparison of some Popular PV Power Forecasting Approaches.....	44
2.3.1 Conclusions of the Comparison between Popular PV Power Forecasting Approaches	46
2.4 PV Cell Modelling Strategy .....	46
2.5 Cell / Module Temperature Model.....	49
2.6 Further Considerations of the PV Model .....	52
2.6.1 Conclusions on PV Modelling .....	54
2.7 Advances in Irradiance Forecasting.....	55
2.7.1 Review of Clear-Sky Irradiance Forecasting .....	55

2.7.2	Review of Irradiance Forecasting considering clouds .....	56
2.7.3	Review of Irradiance Forecasting considering clouds .....	58
2.8	Cloud Cover Modelling and Cloud Classification .....	59
2.8.1	Overview .....	59
2.8.2	Assessment of Cloud Cover .....	65
2.8.3	Identification of Cloud Types .....	66
2.8.4	Conclusions .....	69
2.9	Weather Classification for Mid-term PV Output Forecasting .....	70
2.9.1	Review of Weather Classifications .....	70
2.9.2	Conclusions on Weather Classification .....	71
2.10	Very-short Term PV Power Forecasting .....	72
2.10.1	NWP Approach .....	72
2.10.2	Sky, Satellite Images and Sensor Array Approach .....	72
2.10.3	Random-sequence and Time Series Models .....	78
2.10.4	Conclusions on Very Short-Term PV Forecasting .....	79
2.11	Conclusions on the Literature Review .....	80
3	Irradiance, Cloud, Sun and Shadow Modelling .....	82
3.1	Relevance of the Overall Strategy .....	82
3.2	Clear-Sky Irradiance Modelling Considerations .....	83
3.3	Empirical Irradiance Models .....	89
3.3.1	Kasten-Czeplak (KC) Empirical Model .....	89
3.3.2	Haurwitz Empirical Model .....	89
3.3.3	Berger-Duffie (BD) Empirical Model .....	89
3.3.4	Robledo-Soler (RS) Model .....	90
3.3.5	Laue Model .....	90
3.3.6	Meinel Model .....	90
3.3.7	Denashyar-Paltridge-Proctor (DPP) Model .....	90
3.3.8	Adnot-Bourges-Campana-Gicquel (ABCG) Empirical Model .....	91
3.3.9	Dazhi Empirical Model .....	91
3.3.10	Threlkeld and Jordan (TJ) Empirical Model .....	91
3.4	Physical Models: Case Study of the Meeus, Bird and Hulstrom Model .....	92
3.4.1	Comparison of the some Relevant Empirical Models for Western Australia. ....	96
3.4.2	Initial Regression of Empirical Data .....	100
3.4.3	Refined Regression Made with an Extended Dataset .....	105
3.4.4	Accuracy of the Regression Models .....	108
3.4.5	Conclusions on the Regression Models .....	111
3.5	Cloud, Sun and Shadow Projection Modelling .....	112

3.5.1	Single-point cloud model .....	112
3.5.2	Four-point cloud model.....	120
3.5.3	A possibly more Realistic Cloud Model .....	122
3.5.4	Simple shadow model .....	122
3.5.5	Camera Calibration .....	124
3.5.6	3D rotation matrix.....	129
3.5.7	Solar disk and cloud model.....	133
3.5.8	Estimation of the shading time and duration.....	144
3.5.9	Shading intensity.....	150
3.5.10	Sun-Pixel Angle .....	154
3.6	Conclusions on Irradiance, Cloud, Sun and Shadow Modelling.....	158
4	Cloud/Sky Discrimination and Cloud Tracking.....	159
4.1	Introduction and Prerequisites .....	159
4.2	Hardware Consideration, Camera Settings and Maintenance.....	159
4.2.1	Preparation of a Hardware Setup .....	159
4.2.2	Recommended Choices and Recommended Settings .....	160
4.2.3	Notes on Maintenance Routine and Camera Limitations .....	161
4.3	Clear Sky Pixels / Cloudy Pixels Discrimination .....	162
4.3.1	Preparation of the Input Data .....	162
4.3.2	Chromatic Analysis of the Input Data.....	164
4.3.3	Pixel Discrimination Criterion Assessment Method.....	166
4.3.4	Discrimination Criteria .....	168
4.4	Automatic Cloud Classification and Optical Depth.....	171
4.5	Cloud Tracking and Random Finite Set based Forecasting Approaches .....	172
4.5.1	Probability Hypothesis Filtering (PHD).....	173
4.5.2	Kalman Prediction of Births .....	177
4.5.3	Cardinalised Probability Hypothesis Filtering (CPHD).....	178
4.5.4	State of the Art Methods .....	179
	Comparison and Discussion.....	182
4.5.5	Comparison Results .....	184
4.5.6	Conclusion of the comparison.....	185
5	Data Collection and Analysis.....	186
5.1	Introduction.....	186
5.2	Experimental Setup.....	186
5.3	Quick Chromatic Analysis .....	192
5.4	Detailed Chromatic Analysis of a Clear-Sky Library .....	201
5.5	Cloud Feature Analysis and Automated Cloud Classification.....	202

5.5.1	Clear-sky and Cloud Data Collection .....	203
5.5.2	Clear-Sky and Cloud Feature Characterisation.....	204
5.6	Measurement of the Optical Thickness.....	207
5.6.1	Cloud type 2: Cumulus Clouds .....	208
5.6.2	Cloud type 3 : Stratocumulus Clouds.....	211
5.6.3	Cloud type 4 : Stratus/Altostratus Clouds.....	213
5.6.4	Cloud type 5 : Cirrocumulus/Alto cumulus .....	215
5.6.5	Cloud type 6 : Cirrus/Cirrostratus .....	216
5.6.6	Cloud type 7: Cumulonimbus/Nimbostratus.....	218
5.6.7	Comparison of All Types of Cloud.....	220
5.7	Conclusions.....	221
6	PV Power Control.....	222
6.1	Battery Storage Management.....	222
6.1.1	PV Array Model.....	222
6.1.2	Battery Model .....	260
6.1.3	Battery and Bidirectional DC/DC Converter with Controller.....	265
6.1.4	PV Power Control .....	271
6.2	Voltage Source Converter Control.....	272
6.2.1	VSC Controller Proposed Equipment and Setup .....	273
6.2.2	VSC Controller Program Testing.....	278
6.3	Conclusions.....	287
7	Conclusions and Future Works .....	288
7.1	Conclusions.....	288
7.1.1	Conclusions about the Method.....	288
7.2	Future Works .....	288
7.2.1	Full Validation of the Grid-Connected PV System with Battery Storage.....	288
7.2.2	Further Testing of the Forecasting Method.....	289
7.2.3	More Accurate Clear-Sky Irradiance Modelling.....	289
7.2.4	Automation of the Calibration .....	289
7.2.5	Application to a Large Solar Farm.....	290
7.2.6	Application to a Microgrid and Optimisation of Local Energy Resources.....	290
7.2.7	Hardware implementation.....	291
7.2.8	Transactions .....	291
7.2.9	Scada Communications.....	291
7.2.10	Composite Method.....	291
Appendix A –	Forecasting Method Assessment.....	292
Appendix B –	Camera installation .....	298

Appendix C – Camera Box .....	301
Appendix D – Authorship Contribution Statements.....	302
Appendix E – Copyright Owner Permissions .....	308
Bibliographical References .....	312

---



## LIST OF FIGURES

---

Figure 1.1 – (a) Received DNI and Power Output of a Thin-film Array and (b) Corresponding Ambient Temperature and Relative Humidity.....	30
Figure 1.2 – Seasonal Variations of Global Solar Irradiance for a Given Place on the Surface of the Earth.....	32
Figure 1.3 – Power networks and major photovoltaic farms in Western Australia [40][41] superposed on top of a satellite view of the state shot on the 29 <sup>th</sup> of February 2016 at 5 pm [42], illustrating the geographical disparities of cloud cover over a large territory at a given time. ....	33
Figure 2.1 – I-V curves for various solar radiances with a Kyocera KC200GT [27]. ....	47
Figure 2.2 – P-V curves for various solar radiances with a Kyocera KC200GT [29]. ....	47
Figure 2.3 – I-V curves for different temperatures with a Kyocera KC200GT [29]. ....	48
Figure 2.4 – P-V curves for different temperatures with a Kyocera KC200GT [29]. ....	48
Figure 2.5 – Two-diode model of a single PV cell [50].....	53
Figure 2.6 – Two-diode model of a PV module with partial shading of one cell among a module of n cells [50]. ....	54
Figure 2.7 – Illustration of the various types of clouds.....	62
Figure 2.8 – Temporal resolutions compared to the spatial range of observation for various types of input data.....	73
Figure 2.9 – Typical sky imaging processing flow. ....	74
Figure 2.10 – Satellite images processing flow. ....	75
Figure 2.11 – Expected number of tap-changing operations and ESU throughput with respect to time. ....	79
Figure 2.12 – Irradiance vs. time and corresponding ESU state vs. time. A state of +1 indicates discharging, -1 indicates charging, and 0 indicates standby [12].....	79
Figure 3.1 – Flowchart of an ensemble forecasting method. ....	83
Figure 3.2 – Incidence angle of direct sunlight on a fixed tilt and fixed orientation PV collector [4].	88
Figure 3.3 – Empirical GHI model curves compared with measured GHI every minute on 21/10/2015. ....	100
Figure 3.4 – Normalised GHI in function of the zenith angle $z$ , used for the initial Dazhi-based regression.....	101
Figure 3.5 – Screenshot of the curve-fitting toolbox used for the initial Dazhi-based regression. ....	102
Figure 3.6 – Normalised GHI in function of the zenith angle $z$ , used for the Initial TJ-based regression.....	103
Figure 3.7 – Screenshot of the curve-fitting toolbox used for the initial TJ-based regression. ....	104

Figure 3.8 – Normalised GHI in function of the apparent zenith angle $z$ , used for the refined Dazhi-based regression. ....	106
Figure 3.9 – Screenshot of the curve-fitting toolbox used for the refined Dazhi-based regression. ...	107
Figure 3.10 – Normalised GHI in function of the apparent zenith angle $z$ , used for the refined Trelkeld-Jordan model--based regression. ....	107
Figure 3.11 – Screenshot of the curve-fitting toolbox used for the refined TJ-based regression. ....	108
Figure 3.12 – First regression GHI model curves compared with measured GHI every minute on 21/10/2015. ....	110
Figure 3.13 – Regression GHI model curves compared with measured GHI every minute on 21/10/2015, based on the extended 88-day dataset. ....	111
Figure 3.14 – Cartesian and spherical coordinates of a point object in the sky. ....	112
Figure 3.15 – Cartesian and spherical coordinates of a point object in the sky. ....	113
Figure 3.16 – Projection of a specific point object on the camera sensor. ....	113
Figure 3.17 – A sensor pixel of a Nikon D7100 DSLR camera. ....	115
Figure 3.18 – Typical projection of a circular fisheye lens on a sensor. ....	116
Figure 3.19 – Projection with square shaped pixels (a) and with rectangular shaped pixels (b). ....	116
Figure 3.20 – Position of the projections of point objects on the frame according to their angle $\theta$ with respect to the optical axis. ....	117
Figure 3.21 – Illustration of the roll, pitch, and yaw of a sky camera (a) and their respective deviations on their respective planes (b). ....	118
Figure 3.22 – A 50-km wide mesh in the sky at 3000 m of altitude above the camera. ....	118
Figure 3.23 – Same mesh seen through a 4.5 mm equisolid circular fisheye lens. ....	119
Figure 3.24 – Cartesian and spherical coordinates of a parallelogram object ABCD in the sky. ....	120
Figure 3.25 – Polar coordinates of the image of the parallelogram object on a ground camera sensor. ....	121
Figure 3.26 – Shadow model of a couple of clouds in a sky frame. ....	123
Figure 3.27 – Image referential frame. ....	124
Figure 3.28 – Geometrical approximate method to locate the centre of the projected image. ....	125
Figure 3.29 – Orientation of the reference checkerboard with respect to the camera. ....	126
Figure 3.30 – Camera calibration procedure. ....	126
Figure 3.31 – Retroprojection function of the lens. ....	128
Figure 3.32 – Error of the retroprojection function along the X and Y axes. ....	128
Figure 3.33 – Configuration of an ideal camera and the projected image on the sensor. ....	129
Figure 3.34 – Configuration of a real camera and the projected image on the sensor. ....	130
Figure 3.35 – Cartesian and spherical coordinates of a given point $M'$ within the 3D unit frame $F'$ . ..	131
Figure 3.36 – Cartesian and spherical coordinates of a given point $M$ within the 3D unit frames $F$ and $F'$ . ....	131

Figure 3.37 – Simple rotation of the unit frame around the $w'$ axis by an angle $\alpha$ as seen in the $u'-v'$ plane.....	132
Figure 3.38 – Simple rotation of the unit frame around the $u'$ axis by an angle $\beta$ as seen in the $v'-w'$ plane.....	132
Figure 3.39 – Angular radius of the sun as seen from the earth.....	133
Figure 3.40 – Angular position and angular radius of the sun in the sky.....	134
Figure 3.41 – Position of the projected solar disk on the camera sensor. ....	135
Figure 3.42 – Clear sky picture showing the position mismatch between predicted and actual projection of the centre of the solar disk.....	136
Figure 3.43 – Distance error between the expected projected position of the centre of the solar disk on the camera sensor and its actual position, expressed in pixels, before and after 3D matrix correction. ....	138
Figure 3.44 – Expected projected position of the centre of the solar disk and its actual position on the camera sensor on 28/11/2016 for all the studied hours before (a) and after correction (b). ....	138
Figure 3.45 – Scatter plot of the distance error between the expected and actual centre of the solar disk in function of the number of sky images, with asymptotes.....	140
Figure 3.46 – Example of a position for a cloud in satellite view.....	140
Figure 3.47 – Same cloud seen by the camera with the solar disk shown in yellow. ....	141
Figure 3.48 – First example of cloud tracking and sun cover measurement.....	141
Figure 3.49 – Second example of cloud tracking and sun cover measurement. ....	142
Figure 3.50 – Third example of cloud tracking and sun cover measurement. ....	142
Figure 3.51 – Example of an image of the sky simplified with Photoshop. ....	143
Figure 3.52 – Bonding boxes enclosing binary shapes. ....	143
Figure 3.53 – Shading episode as seen through a sky camera. ....	144
Figure 3.54 – Current position and predicted positions of a single cloud bonding box.....	145
Figure 3.55 – Direct application of the bonding box and centroid forecasting approach and the feature approach in the camera sensor referential.....	146
Figure 3.56 – Example of a reprocessed sky image taken on 28/01/2017 at the GEEP lab at 9:23 am. ....	146
Figure 3.57 – Retro-projection of the reprocessed sky image taken on 28/01/2017 at the GEEP lab at 9:23 am onto the unit sphere, unit altitude plane and actual altitude plane, within the actual 3D world. ....	147
Figure 3.58 – Comparison the bounding boxes within the camera sensor domain (a) and within the unit altitude plane.....	147
Figure 3.59 – Shading Model in $(y, z)$ two-dimensional sky plane or side view of a shading episode. ....	148

Figure 3.60 – Shading model in (x, y) two-dimensional sky plane or top view of a shading episode. .....	149
Figure 3.61 – Shading model as seen through a ground camera of a shading episode. ....	149
Figure 3.62 – Model of an ideal cloud with a perfectly uniform optical thickness and sides. ....	150
Figure 3.63 – A model corresponding to realistic shading conditions. ....	151
Figure 3.64 – Reconstructed vertical cross-section of the cloud passing over the monocrystalline PV array, the times when each section of the clouds passed over the array and the corresponding optical thicknesses. ....	152
Figure 3.65 – Optical depth of the detected small passing cloud. ....	152
Figure 3.66 – Measured and simulated beam intensities. ....	153
Figure 3.67 – 3D unit sphere (a) and 2D projection (b) featuring the sun and a pixel located in a 20 degrees periphery of the sun on 25/01/2017 at 9:19:41 am. ....	154
Figure 3.68 – 3D unit sphere showing the original cone and the rotated cone. ....	155
Figure 3.69 – 3D unit sphere showing the position of the sun and all of the pixels within 10 degrees around it on 25/01/2017 at 7:19:41 am. ....	156
Figure 3.70 – 3D unit sphere showing the position of the sun and all of the pixels within 10 degrees around it on 25/01/2017 at 7:19:41 am. ....	156
Figure 3.71 – 3D positions of the sun every two hours during daytime on 25/01/2017. ....	157
Figure 3.72 – Projected 2D positions of the sun every two hours during daytime on 25/01/2017. ....	158
Figure 4.1 – Edition of a sky image with Photoshop. ....	162
Figure 4.2 – View of the layers stacked to reconstruct an image. ....	163
Figure 4.3 – Various pixel classes. ....	163
Figure 4.4 – Example of a few images from the dataset at different instants k. ....	164
Figure 4.5 – Number of occurrences of each red value (a), green value (b), and blue value (c) for the 'clear-sky pixels' class (in blue) and the 'cloudy pixels' class (in red) of a given sky image. ....	165
Figure 4.6 – Flow chart of the pixel discrimination assessment method. ....	166
Figure 4.7 – Reference clear-sky image (a), masked reference clear-sky image (b), and reference cloud layer (c). ....	168
Figure 4.8 – Comparison between the reference sky (a) and the simple Bayes pixel criterion for sky pixels (b) and between the reference cloud (c) and the simple Bayes pixel criterion for clouds (d). .	169
Figure 4.9 – Number of occurrences of each R/B value (a), R-B value (b), and B-G value (c). ....	170
Figure 4.10 – Flowchart of the implementation of a PHD filter. ....	176
Figure 4.11 – Movement of all extracted features over time. ....	180
Figure 4.12 – Position of the feature 35 over time (a) and after linear regression (b). ....	181
Figure 4.13 – Cloud Signal along regression line for feature #35 at step 28. ....	181
Figure 5.1 – Comparison between pictures of the sky obtained with the AllSkyCam (a) and with the Nikon D7100 (b) at the exact same instant (24/02/2016, 1:22:33 pm) a few metres apart. ....	187

Figure 5.2 – View of the camera setup among other installations at the GEEP lab. ....	188
Figure 5.3 – View of the camera setup mounted on its pole.....	188
Figure 5.4 – Labview-based GEEP client software. ....	189
Figure 5.5 – Overview of the main window of the enhanced DigiCamControl software version 2.0.21.0. ....	190
Figure 5.6 – Time lapse window of the DigiCamControl software.....	190
Figure 5.7 – Original clear-sky image. ....	192
Figure 5.8 – Fully masked clear-sky image. ....	192
Figure 5.9 – Original image containing cumulus clouds. ....	193
Figure 5.10 – Masked cumulus cloud image. ....	193
Figure 5.11 – Recomposed complete masked image.....	193
Figure 5.12 – Chromatic analysis of the R component of clear-sky and cloudy pixels.....	194
Figure 5.13 – Chromatic analysis of the G component of clear-sky and cloudy pixels.....	195
Figure 5.14 – Chromatic analysis of the B Component of clear-sky and cloudy pixels.....	195
Figure 5.15 – B-R chromatic analysis.....	196
Figure 5.16 – B-G chromatic analysis. ....	197
Figure 5.17 – G-R chromatic analysis. ....	197
Figure 5.18 – R/B chromatic analysis.....	198
Figure 5.19 – Algorithm applied to calculate the error in clear-sky pixel detection.....	198
Figure 5.20 – Reference binary sky image featuring clear-sky pixels only.....	199
Figure 5.21 – Clear-sky pixels detected with a Naive Bayes classifier. ....	199
Figure 5.22 – Clear-sky pixels detected with the R/B criterion. ....	200
Figure 5.23 – Clear-sky pixels detected with the B-R criterion.....	200
Figure 5.24 – Clear-sky pixels detected with the multicolour criterion.....	200
Figure 5.25 – Graphical representation of the cloud type frequency distribution for the period under consideration.....	203
Figure 5.26 – Measured and clear-sky GHI on 2/05/2016.....	209
Figure 5.27 – Optical thickness on 2/05/2016. ....	209
Figure 5.28 – Power output on 2/05/2016.....	209
Figure 5.29 – Measured and clear-sky GHI on 19/01/2017.....	210
Figure 5.30 – Optical thickness on 19/01/2017. ....	210
Figure 5.31 – PV Power output on 19/01/2017. ....	210
Figure 5.32 – View of the sky on 19/01/2017 at 9:59 am.....	211
Figure 5.33 – Measured and clear-sky GHI on 17/03/2017.....	212
Figure 5.34 – Optical thickness on 17/03/2017. ....	212
Figure 5.35 – PV Power output on 17/03/2017. ....	212
Figure 5.36 – Measured and clear-sky GHI on 19/05/2016.....	214

Figure 5.37 – Optical thickness on 19/05/2016. ....	214
Figure 5.38 – PV Power output on 19/05/2016. ....	214
Figure 5.39 – Measured and clear-sky GHI on 26/06/2016.....	216
Figure 5.40 – Optical thickness on 26/06/2016. ....	216
Figure 5.41 – PV power output on 26/06/2016.....	216
Figure 5.42 – Measured and clear-sky GHI on 21/04/2016.....	218
Figure 5.43 – Optical thickness on 21/04/2016. ....	218
Figure 5.44 – PV power output on 21/04/2016.....	218
Figure 5.45 – Measured and clear-sky GHI on 10/02/2017.....	219
Figure 5.46 – Optical thickness on 10/02/2017. ....	220
Figure 5.47 – PV power output on 10/02/2017.....	220
Figure 6.1 – Diagram of the GEEP lab 1.5kW monocrystalline PV system.....	223
Figure 6.2 – MATLAB SIMULINK model of the GEEP lab 1.5kW Monocrystalline PV system....	223
Figure 6.3 – DNI and output signals of the 1.5kW monocrystalline array from the GEEP measured on 19/04/2018. ....	224
Figure 6.4 – Linear regression model for the simplified irradiance-current relationship. ....	225
Figure 6.5 – First comparison between the simulated and measured output current (a) and voltage (b). ....	225
Figure 6.6 – Second comparison between the simulated and measured output current (a) and voltage (b).....	226
Figure 6.7 – Diagram of the GEEP lab 1.5kW monocrystalline PV system with its target voltage levels. ....	226
Figure 6.8 – Evolution of the DNI over time when the sunlight dimming factor is 10%. ....	227
Figure 6.9 – Evolution of the voltage at every link over time when the sunlight dimming factor is 10%. ....	228
Figure 6.10 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 10%. ....	228
Figure 6.11 – Evolution of the current at every link over time when the sunlight dimming factor is 10%. ....	229
Figure 6.12 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 10%. ....	229
Figure 6.13 – Evolution of the active power at every link over time when the sunlight dimming factor is 10%. ....	230
Figure 6.14 – Evolution of the DNI over time when the sunlight dimming factor is 20%. ....	231
Figure 6.15 – Evolution of the voltage at every link over time when the sunlight dimming factor is 20%. ....	231

Figure 6.16 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 20%. ..... 232

Figure 6.17 – Evolution of the current at every link over time when the sunlight dimming factor is 20%. ..... 232

Figure 6.18 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 20%. ..... 233

Figure 6.19 – Evolution of the active power at every link over time when the sunlight dimming factor is 20%. ..... 233

Figure 6.20 – Evolution of the voltage at every link over time when the sunlight dimming factor is 30%. ..... 234

Figure 6.21 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 30%. ..... 235

Figure 6.22 – Evolution of the current at every link over time when the sunlight dimming factor is 30%. ..... 235

Figure 6.23 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 30%. ..... 236

Figure 6.24 – Evolution of the active power at every link over time when the sunlight dimming factor is 30%. ..... 236

Figure 6.25 – Evolution of the voltage at every link over time when the sunlight dimming factor is 40%. ..... 237

Figure 6.26 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 40%. ..... 238

Figure 6.27 – Evolution of the current at every link over time when the sunlight dimming factor is 40%. ..... 238

Figure 6.28 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 40%. ..... 239

Figure 6.29 – Evolution of the active power at every link over time when the sunlight dimming factor is 40%. ..... 239

Figure 6.30 – Evolution of the voltage at every link over time when the sunlight dimming factor is 50%. ..... 240

Figure 6.31 – Evolution of the reactive power at the AC Links over time when the sunlight dimming factor is 50%. ..... 241

Figure 6.32 – Evolution of the current at every link over time when the sunlight dimming factor is 50%. ..... 241

Figure 6.33 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 50%. ..... 242

Figure 6.34 – Evolution of the active power at every link over time when the sunlight dimming factor is 50%. .....	242
Figure 6.35 – Evolution of the voltage at every link over time when the sunlight dimming factor is 60%. .....	243
Figure 6.36 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 60%. .....	244
Figure 6.37 – Evolution of the current at every link over time when the sunlight dimming factor is 60%. .....	244
Figure 6.38 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 60%. .....	245
Figure 6.39 – Evolution of the active power at every link over time when the sunlight dimming factor is 60%. .....	245
Figure 6.40 – Evolution of the voltage at every link over time when the sunlight dimming factor is 70%. .....	246
Figure 6.41 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 70%. .....	247
Figure 6.42 – Evolution of the current at every link over time when the sunlight dimming factor is 70%. .....	247
Figure 6.43 – Evolution of the power factor at the AC Links over time when the sunlight dimming factor is 70%. .....	248
Figure 6.44 – Evolution of the active power at every link over time when the sunlight dimming factor is 70%. .....	248
Figure 6.45 – Evolution of the voltage at every link over time when the sunlight dimming factor is 80%. .....	249
Figure 6.46 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 80%. .....	250
Figure 6.47 – Evolution of the current at every link over time when the sunlight dimming factor is 80%. .....	250
Figure 6.48 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 80%. .....	251
Figure 6.49 – Evolution of the active power at every link over time when the sunlight dimming factor is 80%. .....	251
Figure 6.50 – Evolution of the voltage at every link over time when the sunlight dimming factor is 90%. .....	252
Figure 6.51 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 90%. .....	253



Figure 6.52 – Evolution of the current at every link over time when the sunlight dimming factor is 90%.	253
Figure 6.53 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 90%.	254
Figure 6.54 – Evolution of the active power at every link over time when the sunlight dimming factor is 90%.	254
Figure 6.55 – P-V (a) and I-V (b) characteristics of the PSCAD PV array model when receiving a DNI of 1000W/m <sup>2</sup> and operating at 25°C.	256
Figure 6.56 – Partial view of the grid-connected PV system.	257
Figure 6.57 – Evolution of the signals at the PV output over time when the sunlight dimming factor is 90%.	258
Figure 6.58 – Evolution of the signals at the DC link over time when the sunlight dimming factor is 90%.	258
Figure 6.59 – Evolution of the signals at the VSC output over time when the sunlight dimming factor is 90%.	259
Figure 6.60 – Evolution of the signals at the RL feeder output over time when the sunlight dimming factor is 90%.	259
Figure 6.61 – Evolution of the signals at the transformer output over time when the sunlight dimming factor is 90%.	260
Figure 6.62 – Accurate electrical battery model as proposed in [168, 169].	261
Figure 6.63 – Top view of the MATLAB/SIMULINK battery bank model.	261
Figure 6.64 – View of the power part of the MATLAB/SIMULINK battery bank model.	262
Figure 6.65 – State of charge of the battery bank.	263
Figure 6.66 – Open circuit voltage of one single cell in the battery bank.	263
Figure 6.67 – Output voltage of one single cell in the battery bank.	264
Figure 6.68 – Output voltage of the complete battery bank.	264
Figure 6.69 – Diagram of the Dual Active Bridge (DAB) envisaged for this study.	265
Figure 6.70 – PSCAD DAB model connected to a utility grid block with a local load.	268
Figure 6.71 – DC link power when the DAB supplies power to AC grid.	269
Figure 6.72 – DAB voltage reference $V_{ref}$ and actual output voltage $V_{dc}$ when the DAB supplies power to AC grid.	269
Figure 6.73 – Grid Power when the DAB supplies power to AC grid.	270
Figure 6.74 – DC link power when the DAB absorbs some power from the grid.	270
Figure 6.75 – DAB voltage reference $V_{ref}$ and actual output voltage $V_{dc}$ when the DAB absorbs some power from the grid.	271
Figure 6.76 – Grid power when the DAB absorbs some power from the grid.	271

Figure 6.77 – Complete grid-connected PV system with battery and bidirectional DC-DC converter. .....	272
Figure 6.78 – Diagram of the overall system with closed-loop controlled three-phase VSC.....	273
Figure 6.79 – Characterisation of the current transducer LEM HAIS 50-P.....	274
Figure 6.80 – Characterisation of the voltage transducer LEM LV 25-P. ....	274
Figure 6.81 – PSPICE schematic and DC analysis of the level shifter for voltage sensors.....	275
Figure 6.82 – KiCad schematic of the level shifter for voltage sensors.....	276
Figure 6.83 – KiCad layout of the level shifter for voltage sensors. ....	276
Figure 6.84 – PSPICE schematic and DC analysis of the level shifter for the current sensors. ....	277
Figure 6.85 – KiCad schematic of the level shifter for current sensors.....	277
Figure 6.86 – KiCad layout of the level shifter for current sensors. ....	278
Figure 6.87 – Complete overall architecture of the VSC controller. ....	279
Figure 6.88 – Example of reference signal generated by the signal generator block.....	280
Figure 6.89 – Output flag signal corresponding to the reference generated signal given in Figure 6.88. .....	281
Figure 6.90 – Example of result obtained with the integration block. ....	282
Figure 6.91 – Input signal used to test the DFT block. ....	284
Figure 6.92 – Sine modulated input signal. ....	286
Figure 6.93 – Retrieved high frequency component of the sine modulated input signal.....	286
Figure B.1 – View of the compartment enclosing the power supply of the camera. ....	298
Figure B.2 – View of the camera installed in its enclosure box with power and data cables. ....	298
Figure B.3 – Camera box closed and ready to be installed.....	299
Figure B.4 – View of the cabinet containing the computer and the hard drive. ....	299
Figure B.5 – Complete camera setup. ....	300

## LIST OF TABLES

---

Table 2.1 – Wind speed/power forecasting time frames [33, 39, 52].	38
Table 2.2 – Popular methods used for wind speed/power forecasting [39].	39
Table 2.3 – One possible definition of the various horizons for PV power forecasting [61].	44
Table 2.4 – Alternative definition of the various horizons for PV power forecasting [29].	45
Table 2.5 – nRMSE for different forecasting methods. Data collected by Yuehui (2010) [61].	45
Table 2.6 – Various empirical models available in literature to evaluate PV cell temperatures, as listed by Jakhrani et al (2011) [32] and Ciulla et al (2013) [22].	50
Table 2.7 – Mean nominal conversion rate $\eta_0$ by technology [80].	52
Table 2.8 – Possible reference sources of data for PV forecasting.	57
Table 2.9 – Approximate heights of cloud levels.	60
Table 2.10 – Cloud classification for PV forecasting purposes.	60
Table 2.11 – Cloud classification by optical thickness and top pressure.	61
Table 2.12 – Relevant features to process images from sky cameras based on the works of Heinle (2010) [24] and Katzanizidis (2012) [25].	66
Table 2.13 – Additional relevant image processing parameters [25].	69
Table 2.14 – Additional sub-classes [25].	69
Table 2.15 – Simple weather classification based on a textual forecast [49].	71
Table 3.1 – Values recommended by the NOAA (2004) for the atmospheric refraction correction [17].	85
Table 3.2 – Measurement dates of the GHI and ambient temperature.	97
Table 3.3 – Average values of the accuracy metrics for some existing models.	98
Table 3.4 – Second set of measurement dates for the GHI and ambient temperature.	105
Table 3.5 – Average values of the accuracy metrics for the refined regression models.	108
Table 3.6 – Ranking of the top four most accurate empirical models overall over 16 days.	109
Table 3.7 – Ranking of the top four most accurate empirical models.	110
Table 3.8 – Fisheye lens projection types and their respective mapping functions.	113
Table 3.9 – Distance error between the predicted and the actual position of the sun on four different days, expressed in pixels, before and after application of each respective 3D rotation matrix.	137
Table 3.10 – Investigating the optimal time resolution to calculate the 3D rotation matrix on 25/01/2017.	139
Table 3.11 – Investigating the optimal number of sky images to calculate the 3D rotation matrix on 25/01/2017 with a 30 minutes time resolution.	139
Table 4.1 – Criteria used for pixel discrimination.	171
Table 4.2 – Indicative optimal threshold values for a minimal error.	171

Table 4.3 – Relative prediction errors against the benchmark.....	183
Table 5.1 – Chromatic range of clear-sky and cloudy Images.....	194
Table 5.2 – Clear-sky pixel discrimination criteria under consideration.....	196
Table 5.3 – Optimal threshold values.....	199
Table 5.4 – Best relative errors corresponding to the optimal threshold values.....	201
Table 5.5 – Description of the clear-sky dataset.....	201
Table 5.6 – Average values of features for the circumsolar region and for the extrasolar region within the dataset.....	202
Table 5.7 – Cloud type frequency for the period under consideration.....	203
Table 5.8 – Cloud feature characterisation of sky images – red (R) and green (G) features.....	204
Table 5.9 – Continuation of the cloud feature characterisation of sky images from the Nikon D7100 camera – blue (B) feature.....	205
Table 5.10 – Cloud feature characterisation of sky images from the Nikon D7100 camera – other features.....	206
Table 5.11 – Observations made on cumulus clouds.....	208
Table 5.12 – Observations made on stratocumulus clouds.....	211
Table 5.13 – Observations made on stratus and altostratus clouds – First part.....	213
Table 5.14 – Observations made on stratus and altostratus clouds – Second part.....	213
Table 5.15 – Observations made on cirrocumulus and altocumulus clouds – First part.....	215
Table 5.16 – Observations made on cirrocumulus and altocumulus clouds – Second part.....	215
Table 5.17 – Observations made on cirrus and cirrostratus clouds – First part.....	217
Table 5.18 – Observations made on cirrus and cirrostratus clouds – Second part.....	217
Table 5.19 – Observations made on cumulonimbus and nimbostratus clouds – First part.....	219
Table 5.20 – Observations made on cumulonimbus and nimbostratus clouds – Second part.....	219
Table 5.21 – Observations made on all cloud types across the whole dataset.....	220
Table 6.1 – Approximate target voltage levels for the readjusted Simulink model.....	226
Table 6.2 – Summary of the settled signal values of the Simulink PV model for each final value when $T_{cell} = 35^{\circ}\text{C}$ .....	255
Table 6.3 – Summary of the settled signal values of the PSCAD PV model for various DNI values when $T_{cell} = 35^{\circ}\text{C}$ .....	257
Table A.1 – Threshold values that have been found to minimize cloud detection error for every image and in average.....	292
Table A.2 – Time needed to evaluate each criterion based on the picture of every step.....	293
Table A.3 – Relative classification error expressed as a percentage (%) for each criteria.....	294
Table A.4 – Actual and estimated total cloud cover (TCC) expressed as a percentage (%) for each image of the dataset.....	295

Table A.5 – Relative error in the estimation of the total cloud cover by each criterion, expressed as a percentage. .... 296

Table A.6 – Execution Time for each criterion and each image, expressed in seconds. .... 297

## ACKNOWLEDGEMENTS

---

I would like to express my sincerest gratitude to my thesis supervisor and advisor, Prof Arindam Ghosh. Words are insufficient to express how much I owe him. I am extremely grateful to him for helping and trusting me the way he did.

I would like to thank my co-supervisor, Associate Prof Sumedha Rajakaruna, for his continuous support and for giving me access to some useful equipment.

I am grateful to my other co-supervisor, Prof Ba-Tuong Vo, along with his brother Prof Ba-Ngu Vo, for sharing their expertise on filtering and statistical methods.

I also thank Arthur K. Barnes, Juan Carlos Barda, Jonathan.K. Hayes, Paolo Adesso, Roberto Conte, Maurizio Longo, Rocco Restaino, Gemine Vivone, Anna Heinle, Andreas Macke, Panagiotis Tzoumanikas, Andreas Kazantzidis, Cai Chengrui, John Pye, José Zapata, Jan Kleissl, Ben Kurtz, Duka Istvan and Christian Gueymard for their kind cooperation and their assistance.

I would to express my gratitude to Gilbert M. Masters for his courtesy.

I thank Curtin University and in particular the department of Electrical and Computing Engineering for their continuous assistance. In particular, I am very grateful to our former head of Department, Dr Yee-Hong Leung, for believing in my potential and for supporting me. I also acknowledge the financial support of the Australian Research Council (ARC) for providing me with financial support through the Discovery Grant DP110104554.

I thank all of my friends and my current and former fellow higher degree by research (HDR) students, in particular Linh Thi Thuc Tran, Ehsan Pashajavid, Amit Datta and Dilum Hettiarachchi.

I would like to express some very special thanks to Corentin Riffart and Muaz Rahman for their help.

I also very grateful to my spouse, Stephanie, for her patience and all her other natural qualities.

Finally, last but not least, I would like to thank Margaret Pittuck, Robyn Cornwell. I also wish to thank the technical team for the outstanding quality of their work and their amazing support, namely Mark Fowler, Nick King, Zibby Cielma and Russell Wilkinson.

## ABSTRACT

---

An important bottleneck to the large scale development of utility-scale photovoltaic farms is the output power quality issues due to swift cloud cover changes. Cloud-induced sunlight disruption may cause significant power output fluctuations in solar power systems, jeopardizing power quality. Forecasting the power output with a very short horizon below 30 seconds can be seen as a cloud cover forecasting problem solved by processing images from a sky camera, which may involve classifying pixels and calculating a 3D rotation matrix. After a pre-processing stage that identifies clouds with a criterion based on the red-green blue (RGB) values of each pixel, a Probability Hypothesis Density (PHD) filter and a more advanced Cardinalised Probability Hypothesis Density (CPHD) filter can be used to an unknown and varying number of clouds. Both filters have been compared for performance by using an original centroid-based cloud tracking method and both have also been compared to the state of the art optical flow. It has been found that both filters are suitable for very-short term irradiance forecast and are fast enough for implementation. Both filters have also proven to be more appropriate for very-short forecast than optical flow.

Also, some cloud types have been identified as causing irradiance peaks on top of shading. After a detailed study of the effects of each cloud type on the global horizontal irradiance, it has been observed that rain clouds cause the most dramatic shading, for the longest period of time. However, cumulus and stratocumulus clouds have been identified as responsible for the most intense irradiance peaks, whereas cirrus and cirrostratus cause the smallest irradiance peaks.

To mitigate the negative effects of both shading and peaks of irradiance on the consistence and quality of photovoltaic output, a battery system with bidirectional converter with a proper controller and integration of the photovoltaic power has been envisaged and modelled.

## STATEMENT OF ATTRIBUTION BY OTHERS

---

The purpose of this statement is to summarise and clearly identify the nature and extent of the intellectual input provided by the candidate and any co-authors.

Article 1 title: Very short-term photovoltaic power forecasting with cloud modeling: A review

Article 2 title: A Comparative Study of Clear-Sky Irradiance Models for Western Australia

Article 3 title: Intra-Hour Cloud Tracking Based on Probability Hypothesis Density Filtering

Attribution of Prof. Arindam Ghosh:

Article	Literature review	Concept design	Software modelling	Experimental results	Data analysis	Discussion	Paper writing	Thesis Chapter
1		✓				✓	✓	2
2		✓				✓	✓	3
3		✓				✓	✓	4

Attribution of A/Prof. Sumedha Rajakaruna:

Article	Literature review	Concept design	Software modelling	Experimental results	Data analysis	Discussion	Paper writing	Thesis Chapter
1		✓				✓		2
2		✓				✓		3
3		✓				✓		4

Attribution of Prof. Ba-Tuong Vo:

Article	Literature review	Concept design	Software modelling	Experimental results	Data analysis	Discussion	Paper writing	Thesis Chapter
2						✓		3
3		✓	✓			✓		4

Attribution of Corentin Riffart:

Article	Literature review	Concept design	Software modelling	Experimental results	Data analysis	Discussion	Paper writing	Thesis Chapter
2						✓		3
3			✓			✓		4



[Student]



[Supervisor]



# NOMENCLATURE

---

$GHI$	Global Horizontal Irradiance.
$DNI$	Direct Normal Irradiance.
$NRMSE$	Normalised Root Mean Square Error.
$MAE$	Mean Absolute Error.
$MBE$	Mean Bias Error.
$RGB$	red green blue.
$PV$	Photovoltaic.
$PHD$	Probability Hypothesis Density.
$CPHD$	Cardinalised Probability Hypothesis Density.
$\rho$	Pearson's Correlation Factor.
$k$	Discrete time.
$\Delta k$	Discrete period of time.
$\zeta$	Previous state.
$x$	Current state.
$Z_k$	Measurement.
$J_k$	Total number of targets at time $k$ .
$J_{max}$	Maximal number of targets.
$i$	A specific single target among the total of $J_k$ targets.
$w_k^{(i)}$	Weight
$m_k^{(i)}$	Mean value of the cloud's centroid coordinates
$P_k^{(i)}$	Covariance
$v$	Intensity
$p_{nD}$	Probability of non detection
$F_k$	Transformation matrix.

$f_{k k-1}(x \zeta)$	Single target transition density.
$p_{S,k}(\zeta)$	Probability of target existence.
$p_{S,k}$	Probability of target survival at time $k$ .
$\gamma_k(k)$	Intensity of spontaneous births.
$p_{\Gamma,k}(k)$	Cardinality distribution of births.
$g_k(z x)$	Single target measurement likelihood.
$p_{D,k}(x)$	Probability of the target detection.
$K_k(z)$	Measurement of clutter intensity.
$P_{K,k}(z)$	Cardinality distribution of clutter.

# 1 INTRODUCTION AND PROBLEMATIC

---

## 1.1 MAIN OBJECTIVES OF THIS STUDY

As of now, photovoltaic (PV) resources are deemed non-dispatchable due to the highly stochastic nature of sunlight. This issue greatly limits the development of PV installations and deters investments on a global scale. This thesis endeavours to address this problem by controlling and improving the quality of PV power output by forecasting that power systems within a very short-term timeframe in the presence of passing clouds. An exhaustive methodology has been investigated in that regard, with an initial limitation to a small geographic area, corresponding to a single PV array. Prior to being extended to a larger area, the methodology in question seeks to tackle the following goals:

Firstly, an appropriate forecasting horizon in terms of temporal and spatial resolution is selected based on the information gathered through an exhaustive investigation and comparison of the various approaches and techniques used in literature to forecast PV power in very short-term. A suitable approach and an appropriate source of data also need to be chosen.

Secondly, the clear-sky irradiance needs to be modelled to match with experimental observations. The purpose is to forecast the global horizontal irradiance for a given place and time in the complete absence of clouds and aerosols with a satisfactory accuracy prior to superposing the effects of clouds, for a realistic and exhaustive real-time forecast.

Thirdly, it is necessary to pre-process the images from a sky camera in order to tell apart the pixels corresponding to clear sky from the pixels corresponding to clouds. Empirically, the optimal criterion depends heavily on the camera sensor that is used. It is thus necessary to define a process to determine what criterion is best for a specific sky camera.

Finally, the last objective is to track the position of clouds with respect to the sun within a series of sky images with an accurate and time-efficient statistical method. The position of the sun needs to be predicted in every frame and a comparison of several tracking methods is necessary at this stage.

## 1.2 CONTRIBUTIONS OF THIS STUDY

In an effort to fulfil the objectives mentioned in the previous section, this work provides the reader with the following original contributions:

- An exhaustive literature review on short-term and very-short photovoltaic forecasts, with a comparison of both tools and methods as well as some recommendations. This literature review does not only cover the sole needs of this study but also provides a solid guidance to other researchers seeking to work in a similar area. An article has been published in a reputable peer-reviewed journal from Elsevier in that regard [1].
- A comparison of some commonly and widely used empirical have been compared to 1-minute measurements of the global horizontal irradiance (GHI) made with an affordable weather station. A linear regression has been made on the basis of the two models that were the closest fits with respect to the measured values, leading to the design of two new models that specific fit insolation conditions of Western Australia. This work has been presented during an IEEE Power and Energy Society General Meeting Conference [2].
- An original method has been designed to pre-process the images supplied by a sky camera. In particular, this method is used to assess the accuracy of various clear sky/cloudy pixel discrimination criteria. It also finds the optimal thresholds for a specific camera sensor. As previously mentioned, this is particularly important since results may vary dramatically from a camera sensor to another. The method in question has been introduced in an article published in the peer-reviewed IEEE Transactions on Sustainable Energy [3].
- Identifying cloud types and studying their effect on the reduction of light intensity when they mask the sun.
- Defining a custom method to identify the tilt angle and the deviation of the camera with respect to the true North. This step is absolutely necessary to determine accurately the position of the centre of the solar disk over time and thus forecast sunlight dimming.
- Tracking cumulus clouds within a series of sky images with an accurate and time-efficient statistical method. A few methods have been compared in speed and accuracy [3].
- An example of practical utilisation of the forecast to control photovoltaic power output.

### **1.3 ADVANTAGES OF PV GENERATION**

Solar power is largely underexploited globally. However, the sun can be seen as a gigantic natural fusion nuclear reactor [4] that provides the Earth with far more energy that humanity needs [5][6] or will most probably ever need in the future [7]. It is a completely free and clean source of energy since it does not need any maintenance nor any fuel supply and it does not produce any waste. The solar resource is also at the origin of many other sustainable resources like wind energy [8]. It is even reported that it is the first source of energy of the planet, exceeding the second largest energy resource by a factor 104 [9]. Also, the Sun will continue to send a steady quantity of solar energy to the surface of the Earth for a period of time stretching far beyond the scale of human history. All of these factors make solar

energy an excellent candidate to tackle economic, environmental and public health challenges of our developing energy crisis.

## **1.4 FAST INCREASING POPULARITY OF PV GENERATION**

PV generation is advantageous and valuable in various markets, as its peak generation matches with the timeframe of higher load demand. M. Morjaria et al. [10] mention that the cost of PV generation has significantly plummeted, increasing greatly the competitiveness of PV power. These authors have even seen that PV-generated electricity has reached grid parity in a large variety of markets, with a price equal or lower with respect to electromechanical non-renewable electricity. As a result, the number and the size of utility-scale PV farms have skyrocketed. As of today, large capacity utility-scale grid-connected PV farms already account for 38% of the overall PV generation [11]. Manz et al. [12] have reported that some solar power plants are “planned to exceed several hundred megawatts of capacity”. To illustrate that growth, the global capacity of utility-scale photovoltaic (PV) plants has exceeded an estimated value of 60 GW<sub>AC</sub> by the end of 2015 with an extra 87 GW<sub>AC</sub> capacity commissioned and in development ever since then [13]. Australia alone accounts for 5.4 GW of PV capacity as of September 2016, part of which 692 MW has been installed in the preceding 12 months [14].

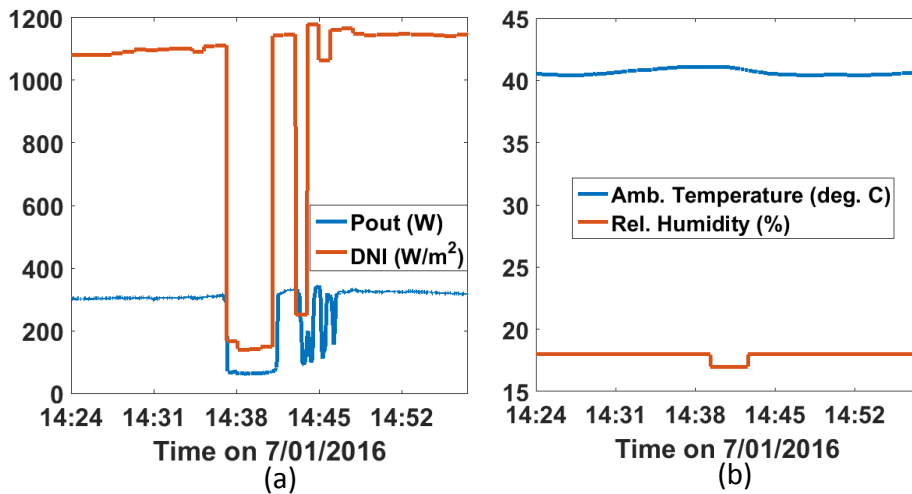
## **1.5 CHALLENGES OF PV GENERATION**

Nonetheless, despite its increasing popularity, solar power generation is considered unreliable because of its direct dependence on stochastic weather conditions, which raises grid integration concerns [15] in particular because of the difficulty to dispatch that energy. As a matter of fact, it is important to smoothen the PV power output prior to its integration into the grid, for quality purposes [10]. This may be achieved through the adequate control of a battery energy storage system (BESS) [16] and inverters [17-19]. However, battery banks may be costly and ancillary generators require some time to ramp up if they are needed. Therefore, a forecast of PV power fluctuations helps to minimize costs and optimize efficiency.

In particular, for the very large number of small islanded systems in regional Western Australia, power continuity in local power generation is vital. Such systems need to be autonomous, which means that intermittency cannot be mitigated with a connection to another system. Instead, local reserves must be managed in an optimized manner. This means such power systems cannot afford to rely on a large share of solar power if they do not have enough visibility over future generation [20][21][22]. An accurate very-short term irradiance forecasting method can play a key role in optimizing control systems. It can permit to control effectively and efficiently local battery banks or auxiliary energy resources. It can also help to manage ramp rates to guarantee PV power output continuity [16, 23, 24].

## 1.6 SUNLIGHT FLUCTUATIONS AND THEIR EFFECT ON PV POWER

Severe swift power output oscillations are mostly due to sudden changes in solar irradiance [25, 26] because irradiance is both the main contributor to the photovoltaic effect and also the fastest changing physical quantity [27]. By extension, PV power forecasting can be seen as monitoring the cloud cover [11, 28-30].



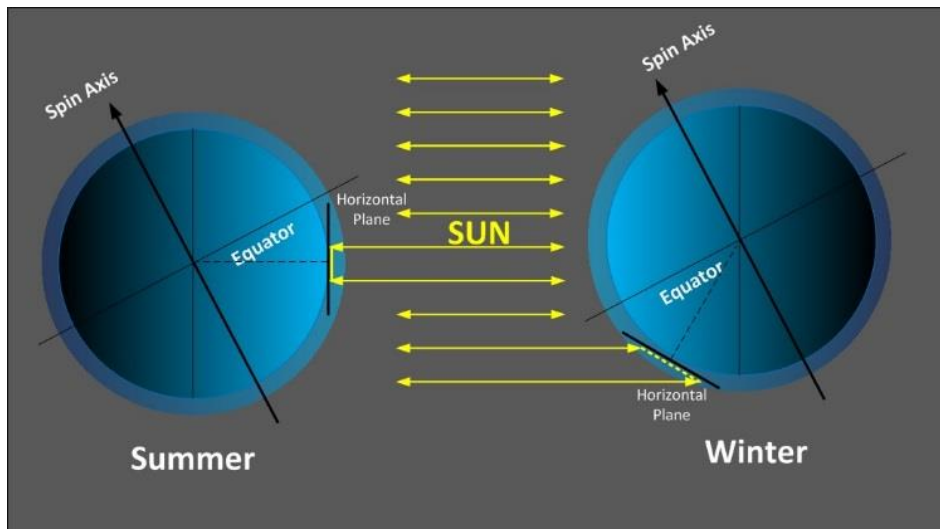
**Figure 1.1 – (a) Received DNI and Power Output of a Thin-film Array and (b) Corresponding Ambient Temperature and Relative Humidity.**

As pointed out by Ramakumar, Jewel and Hill [31, 32] and Chengrui et al. [33], broken cumulus clouds are the cloud type that affects irradiance and thus PV generation the most severely. They can cause AC power fluctuations of small 2kW as well as large 1.6MW PV systems in excess of 50% of their rated capacity in less than 10 seconds. [26]. This kind of effect is illustrated in Figure 1.1 (a), where the power output (in blue) of a 360W amorphous thin film array located at Curtin University, Perth, Western Australia, dropped by 76% in a matter of 12 s. The cause was an opaque cumulus cloud with an optical thickness  $\tau$  assessed between 0.36 and 0.9. It is visible that the power output closely follows the fluctuations of the direct normal irradiance (DNI) (in red), with some extra oscillations due to the maximum power point tracking (MPPT) system. In contrast, the ambient temperature (in blue) in Figure 1.1 (b) and the relative humidity (in red, in the same figure) did not vary as dramatically as the solar irradiance over the exact same period of time. For its part, ground wind speed may vary even faster than solar irradiance. However, wind speed on the surface of the ground only plays a secondary role in PV power generation in helping the dissipation of heat and by partially affecting the PV cell temperature. With that said, most of PV cell models neglect that effect [34]. It is in comparison more relevant to focus on cloud velocity, which may significantly differ from ground wind in terms of both speed and direction [27, 35].

## 1.7 THE DYNAMICS OF PV GENERATION

In terms of dynamics, PV plants' response to grid system disturbances conceptually differs from more traditional synchronous machines. It is however possible to improve ramp-rate controls along with the ability to dispatch PV power through sophisticated control algorithms. An adequate control of PV power can also provide the grid with a stable, reliable and effective operation. Additionally, to manage the variable generation of PV resources, grid operators need short-term forecasting methods as well as some support from alternative compensatory controllable resources. Indeed, the rapid progress and increasing accuracy of short-term solar generation forecast guarantees continued efficient and reliable system operations.

Although forecasting PV power may seem straight-forward at first glance, it must be noted that the dependence on weather conditions is an important obstacle to tackle. Indeed, since the emergence of PV power generation, the abundant, free solar energy has been difficult to harvest and dispatch due to its entire dependency on uncertain, intermittent solar radiance. The stochastic discontinuities of sunlight intensity during the day create PV power fluctuations. Additionally, sunlight is not always available for harvest, as it is the case for instance during night time. Also, changing climatic conditions may result in changing PV power output. As a result, the uncertainty of solar power makes it difficult to integrate into the grid at a high penetration level. This can be an important issue if solar generation is to become a major source of energy in the future [12]. While storing solar energy has long been seen as a solution to this problem, a precise forecast of the available energy is necessary for the previously mentioned control. For the sake of simplicity, the solar radiance in the complete absence of clouds in the atmosphere can be modelled analytically. Indeed, the rotation of the Earth causes a daily disruption of the flux of extra-terrestrial energy received by the top of the atmosphere. Additionally, the tilt of the spin axis of our planet with respect to the Earth's plane of orbit around the Sun makes the global horizontal irradiance, or the rate of total incoming solar energy received on a unitary horizontal area will be higher in summer than in winter [36].



**Figure 1.2 – Seasonal Variations of Global Solar Irradiance for a Given Place on the Surface of the Earth.**

A unitary horizontal area is defined as measuring 1 square kilometre located on a horizontal plane on the Earth's surface. This effect is more dramatic for regions away from the equator, as illustrated in Figure 1.2. Additionally, the incidence angle of sunlight when it goes through the atmosphere plays an important role.

Besides those obstacles to a consistent harvest, the light intensity emitted by the Sun varies on a regular basis, through magnetic activity cycles. On the other hand, it is not possible to model analytically the intermittent variations or disruptions of sunlight due to the presence of clouds and aerosol in the atmosphere. Yet, the motion of clouds has a dramatic impact on the performance of a solar generator [37] and thus need to be forecasted to avoid undesired issues and costs.

Due to changes in the cloud cover, some significant drops in PV generated power output may occur, in the range of up to 70% in a matter of 5 seconds [11, 38]. This may result in episodes of grid instability if not properly compensated. Yet, the cost of such weather-related power outages is far from negligible. As an example, in the comparable case of wind generators, errors in power prediction may cost up to 10% of the income of selling the generated energy [39].



## 1.8 THE CASE STUDY OF PV GENERATION IN WESTERN AUSTRALIA

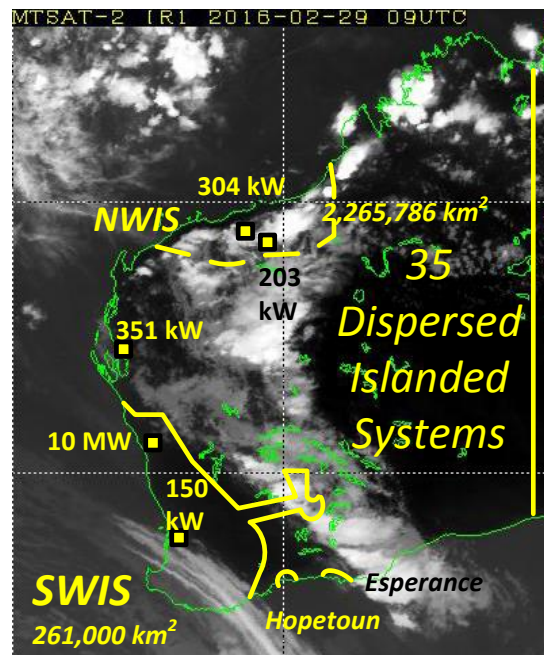


Figure 1.3 – Power networks and major photovoltaic farms in Western Australia [40][41] superposed on top of a satellite view of the state shot on the 29<sup>th</sup> of February 2016 at 5 pm [42], illustrating the geographical disparities of cloud cover over a large territory at a given time.

Only four interconnected systems exist in the state of Western Australia: Hopetoun, Esperance, the North-West Interconnected System (NWIS) [43], and the South-West Interconnected System (SWIS) [44][45][46]. The major PV farms have been placed on the map of Figure 1.3 with their installed capacities. Also, the map features a satellite view of the cloud cover conditions. On the one hand, small regional dispersed islanded service areas may suffer from the previously mentioned problems that have been identified by Jewel, Ramakumar and Hill. On the other hand, it is visible that potentially all the PV farms connected to an interconnected grid such as the SWIS or the NWIS can simultaneously experience significant cloud cover fluctuations overhead. This makes PV integration difficult as the power output of a significant number of PV farms may oscillate simultaneously, under the effect of distinct cloud patterns, and in a fashion that may be difficult to predict precisely. This problem may grow if more PV farms are connected in the future. This is also true for dispersed residential PV generation, which is not represented in Figure 1.3. Including residential sources, solar generation represented 336 MW in the SWIS in 2014, i.e. 6% of the total capacity of the grid, which was 5.5 GW [47].

An accurate very-short term irradiance forecasting method could play a key role in optimizing control systems aiming to maintain a constant output voltage with the help of inverters or a Volt/VAR control system responsible for the actions of tap changing transformers, voltage regulators, switched capacitor banks [15] or other similar systems [48-51].

## **1.9 SOLAR POWER FORECASTING AS A SOLUTION AND SIMILARITIES WITH WIND FORECASTING**

Generally speaking, power forecasting is a smart option to tackle the problems that have been previously mentioned. Power forecasting can help to manage the ramp rate and running time of compensatory systems to ensure both power continuity and low operational costs. Diagne, Boland et al. (2013) [29] have argued in that sense, stating that both an efficient integration of the variable energy output of PV systems and a better quality service require an accurate and reliable forecast of irradiance. Indeed, reliable forecast information on the solar resource is needed to adjust the power to the expected load profile. In other words, PV forecasting may help to better manage the electrical balance between power generation and consumption. However, PV power forecasting is still a relatively recent research topic. Forecasting PV power requires many weather parameters to be taken into consideration. In comparison, simpler wind power forecasting has been already better studied during the past years and is at a more mature stage of research. As a result, PV power forecasting could be inspiring regarding the methodologies already in use in wind power forecasting. This idea is supported by Manz et al., who state in [12] that “the integration of large-scale PV plants in the transmission system can follow the successful model already established by wind integration, with the consequential impact of variability treated in the same manner”. Consequently, the methodology used in wind power forecasting has been judged relevant and inspirational for the purpose of PV power forecasting. This is the reason why we decided it was worth examining the approaches used in wind power forecasting within the next chapter.

### **1.10 OUTLINE OF THIS THESIS**

In total, this thesis is composed of seven chapters. The various chapters outline the overall methodology used to address the challenges posed by PV generation.

In this chapter, which is Chapter 1, the reader has already been provided with an overview of the problems encountered by PV generation. In introduction to solar power forecasting has also been given.

The state of the art of solar power forecasting broadly speaking is studied in Chapter 2. As previously mentioned, because of its similarities with solar power forecasting and since it is more mature topic, wind power forecasting is studied as a source of inspiration. The performance metrics and the definition of forecasting horizons used in wind power forecasting are reviewed and compared with those used in solar power forecasting. The various forecasting horizons in use for solar power forecasting are reviewed and compared. All the prerequisites for solar power forecasting such as PV cell temperature modelling, PV cell voltage and current modelling, clear-sky irradiance modelling, irradiance forecasting, cloud cover, cloud classification are also covered in that chapter. Finally, an

exhaustive review on the advances in very short-term PV power forecasting is given at the end of the chapter.

A detailed view on irradiance modelling is given in Chapter 3, first in the absence of clouds and then when taking clouds into consideration. In the complete absence of clouds and aerosols, the position of the sun in the sky and the extra-terrestrial irradiance are very regular and can be quite easily predicted with some relatively simple analytical models. The generic astronomical parameters that define the solar resource at any given location on the globe are reviewed at the beginning of the chapter. Following this, a large number of empirical irradiance models based on measurements made in different regions on Earth are reviewed, compared with one another and also with more complex physical models. In addition, the empirical models are compared with 1-minute measurements of the irradiance during all four seasons. On the basis of commonly used performance metrics, the two best fits are selected and used as a basis for a linear regression. In the same chapter, the positions of both the sun and clouds are studied as seen from a sky camera installed on the ground. The shading effect of clouds is investigated and a calibration method is considered for sky cameras.

Chapter 4 describes in depth the steps of the method used to track clouds in the images of a sky camera. First, the images are pre-processed. They are masked and a criterion needs to be applied to identify sky pixels. Since various criteria exist in literature and since the threshold values of each criterion are unique for any camera sensor, a novel method has been designed and used to assess the accuracy of each criterion and find optimal threshold values. A couple of implementation of Random Finite Set (RFS), namely a Probability Hypothesis Density (PHD) filter and a Cardinalised Probability Hypothesis Density (CPHD), are introduced and used on a simplified data set to track clouds. Some metrics are used to assess the performance of the RFS implementation against an optical flow based state of the art method. A full discussion of the advantages of the RFS implementation is given.

In Chapter 5, the discrimination criterion between clear sky pixel and cloud pixels is reassessed to suit a new camera that been used to observe the sky over an extended period of time. As a result, a clear-sky library and a cloud library have been collected with a scheduled time lapse. The chromatic properties of those two libraries have been studied in details to characterise clear-sky conditions as well as each cloud type. All features have been looked at to set up an automatic cloud classification. The frequency of the optical thickness of each cloud type has been studied as well by comparing observations and the clear sky model obtained by linear regression in Chapter 3.

Chapter 6 explores a strategy to control PV power in spite of short-term sunlight fluctuations. An accurate and experimentally verified model has been designed for a small grid-connected PV system with maximum power point tracking (MPPT) and voltage source controller (VSC). The effect of various levels of sunlight disruption has been studied on this model to determine a mitigation. A state-of-the-

art battery model has been implemented. Additionally, a dual active bridge and its controller have been designed as an interface between the grid-connected PV system and the battery. The behaviour of the complete system has been tested with various sunlight disruptions through simulation studies. In complement to this analysis, the hardware implementation of a VSC controller is presented with a strategy to optimise its real-time functionality.

Finally, Chapter 7 concludes the study and discusses the future works that are envisaged to further complete and enrich this study.

## **2 STATE OF THE ART OF SHORT-TERM SOLAR POWER FORECASTING**

---

This chapter intends to present exhaustively the state of the art in the field of short-term power forecasting. The purpose is to study the latest advances in that area and compare approaches and tools in order to choose our own methodology. Some conclusions and recommendations are also provided to the reader at the end of each section. This chapter is based on the article “Very short-term photovoltaic power forecasting with cloud modeling: A review” published by the author in the peer-reviewed journal “Renewable and Sustainable Energy Reviews” from Elsevier.

The chapter reviews some relevant methods used by the more advanced field of wind power forecasting. We also review some tools, metrics and definitions that are commonly used in that field as well as in solar power forecasting. After studying a comparison of various approaches of solar power forecasting, we look at various PV cell models, in particular PV cell temperature models. We also study clear-sky irradiance, the different types of clouds and the effects of clouds on the clear-sky irradiance. Finally, we will focus on the various approaches used in very-short term PV power forecasting.

### **2.1 RELEVANT METHODS USED IN THE REFERENCE FIELD OF WIND POWER FORECASTING**

Solar power forecasting is a relatively new field of research that lags behind wind power forecasting in terms of progress. Wind forecasting can consequently be taken as a good reference, as it has been further developed over the past years.

#### **2.1.1 Analysis of the Reference Field of Wind Power Forecasting**

Comprehensive reviews conducted by S.S. Soman et al (2010) [39] and by S.M. Lawan et al (2014) [52] both give a good overview of the different horizons that are taken into account in that field. Soman et al classified wind forecasting techniques into different categories according to their temporal resolution as well as according to their approach. Some studies like the one conducted by Espino and Hernandez (2011) [53] have compared several approaches within various time scales in order to define which method fitted best for each horizon. The paper focuses on short (from 10 to 240 minutes) and very-short term wind forecasting using a statistical approach based on support vector regression (SVR). A quick overview of the different time horizons for wind speed and power forecasting is given in Table 2.1 [39, 52]. These categories are only a popular set and not the sole classification used in literature. As illustrated in

, the very short-term horizon is also known as ‘nowcasting’ [53] and ranges from a few seconds to 30 minutes ahead. It is the shortest timeframe considered in prediction, useful for immediate actions like electricity market clearing and pricing. Only little literature is available for the very-short term timeframe considering both wind and PV power forecasting.

**Table 2.1 – Wind speed/power forecasting time frames [33, 39, 52].**

<i>Time horizon</i>	<i>Range</i>	<i>Possible purpose</i>
<i>Very Short-term / Nowcasting</i>	<i>A few seconds to 30 minutes</i>	<i>Electricity market clearing</i>
<i>Short-term</i>	<i>30 minutes to 6 hours</i>	<i>Economic load dispatching</i>
<i>Medium-term</i>	<i>6 hours to 1 day</i>	<i>Operational security</i>
<i>Long-term</i>	<i>More than 1 day</i>	<i>Maintenance scheduling to optimize the operating cost</i>

The actual short-term forecasting horizon is comprised between 30 minutes and 6 hours ahead. This is the timeframe used for economic load dispatch as 30 minutes is the usual timeframe for grid operations and operating reserve. Most of the research related to wind forecasting has been done in this time scale. The medium-term horizon corresponds to a time scale ranging between 6 hours and 1 day whereas the long-term horizon exceeds 1 day ahead.

These previously defined limits of forecasting terms are however not very strict and may vary depending on the application of the prediction model. For example, in the application field of meteorology, the US National Weather Service considers 0 to 3 hours ahead forecasting as nowcasting [54]. For some other agencies, forecasts up to six hours can also be called nowcasting.

Besides various prediction time frames, several wind speed and power prediction models have already been used in scientific literature, the most popular of which are listed in

. When using the persistence model, also known as ‘naïve predictor’, it is assumed that the physical quantity that is forecast, wind speed or direction, as it happens, remains unchanged during a time increment  $\Delta t$ .

This means that for example the wind vector  $\vec{v}$  is the same at time  $t+\Delta t$  as it was at time  $t$ . This approach thus implies a very strong correlation between present and future values. This prediction method has been reported to be the most accurate for very short-term and short-term forecasts. This approach is even commonly used by meteorologists as a reference benchmark tool to assess the performance of another wind forecasting method.

The persistence model can advantageously supplement the physical model, which is mainly elaborated using Numerical Weather Prediction (NWP). This approach consists in taking into consideration a detailed physical description of the atmosphere. Indeed, the future wind speed and direction both strongly depend on other current meteorological quantities such as ambient temperature, solar irradiance, atmospheric pressure, relative humidity, and dew point. The topology of the terrain is also

considered when using a NWP for wind forecasting purposes. Due to the difficulty to gain live data and due to high costs, NWPs are run only a few times per day on powerful hardware resources, using complex mathematical models commonly based on kinematic physical equations.

**Table 2.2 – Popular methods used for wind speed/power forecasting [39].**

<i>Approach</i>	<i>Subclass/Tools</i>	<i>Examples</i>	<i>Characteristics</i>
<i>Persistence Method</i>		$P(t+\Delta t)=P(t)$ if $k$ is small $\vec{v}(t + \Delta t) = \vec{v}(t)$	<i>Reference benchmark.</i> <i>Very accurate for very short-term and short-term wind forecasting.</i>
<i>Physical Approach/ Numerical Weather</i>	<i>Numerical Weather Prediction (NWP)</i>	<i>Global Forecasting System</i>	<i>Use of meteorological data.</i> <i>Accurate for long term wind forecasting.</i>
<i>Statistical Methods</i>	<i>Neural Network (NN)</i>	<i>Artificial Neural Network (ANN)</i> <i>Multilayer perceptron</i>	<i>Accurate for short-term</i>
	<i>Other techniques</i>	<i>Fuzzy Logic</i> <i>Genetic Algorithm</i> <i>Markov Chain</i> <i>SVR</i>	<i>Good for short-term</i>
	<i>Time-series Models</i>	<i>ARMA</i> <i>ARMAX</i> <i>ARIMA</i> <i>AR</i>	<i>Accurate for short-term</i>
<i>Hybrid Structures</i>	<i>ANN + Fuzzy Logic</i> <i>= ANFIS</i>		<i>ANFIS is very good for very-short term forecast</i>
	<i>NWP + NN</i>		<i>NWP+NN is very accurate for medium and long term</i>

Consequently, this model is not appropriate for relatively short-term forecasts and is only accurate for medium and long-term horizons, in excess of 6 hours. It actually performs best with time scales comprised between 48 hours and 172 hours ahead.

Statistical approaches consist in training a model with measurement data and then comparing the predicted values with near-past actual values to fine-tune predicted parameters. Such models are accurate for relatively short-term horizons. They can be divided into two main subclasses: Time-series

based approaches and Artificial Neural Network (ANN) based methods. Other statistical methods like fuzzy logic and genetic algorithm can also be used.

ANN models are powerful nonlinear data driven methods that are easy to model, inexpensive and less time consuming than other methods. Their prediction model is based on patterns rather than on any predefined mathematical model. However, the accuracy of such prediction models quickly drops as the time horizon stretches. Several variants of Neural Networks are used like feed-forward neural networks (FNNs), multi-layer perceptrons (MLP), recurrent neural networks (RNNs) and time delay neural networks (TDNN) to name a few. More details are given in the next section.

Furthermore, some additional artificial intelligence (AI) techniques can also be successfully applied for short-term predictions, like for example genetic algorithms (GA), fuzzy logic, Markov chains (MC), etc.

Time series based forecasting methods have been very popular for wind prediction in the past years. The most commonly used model is the Auto-Regressive Moving Average (ARMA), which is an improvement of the regular predictive auto-regressive method (AR). However, this method can only be used for stationary time series. Several variations exist like Auto-Regressive Integrated Moving Average (ARIMA) and Auto-Regressive Moving Average with exogenous inputs (ARMAX). The coupled autoregressive and dynamical system (CARDS) is also a popular method. Other models like Bayesian Model Averaging (BMA) or Support Vector Regression (SVR) [53] are also applied.

Model performance is often assessed by the means of Root Mean Square Error (or deviation) (RMSE), normalized (or standardized, or relative) Root Mean Square Error (nRMSE or rRMSE), Mean Absolute Error (MBA) or Mean Bias Error (MBE) or other means.

The root mean square error (RMSE) and normalized root mean square error (nRMSE) are commonly used to evaluate the performance of an approach or a model. If we consider the number  $N$  of samples, both errors are respectively defined by equations (2.1) and (2.2) [55]:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (forecast - measurement)^2} \quad (2.1)$$

$$nRMSE = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (forecast - measurement)^2}}{\frac{1}{N} \sum_{i=1}^N measurement} \quad (2.2)$$

The nRMSE is nothing but the RMSE divided by the range of measurement. It is usually expressed as a percentage.

The mean bias error (MBE) and mean absolute error (MAE) are other statistical quantities used to measure how close forecasts or predictions are to the eventual outcomes. The MBE is an average of the errors while the MAE is an average of the absolute errors. Both are respectively given by equations (2.3) and (2.4):



$$MBE = \frac{1}{N} \sum_{i=1}^N (forecast - measurement) \quad (2.3)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |forecast - measurement| \quad (2.4)$$

Another important metric is the Pearson's correlation coefficient, which quantifies how well a model matches practical data in shape. A perfect match is given by a result of 1. This coefficient is usually denoted by  $\rho$  and consists in the covariance of both the modeled and the measured GHI divided by the standard deviation of both GHI:

$$\rho = \frac{cov(GHI_{Modeled}, GHI_{Measured})}{\sigma(GHI_{Modeled}) \cdot \sigma(GHI_{Measured})} \quad (2.5)$$

RMSE shows the overall accuracy of a model while the MBE demonstrates if the model is generally high or low and the MAE shows how far off the model is. The accuracy and offsets measured by these metrics may be easily changed by altering some coefficients in the model.

Additionally, it is also common practice to combine various approaches, such as for instance the physical and statistical approaches. Alternatively, various horizons, such as short-term and medium-term, may also be combined. Such combinations are referred to as hybrid methods and aim to optimize the short-term prediction values. Indeed, many researchers use a combination of methods. They unanimously agree that hybrid methods outperform the individual methods [29]. Amongst the large number of possible combinations, mixing together NNs and fuzzy logic leads to the popular approach of Adapted Neuron Fuzzy Inference (ANFIS). Another popular technique consists in correlating the wind speeds at different sites.

All the discussed time scales and methods are also valid for PV power forecasting. Indeed, wind speed and direction is one factor amongst other meteorological quantities that greatly influence the prediction of PV power. As a matter of fact, a fast frontal wind may cool down the cell temperature within a PV array, which in turn increases the voltage output and by extension also the power output.

### 2.1.2 Conclusions on What Can Be Learned From the Field of Wind Power Forecasting

Although the following considerations are valid for wind power forecasting, they can also be used as a guideline for solar power forecasting, which is more complex since more variables are taken into account. As a first step, the forecasting horizon is defined according to the grid operation. Secondly, an approach is chosen on the basis of the horizon. Studies carried on wind power forecasting show that the persistence approach is the best choice for very short-term horizons. On the other hand, time-series and ANN-based approaches are only reliable for short term horizons. NWP, for their part, prove to be only accurate for medium to long horizons.

## 2.2 COMPLEMENTARY INFORMATION ON STATISTICAL METHODS FOR PV FORECASTING

Tools like machine learning methods are very effective to deal with non-linear systems. Some examples of commonly used methods are [5]:

- Expert Systems (ES)
- Artificial Neural Network (ANN)
- Genetic Algorithm (GA)
- Fuzzy Logic (FL)
- Computer Vision (CV)
- Hybrid Systems (HS)
- Many other methods

### 2.2.1 Review of the Statistical Methods for PV Forecasting

ESs interpret data to deal with knowledge processing and make decisions for complex problems.

Genetic Algorithms utilize a fixed-sized population containing individual possible solutions to a given problem, which evolve in time. Genetic operators such as selection, crossover and mutation are applied to eliminate the poorest solutions and create new solutions from selected existing ones. These algorithms are very popular in machine learning.

ANNs are based on the emulation of a biological brain. It has the capacity to be trained and to learn. The architecture consists in multiple layers of interconnected artificial neurons. The neurons communicate through synapses. Each artificial neuron (AN) compares their inputs with a threshold value to produce an output. The weight of each input can be adjusted during a preliminary learning phase [56]. The ANNs support both supervised and unsupervised learning.

A multi-layer perceptron (MLP) is an ANN consisting in several fully interconnected layers of ANs. It is trained with a supervised learning technique. MLPs are the most popular form of ANNs. They consist in an input layer, at least one hidden layer and an output layer. The hidden layers often use a hyperbolic tangent function of the input vector  $x$ , as shown in equation (2.6) [29].

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.6)$$

Considering a MLP that processes a number  $n$  of input vectors  $X$ , is characterized with  $h$  hidden neurons and a single linear output  $y$ , the relationship between the input and the output is given by (2.7) [29].

$$y = \sum_{j=0}^h \left[ W_{ij} f \left( \sum_{i=0}^n W_{ji} X_i \right) \right] \quad (2.7)$$

When used for one day-ahead forecasting, MLPs may lead to very accurate results. For example, Mellit and Pavan (2010) [57] have successfully used a MLP to forecast the solar radiance 24 hours ahead with a 98% precision during sunny days and less than 95% during cloudy days.

Fuzzy Logic is based on the utilization of logical statements and binary operators to model problem treated with a human-like reasoning. It is mainly used in control engineering.

These artificial intelligence methods have been successfully applied for the purpose of PV power output or irradiance forecasting [57]. According to Mellit et al (2010), MLP and Radial Basis Functions (RBF) networks as well as fuzzy logic are particularly suited in the case of precise forecasts based on a significant number of meteorological and geographical input data. In other words, such methods are effective when the forecast irradiance is expressed as follows:

$$G(t + 1) = f \left( T_{amb}(t), RH(t), TCC(t), w_{speed}(t), P(t), Lat, Long, G(t) \right) \quad (2.8)$$

In equation (2.8),  $T_{amb}$  is the ambient temperature,  $RH$  is the relative humidity,  $TCC$  is the total cloud cover,  $w_{speed}$  and  $w_{direction}$  are respectively the wind speed and direction,  $P$  is the atmospheric pressure,  $Lat$  and  $Long$  are respectively the latitude and longitude, and  $G$  is the irradiance.

On the other hand, recurrent Neural Networks (RNN), wavelet and wavelet-fuzzy networks are more suitable for forecasts based on past observed data only. In mathematical terms, the forecast of the irradiance, for instance, is then expressed as:

$$G(t + 1) = f(G(t), G(t - 1), G(t - 2), \dots) \quad (2.9)$$

Alternatively, ANFIS methods are suitable when the forecast irradiance is modelled as a combination of the two previous models:

$$G(t + 1) = f \left( G(t), G(t - 1), G(t - 2), \dots, T_{amb}(t), RH(t), TCC(t), w_{speed}(t), P(t), Lat, Long \right) \quad (2.10)$$

Computer vision deals with image processing and motion analysis. Many systems of this type have been developed for a large number of applications. However, the strategy employed depends mainly on the system that is analysed.

Other methods may include Markov Random Filters (MRF), Probability Hypothesis Density filters (PHD), k-nearest neighbour (kNN) or State Vector Machine (SVM).

The kNN approach is used in particular by Heinle et al [58] and Kazantzidis et al [59] to classify clouds into cloud classes according to some of their features. This operation is performed by majority vote. However, kNNs can be slow to run and may require a large amount of memory.

A SVM is a nonlinear model that demonstrates faster computing features than an ANN as well as good convergence [60]. It requires nonetheless a large amount of data for regression.

For the purpose of solar power forecasting, Diagne, Boland et al (2013) recommend the use of statistical methods for short forecasting horizons (intra-hour or intra-day), hence for high temporal resolutions

(less than 10 hours) and high spatial resolutions (up to 500 meters). Even though hybrid models are more robust, single methods can present a particular strength within for a specific resolution. Indeed, ANNs may be recommended for temporal resolutions between 30 minutes and 10 hours as long as the targeted spatial resolution is between 10 and 500 metres. The persistence method is the best choice for very fine resolutions below 10 minutes and 10 meters. On the other hand, ARIMA, ARMA and CARDS are suitable for temporal resolutions from a few seconds to 1 hour and spatial resolutions up to 12 meters.

### 2.2.2 Conclusions of the Review of Statistical Methods Used for PV Forecasting

A large number of statistical methods are available. ANNs can be accurate, especially when they are combined with other artificial intelligence methods like genetic algorithms. However, they show slow computational features. On the other hand, SVM-based method can be faster but require more data.

## 2.3 COMPARISON OF SOME POPULAR PV POWER FORECASTING APPROACHES

The definition of forecasting horizons for PV power forecasting may differ when compared to wind forecasting, and so does the choice of approaches.

As an example, a statistical and a physical approach have been confronted to each other in a study conducted by Yuehui et al [61] by considering the case study of a 1 MW PV farm comprising 6006 panels. Firstly, this quick review of PV power forecasting methods defines the various time scales used in power system operation with different limits, compared to wind forecasting. The new definitions of the various horizons are summarized in Table 2.3.

The very short-term horizon ranges from several minutes to several hours, aiming for intraday real-time control. The short-term horizon ranges from a few hours to up to 3 days and is used for day-ahead economic dispatch. As an illustration of short-term horizon, Diagne, Boland et al [29] mention that load patterns need to be forecast 2 days ahead for the purpose of scheduling power plants and for planning transactions in the electricity market. The long-term forecast is defined as ranging from a week to a year and targets grid balance operations.

**Table 2.3 – One possible definition of the various horizons for PV power forecasting [61].**

<i>PV Horizon</i>	<i>Time resolution</i>	<i>Possible purpose.</i>
<i>Very Short-term / Nowcasting</i>	<i>A few minutes to several hours</i>	<i>Real-time control</i>
<i>Short-term</i>	<i>Up to 3 days</i>	<i>Economic dispatch</i>
<i>Long-term</i>	<i>A week to a year</i>	<i>Grid balance</i>

**Table 2.4 – Alternative definition of the various horizons for PV power forecasting [29].**

<i>PV Horizon</i>	<i>Time resolution</i>	<i>Possible purpose</i>
<i>Intra-hour</i>	<i>15 minutes to 1 hours</i>	<i>Ramp rates control, variability management</i>
<i>Intra-day</i>	<i>1 to 6 hours</i>	<i>Load following</i>
<i>Day ahead</i>	<i>1 to 3 days</i>	<i>Transmission scheduling</i>

This classification of PV forecasting horizons is not unique and Kostylev and Pavlovski (2011) also distinguish 3 forecasting horizons: intra-hour, intra-day and day ahead. This alternative discrimination of horizons has been used by Diagne, Boland et al and is exposed in Table 2.4 [29].

The effect of irradiance intermittency can be attenuated through the use of energy storage, spinning reserve and demand response [38]. All these methods need an accurate forecast on various timescales to be optimally operated. Indeed, day-ahead forecasting is required to determine pricing in the market. In contrast, sub-hour-ahead forecast is useful to schedule spinning reserves.

The comparison focuses on short-term horizon and compares the results of a physical approach with the ones of a statistical method by using the same days over a period of one month during each season of the year in a place located in the northern hemisphere. The physical model that is utilized for PV generation consists in a clear sky model and diode PV model. The statistical method uses a neural network (NN).

The measured output power of one representative month of each season is compared with the clear sky model power prediction, adjusted with NWP data. Measurements have been made in the northern hemisphere.

Unsurprisingly, the best match is achieved during clear days of each month whereas the biggest mismatch is obtained during the month of October and December, which have been chosen to represent respectively autumn and winter of Northern Hemisphere. Both seasons are characterized by a measured power output much lower than the prediction. Overall, the nRMSE values range between 11.48% and 16.58%, as shown in Table 2.5.

**Table 2.5 – nRMSE for different forecasting methods. Data collected by Yuehui (2010) [61].**

<i>Month/Season</i>	<i>Physical Model (NWP input)</i>	<i>Physical Model (measured G and T inputs)</i>	<i>Statistical Approach (NN)</i>
<i>March/Spring</i>	<i>11.48</i>	<i>8.78</i>	<i>11.05</i>
<i>June/Summer</i>	<i>11.54</i>	<i>5.17</i>	<i>10.45</i>
<i>October/Autumn</i>	<i>16.58</i>	<i>4.93</i>	<i>15.49</i>
<i>December/Winter</i>	<i>15.33</i>	<i>5.61</i>	<i>12.62</i>
<i>Overall Year</i>	<i>12.45</i>	<i>5.51</i>	<i>10.5</i>

When the actually measured values of temperature and irradiance are additionally fed to the physical model, the precision increases dramatically and the nRMSE values drop down to 4.93%-8.78%. This makes the biggest discrepancy occur in spring for the cloudy month of March. This enhanced precision validates the physical model, provided it is fed with sufficient weather data.

In Table 2.5, the central column corresponds to a physical model using the measured irradiance  $G$  and temperature  $T$  as inputs. The nRMSE of the predictions from the best trained NN model are shown in the rightmost column of Table 2.5. The optimal NN is set to have 11 neurons and is fed with several weather data such as the ambient temperature, the cloud cover, the irradiance and the position of the sun. Compared to the NWP physical model predictions, the results appear to be significantly better for the wintry month but only slightly more accurate for all other seasons. Overall, there is only little improvement with respect to the NWP physical model. The study concludes that both approaches are valid for the considered horizon and that the precision mainly depends on the data input itself.

Another study, carried out by Guarnieri et al (2008) [62] has demonstrated that the ANNs using training data can reduce the nRMSE of daily average global horizontal irradiance by 15% compared to a 12-18 hours ahead irradiance forecast based on a NWP approach as pointed out by Diagne et al in their review paper [29].

### **2.3.1 Conclusions of the Comparison between Popular PV Power Forecasting Approaches**

For solar power forecasting purposes, very short-term may be defined on a longer time frame than for wind power forecasting. Indeed, this horizon may be in range of several minutes rather than in range of several seconds. However, a forecasting horizon beyond very short-term is required for power dispatch and grid balance. For such horizons, the input data proves to be actually more important than the chosen method. Nonetheless, neural network based forecasting methods are more accurate than NWP-based approaches for periods of the year characterized by heavy rain and snow falls like winter.

## **2.4 PV CELL MODELLING STRATEGY**

Solar irradiance is the main driving parameter that influences PV power output [33]. The effect of the irradiance on the current output and on the power output for a fixed temperature are respectively illustrated in Figure 2.1 and Figure 2.2 with the example of a mono-junction polycrystalline module.

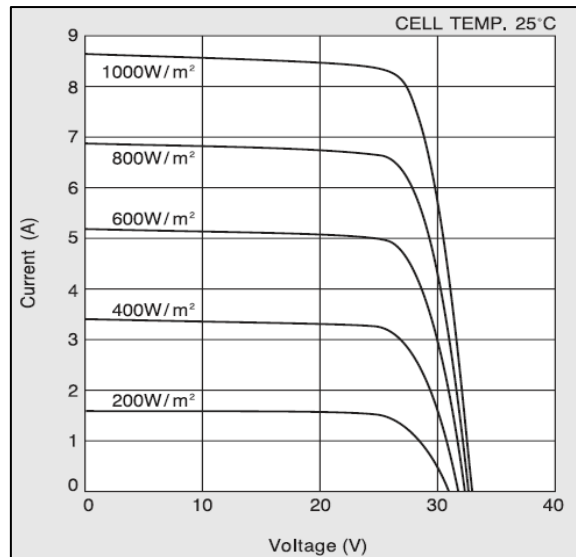


Figure 2.1 – I-V curves for various solar irradiances with a Kyocera KC200GT [27].

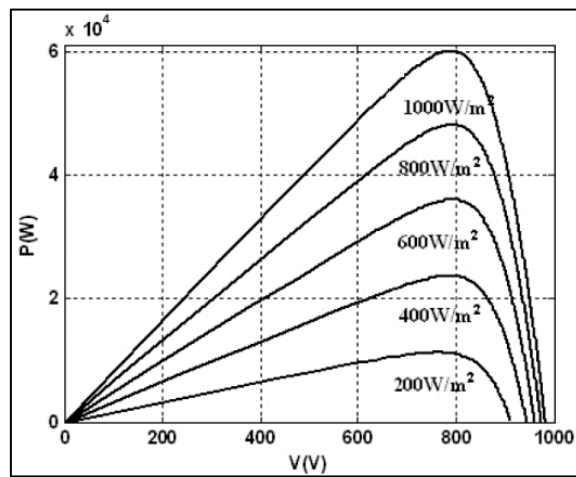
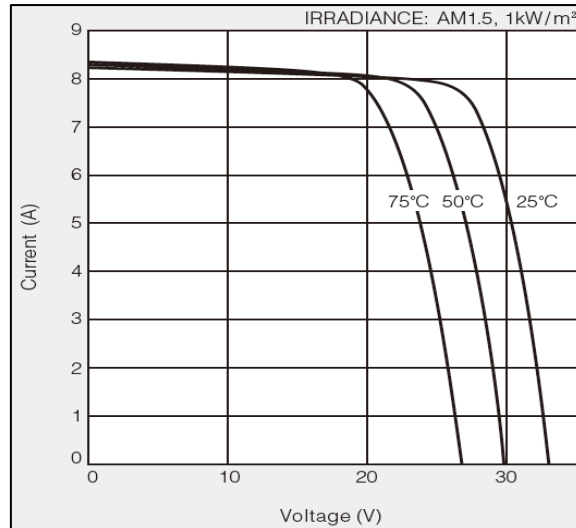
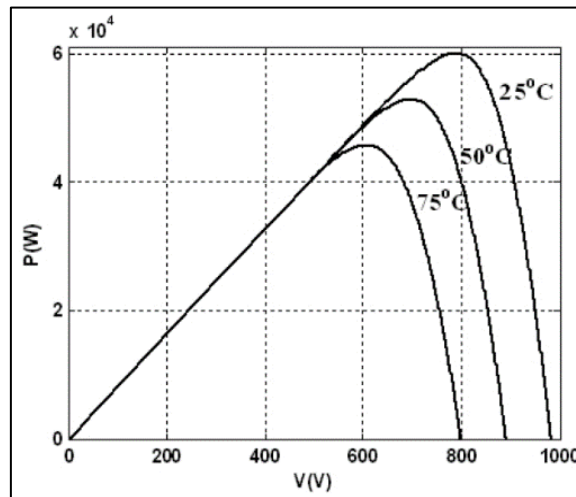


Figure 2.2 – P-V curves for various solar irradiances with a Kyocera KC200GT [29].



**Figure 2.3 – I-V curves for different temperatures with a Kyocera KC200GT [29].**



**Figure 2.4 – P-V curves for different temperatures with a Kyocera KC200GT [29].**

Although high-efficiency multi-junction PV cells are under development, the large majority of the current PV installations use mono-junction modules. Also, most of the current studies also concentrate on mono-junction modules.

Consequently, the focus of this study has also been restricted to mono-junction modules. Additionally, the cell temperature and by extension the ambient temperature play an important role as well, though not as important as the solar radiance.

Knowing the operating temperature of photovoltaic modules is necessary to improve the energy performance. The reason is because the operating temperature  $T_c$  greatly affects the energy efficiency of the conversion of light into electricity. This is visible on the example of current voltage and power voltage characteristic curves of a PV module respectively shown in Figure 2.3 and Figure 2.4.



Amongst several empirical formulas, Evans' expression of efficiency summarizes the effect of both cell temperature and solar irradiance [56, 63]:

$$\eta = \eta_{ref} \cdot \left[ 1 - \beta \cdot (T_c - 25) + \gamma \cdot \log_{10} \left( \frac{G}{G_{ref}} \right) \right] \quad (2.11)$$

In equation (2.11),  $\beta = 4.5 \cdot 10^{-3} \text{ } ^\circ\text{C}^{-1}$  and  $\gamma = 0.12$  for crystalline silicon modules.

The cell temperature can either be modelled or alternatively assessed through the utilization of adaptive techniques such as artificial neural networks (ANNs). An adaptive technique can actually be used as an advantageous tool for a fast and precise forecast of the cell temperature and to guarantee high performance. Only 10 to 20% of sunlight is converted into a current by mono-junction modules whereas the rest is converted into heat that is partially transferred to the surrounding environment according to a specific coefficient. Consequently, the temperature of a module automatically rises as soon as it absorbs some irradiance, which in turn affects the operating point and leads to a diminished electrical conversion of light [56]. Yet, a comprehensive analytical model would be excessively complex to use for calculations.

This is why the various temperature cell models found in literature assume some simplifications and do not consider accurately the full complexity of the irradiance absorption effect. However, an artificial neural network can help to deal with such complexity in a timely efficient manner. In addition, it must be noted that the current-voltage curves provided by manufacturers in their datasheets cannot be used to reliably assess the thermo-electrical behaviour of a PV module as they are given for a constant temperature or a constant irradiance. Indeed, datasheet curves do not consider actual conditions, which involve cell temperature and irradiance changes.

## 2.5 CELL / MODULE TEMPERATURE MODEL

Various empirical models are available in literature [34, 56]. Table 2.6 comprises some of them. In this table, the instantaneous solar insolation is referred as  $G$  ( $\text{kW}/\text{m}^2$ ) whereas  $T_{amb}$  is the ambient temperature given in  $^\circ\text{C}$ . The NOCT conditions are defined as  $G_{NOCT} = 800 \text{ W}/\text{m}^2$ ,  $w_{speed} = 1 \text{ m}/\text{s}$  and  $T_{c, NOCT} = 20^\circ\text{C}$ .

**Table 2.6 – Various empirical models available in literature to evaluate PV cell temperatures, as listed by Jakhrani et al (2011) [32] and Ciulla et al (2013) [22].**

<i>Year</i>	<i>Authors</i>	<i>Empirical Model</i>	<i>Comments</i>
1976	Ross [64]	$T_c = T_{amb} + k \cdot G$	$k = 0.02-0.04^\circ\text{C}\cdot\text{m}^2/\text{W}$
1980	Rauschenbach [65]	$T_c = T_{amb} + \frac{G}{G_{NOCT}} \cdot (T_{c,NOCT} - T_{amb,NOCT}) \cdot \left(1 - \frac{\eta_m}{\tau\alpha}\right)$	$w_{speed} > 1\text{m/s}$ and constant heat loss coefficient $U_L$
1983	Risser-Fuentes [66]	$T_c = 3.81 + 0.0282 \cdot G + 1.31 \cdot T_{amb} - 165 \cdot w_{speed}$	Model verified with MPPT
1985	Schott [67]	$T_c = T_{amb} + 0.028 \cdot G - I$	Model verified for $1\text{m/s} < V_w < 1.5\text{m/s}$ and $0^\circ\text{C} < T_{amb} < 35^\circ\text{C}$
1985	Servant [68]	$T_c = T_{amb} + \alpha \cdot G \cdot (1 + \beta \cdot T_{amb}) \cdot (1 - \gamma \cdot w_{speed}) \cdot (1 - 1.053 \cdot \eta_m)$	$A=0.0138$ , $\beta=0.031$ , $\gamma=0.042$ , $T_{amb}$ is given in $^\circ\text{C}$ with a constant wind speed $w_{speed}=1\text{m/s}$
1986	Ross-Smokler [69]	$T_c = T_{amb} + \frac{G}{G_{NOCT}} \cdot (T_{c,NOCT} - T_{amb,NOCT})$	only valid for free standing modules with $w_{speed}=1\text{m/s}$ and constant $U_L$
1990	Lasnier-Ang [70]	$T_c = 30.006 + 0.0175 \cdot (G - 300) + 1.14 \cdot (T_{amb} - 25)$	For p-Si only, $T_{amb}$ is given in $^\circ\text{C}$ whereas the wind speed and heat loss coefficient are not taken into account
2000	Hove [71]	$T_c = T_{amb} + \frac{G \cdot (\tau\alpha - \eta)}{U_L}$	$T_{amb}$ is given in $^\circ\text{C}$ , the ratio $\tau\alpha/U_L$ (with the transmittance $\tau$ , absorbance product $\alpha$ and heat loss coefficient $U_L$ ) is determined experimentally by assuming that the efficiency $\eta = 0$
2004	Krauter [72-74]	$T_c = T_{amb} + k \cdot G$	$k = 0.0058$ for lower modules, $0.012$ for upper modules and $0.03$ for usual modules
2005, 2007	Mondol et al. [75]	$T_c = T_{amb} + 0.031 \cdot G - 0.058$	$T_{amb}$ is given in $^\circ\text{C}$ , $w_{speed} > 1\text{m/s}$ with constant heat loss coefficient $U_L$
2006	Duffie-Beckman [76]	$T_c = T_{amb} + \frac{G}{G_{NOCT}} \cdot \left(\frac{9.5}{5.7 + 3.8 \cdot w_{speed}}\right) \cdot (T_{c,NOCT} - T_{amb,NOCT}) \cdot (1 - \eta_m)$	The product $\tau \alpha$ of transmittance $\tau$ , absorbance product $\alpha$ is taken as $0.9$ whereas the heat loss coefficient $U_L$ is associated with wind speed
2007	Chenni et al [37]	$T_c = 0.943 \cdot T_{amb} + 0.028 \cdot G - 1.528 \cdot w_{speed} + 4.3$	The coefficient of heat loss $U_L$ is not taken into account.

In any case, the two main parameters always remain the irradiance and the ambient air temperature. Amongst popular models, the Ross and Smokler formula from 1986 is widely used [77]:

$$T_c = T_{amb} + \frac{G}{G_{NOCT}} \cdot (T_{c,NOCT} - T_{amb,NOCT}) \quad (2.12)$$

It must be noted that the Ross and Smokler model is only valid for a free standing module.

Another more precise example that considers 3 weather parameters is Chenni et al's model. The coefficients are average values of measurements performed on 6 different PV technologies (amorphous silicon, monocrystalline silicon, copper indium diselenide, EFG-polycrystalline silicon, polycrystalline silicon, and cadmium telluride) located on two different geographical sites [78]:

$$T_c = 0.943 \cdot T_{amb} + 0.028 \cdot G - 1.528 \cdot w_{speed} + 4.3 \quad (2.13)$$

Nevertheless, this still remains an imperfect model as it does not take heat losses into account.

The approach of assessing and predicting the cell/module temperature is particularly interesting for hybrid PV/thermal systems containing a fluid cooling the PV modules. Controlling the temperature of the PV modules help to increase the voltage output. Additionally, the heat that is extracted from the module can be in turn exploited to improve the efficiency of the complete system. As mentioned by Ciulla, Lo Brano, and Moreci [56], best results are achieved through the employment of a multilayer perceptron network. In this publication, the MLP is composed of 2 non-linear threshold function blocks, 2 weight layers, 1 hidden layer, 1 error criterion block and 2 input sources that consider the following parameters:

- ambient temperature  $T_{amb}$
- wind speed  $w_{speed}$
- irradiance  $G$
- electrical power output of the module  $P_{elec} = I \cdot V$
- short circuit current of the module  $I_{SC}$
- open circuit voltage of the module  $V_{OC}$

The weather data comes from a weather station and the actual temperature of the panels of the experimental setup is measured with thermocouples. The electrical data are measured as well.

As a result, the confidence band of the evaluation is narrower than  $\pm 1^\circ\text{C}$ . For different cell temperature correlations, the mean absolute error (MAE) obtained with the ANN is around 0.1–0.2, according to the type of PV panel. This is much below the MAE of the analytical empirical correlations like Servant, Duffie-Beckman and Hove. These models display an MAE between 4 and 16, making the ANN

superiorly accurate. Additionally, the authors suggest that the very short learning time enables to implement the ANN on a real-time system to evaluate the cell temperature in run-time conditions.

## 2.6 FURTHER CONSIDERATIONS OF THE PV MODEL

Knowing the operating temperature of the cell helps to determine the maximum solar power per m<sup>2</sup> as expressed in equation (2.14) [79].

$$P_m = P_p \cdot \eta_t \cdot \frac{G}{G_{STC}} \quad (2.14)$$

Where  $P_p$  is the nominal power (W) given by the manufacturer at standard test conditions (STC) solar irradiance 1000 W/m<sup>2</sup>, air mass of 1.5 and cell temperature of 25°C,  $G_{STC}$  is the reference irradiance (1000 W/m<sup>2</sup>) and the normalized efficiency  $\eta_t$  is defined by equation (2.15).

$$\eta_t = 1 + \gamma_{mp} \cdot (T_c - T_{STC}) \quad (2.15)$$

In equation (2.15),  $\gamma_{mp}$  is the maximum power thermal coefficient, typically about  $-0.5\%/^{\circ}\text{C}$  for crystalline modules. The efficiency rate  $\eta = \eta_0 \cdot \eta_t$  is defined on the basis of  $\eta_0$ , which is the nominal conversion rate in STC and for which the value depends on the PV technology.

**Table 2.7 – Mean nominal conversion rate  $\eta_0$  by technology [80].**

<i>Monojunction technology</i>	<i>Monocrystalline</i>	<i>Polycrystalline</i>	<i>Amorphous thin film</i>
$\eta_0$ (%)	14.96	14.36	10.49

The average nominal efficiency of the three main categories of mono-junction PV technologies is summarized in Table 2.7 [80]. Monocrystalline PVs have the highest nominal conversion rate whereas thin film amorphous silicon systems have the lowest.

The maximum power of an array characterized by an area  $Area$  is given by equation (2.16) [81].

$$P_{m,Area} = P_m \cdot Area \quad (2.16)$$

The cell temperature also both affects current and voltage [8]:

$$I = I_{NOCT} \cdot \frac{G}{G_{NOCT}} \cdot [1 + \alpha \cdot (T_c - T_{c,NOCT})] \quad (2.17)$$

$$V = V_{NOCT} \cdot [1 + \beta \cdot (T_c - T_{c,NOCT})] \cdot \left[1 + \delta \cdot \ln\left(\frac{G}{G_{NOCT}}\right)\right] \quad (2.18)$$

In equation (2.17) and (2.18), the coefficients  $\alpha$ ,  $\beta$  and  $\delta$  must be determined from an initial condition.

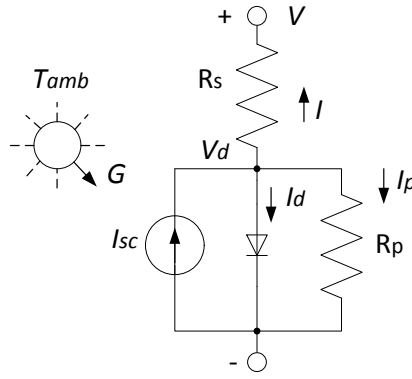


Figure 2.5 – Two-diode model of a single PV cell [50].

The two-diode model is the most precise for a photovoltaic cell [82] as shown in Figure 2.5, where:

$$V = V_d - R_s \cdot I \quad (2.19)$$

$$I = I_{sc} \cdot I_0 \cdot \left\{ e^{\frac{q \cdot V_d}{k \cdot T}} - 1 \right\} - \frac{V_d}{R_p} \quad (2.20)$$

It should be noted that at 25°C,  $q \cdot V_d / k \cdot T = 38.9$ .

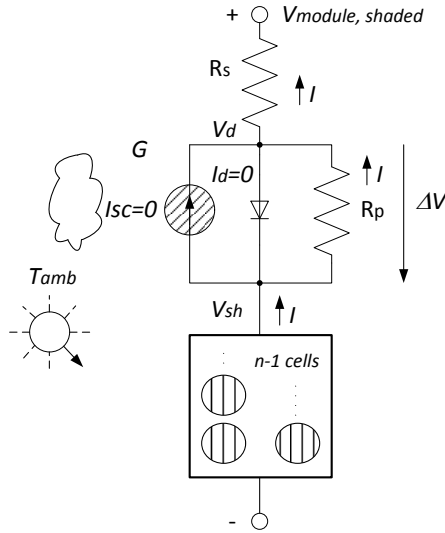
Additionally, the series resistance  $R_s$  is typically significantly smaller than the parallel resistance  $R_p$ .

When  $n$  cells are arranged as a module with all cells equally insolated, the voltage of the module becomes:

$$V_{module,insolated} = n \cdot (V_d - R_s \cdot I) \quad (2.21)$$

However, if only one of the cells gets shaded, the parallel resistance of the shaded cell will cause a dramatic voltage drop:

$$\Delta V = I \cdot (R_p - R_s) \cong R_s \cdot I \quad (2.22)$$



**Figure 2.6 – Two-diode model of a PV module with partial shading of one cell among a module of  $n$  cells [50].**

This effect is illustrated in Figure 2.6. In that case, the output voltage then becomes:

$$V_{module,1shadedcell} = \left(\frac{n-1}{n}\right) \cdot V_{module,insolated} - \Delta V \quad (2.23)$$

The voltage drop compared with a fully insolated module is

$$V_{module,1shadedcell} - V_{module} = \frac{V_{module,insolated}}{n} \cdot -\Delta V \quad (2.24)$$

Using shunt diodes to compensate this effect is a common practice. It is also possible to work on the design to attenuate the effects of partial shading.

### 2.6.1 Conclusions on PV Modelling

Ross and Smokler’s expression of the cell temperature is very popular and widely used by manufacturers in their datasheets. However, Chenni et al’s formula is the most precise to date although it can still be improved as it does not take the coefficient of heat dissipation into consideration. It must also be noted that the 2-diode model is the most precise to date for a mono-junction PV cell. From the point of view of the nominal conversion rate, monocrystalline PV cells and modules are the most productive, most expensive and also most sensitive to temperature gradient amongst mono-junction modules. Thin film modules are on the contrary the least efficient but cheapest and most robust ones.

## 2.7 ADVANCES IN IRRADIANCE FORECASTING

### 2.7.1 Review of Clear-Sky Irradiance Forecasting

Using the output power of fast growing utility-size solar photovoltaic (PV) farms as base load supplies arises serious challenges that are yet to be overcome. In particular, managing the fluctuations and uncertainty of solar power has been identified as a major problem [24]. The purpose of power smoothing is to mitigate large power fluctuations due to clouds passing over a PV array. Such an event may induce PV power output losses in range of 70% in a matter of 5 seconds [11, 83]. In contrast, the utility operators of grids with high renewable energy penetration only tolerate ramp rates up to 10% per minute [26]. On the other hand, peak shifting may be used to shift the excess peak load generation to hours when the demand cannot be met with instantaneous PV power, such as during mornings, evenings or night times. It has been long thought that both problems could be addressed with the help of an adequately sized battery energy storage systems (BESS) [11, 83]. Prior to forecasting the instantaneous irradiance under realistic conditions at any given place or time, it is necessary to accurately model the clear-sky Global Horizontal Irradiance (GHI). After obtaining the clear-sky GHI, perturbations due to cloud cover, aerosols, water vapour, ozone, and dust in the atmosphere and on PV modules are added to the initial state. It must be noted that the orientation of the PV receiver may affect the incidence angle  $\theta$  of the Direct Normal Irradiance (DNI), which is an important component of the GHI.

Diagne, Boland et al have stated that forecasting global horizontal irradiance GHI is the first and necessary step for most of PV power prediction models [29]. They also have stated the fact that an accurate and efficient irradiance forecast can help grid operators to maintain grid balance. Two steps are needed for a comprehensive irradiance forecast. First, a clear sky irradiance needs to be modelled based on astronomical facts. The actual irradiance can be forecast in a second time by considering the effects of clouds on the clear sky solar radiance.

Some more accurate models have been defined on a physical basis. These clear-sky irradiance models have not been designed to fit one particular region of the world but require the measured values of many specific atmospheric parameters. For example, the intermediate precision models by Kasten [84] and Ineichen and Perez [85] both need the Linke turbidity [86, 87], which can be defined as the total optical thickness of the atmosphere. Some even more precise and thus more complex models like for instance Bird [88, 89], MAC [90, 91], Iqbal [92], ASHRAE [93], and REST2 [94, 95] need a larger number of atmospheric input parameters, like the optical depth of the ozone layer in the atmosphere, or the ground albedo. Such data may be provided by AERONET [96].

Some simpler empirical models can be obtained through regression of local measurements. Sandia National Laboratories [97] have listed several empirical approaches. Masters also mentions one model

[98]. Such models are based on irradiance measurements made in a specific part of the world and may not always be suitable to model the clear-sky irradiance away from the original measurement site [99]. Most of these models are a simple function of the zenith angle  $z$ , expressed in degrees.

The equations of the empirical clear-sky irradiance models cited by Sandia will be shown in further detail in the next chapter. They have been originally developed for various geographical regions mainly located in the northern hemisphere. For example, the Kasten-Czeplak (KC) GHI model [100] is based on continuous records made at Hamburg, Germany, between 1964 and 1973. The Haurwitz GHI model [101, 102] has been designed on the basis of data gathered by the Meteorological Observatory of Harvard University in the Northeast of the United States. The Threlkeld and Jordan (TJ) [98, 103] model has also been designed the United States and requires the day number. The Berger-Duffie (BD) GHI model [104] and the Adnot-Bourges-Campana-Gicquel (ABCG) [105] are both based on data collected in France. The Robledo-Soler (RS) GHI model [106] has been defined on the basis of some data obtained during the period from 1994 to 1995 in Madrid, Spain. The Laue DNI model [107] calculates the DNI thanks to data collected in California. Besides the zenith angle, the Laue model also takes the air-mass ratio as well as the site elevation as inputs. The Meinel DNI model [108] can be seen as a simplified variant of the Laue model where the site elevation is considered to be at sea level (0 m ASL).

It must be added that some empirical models can sometimes be a good fit for an area that is very different from its area of origin. This is for example the case for the ABCG model that inspired of the Dazhi model, which consists in a regression of measurements of the clear-sky GHI made in Singapore [109, 110]. That means at least one of these models can possibly be a good fit for Western Australia.

Also, empirical models are very relevant for very-short-term irradiance forecasting because of they are fast to compute and do not require any sophisticated measurements. This why many popular empirical models will be compared in the next chapter to find the best match with clear-sky irradiance conditions in Western Australia.

This topic is covered in further details within the next chapter.

### **2.7.2 Review of Irradiance Forecasting considering clouds**

The precision of a PV power forecast depends greatly on the precision of the solar radiance forecast. So conclude E. Lorenz et al [111]. In their study, these authors have managed to use one-day ahead irradiance forecasts at the scale of a country to predict regional PV power output up to three days ahead. Overall, the method consists in considering a regional observation, which greatly improves forecasts with respect to a single site observation. It must be noted that the precision of this method depends on the size of the observed region. Indeed, the larger the region where the PV systems are distributed, the more precise the forecast. As an illustration, the overall nRMSE for a day ahead for a single PV site is



37%. This value is decreased to 19% for a region that is 3° latitude high and 3° longitude wide in size. The nRMSE is further reduced to 13% for the whole of Germany, which corresponds of an area of size 9° × 10°. To describe the methodology that is used, a global numerical weather prediction (NWP) model of resolution 3 hours and 0.25° × 0.25° (which is approximately 27 km × (16-19) km in Germany) from the European Centre for Medium-Range Weather forecasts (ECMWF) is used as the source of forecasted irradiance data. A network of 11 PV systems with a resolution of 200 × 120 km is also used to measure the power output. Finally, 200 weather stations spread over Germany are utilized to measure the irradiance over a period of 10 months. A spatial interpolation technique on a region of 100 km × 100 km is first utilized to refine the ECMWF global irradiance forecasts for a specific given site. Additionally, a temporal interpolation is performed by combining the 3-hourly forecast data  $G_{forecast,3h}$  provided by the ECMWF with a clear sky model. The 3-hourly mean value clear sky index  $k_{3h}$  is calculated as  $k_{3h} = G_{forecast,3h} / G_{global\ irradiance,3h}$ . The hourly mean value  $k_{1h}$  is then derived by linear interpolation. Alternatively, the mean hourly irradiance can be directly interpolated from the original 3-hourly data  $G_{forecast,3h}$ , but this method has proved to be less accurate in situations without clouds or with only a few. A correction may be necessary in some sky configurations. For example, in the case of the German region, the forecast irradiance is underestimated in overcast situations with low irradiance ( $k < 0.2$ ) and overestimated in the case of variable cloud cover ( $0.3 < k < 0.8$ ). The values are thus replaced by the clear sky model irradiance in the case of a total cloud cover below 0.03.

In addition, a situation-specific bias is introduced for correction purposes. Other studies also use linear time series of global irradiance, possibly normalized with respect to a clear sky model, as an input with a two stage approach. During the first stage, either a PV system network [112] or satellite data [55] are utilized to forecast the irradiance. In a second stage, some autoregressive models such as autoregressive moving average (ARMA) [55] or autoregressive with exogenous input (ARX) [112] are also used.

**Table 2.8 – Possible reference sources of data for PV forecasting.**

<i>Type of forecast</i>	<i>Very short-term/Grid management</i>			<i>Utility Applications</i>
<i>Approaches based on</i>	<i>Sky camera</i>	<i>Sensor network</i>	<i>Satellite</i>	<i>NWP</i>
<i>Temporal resolution</i>	<i>Seconds-minutes</i>	<i>Minutes</i>	<i>Minutes-hours</i>	<i>Day-days</i>
<i>Usage</i>	<i>cloud cover and cloud classification</i>	<i>Benchmark and only direct mean of measurement</i>	<i>cloud cover</i>	<i>Forecast of atmospheric conditions (ambient temperature, wind speed)</i>

However, such methods only focus on the irradiance and neglect the effect of temperature and wind on the performance of the systems. Moreover, sudden changes in cloud cover may not be described in the model. As a matter of fact, the wind plays an important role in the abrupt changes of irradiance due to the passing of clouds. With an increasingly fine temporal resolution, the swift changes of cloud cover play a more and more important role.

Independently from the irradiance forecasting method, the source of data can be chosen on the basis of the desired temporal resolution. A finer resolution can always be obtained through interpolation.

Such temporal interpolations can be satisfactory during periods when the sky is completely clear or overcast. However, they may lead to significant errors in situations during which the sky cover varies stochastically, as it is the case in sub-hourly forecasts. Table 2.8 gives a broad overview of the various sources of data that may be used to forecast solar radiation. Hassanzadeh et al (2010) suggest that satellite data is well adapted for very short-term forecasting, whereas numerical weather prediction (NWP) data are better suited for long term forecasting, beyond two days. However, for fine temporal resolutions, real-time measured irradiance data can dramatically improve the performance of now-casting. In particular, Bacher et al (2009) have demonstrated that direct measurements of the irradiance prevail in forecasts up to 2 hours ahead whereas NWP inputs are more valuable beyond that horizon.

On top of the input source, the targeted temporal resolution also defines the choice of forecasting methodology. For example, grid-connected photovoltaic plants require a horizon of at least 24 hours for integration purposes. Mellit et al [57] have chosen to use an ANN-based approach for such a 24-hour irradiance forecast. More specifically, this team has utilized a multilayer perceptron to process the daily average irradiance and daily average temperature. The data consists in measurements taken in Italy over the course of several months. The choice of a method based on artificial intelligence is justified by the nature of the input data, as it is a time series containing several meteorological variables. With the help of a cross validation method, the team has managed to evaluate a performance above 98% during sunny days but below 95% during cloudy days. The authors suggest that adding more input parameters such as cloud, pressure, and wind speed and sunlight duration could possibly improve the accuracy of the forecast.

### **2.7.3 Review of Irradiance Forecasting considering clouds**

A clear sky model requires two main sub-models. The position of the sun needs to be precisely defined. The solar position algorithm developed by Reda and Andreas [113] is the most precise to date. Additionally, the direct or global horizontal solar radiance also needs to be modelled. The Meeus and Bird and Hulstrom [114, 115] model is the most accurate with respect to measured insolation values. After modelling the irradiance in absence of clouds, the solar radiance must be modelled by considering the effect of clouds. Doing so requires to follow a methodology. The temporal resolution is chosen in

accordance with the grid application that is sought. The selected temporal resolution influences the choice of the source of input data. A sky camera is the best source of input data for very short-term forecasting, whereas sensor networks and satellite images are both very suitable for short term. NWP data only performs well for longer horizons. In addition, the temporal resolution also conditions the forecasting method. For very-short horizons, persistence-based methods are the best, time series and ANN methods are the most appropriate for short term whereas an NWP approach is well-fitted for longer terms.

## **2.8 CLOUD COVER MODELLING AND CLOUD CLASSIFICATION**

### **2.8.1 Overview**

Using energy storage units (ESU) has long been proposed as a solution to dispatch PV power. Frequency domain has been previously utilized to specify PV power. Yet, this approach is not suitable for sizing an ESU. It is indeed necessary to know the peak energy amount corresponding to a charge or discharge of the battery [13]. Therefore, modelling and predicting cloud-related PV power intermittency is required for an efficient ESU control strategy with a deterministic scheduling approach. Studies have proven that two parameters can strongly impact the generation of PV power: the cloud cover of the sky and the cloud type or class [59]. The cloud cover can either be expressed as a percentage or in octas (also spelled oktas), which is a unit corresponding to 1/8th of the sky, i.e. 12.5% of the entire sky. Indeed, the size, thickness and altitude of the clouds covering the sky can greatly influence the direct solar irradiance received by a photovoltaic module [33, 58]. As a matter of fact, both the total cover of the sky and the altitude of the cloud define the size of the shaded area on the ground, whereas the thickness of the cloud is linked to its opacity and by extension to its ability to block sunbeams. Consequently, considering a precise sky model is very relevant for PV power forecasting purposes. Such aspects are covered in Chapter 4.

The cloud cover can be measured by analysing the individual pixels of digital pictures obtained from a sky camera. Clear sky pixels are characterized with a high saturation of colours and a predominance of blue. As opposite to cloud-free pixels, cloudy pixels display dull colorations featuring different shades of whites, possibly tinted of colourful hues ranging from red to blue.

These characteristics have been notably used by Martins et al (2003) and Souza-Echer et al (2006) to define a method to estimate the cloud coverage [59]. Alternatively, in the absence of a sky camera, the cloud cover can also be defined as the percentage of the ground area that is covered by cloud shadow [33]. This last approach implies however the use of an accurate cloud shadow model. This is an approach that can be used for example with a regional network of dispersed ground sensors. After

correction of shading due to permanent obstacles that are not clouds, the clear-sky index  $K$  of a location can be defined as:

$$K = \frac{G}{G_{clear}} \quad (2.25)$$

In equation (3.42),  $G$  is the instantaneous irradiance and  $G_{clear}$  the predicted clear sky irradiance [38].

By convention, there is a total of 27 identified types of clouds, equally divided within three altitude levels also called étages. Each étage thus contains 9 types of clouds.

**Table 2.9 – Approximate heights of cloud levels.**

<i>Level</i>	<i>Polar Region</i>	<i>Temperate Region</i>	<i>Tropical Region</i>
<i>High Clouds</i>	<i>10,000-25,000 feet</i> <i>(3-8 km)</i>	<i>16,500-40,000 feet</i> <i>(5-13 km)</i>	<i>20,000-60,000 feet</i> <i>(6-18 km)</i>
<i>Middle Clouds</i>	<i>6,500-13,000 feet (2-4 km)</i>	<i>6,500-23,000 feet (2-7 km)</i>	<i>6,500-25,000 feet (2-8 km)</i>
<i>Low Clouds</i>	<i>Surface-6,500 feet</i> <i>(0-2 km)</i>	<i>Surface-6,500 feet</i> <i>(0-2 km)</i>	<i>Surface-6,500 feet</i> <i>(0-2 km)</i>

**Table 2.10 – Cloud classification for PV forecasting purposes.**

<i>#</i>	<i>Cloud Classification</i>	<i>Level of the cloud</i>	<i>Altitude range of obstruction</i>
<i>1</i>	<i>Clear Sky (Cl)</i>	<i>Not applicable</i>	<i>Not applicable</i>
<i>2</i>	<i>Cumulus (Cu)</i>	<i>Low</i>	<i>Up to 6,500ft/2,000m</i>
<i>3</i>	<i>Stratocumulus (Sc)</i>	<i>Low</i>	<i>Up to 6,500ft/2,000m</i>
<i>4</i>	<i>Stratus-altostratus (St-As)</i>	<i>Middle - Low</i>	<i>Up to 20,000ft/6,100m</i>
<i>5</i>	<i>Cirrocumulus-altocumulus (Cc-Ac)</i>	<i>Middle - High</i>	<i>From 6,500ft/2,000m</i>
<i>6</i>	<i>Cirrus-cirrostratus (Ci-Cs)</i>	<i>High</i>	<i>Above 20,000ft/6,100m</i>
<i>7</i>	<i>Cumulonimbus-nimbostratus (Cb-Ns)</i>	<i>Low-middle-high</i>	<i>Very low base below 6,500ft/2,000m and high top potentially above 20,000ft/6,100m</i>

It must be noted that the elevation of the base of each type of cloud depends on local atmospheric conditions. This elevation varies according to the season and the location. Table 2.9 illustrates some approximate heights [116]. A cloud scatters a portion of the sunlight intensity depending on its own optical depth  $\tau$ , as shown in Equation (2.26) [117].

$$\frac{I}{I_0} = e^{-\tau} \quad (2.26)$$

In Equation (2.26),  $\tau$  is the specific optical depth or optical thickness of the cloud,  $I$  is the light intensity transmitted by the cloud while and  $I_0$  is the original incident light intensity received on top of the cloud.

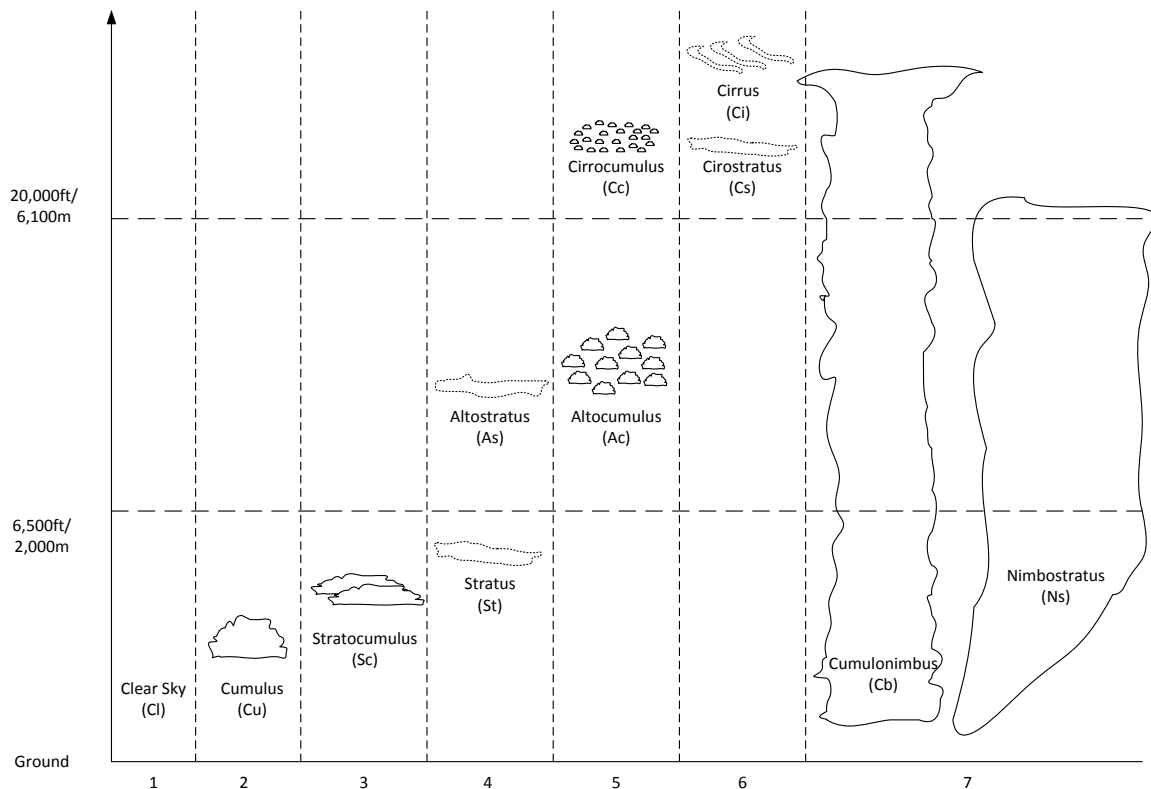
An example shown by Raschke and Cox is that an atmospheric depth of 6 would reduce a direct beam intensity of 1356 W/m<sup>2</sup> at the top of the atmosphere to only 3.3 W/m<sup>2</sup> at ground level [117].

From a shading point of view, it may not be relevant to deal with the whole complexity of all 27 types of clouds. Indeed, several types may cause similar shading effect. Heinle et al have managed to automatically classify clouds into the 7 following categories [58, 59] summarized in Table 2.10.

**Table 2.11 – Cloud classification by optical thickness and top pressure.**

Cloud Top Pressure (mb)	50	Cirrus	Cirrostratus	Deep Convection	
	440	Alto cumulus	Altostratus	Nimbostratus	
	680	Cumulus	Stratocumulus	Stratus	
		0	3.6	23	379
		Optical thickness			

Some values of optical thickness defined by the International Satellite Cloud Climatology Project (ISCCP) [118] are given in Table 2.11. This quantity is directly related to the opacity of the clouds. However, the average values of the optical depth or thickness for each type of clouds are not available as yet. A study is currently in progress to define simple and reliable values that could be used to define the transparency of clouds. Indeed, the type and thickness of cloud directly both affect the beam component of sunlight as mentioned by Chengrui et al [33]. The 7 categories of clouds mentioned in Table 2.10 are shown in Figure 2.7. Clouds can be categorized in four main categories according to their altitude level.



**Figure 2.7 – Illustration of the various types of clouds.**

Low level clouds occur below 6,500 feet i.e. 2,000 meters and are either named *cumulo-* (meaning ‘heap’) or *strato-* (which means ‘layer’) [119]. Stratus, cumulus and stratocumulus clouds are the only types that exclusively populate the lower level. Stratus (St) clouds are composed of a uniform and diffuse flat layer of grey cloud cover [120]. As they are partially transparent, the outline of the sun is clearly visible when seen through such clouds. This type of cloud may remain stationary from several hours to a couple of days but some layers may break up or dissipate, partially revealing some clear sky. They develop horizontally. In contrast, cumulus (Cu) clouds are detached and dense with a sharp outline. They grow vertically and eventually mature as cumulonimbus clouds prior to eventually dissipating. They are characterized with flat horizontal bottoms and rounded tops that resemble cauliflowers. The sunlit parts of such clouds are bright white while the base is darker. They usually develop during clear sky conditions, due to diurnal convection. Low-altitude cumulus clouds may move fast and cause significant sunlight discontinuity as these clouds have a clearly defined edge and cast a deep shadow [33]. On the other hand, stratocumulus (Sc) clouds are a hybrid type composed of several individual cumulus clouds distributed in characteristic stratus layers. They consist in a mix of thick and thin layers of grey or whitish patches. They also almost always have dark rounded masses resembling a honeycomb.

Mid-level clouds occur between 6,500 and 20,000 feet i.e. between 2,000 and 6,100 meters. Their names are given the prefix *alto-* (meaning ‘high’). The two main types are altostratus and altocumulus.

Altostratus (As) clouds are similar to stratus in structure and can be associated with them. They present grey or bluish layers of fibrous cloud sheets that may totally cover the sky. They are very thin and commonly reveal the sun as if it was seen through ground glass. They will thus not cause any significant decrease of insolation. They do not create any halo phenomena and do not cast any shadow on the ground. Similarly to stratocumulus clouds in low levels, altocumulus (Ac) clouds are constituted of clusters of cumulus type features in middle levels. They are the most common mid-level clouds, composed of thin semitransparent white or light grey layers of rounded shapes that may be partly diffuse and fibrous. Several layers may appear, often in presence of other types of clouds. A corona appears when such clouds pass in front of the sun. This ring has blue in the inside and red on the outside.

High level clouds appear above 20,000 feet i.e. 6,100 metres and are given the prefix cirro- (which means ‘wisp of hair’). Cirrus clouds (Ci) are wispy and feathery as they are entirely composed of ice crystals. They consist in delicate separate hair-like semitransparent white silky filaments and they hardly diminish the insolation. Cirrostratus (Cs) clouds form a widespread layer of whitish transparent veil-like clouds similar to low level stratus clouds. They almost always cover the whole sky with an extensive sheet cover. The sun produces a halo when hidden by a layer of cirrostratus.

Cirrocumulus (Cc) clouds are layered with small cumulus-like clouds. They correspond to a degraded stage of cirrus or cirrostratus. They are composed of more or less regularly arranged very small thin white grains, forming a ‘mackerel sky’ or ‘buttermilk sky’.

Finally, some enormous clouds may be present in all three levels of the sky. This is the case for cumulonimbus (Cb) clouds, which correspond to a mature stage of vertical development of cumulus cloud. These dense clouds are responsible for storms, hail and heavy rains. The base is very dark and the overall cloud is very large and extremely opaque. The upper part is usually smooth and flat, reminding the shape of an anvil. Like cumulonimbus clouds, nimbostratus (Ns) clouds also occupy a large range of altitude, although not necessarily as high.

They are dark grey, very opaque and they obscure the sky, causing very low insolation. Nimbostratus clouds result from thickening dense stratus, altostratus or stratocumulus clouds. They cause episodes of continuous rain or snow.

Regarding the impact of a specific type of cloud on the solar irradiance, it must be noted that to the knowledge of authors like Cai et al [33], no universal algorithm can deal with all types of clouds due to the complexity of a comprehensive atmospheric model. Indeed, many occurrences may simultaneously happen on various levels of altitude, caused by unrelated parameters. The hereinabove mentioned authors recommend to focus on low level cumulus clouds as they are the type of clouds that are responsible for most of fluctuations. The same authors also propose to model the solar irradiance in a

probabilistic manner. Climate science has established that the shape of cumulus cloud shadows contours have a fractal structure [11, 33].

Cai et al use a midpoint displacement algorithm to model this structure. The algorithm used result in a 3-dimensional fractal output, which is only a mathematical artefact that enables to generate cloud shadows by intersecting a cross section of this output with a horizontal plane. Barnes et al [11] point out that the size  $x$  a cumulus clouds shading a ground area for a period of time  $t$  follows the power law distribution where  $A$  and  $b$  are distribution parameters:

$$f(x) = A \cdot x^b \quad (2.27)$$

It has been inferred that the clear and shaded conditions distribution of the same ground area follow a generalized Pareto distribution where  $k$ ,  $\sigma$  and  $\theta$  are parameters of the distribution. The duration of the clear condition is truncated based on the length of the day. The distribution is as follows:

$$f(t|k, \sigma, \theta) = \frac{1}{\sigma} \cdot \left[ 1 + k \cdot \frac{(t - \theta)}{\sigma} \right]^{-1 - \frac{1}{k}} \quad (2.28)$$

In order to evaluate the area shaded by a cloud, it matters to know the altitude of the cloud base. This can be either done by using a ceilometer or indirectly by calculating the lifted condensation level. A laser ceilometer uses a laser beam that is reflected back by the base of clouds.

By using either time travel method or triangulation, the device can measure the distance of a light beam. Generally speaking, the altitude range for such a device is 25,000 feet (7,620) metres and clouds above this height are not detected. Alternatively, knowing the dew point temperature, the ambient temperature and the altitude of the weather station with respect to the sea level, it is possible to evaluate the altitude of the base of low level clouds like cumulus clouds. The lifted condensation level above sea level (ASL) corresponding to this altitude can be evaluated in meters with the following expression [121]:

$$LCL = 400 \cdot 0.3048 \cdot (T_{amb} - T_{dew}) + h \quad (2.29)$$

In equation (2.29),  $LCL$  is given in meters,  $h$  is the altitude above sea level (ASL) of the weather station expressed also in metres and  $T_{amb}$  and  $T_{dew}$  are respectively the ambient and the dew point temperature expressed in °C, at the location of the weather station.

Furthermore, identifying accurately the type of cloud and classifying clouds also helps to get an idea of the altitude range. This is mainly useful for high level clouds.



A full classification of clouds helps to evaluate the cloud cover and as well as the type of clouds above a specific site.

### 2.8.2 Assessment of Cloud Cover

It is necessary to process images from a whole sky camera to evaluate the cloud coverage. This can be performed through a pixel by pixel classification and sorting pixels as either ‘*clear sky*’ or ‘*cloudy*’ pixels. As a first step, an image-mask must be applied onto the images in order to filter out any irrelevant information that are not related to the sky, like the ground or surrounding infrastructures. The solar disk and visible sunrays may also bias the estimation of cloud coverage and must be eliminated from the image as well. This can be done by using the fact that the related pixels in the image are usually completely white, revealing an overexposure of the camera’s sensor. In addition, a whitening effect of cloud-free pixels around the solar disk may lead to a misclassification of these pixels. After filtering all irrelevant pixels, the remaining pixels need to be classified as ‘*clear sky*’ or ‘*cloudy*’. An interesting property of the atmosphere can be used at this effect. As a matter of fact, in a clear atmosphere, which means in absence of natural (like fog or mist) or artificial (like haze or fog) aerosols, air molecules scatter more blue light than red light, making the sky appear blue to the human eye.

In contrast, clouds scatter blue and red light in similar proportion, leading us to see them light grey or even white [58, 122]. Pfister et al (2003) and Long et al (2006) have developed a method to process colour sky images by comparing the RGB images with some blue and red thresholds in order to estimate the total cloud coverage. By using this method, pixels can be classified as cloudy or not cloudy according to the value of the red/blue (R/B) signal ratio. According to Long et al’s method, pixels with values above 0.6 are classified as cloudy. However, some problems may arise to detect thick clouds or to classify circumsolar pixels.

To tackle these issues, Heinle et al [58] have proposed to use the difference  $R - B$  as a criterion instead of the ratio R/B. The value  $R - B = 30$  is used by this team as an optimal reference threshold. Kazantzidis et al [59] could also check that the difference  $R - B$  outperformed R/B for a sufficient set of images. Nevertheless, this team also discovered that a multi-colour criterion combining R, G and B thresholds ( $B < R + Th_1$  &  $B < G + Th_2$  &  $B < Th_3$ ) outperforms the R-B difference method in some common cases of cloud cover. The threshold values  $Th_1 = 20$ ,  $Th_2 = 20$ , and  $Th_3 = 60$  provided by the authors are completely indicative and may drastically vary accordingly with the camera sensor.

However, the discrimination of thin cirrus or cumulus clouds or even clouds close to the sun is problematic and a source of uncertainty, leading to significant errors such an underestimation of the cloud cover. It must however be noted that this method has not been proven yet to be accurate when the sun is covered by clouds.

### 2.8.3 Identification of Cloud Types

**Table 2.12 – Relevant features to process images from sky cameras based on the works of Heinle (2010) [24] and Katzanidis (2012) [25].**

Type of Feature	Name	Expression	Comment
Spectral	Mean (R)	$M_R = \frac{1}{N} \cdot \sum_{i=0}^{N-1} p_{R,i}$	Useful to define the average tonal values
	Mean (B)	$M_B = \frac{1}{N} \cdot \sum_{i=0}^{N-1} p_{B,i}$	
	Standard Deviation (B)	$SD_B = \sqrt{\frac{1}{N-1} \cdot \sum_{i=0}^{N-1} (p_{B,i} - M_B)^2}$	These features give some information about the tonal variations $E_B = \sum_{a=0}^{G-1} \sum_{b=0}^{G-1} P_{dist}(a,b) \cdot \log_2(P_{dist}(a,b))$
	Skewness (B)	$SK_B = \frac{1}{N} \cdot \sum_{i=0}^{N-1} \left( \frac{p_{B,i} - M_B}{SD_B} \right)^3$	
	Difference (R-G)	$D_{RG} = M_R - M_G$	
	Difference (R-B)	$D_{RB} = M_R - M_B$	
	Difference (G-B)	$D_{GB} = M_G - M_B$	
Textural	Energy (B)	$M_B = \sum_{a=0}^{G-1} \sum_{b=0}^{G-1} P_{dist}(a,b)^2$	This shows the homogeneity of monochromatic level differences
	Entropy (B)	$E_B = \sum_{a=0}^{G-1} \sum_{b=0}^{G-1} P_{dist}(a,b) \cdot \log_2 \left( \sum_{b=0}^{G-1} P_{dist}(a,b) \right)$	This measures how random the monochromatic level differences are.
	Contrast (B)	$C_B = \sum_{a=0}^{G-1} \sum_{b=0}^{G-1} (a-b)^2 \cdot P_{dist}(a,b)$	This measures local variations of monochromatic level differences
	Homogeneity (B)	$H_B = \sum_{a=0}^{G-1} \sum_{b=0}^{G-1} \frac{P_{dist}(a,b)}{1 +  a-b }$	This measures how similar neighbour pixels are in terms of monochromatic levels.
Total Cloud Cover	Cloud Cover (R,G,B)	$CC = \frac{N_{CloudyPixels}}{N}$	This is a measure of the average cloud cover, useful to distinguish stratiform clouds

Heinle et al (2010) [58] have automated the classification of clouds according to their type as a further step following the clear/cloudy classification. On the basis of their Fisher Distances, a technique used

in satellite imagery (Pankiewicz, 1995), 12 features have been selected to classify clouds. Seven of these features are spectral while 4 are textural and 1 is simply the total cloud cover. Indeed, the spectral features of an image do not provide any information about the spatial distribution of clouds.

As a matter of fact, solely using colours proves to be insufficient to distinguish altocumulus from stratocumulus clouds. This is why textural features must also be taken into consideration. In addition, the cloud cover is also used as a feature to tackle the issue of spatial distribution. In particular, this helps to distinguish stratiform clouds from the rest.

The colour information of each cloud image is partitioned into a set of three monochromatic (also referred to as Grey Level) red (R), green (G) and blue (B) data sets. Each monochromatic partial image is treated separately as levels of R or G or even B. The average colour and the tonal variation of the image both provide some useful information to distinguish thick and thin clouds, or even high altitude bright clouds from the others. The colour component B proved to have the highest separation power because it is the dominant colour of the sky.

Heinle et al (2010) [58] have automated the classification of clouds according to their type as a further step following the clear/cloudy classification. On the basis of their Fisher Distances, a technique used in satellite imagery (Pankiewicz, 1995), 12 features have been selected to classify clouds. Seven of these features are spectral while 4 are textural and 1 is simply the total cloud cover. Indeed, the spectral features of an image do not provide any information about the spatial distribution of clouds.

As a matter of fact, solely using colours proves to be insufficient to distinguish altocumulus from stratocumulus clouds. This is why textural features must also be taken into consideration. In addition, the cloud cover is also used as a feature to tackle the issue of spatial distribution. In particular, this helps to distinguish stratiform clouds from the rest.

The colour information of each cloud image is partitioned into a set of three monochromatic (also referred to as Grey Level) red (R), green (G) and blue (B) data sets. Each monochromatic partial image is treated separately as levels of R or G or even B. The average colour and the tonal variation of the image both provide some useful information to distinguish thick and thin clouds, or even high altitude bright clouds from the others. The colour component B proved to have the highest separation power because it is the dominant colour of the sky.

comprises some details related to the 12 hereinabove mentioned features. Each feature is calculated for a colour that is abbreviated between the brackets within the name of the feature. In the expressions,  $N$  represents the total number of pixels in the image under consideration,  $p_{R,i}$  and  $p_{B,i}$  are respectively the monochromatic intensity value of red and blue for an individual pixel referred to with the number  $i$  so that  $0 < i < N - 1$ . The variables  $a$  and  $b$  represent monochromatic intensity levels of two pixels of the

same image. The values of these variables is comprised between 0 and  $G = 256$ . The value  $P_{dist}(a, b)$  is the pixel distance between both pixels, using a Grey Level Co-occurrence Matrix (GLCM).

A very accurate k-nearest-neighbour (kNN) classifier is used to discriminate sky images into one of the seven cloud conditions on the basis of these features.

Heinle and al. used a sample of 1500 reference images for training purposes, containing about 200 independent images per cloud class, with a wide variety of cloud forms. After training the system and processing the input images, the Manhattan Distance is calculated to assess the difference between the normalized feature vector  $f$  of each element and the reference feature vector  $f_{ref}$ . On this purpose, the following formula is used:

$$d(f, f_{ref}) = \sum_{i=0}^{dim} |f_i - f_{ref,i}| \quad (2.30)$$

The cloud visualized in the input images is classified according to the majority vote of the  $k$  closest matches to each known cloud class. The parameter  $k$  has been set as 3 by Heinle et al, leading to an average performance of 96% elements classified correctly over all 7 cloud classes.

The performance is measured by using a Leave-One-Out Cross-Validation (LOOCV).

The authors noticed that the only difference between cumulus and altocumulus clouds is the size of individual cloudlets.

Additionally, some noticeable misclassification errors appear due to a confusion between stratiform clouds and thick rain clouds, such as nimbostratus. This is due to the fact that stratiform clouds such as stratus and stratocumulus can build up to form thick rain clouds that still display the same characteristics. Moreover, the presence of rain drops on the camera dome may distort images and lead to some further misclassifications.

To address these problems, Kazantzidis et al (2012) [59] improved this classification method by introducing subclasses and by proposing a method to detect the presence of raindrops over the camera lens. A circle factor (CF) is used to evaluate how close to a perfect circle the fish-eye image of the camera is. In the presence of rain drops, the contour of image is distorted, leading to a circle factor of about  $CF = 0.6$  whereas this factor is close to  $CF = 0.9$  with a dry dome. This factor does not only help to prevent a misinterpretation of texture values but can also detect the presence of rain clouds. Furthermore, as there is only small differences of entropy feature among the various cloud classes, Kazantzidis et al have decided not to use this feature. A set of 1050 images have been used to train the algorithm after visual inspection, while 1500 images are used for testing. A minimum of 150 images

per cloud class has been used for the training, with only one cloud class per image. In order to tackle the problem of variability of cloud within the same cloud class, three new parameters have been introduced, as per Table 2.13. According to the values of these parameters, three tiers are defined. These new parameters lead to the creation of a certain number of new subclasses for each cloud class as per Table 2.14.

**Table 2.13 – Additional relevant image processing parameters [25].**

<i>Criterion</i>	<i>Tier 1</i>	<i>Tier 2</i>	<i>Tier 3</i>
<i>Solar Zenith Angle <math>\phi_s</math> (degrees)</i>	<40	40-65	>65
<i>Fraction of Cloud Coverage (octas)</i>	<3	3-6	7-8
<i>Visible Fraction of the Solar Disk (%)</i>	<20	20-80	>80

**Table 2.14 – Additional sub-classes [25].**

Cloud Class	St-As	Cb-Ns	Cu	Cc-Ac	Sc	Ci-Cs	Cl
Number of sub-classes	4	6	8	9	10	12	1

The accuracy of the overall classification ranges between 78% and 95%. In particular, cumulus clouds are detected with a success rate of about 92%. Despite these innovations, a small confusion remains between the cloud classes St-As and Cb-Ns. The misclassification of St-As engenders the lowest performance of 78%. However, a correct classification of clouds is very relevant for PV power forecasting when the optical depth of each cloud class is known.

## 2.8.4 Conclusions

The shorter the forecasting timeframe, the more dramatic the impact of clouds on the power output of PV systems. Sky imagers are the best source of input images is indeed the only one that can help to distinguish the different types of clouds. Considering the dynamics of clouds, a sampling period of approximately 1 second between each consecutive images is ideal to extract an accurate cloud motion vector. Adequate post-processing of the images is necessary to determine the cloud cover as well as the type of clouds present in the sky. A multicolour criterion combining different thresholds for R, G and B data is the optimal method for this step. The images from a sky camera can also possibly be used to estimate the size of clouds are therefore the area of the cast shadow. After identifying the type of a cloud masking the sun by comparison with a database, it is also possible to infer the opacity of that cloud and thus the instantaneous irradiance. For this purpose, it is indeed vital to be able to identify each type of cloud among a set of 7 classes. Heinle's method performs very well to identify clouds accurately. It consists in working on monochromatic R, G and B images. A set of 12 features are extracted and

discriminated through a kNN classifier with k ideal set to 3. A risk of misclassification nonetheless remain leading to a possible confusion between stratiform clouds and rain clouds, which both scatter sunlight differently. A solution proposed by Katzantzidis consists in introducing 3 more criteria and subclasses. However, considering the effect of shading on PV systems, a sensor network remains the ideal option as it measures directly the area and shape of the shadow cast by clouds on the surface. For this reason, such networks are often utilized as reference data. Pyranometer networks are precise but expensive although it is possible to use an existing PV network to cut costs. The size and configuration of the network must nonetheless match with the required spatial resolution. Independently from the source of input data, the altitude of the base of low altitude clouds, which are responsible for the fastest variations, can be assessed from the ambient temperature, the dew temperature and the elevation of the weather station. The full topic is covered in further details within Chapter 4.

## **2.9 WEATHER CLASSIFICATION FOR MID-TERM PV OUTPUT FORECASTING**

### **2.9.1 Review of Weather Classifications**

For many utility applications including power dispatch and grid balance, a short-term to long-term horizon suffices, as previously mentioned in Table 2.4. In this instance, PV power forecasting methods consist in focusing on a period of time that ranges between 30 minutes and several days ahead. Such methods may mix statistical and physical approaches with either PV networks or numerical weather prediction (NWP) as an input. Peder Bacher et al [112] have demonstrated that direct solar power observations are the most appropriate source of input data for forecasts up to 2 hours ahead whereas NWP is better suited beyond the horizon of 2 hours.

Weather classification is widely used for forecasting purposes for horizons exceeding 24 hours. For example, a day-ahead or week-ahead PV power output profile can be obtained with a classification of simple daily weather patterns such as rainy, cloudy, sunny (clear sky) and foggy days. Historical data of solar irradiance (for an indirect forecasting approach) [81] or PV output (for a direct forecasting approach) [60, 123] and meteorological forecasts are used as input data on this purpose.

These two frequently used PV output forecasting methods can play a pivotal role to size the storage batteries required to complete the PV system under consideration. As an illustration, the case study of a sun-powered house in the Tokyo area, Japan, within the government supported smart city project, as described by Niimura et al (2012) the text weather forecast from 9 years of historical data has been used to establish the classification [81]. Table 2.15 is an example of how the weather conditions can be reinterpreted in simple terms, based on various key words that may appear within the descriptive text of a weather forecast. Such a simplification of weather patterns is very suitable to forecast the power generation from a day to a week ahead. The hourly profile of insolation is obtained based on the values that are statistically the most likely considering the time and weather conditions. The most likely

insolation values from 4 am to 8 pm under various weather conditions are obtained from the probability distribution of insolation, which is based itself on hourly solar radiance data.

**Table 2.15 – Simple weather classification based on a textual forecast [49].**

<i>Key words</i>	<i>Classification</i>
<i>Clear</i>	
<i>Fair</i>	<i>Sunny</i>
<i>Slightly overcast</i>	
<i>Heavily overcast</i>	
<i>Cloudy</i>	<i>Cloudy</i>
<i>Rain/Snow</i>	<i>Rainy</i>
<i>Fog/Mist</i>	<i>Foggy</i>

The estimated radiation of a specific day and at a specific hour  $t$  is then defined based on the weather forecast by  $g(t) = G_{max} \cdot g_{most}(t) \cdot P_{precipitations}$  where  $G_{max}$  is the maximum monthly insolation,  $g_{most}(t)$  is the most likely normalized value of solar radiation at the hour  $t$  and  $P_{precipitations}$  is a factor that indicates the probability of rain. Finally, in that case study, the maximum seasonal Mean Absolute Error (MAE) of the forecast is about 0.33 for seven consecutive days. The major prediction errors occur during spring and summer in the northern hemisphere. The authors conclude that the major of the errors cause is the 80% of uncertainty of weather forecasts.

A support Vector Machine can be used for the same purpose of weather classification based PV power output forecasting for a one-day horizon. For example, Jie Shi and al. [60] have selected a simple weather model out of four possibilities (sunny, cloudy, rainy, foggy) on the basis of historical data and weather forecast of the next day. The historical input data are similar to what has been used by Niimura. The performance is best for sunny days with mean Relative Error (MRE) of 4.85% whereas the MRE is 12.42% for cloudy days, with an average of 8.64% for all models.

### **2.9.2 Conclusions on Weather Classification**

For the purpose of power dispatch, SVM and weather classification prove to be extremely accurate for sunny or uniformly overcast days. Such an approach lacks precision for cloudy days though, because of the dynamic variation of the cloud cover. It can be applied when an average daily value of the solar radiance is sought but some significant errors may be expected during cloudy days.

## **2.10 VERY-SHORT TERM PV POWER FORECASTING**

Accurate very-short term PV power output forecasting methods are required for safety and efficiency reasons. In particular, utility operators require a ten-minute warning to bring spinning reserves online. Besides ramp rates issues, an accurate forecast within a solar-aware smart grid can also help to prevent keeping industrial-scale loads such as water-pumps on when not needed. Additionally, a precise prevision enables to plan when to turn such loads on. Nevertheless, the main problem with very short-term PV power forecasting is the swift intermittency of sun radiation, which is essentially due to cloud motions. Tracking and forecasting such cloud movements is therefore necessary in order to deal with this problem. However, when it comes to deterministic prediction of PV output power, adequate temporal and space resolutions are critical [11, 33]. A first group of studies have favoured the use of a numerical weather prediction (NWP) model whereas a second group prefers to use sky imagers, or geographically dispersed sensors or even satellite images to track clouds. A third group utilizes a random-sequence and time series models.

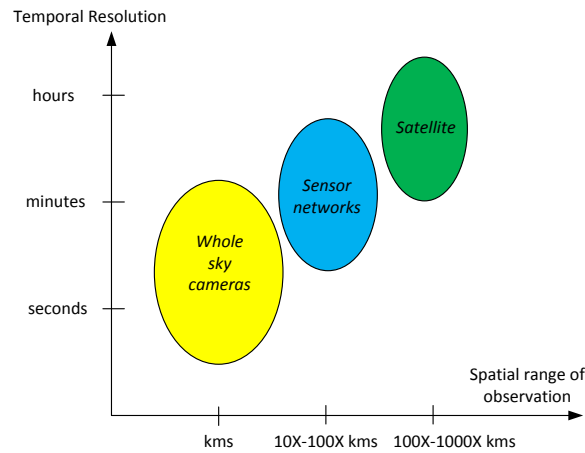
### **2.10.1 NWP Approach**

NWP models can provide very valuable outputs such as forecasts of cloudy conditions, wind direction, wind speed and even sometimes irradiance with a large spatial and temporal resolution. Such a resolution fits well with the day-ahead forecasts needed to determine energy pricing or for energy dispatch purposes. However, the accuracy of such an approach is quite variable. In the case of a weather research and forecasting (WRF) numerical atmospheric model that would be made 30 hours in advance, the best result seem to be obtained in the occurrence of large-scale cirrus clouds or in the event of an entirely covered skies. Such a model may otherwise sometimes fail to predict the presence of clouds during a day or be mistaken in the timing of clouds. Overall, this approach is not suited to deal with small clouds in a partially cloudy day. Additionally, substantial computation time is required for NWP models. Indeed, 2 hours of computation may be needed to provide only 2 minutes of forecast [38].

### **2.10.2 Sky, Satellite Images and Sensor Array Approach**

For studies that use image processing and sensor analysis, three variables must be taken into account when tracking clouds: position, velocity (as a vector) and size [124].





**Figure 2.8 – Temporal resolutions compared to the spatial range of observation for various types of input data.**

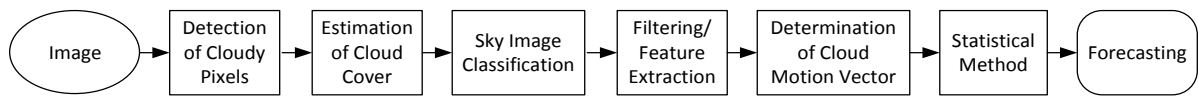
A few bottlenecks faced by this approach are respectively the effectiveness of the feature extracting filter and the associated extensive computation time. Additionally, time resolution may sometimes be too coarse for a reliable forecast of power variation caused by clouds. This is because cloud-induced fluctuations of power output happen in matter of seconds. Figure 2.8 gives a rough idea of the temporal resolution with respect to the area that is observed for three different categories of input data. The range of observation must not be confused with the spatial resolution. On the other hand, the spatial resolution may vary a lot among each type of input data source and the format of the spatial resolution may differ too much to make comparisons. For instance, sky imagers are commonly used to observe the base of clouds and are characterized by a resolution expressed in pixels. In contrast, satellites observe the top of clouds or the ground with a precision expressed in metres or kilometres. For their part, neighbouring PVs or sensors of weather stations may often be dispersed in an irregular manner.

An important challenge is to determine the velocity of clouds. It might be challenging because ground-based measurements of the wind are not reliable for this purpose [38]. Indeed, the wind direction and speed may vary according to the altitude. The only relevant wind measurement would need to be done at the level of clouds.

### **2.10.2.1 Sky Images**

In most cases, a fish-eye lens is mounted on whole sky cameras to give a panoramic field view of 180 degrees in every direction, so that the entire sky can be observed. However, the spatial resolution is limited to a local range consisting in a radius of a few kilometres around the camera. On the other hand, the time resolution of whole sky cameras is excellent as the time between two consecutive shots can be less than a second, which virtually enables real-time observations. This basically gives a significant advantage to whole sky imagers for a detailed examination of the cloud cover.

Additionally, sky cameras can be used to identify the types of clouds and small-scale variations of cloud cover over a specific spot. Combined with a solar radiation sensor, this enables to uncover the effect of specific types of clouds on the insolation.



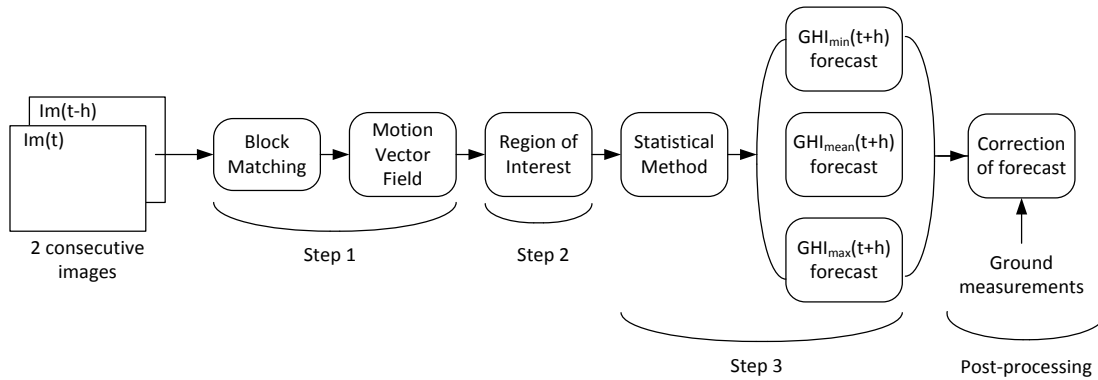
**Figure 2.9 – Typical sky imaging processing flow.**

A 1-second sampling period is a popular choice for taking and storing snapshots of the sky [125] but slower periods ranging from 15 seconds [58] to up to 5 minutes [30, 59, 126, 127] are also considered. The resolution of sky camera images vary within the range of 8 to 30-bit JPEG 640×480 [30, 59, 125, 127], 512×512 [126], 1024×1280 [28], or up to 3648×2736 pixels [58]. One or two axis sun trackers as well as sun-masking techniques can be sometimes added. It must be noted that processing colour images is an important step to classify clouds. Colour [126] or infrared [127] filters can possibly be applied on the camera. In any case, due to the high amount of data, sky images must be automatically classified with high accuracy. The classification is usually either based on a binary clear/cloudy basis or on an identification of cloud types. Various filtering methods are applied to the images after classification. The extraction of features lead to the determination of a cloud motion vector including cloud velocity, which is used in turn for prediction purposes. An example of complete processing flow is given in Figure 2.9. It must be noted that the methods used in each step of the flow may vary, especially the statistical method used prior to forecasting.

Optical flow based on Lucas-Kanade method has been successfully used by Wood-Bradley et al [125] to process whole sky images with a prediction error from 5.3% (1 second off) to 21.2% (7 seconds off) in the estimation of the obstruction time of the sun by a cloud, with times to shading ranging from 20 to 40 seconds. Based on JPEG images, the method follows a flow of cloud detection, cloud movement tracking, cloud movement forecasting, and image processing through OpenCV functions. Unfortunately, the excessive computational time of this method may not permit to implement it online.

### **2.10.2.2 Satellite Images**

Satellite images offer a wide spatial range stretching over a few thousands of kilometres. The time resolution of forecasts based on velocimetry from satellite images is often comprised between 1 and 6 hours, which however may not be well suited for the observation of small-scale or fast cloud motion [38].



**Figure 2.10 – Satellite images processing flow.**

As an example, forecasts based on GOES satellite images have a 10 km spatial resolution and are updated every hour. Because of their limited temporal resolution, satellite-based forecasts can sometimes be outperformed by forecasts based on persistence models [124]. Additionally, low-resolution large-scale cloud observation from satellites does not permit cloud-radiation studies, e.g. determining the impact of a specific type of cloud on the solar irradiance reaching the ground. Nevertheless, satellite images are very useful to forecast Global Horizontal Irradiance (GHI) on a relatively short-term basis, for horizons shorter than 6 hours [128]. In particular, for up to 5 hours ahead GHI forecasts, satellite derived methods outperform considerably NWP based forecasts. It has even been demonstrated that satellite based methods are the best option for a temporal horizon comprised between 30 minutes and 6 hours [129]. This makes satellite images valuable to maintain dynamically the balance between solar energy supply and consumption. Satellite observation generally involve cloud motion vector field extraction from the images.

Some high quality services such as EUMETCAST [130] can deliver satellite images on a 15-minute basis for Europe, Africa and South America, which is an excellent temporal resolution. In particular, Dambreville et al (2014) have used an estimation of the GHI forecast provided by HelioClim-3 images with a high resolution of 3 km × 3 km per pixel sampled every 15 minutes. In comparison, Hammer, Heinemann, Lorenz and Lückehe (1999) could observe cloudiness with 30 minutes precision and a resolution of 2.5 km × 2.5 km from METEOSAT satellite. Based on a year-long database, Dambreville et al have used a 3-step method, which consists in the following. Firstly, a block matching algorithm is used to extract the cloud motion vector field after comparing two consecutive satellite images  $Im(t-h)$  and  $Im(t)$  where  $t$  is the time and  $h = 1$  corresponds to a 15 minutes sampling period. This cloud motion vector field represents the global movement observed between the images. As a second step, a region of interest (ROI) is found in the upstream direction. A third step consists in using statistical method or trained conversion within the ROI to define a minimum, a maximum and a mean forecast of the GHI at time  $t+h$ . Finally, the forecast is compared to ground measurements and corrected if needed. This flow is illustrated in Figure 2.10.

During the first step, two consecutive images are divided into square-shaped pixels blocks characterized by a side  $b = 20$  pixels. A research horizon  $r_h = 10$  pixels (corresponding to a wind speed of 120 km/h) is defined as the furthest possible translation (diagonally) of one block between the first and the second image. A block in the second image  $Im(t)$  is identified as the same block from the first image  $Im(t-h)$  through a maximum correlation coefficient. The vector of each block is defined by the difference of the coordinate of the block centre between both images.

The global motion vector is drawn from the average of the vectors within the 9 closest blocks from the site that is considered.

The region of interest is calculated on the basis of a window size expressed by:

$$w_s = 1 + 2 \cdot T_s \quad (2.31)$$

In equation (2.31),  $T_s$  is the sampling period.

The atmospheric transmission  $k(t) = GHI(t,p)/GHI_{clearSky}(t,p)$  at the instant  $t$  is defined as the ROI's GHI normalized with respect to the clear sky  $GHI(t,p)$  at the same time  $t$  and for the same pixel  $p$ , based on the European Solar Radiation Atlas model. The value is defined as  $k = 1$  in clear sky conditions and  $k = 0$  in case of complete obstruction. The average forecast value for  $k$  is used to get the average forecast GHI, on the basis of a wind persistence hypothesis. The average forecast GHI is:

$$\overline{GHI}(t + T_s, p) = \bar{k}(t + T_s) \cdot GHI_{clearSky}(t + T_s, p) \quad (2.32)$$

Similarly, the minimum and maximum values of the forecast  $k$  lead to a minimum and maximum forecasted GHI. However, comparisons with the ground measurements obtained with a sensor network show systematic errors. This is why a post-processing stage is introduced for correction. The performance of the forecast has been assessed against ground measurements. With an nRMSE between 23.5% and 30%, the corrected forecast outperforms slightly the persistence model for horizons between 15 and 45 minutes.

Variability poses a serious tracking and forecasting problem. As clouds are not plastic but continuously vary in shape and number, cloud tracking can be seen as a multi-target tracking (MTT) problem. A framework able to deal with variable number of targets is thus required. A probability hypothesis density filter (PHD) filter has been successfully applied to track cloud features (such as position, size and motion vectors) by Paolo Adesso, Roberto Conte, Maurizio Longo, Rocco Restaino and Gemino Vivone [124].

To complete the flow, a time-dependent penalty term feeds a Maximum a Posteriori Markov Random Fields (MAP-MRF) algorithm used for classification. Features are extracted after classification and used to update the multi-target state. The penalty term considers the classification of previous acquisitions for the current image classification, reducing the risk of misclassification. The purpose of such a method is to reach a trade-off between classification accuracy and reduced computation. The entire flow except the MAP-MRF takes less than a second to compute on single-core 3.4 GHz Intel CPU for each image. Nevertheless, the MAP-MRF takes 82 seconds to run in the same conditions. One of the main benefits is an increased classification accuracy compared with other methods. In particular, the accuracy has been significantly improved around the cloud borders, which often tends to generate a high misclassification rate, due to the fact that edges fade into the background with low contrast.

### **2.10.2.3 Sensor Network**

Geographically dispersed sensors or a largely distributed PV network can be another source of observation data for cloud cover. As it enables direct measurements, it is often used as a reference to assess the performance of an irradiance forecasting method. It provides with an intermediate and probably complementary spatial and temporal resolution with respect to sky imaging and satellite views. In particular, this approach offers a better resolution than forecasts based on satellite images. Indeed, the distributed range of the 83 residential rooftop PV network used by Jayadevan, Lonij, Cronin et al [38, 83, 131] covers an area of  $50 \text{ km} \times 50 \text{ km}$  with an average spacing of 3 km between two neighbouring sensors. It records average AC power normalized by the system rating at 15-minutes intervals. A disadvantage of dealing with such a sensor network is the installation cost.

This problem can be tackled by using an existing infrastructure, as Lonij et al did. Unlike with sky or satellite images, the sun radiation is directly measured. There is thus no need to know about the optical properties of clouds or to use radiative models [38]. The infrastructure in question consists in a network of residential PV modules. The main limitation in precision is due to the relative small size of the network compared with time resolution. The sampling rate is also too low. Moreover, the same publication mentions that cloud edge velocity may not always be the same as wind velocity, which may lead to significant errors as a result. It must also be noted that Weather Research and Forecasting (WRF) numerical atmospheric models are better at forecasting the impact of slowly varying cirrus (high altitude, 9 to 12 km) clouds for up to 50 hours ahead whereas PV network is better to forecast the effects of quickly varying cumulus (low altitude, up to 5 km) clouds.

### **2.10.2.4 Hybrid Solutions**

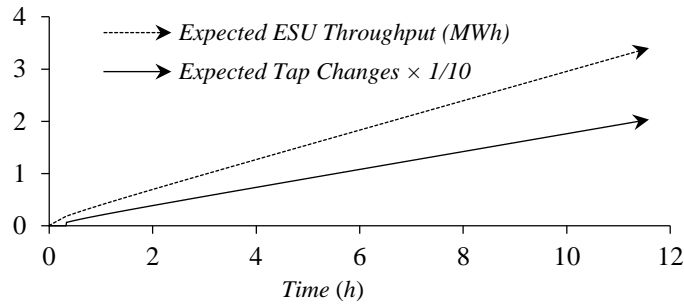
Satellite and sky images can be combined to optimize the resolution of nowcasting of clouds in near-medium future. This is a solution used by Gonzalez et al [127] by interpolation. This team also use an

ANN to detect clouds. A k-nearest-neighbour (kNN) algorithm is often used to classify cloud types [58].

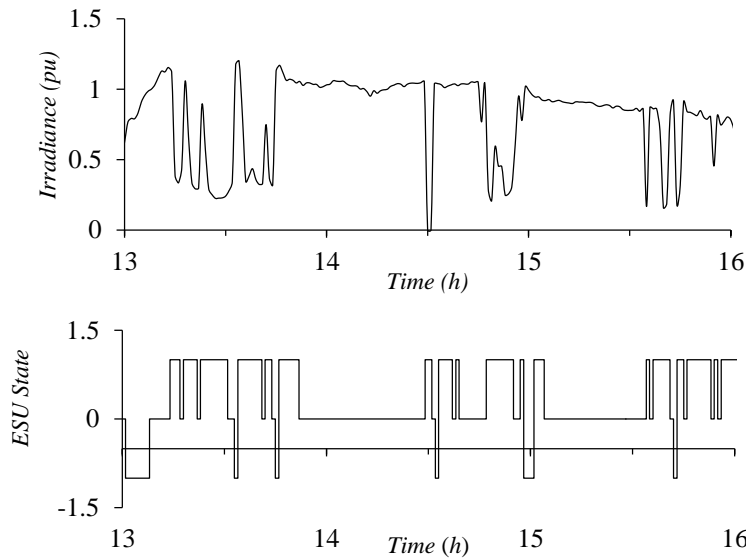
Even all three approaches (NWP, satellite and sensor network) can be combined together to optimize even further the overall temporal and spatial resolution of the forecast. This is done by Lonij, Cronin et al [38] to perform a 45-minute forecast.

### 2.10.3 Random-sequence and Time Series Models

Considering the third group, the methodology involves acquiring and processing power output data commonly coming from a couple of 200-300 W PV panels with MPPT ability as well as irradiance data fed by a sensor like a pyronometer. Sampling these data at a 1-second interval is the most popular practice [11, 33]. However, it must be noted that high solar angles of incidence around sunrise time and sunset time usually result in a poor accuracy of the measurement of the power output, leading to corrupted samples. Nevertheless, little generation is expected during those times. Considering a dataset of valid samples, a threshold is usually defined to separate clear and shaded periods. Cai et al [33] consider the irradiance to define a constant threshold of 85% of the mean direct normal irradiance (DNI)  $G_{clear,avg}$ . Any period of time  $T_{shaded}$  characterized by a DNI below this threshold can be considered as shaded. With the help of a sky camera, it has been observed by this team that the most dramatic variations of the average DNI level  $G_{shaded}$  normalized with respect to  $G_{clear,avg}$  occur during short shaded periods of time, shorter than 200 s. Within these short shaded periods, the normalized irradiance may vary in range from 5% to 90% of the mean  $G_{clear,avg}$  DNI. With the help of a sky camera, it has been established that shaded periods longer than 200 s may be due to opaque and slow-moving cumulus clouds. Similarly, Barnes et al [11] also classify the clear-sky irradiance profile into clear and shaded data by using a classification algorithm. This algorithm consists in performing a second-order curve fit on the irradiance profile completed by normalization of the irradiance profile and K-means clustering with two centroids. The probability for a PV generator to be shaded or fully insolated in the future is a crucial information to size the energy storage unit (ESU) needed to mitigate the effect of sunlight disruptions. It must be noted that only the discharge power of the ESU is taken into account to assess its size. The likelihood for a PV generator to be shaded is calculated on the basis of the time series of binary states (either clear or shaded) considering the duration of each state. A semi-Markov discrete-time process has been modelled and applied. The maximum number of tap changes per day as well as the minimal clear-sky output power that is required to induce a tap change need to be established based on the lifetime of system elements and load-flow analysis.



**Figure 2.11 – Expected number of tap-changing operations and ESU throughput with respect to time.**



**Figure 2.12 – Irradiance vs. time and corresponding ESU state vs. time. A state of +1 indicates discharging, -1 indicates charging, and 0 indicates standby [12].**

For the specifically studied day, ESU throughput has been established as 2.625 MWh with 20 tap-changes per day. Nonetheless, this method is used by Barnes et al for selecting an offline control strategy. The authors would recommend to consider a live feed of weather data to improve the performance. The results presented by the same authors related to this method are shown in Figure 2.11 and Figure 2.12 [11] (courtesy of A.K. Barnes, Dr. J.C. Barda and J.K. Hayes, university of Arkansas).

#### **2.10.4 Conclusions on Very Short-Term PV Forecasting**

An optical flow based forecast is very accurate but takes too long to compute. Hence, such a method can only be used off-line. Since the spatial resolution plays a pivotal role in irradiance forecasting, sky cameras can provide with the ideal space resolution to define very-short term cloud motion vectors. Sky cameras are also needed for cloud classification. A persistence approach also proves to outperform satellite-based methods. However, sensor networks are the reference source of input data since they can measure directly the irradiance and do not require complex cloud models. Such networks are nonetheless costly and may not have the adequate temporal and spatial resolution to track clouds

precisely. Consequently, sky images sampled every second combined with a sensor network with a refresh of data every few minutes is reported to be an ideal combination for an optimal accuracy of irradiance forecasting. Indeed, the correlation of both systems enable to uncover the effect of shading of each type of cloud. Satellite images may also considered in addition to increase the spatial resolution and to observe the top of clouds.

## **2.11 CONCLUSIONS ON THE LITERATURE REVIEW**

As of now, PV power remains not fully predictable and hence relatively expensive and difficult to dispatch. However, a few approaches can be envisaged to improve the possibility to dispatch PV energy, to optimize its generation and to plan its storage, making it a more reliable energy resource. In particular, controlling the energy storage can mitigate the discontinuity of PV generation over various periods of time. However, an accurate forecast of the PV power output is necessary to control a system containing PVs, energy storage and additional machines. The temporal and spatial definition of the system must be defined as a first step, accordingly with the intended grid operation. In particular, to guarantee power continuity and safety, a very fine resolution is needed. Nonetheless, various approaches can address the problem, the persistence model tends to be the most accurate approach for very short-term forecasting. ANNs and SVM are appropriate approaches for short-term horizons while NWP is better suited for longer horizons. Although a probabilistic method based on historical data may be valuable for very long term forecasts, such an approach cannot take into consideration the complex variations of the cloud cover causing short-term sunlight disruptions. Only a deterministic atmospheric modelling approach can deal with the stochastic changes of solar radiance during the day. Within this type of model, NWP data-based models are well adapted for day ahead forecasts but suffer from a too coarse temporal resolution. On the other hand, sky imagers, geographically dispersed sensors or satellite images are excellent sources of data for cloud identification and tracking. Satellite images give a large-scale overview of cloud events, possibly with a good resolution. Sky imagers are a precious tool to identify cloud types and anticipate the impact of the shading on PV power generation. A hybrid system combining at least a sensor network and a sky camera performs better than separate systems. Even combining all three solutions can outweigh the drawbacks of all individual solutions. Low altitude clouds like cumulus clouds may cause fast occurring power discontinuity whereas thick rain clouds like nimbostratus clouds may cause long-lasting and severe power outage. Finally, it has been found that an offline observation of the effect of cloud-related shading on solar power generation over an adequate period of time can help to calculate the size of the required energy storage. This can notably be successfully done by using random sequence and clear/shaded classification. Some further improvements can be sought in the future. In particular, it would be valuable to elaborate algorithms that can calculate cloud coverage and classify clouds using online data and a fine sampling period. In



addition, measuring precisely the effects of each type of cloud on the solar irradiance could greatly help. Finally, finding a method do distinguish reliably rain clouds from stratiform clouds would be valuable.

## 3 IRRADIANCE, CLOUD, SUN AND SHADOW MODELLING

---

A comprehensive state of the art of solar power forecasting has been presented in the previous chapter. This review has helped define the whole strategy to achieve to forecast solar power output through cloud tracking. The overall strategy is presented in the next section. In that strategy, it can be seen that the clear sky irradiance plays a very important role since the accuracy of the clear-sky irradiance determines the accuracy of the whole forecast. Thus, we will focus first on the solar radiation models in the absence of clouds. In particular, we will review some astronomical considerations that determine the position of the sun and the extra-terrestrial beam intensity. Since our application focus on very short term solar forecasting, it makes sense to consider the fastest irradiance models to compute, which are empirical models. We will thus compare a large number of simple irradiance empirical models with one another, as well as with irradiance measurement made at the location of our PV site. We will also study how to model clouds and their cast shadow. In particular, we will see how clouds appear through the lens of a sky camera installed on the ground. In addition, we will calibrate and characterise the lens of a sky camera that has been used for this study.

It must be noted that this chapter is strongly based on a publication presented during the IEEE Power and Energy Society General Meeting that took place in Boston, USA, in July 2016 [2].

### 3.1 RELEVANCE OF THE OVERALL STRATEGY

To mitigate power discontinuities, a BESS needs to be charged when the generation exceeds the demand and to be discharged in the opposite case, through the use of a power converter and a controller. It is safe to say that the effectiveness of the power smoothing method depends on the reactivity of the BESS controller. A forecast of the sky conditions, and thus of the irradiance, may greatly help. Since solar irradiance is the physical quantity that influences the most the output power of PV arrays, the precision of the irradiance forecast may in turn dictate the accuracy of the power output smoothing [29].

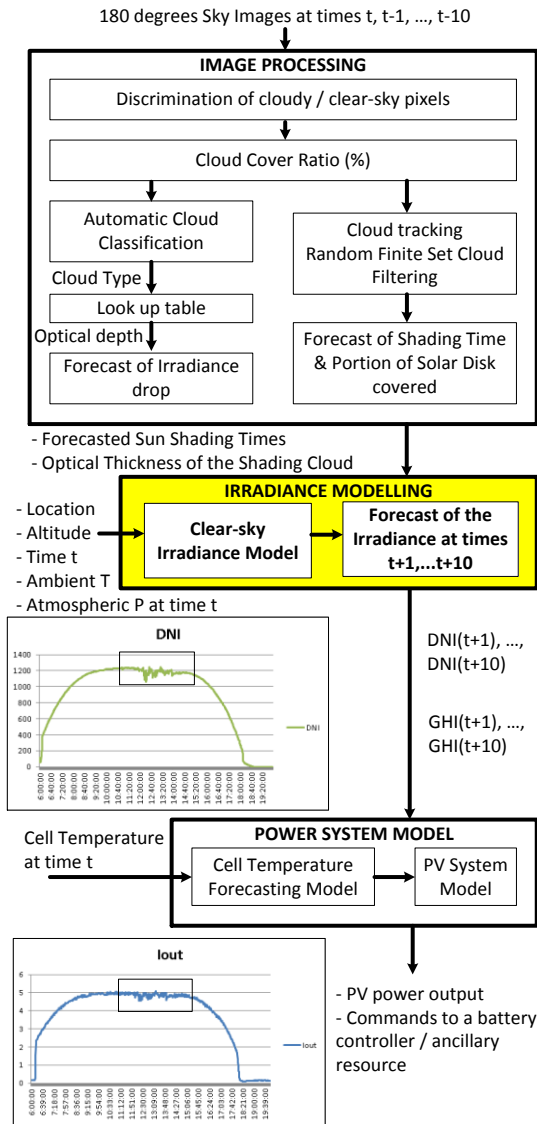


Figure 3.1 – Flowchart of an ensemble forecasting method.

The overall strategy intended for power discontinuity mitigation is presented in Figure 3.1. It and uses a sky camera to observe the cloud cover. The strategy basically consists in three main parts there are related to the processing of sky images, forecasting of irradiance, and modelling of PV system.

### 3.2 CLEAR-SKY IRRADIANCE MODELLING CONSIDERATIONS

The stochastic variations of global horizontal irradiance cannot be dealt with without a proper model of the clear sky solar radiation. Astronomical considerations are necessary to implement this model [7]. Yet, such considerations relate mainly to cyclic variations. The calendar day number of the year, referred to as  $n$  is a fundamental parameter. The number  $n$  hence represents how many days have elapsed since the 1<sup>st</sup> of January of the current year.

For any day of the year, the declination of the Earth  $\delta$  can be simply expressed with Equation (3.1):

$$\delta = 23.45 \cdot \sin \left[ \frac{360 \cdot (n - 81)}{365} \right] \quad (3.1)$$

Another expression of the declination of the Earth  $\delta$  has been formulated by in the form of Fourier series by Spencer in 1971 as given in Equation (3.2) [99].

$$\begin{aligned} \delta = & 0.006918 - 0.399912 \cdot \cos(T_d) + 0.07257 \cdot \sin(T_d) - 0.006758 \\ & \cdot \cos(2 \cdot T_d) + 0.000907 \\ & \cdot \sin(2 \cdot T_d) - 0.002697 \cdot \cos(3 \cdot T_d) + 0.001480 \cdot \sin(3 \cdot T_d) \end{aligned} \quad (3.2)$$

$T_d$  is the angular fraction of the year calculated by Spencer as per equation (3.3).

$$T_d = \frac{2 \cdot \pi \cdot (J_{day} - 1)}{365} \quad (3.3)$$

In equation (3.3),  $J_{day}$  is the Julian ephemeris day. The Julian calendar starts on the 1<sup>st</sup> of January – 4712 at Greenwich mean noon and is calculated as per equation (3.4) according to the Gregorian day with decimal time  $t_{dd}$ , month  $t_{mth}$ , and Gregorian year  $t_{yr}$ .

$$\begin{aligned} J_{day} = & [365.25 \cdot (t_{yr} + 4716)] + [30.6001 \cdot (t_{mth} + 1)] + t_{dd} + 2 - \left\lfloor \frac{t_{yr}}{100} \right\rfloor \\ & + \left\lfloor \frac{t_{yr}}{400} \right\rfloor - 1524.5 \end{aligned} \quad (3.4)$$

The Julian day is adjusted from the local time zone  $h_{tz}$  to Greenwich Mean Time expressed as per equation (3.5).

$$J_{day,adjusted} = J_{day} - \frac{h_{tz}}{24} \quad (3.5)$$

The altitude angle of the sun  $\beta$  is, expressed in degrees:

$$\beta = \sin^{-1}[\cos(Lat) \cdot \cos(\delta) \cdot \cos(H) + \sin(Lat) \cdot \sin(\delta)] \quad (3.6)$$

The zenith angle  $z$  is the complementary angle of the altitude angle  $\beta$ :

$$z = \cos^{-1}[\cos(Lat) \cdot \cos(\delta) \cdot \cos(H) + \sin(Lat) \cdot \sin(\delta)] \quad (3.7)$$

In equation (3.6) and (3.7),  $\delta$  is the declination of the Earth,  $H$  is the hour angle and  $Lat$  is the latitude.

$$H = (SolarTime - 12) \cdot 15 \quad (3.8)$$

$$\text{SolarTime} = \text{LocalTime} + 4 \cdot (\text{StandardMeridian} - \text{LocalMeridian}) + \text{EoT} \quad (3.9)$$

$$\text{EoT} = 9.87 \cdot \sin(2 \cdot x) - 7.53 \cdot \cos(x) - 1.5 \cdot \sin(x) \quad (3.10)$$

$$x = \frac{360}{365} \cdot (n - 81) \quad (3.11)$$

The hour angle  $H$  (3.8) requires a conversion of the civil time into solar time (3.9) by the use of the equation of time (3.10) and (3.11). The meridians are expressed in degrees. The local time meridian is defined according to the time zone of the considered location. For example, the time zone for Perth, Western Australia, is UTC+8 and the corresponding local meridian is 120° East. The coefficient 4 min/degree features the rotation speed of the Earth. Hence, the solar noon can be deduced from the previous equation if the solar time is 12 PM.

The hour angle is defined as  $H = 15^\circ \cdot h$ , where  $h$  is the decimal number of hours before solar noon, so that  $H$  is equal to + 15° at 11:00 A.M. solar time and  $H = - 15^\circ$  at 1:00 P.M. solar time.

The altitude angle of the sun  $\beta$  might need to be corrected, as mentioned by Annear and Wells in their comparison of clear-sky solar radiation models [99]. An atmospheric refraction term is added, as shown in equation (3.12).

$$\beta_{corrected} = \beta + CF \quad (3.12)$$

The value of this correction term depends on the value of the altitude angle  $\beta$  as per Table 3.1.

**Table 3.1 – Values recommended by the NOAA (2004) for the atmospheric refraction correction [17].**

$\beta$	Correction factor $CF$
85° to 90°	0
5° to 85°	$\frac{1^\circ}{3600} \left[ \frac{58.1''}{\tan(\beta)} - \frac{0.07''}{\tan^3(\beta)} + \frac{0.000086''}{\tan^5(\beta)} \right]$
-0.575° to 5°	$\frac{1^\circ}{3600} \left[ 1735'' - 518.2'' \beta + 103.4'' \beta^2 - 12.79'' \beta^3 + 0.711'' \beta^4 \right]$
< -0.575°	$\frac{1^\circ}{3600} \left[ \frac{-20.774''}{\tan(\beta)} \right]$

At a given hour of the day, the solar azimuth angle is defined in degrees by Equation (3.13).

$$\varphi_s = \sin^{-1} \left[ \frac{\cos(\delta) \sin(H)}{\cos(\beta)} \right] \quad (3.13)$$

It must be noted that  $|\varphi_S| > 90^\circ$  unless  $\cos(H) \geq \frac{\tan(\delta)}{\tan(Lat)}$ , in which case  $|\varphi_S| \leq 90^\circ$ .

At solar noon, the altitude angle of the sun becomes  $\beta_N$ , expressed in degrees ( $^\circ$ ) by Equation (3.14).

$$\beta_N = 90^\circ - Lat + \delta \quad (3.14)$$

In the case of fixed orientation, the PV collector tilt angle should be set up according to  $\beta_N$ :

$$Tilt = 90^\circ - \beta_N \quad (3.15)$$

There are few alternative ways to determine the position of the sun. A very exhaustive report has been published by Sandia National Laboratories [97] to describe clear sky models of various complexities and precisions. The report mentions that NREL has developed the most accurate Solar Position Algorithm (SPA) to date, which can be used as the current reference. Based on location, date, and time inputs over a period of 8000 years, this algorithm show as little uncertainties as  $\pm 0.0003$  degrees and has been described by Meeus [114] as well as by Reda and Andreas [113]. NREL has provided an online calculator version of this algorithm [132].

Some parameters are often required to compute these empirical models. Among them, there is the extraterrestrial irradiance  $I_0$  that reaches the outer surface of the Earth's atmosphere may be expressed through a Fourier series [133]:

$$I_0 = I_{SC} \cdot [1.00011 + 0.034221 \cdot \cos(x) + 0.00128 \cdot \sin(x) - 0.000719 \cdot \cos(2 \cdot x) + 0.000077 \cdot \sin(2 \cdot x)] \quad (3.16)$$

The solar constant  $I_{SC}$  is commonly used with the value  $I_{SC} = 1366.1 \text{ W/m}^2$ . The quantity  $x$  has been previously defined in Equation (3.11).

The air mass ratio is another important parameter. The expression used by the Laue and Meinel models is the simple following:

$$AM = \frac{1}{\cos(z)} \quad (3.17)$$

This expression of the air mass ratio is nonetheless not unique as many others exist [97]. For example, the expression  $m$  formulated by Masters [98] for the TJ model is another possible model among many others for the air mass ratio:

$$m = \sqrt{(708 \cdot \sin(\beta))^2 + 1417} - 708 \cdot \sin(\beta) \quad (3.18)$$

The Sun-Earth distance varies slightly by a factor of  $\pm 1.7\%$  throughout the year. This is because the trajectory of the Earth is not perfectly circular but slightly elliptical. The distance can be expressed in km in function of the calendar day number with Equation (3.19):

$$d = 1.5 \cdot 10^8 \cdot \left\{ 1 + 0.017 \cdot \sin \left[ \frac{360 \cdot (n - 93)}{365} \right] \right\} \quad (3.19)$$

As a matter of fact, the extra-terrestrial radiation emitted by the sun follows a few cycles. As per observation ranging from 1699 to 2008, some sunspot activities occur every 11 years in average, to which some rapid irradiance variations with even larger amplitude are superimposed [9]. The extra-terrestrial solar insolation is defined as the solar flux from outer space entering radially through the atmosphere, on any point at the top of the atmosphere. When a sunspot peak happens, the extra-terrestrial solar irradiance may increase by about 1.5% [7]. By neglecting the effect of sunspots, the extra-terrestrial solar radiation varies by about 6.8% during a year and can be expressed in  $\text{W/m}^2$  by:

$$ET = SC \cdot \left[ 1 + 0.034 \cdot \cos \left( \frac{360 \cdot n}{365} \right) \right] \quad (3.20)$$

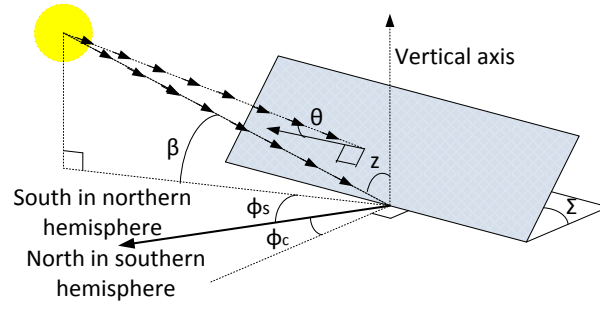
Where  $SC$  is the solar constant and  $n$  is the day number during the year.

The value of  $SC$  is actually not constant and varies with time. This value fluctuates about  $\pm 1 \text{ W/m}^2$  around a mean value during a typical 11-year cycle [134]. The latest and most accurate value is  $1360.8 \pm 0.5 \text{ W/m}^2$  as per the latest measurement performed by NASA SORCE (Solar and Climate Experiment) during the 2008 solar minimum period. This new measurement is much lower than the canonical value of  $1365.4 \pm 1.3 \text{ W/m}^2$  defined in the 1990s [9]. Often  $1366.1 \text{ W/m}^2$  is used as an average solar constant, as defined by Gueymard in 2004 [134]. During the same year 2004, NASA defined the average of all the minimum and maximum data values collected from the 1970s as  $1367.4 \text{ W/m}^2$  [99]. This is why the value of  $1367 \text{ W/m}^2$  may also be often used in literature.

The global horizontal irradiance (GHI) is the total radiation that reaches the surface of the Earth on a flat horizontal plane. It is composed of three components, namely the direct normal irradiance (DNI), the diffuse horizontal irradiance (DHI) and the reflective irradiance, which is often neglected.

$$GHI = DNI + DHI + \text{reflective} \quad (3.21)$$

The DNI is also often referred to as beam radiance, noted  $I_B$ . Likewise, the DHI may be also known as diffuse horizontal irradiance, which is noted  $I_{DH}$ . The DNI and DHI are given by the various models available in literature that will be covered further down in this chapter.



**Figure 3.2 – Incidence angle of direct sunlight on a fixed tilt and fixed orientation PV collector [4].**

Let us consider a single PV module as shown in Figure 3.2. The collector surface is oriented with a fixed angle  $\Phi_c$  with respect to north in the southern hemisphere (or south in the northern hemisphere) and with a tilt  $\Sigma$  with respect to the ground. If the sun is positioned with a zenith angle  $z$  and an azimuth angle  $\Phi_s$ , then the direct beam sunrays strike the collector with an angle  $\theta$  so that:

$$\cos(\theta) = \cos(90^\circ - z) \cdot \cos(\Phi_s - \Phi_c) \cdot \sin(\Sigma) + \sin(90^\circ - z) \cdot \cos(\Sigma) \quad (3.22)$$

The Global Collected Irradiance ( $GCI$ ) is the sum of the collected direct or beam component of light on the collector  $I_{BC}$  and the collected diffuse radiation  $I_{DC}$  (3) [98]. Using tracking devices enable to optimize the harvest of solar energy. However, the orientation of the PV collector with respect to the position of the sun must be carefully studied. If the direct beam of the sun strikes the PV collector with an incidence angle  $\theta$  with respect to a normal vector to the collector face and with an amplitude  $DNI$ , the global collected irradiance  $GCI$ , the collected beam irradiance  $I_{BC}$ , the collected diffuse irradiance  $I_{DC}$  and the reflected irradiance  $I_{RC}$  received by the collector respectively are:

$$GCI = I_{BC} + I_{DC}(I_{RC}) \quad (3.23)$$

$$I_{BC} = DNI \cdot \cos(\theta) \quad (3.24)$$

$$I_{DC} = DHI \cdot \left( \frac{1 + \cos(\Sigma)}{2} \right) \quad (3.25)$$

$$I_{RC} = \rho \cdot (I_{BH} + I_{DH}) \cdot \left( \frac{1 + \cos(\Sigma)}{2} \right) \quad (3.26)$$

The reflective component  $I_{RC}$  of the  $GCI$  is often neglected. This component depends on the ground reflectivity or albedo  $\rho$  as well as on the orientation of the PV collector. Typically,  $\rho$  ranges between approximately 0.1 for bitumen and gravel to 0.8 for fresh snow. The reflectance of grass is about 0.2. Considering a fixed collector with a tilt angle  $\Sigma$ , the reflective irradiance is defined as by equation (3.26).



As mentioned in the previous chapter, several models have been developed to model either or all of the irradiance components. There are basically two categories of irradiance models: the highly accurate physical models and the simpler empirical models.

### 3.3 EMPIRICAL IRRADIANCE MODELS

The irradiance of most empirical models is mainly based on the zenith angle  $z$  of the sun or its complementary angle, the altitude angle  $\beta = 90^\circ - z$ , respectively defined by (3.6) and (3.7). The zenith angle and the altitude angles are expressed in degrees in this studies. However, most of empirical models have been defined based on measurements made in the Northern Hemisphere, as it will be shown further down.

The clear-sky GHI may either be forecast through complex physical models or through simple empirical models. Physical clear-sky irradiance models present the advantage to be independent from the site location. However, they require many specific atmospheric input parameters. As an alternative, researchers may use the best fitting empirical models for the local conditions. This chapter aims to present an approach in comparing empirical GHI models that can be repeated and transposed to any other any location on the globe. The measured insolation in Western Australia has been used as the benchmark to compare a set of ten empirical models.

#### 3.3.1 Kasten-Czeplak (KC) Empirical Model

The Kasten-Czeplak (KC) GHI model [100] is based on continuous records made at Hamburg, Germany, between 1964 and 1973. The GHI is the following:

$$GHI = 910 \cdot \cos(z) - 30 \quad (3.27)$$

#### 3.3.2 Haurwitz Empirical Model

The Haurwitz GHI model [101, 102] has been designed on the basis of data gathered by the Meteorological Observatory of Harvard University in the northeast of the United States. The GHI is expressed like so:

$$GHI = 1098 \cdot \cos(z) \cdot e^{\frac{-0.057}{\cos(z)}} \quad (3.28)$$

#### 3.3.3 Berger-Duffie (BD) Empirical Model

Berger-Duffie (BD) GHI model [104] is based on data collected in France and is formulated as:

$$GHI = I_0 \cdot 0.70 \cdot \cos(z) \quad (3.29)$$

### 3.3.4 Robledo-Soler (RS) Model

The Robledo-Soler (RS) GHI model [106] has been defined on the basis of some data obtained during the period from 1994 to 1995 in Madrid, Spain. The GHI is expressed as a function of both the azimuth angle and its complementary altitude angle:

$$GHI = 1159.24 \cdot [\cos(z)]^{1.179} \cdot e^{[-0.0019 \cdot \beta]} \quad (3.30)$$

### 3.3.5 Laue Model

The Laue DNI model [107] calculates the DNI thanks to data collected in California. Besides the zenith angle  $z$ , the Laue model also takes the air-mass ratio  $m$  as per (3.32) as well as the site elevation  $h$  as inputs:

$$DNI = I_0 \cdot [(1 - 0.14 \cdot h)] \cdot 0.7^{m^{0.678}} + 0.14 \cdot h \quad (3.31)$$

$$m = \frac{1}{\sin(\beta)} = \frac{1}{\cos(z)} \quad (3.32)$$

The diffuse radiation used in this study is the same as for the Denashyar-Paltridge-Proctor model.

### 3.3.6 Meinel Model

The Meinel DNI model [108] can be seen as a simplified variant of the Laue model where the site elevation is considered to be at sea level or 0 m above sea level (ASL). The air-mass ratio  $m$  is the expression as with the Laue model (3.32).

$$DNI = I_0 \cdot 0.7^{m^{0.678}} \quad (3.33)$$

Like for the Meinel model, the diffuse radiation used in this study is the same as for the Denashyar-Paltridge-Proctor model.

### 3.3.7 Denashyar-Paltridge-Proctor (DPP) Model

The direct beam component of Denashyar-Paltridge-Proctor (DPP) model is based on some data collected across all of Australia [135]. However, the diffuse radiation of the DPP model as referred by Sandia is based on measurements made in Tehran, Iran, at a relatively high altitude (1191 m). The diffuse part may thus not be applicable to a low altitude location in Australia like Perth, Western

Australia [136]. However, in the absence of a more appropriate diffusion model, this expression has been currently retained.

$$DNI = 950.2 \cdot [1 - e^{-0.075 \cdot \beta}] \quad (3.34)$$

$$DHI = 14.29 + 21.04 \cdot \beta \cdot \left(\frac{\pi}{180^\circ}\right) \quad (3.35)$$

$$GHI = DNI \cdot \cos(z) + DHI \quad (3.36)$$

### 3.3.8 Adnot-Bourges-Campana-Gicquel (ABCG) Empirical Model

The Adnot-Bourges-Campana-Gicquel (ABCG) [105] is based on data collected in France.

$$GHI = 951.39 \cdot [\cos(z)]^{1.15} \quad (3.37)$$

### 3.3.9 Dazhi Empirical Model

Some empirical models can sometimes be a good fit for an area that is very different from its area of origin. This is for example the case of the Dazhi model, which is based on the ABCG model and which consists in a regression of measurements of the clear-sky GHI made in Singapore [109, 110]. The result of the regression is as follows:

$$GHI = 0.8298 \cdot I_0 \cdot [\cos(z)]^{1.3585} \cdot e^{-0.00135} \quad (3.38)$$

### 3.3.10 Threlkeld and Jordan (TJ) Empirical Model

The model developed by Threlkeld and Jordan (TJ) (1958) is based on data collected in a moderately dusty atmosphere with an average atmospheric vapour content for the United States [98, 103], as mentioned by Gilbert M. Masters [7].

#### 3.3.10.1 Direct Beam Radiation

$$I_B = A \cdot e^{-k \cdot m} \quad (3.39)$$

$$I_{BH} = I_B \cdot \sin(90^\circ - z) \quad (3.40)$$

Where  $I_B$  is the direct beam radiation, normal to the rays whereas  $I_{BH}$  is the horizontal beam solar radiance. The horizontal beam irradiance is defined as the solar beam radiation reaching the surface of the Earth on a flat horizontal plane.  $A$  is the apparent extra-terrestrial solar flux,  $k$  is the optical depth of the atmosphere and  $m$  is the air mass ratio. Both parameters have been defined by the American Society of Heating, Refrigerating and Air-conditioning Engineers (ASHRAE) [7].

$$A = 1160 + 75 \cdot \sin \left[ \frac{360}{365} \cdot (n - 275) \right] \quad (3.41)$$

$$k = 0.174 + 0.035 \cdot \sin \left[ \frac{360}{365} \cdot (n - 100) \right] \quad (3.42)$$

Recently, G.M. Masters has amended the expression of the air mass ratio with a more accurate equation that takes the spherical nature of the atmosphere into account [98]:

$$m = \sqrt{(708 \cdot \sin(\beta) + 1417)^2 + 1417} - 708 \cdot \sin(\beta) \quad (3.43)$$

### 3.3.10.2 Diffuse Radiation

Assuming the sky to be isotropic, the diffuse radiation on a horizontal plane on the surface of the Earth is proportional to the direct beam radiation:

$$I_{DH} = C \cdot I_B \quad (3.44)$$

$$C = 0.095 + 0.04 \cdot \sin \left[ \frac{360}{365} \cdot (n - 100) \right] \quad (3.45)$$

In equation (3.45), C is referred to as the sky diffuse factor.

## 3.4 PHYSICAL MODELS: CASE STUDY OF THE MEEUS, BIRD AND HULSTROM MODEL

Annear and Wells have compared several models to calculate the position of the sun and the irradiance [99]. Such comparisons lead to the conclusion that the more complex the model, the more accurate it is. On the other hand, the simpler the model, the faster it is to compute.

Based on the same comparison of solar models, the Meeus (1999) and Bird and Hulstrom (1981) model has proven to be the most accurate option with RMS errors in average 6 W/m<sup>2</sup> lower than the other tested methods for 17 sites in North America over a period of 2,726 clear days. The Meeus and Bird and Hulstrom model may thus be recommended for optimal accuracy although it is only one model among many other accurate physical clear-sky models, like for instance MAC [90, 91], Iqbal [92], ASHRAE [93], and Christian Gueymard's REST2 [94, 95] and many others, as already mentioned in the previous chapter. These models need a larger number of atmospheric input parameters, like the optical depth of the ozone layer in the atmosphere, or the ground albedo. Such data may be provided by AERONET [137]. However, data are scarce and often outdated for Perth and Western Australia generally speaking. Due to the lack of access to relevant input parameters, no physical model has been considered for this study and empirical models have been preferred for very-short forecasts.

Just to give a taste of the complexity of physical clear-sky models, the Meeus and Bird and Hulstrom model will be specifically analysed. This model requires the ratio of forward-scattered irradiance to the total scattered irradiance, the aerosol absorption, and the atmospheric turbidity as input parameters.

Wunderlich (1972) has formulated a precise first order approximation of the optical air mass ratio  $m$  on the basis of an expression found by Kasten (1964) modified in the light of the variations of atmospheric pressure according to the altitude  $z$  presented by List (1958). This expression of air mass ratio (3.46) is used by Meeus and Bird and Hulstrom, with the sun altitude angle  $\beta$  calculated as per equation (3.6).

$$m = \frac{\left[ \frac{(288 - 0.0065 \cdot z)}{288} \right]^{5.256}}{\sin(\beta) + 0.1500 \cdot (\beta + 3.885)^{-1.253}} \quad (3.46)$$

Klein (1948) has identified two distinct effects of dust on the solar radiation that are the scattering  $d_s$  and the absorption  $d_a$ . However, the absorption is negligible and the total effect of dust  $d$  equals the scattering effect  $d_s$ , which means  $d = d_a + d_s \approx d_s$ . Kimball (1930), Bolsenga (1964), along with Klein, have established a set of values for the effect of dust according to the environment. The dust values  $d$  range from 0 to 0.08 in remote areas and from 0.03 to 0.13 in urban areas.

In order to calculate the total irradiance, several parameters need to be taken into account. The relative position of the Earth if respect to the sun is one of them. Please note that all the following astronomical considerations are drawn from Meeus (1999). Based on the mean anomaly of the sun  $MeanAn$  and on the centre of the sun  $CentSun$ , the true anomaly of the sun expressed in degrees is by equation (3.47).

$$v = MeanAn + CentSun \quad (3.47)$$

Considering the current Julian century  $t$ , the current eccentricity of the Earth's orbit  $e$  is expressed without dimensions as:

$$e = 0.016708634 - t \cdot (0.000042037 + t \cdot 0.0000001267) \quad (3.48)$$

The current distance between the Earth and the Sun expressed in astronomical units (AU) both depend on  $v$  and  $e$ :

$$r = \frac{1.000001018 \cdot (1 - e^2)}{1 + e \cdot \cos(v)} \quad (3.49)$$

The eccentricity correction, which is dimensionless, depends on the current Earth-Sun distance  $r$  compared to the average distance  $r_0$ , which is 1 AU.

$$E_0 = \left(\frac{r}{r_0}\right)^2 \quad (3.50)$$

This leads to a new expression of the apparent extra-terrestrial irradiance  $A$  that is function of the solar constant  $SC$  and of the uncorrected sun altitude angle  $\beta$ :

$$A = SC \cdot E_0 \cdot \sin(\beta) \quad (3.51)$$

This model takes into account 4 different causes of sunlight scattering.

The water vapour suspended in the atmosphere participates to scatter sunlight. Indeed, depending on the dew-point temperature  $T_{dpt}$ , the mean hourly content of water that can potentially precipitate  $wc$ , expressed in cm, as per Reitan (1963) and Bolsenga (1965), is given by equation (3.52).

$$wc = e^{(-0.0592 + 0.06912 \cdot T_{dpt})} \quad (3.52)$$

All the following formulas are based on the expression of the air mass ratio  $m$  given by equation (3.46). The transmittance of water vapour  $X_w$  in cm is as follows:

$$X_{wc} = wc \cdot m \quad (3.53)$$

Bird and Hulstrom (1981) have determined the transmittance of water vapour  $T_w$  as the dimensionless expression given by (3.54):

$$T_w = 1 - \frac{2.4959 \cdot X_w}{(1 + 79.034 \cdot X_w)^{0.6828} + 6.385 \cdot X_w} \quad (3.54)$$

The ozone content in the atmosphere also scatters sunlight.

The ozone content value  $U_0$  has been empirically measured as 0.31 cm for mid-latitude summer atmospheric conditions.

Still according to Bird and Hulstrom (1981), using (3.46), the amount of ozone in a slanted path  $X_0$  is:

$$X_0 = U_0 \cdot m \quad (3.55)$$

As a result, the transmittance of ozone content is given:

$$T_0 = 1 - 0.1611 \cdot X_0 \cdot (1 + 139.48 \cdot X_0)^{-0.3035} - \frac{0.002715 \cdot X_0}{1 + 0.044 \cdot X_0 + 0.0003 \cdot X_0^2} \quad (3.56)$$

Uniformly mixed gases like carbon dioxide and oxygen are responsible for scattering sunlight as well. The transmittance of such gases is the following:

$$T_{UM} = e^{-0.0127 \cdot m^{0.26}} \quad (3.57)$$

In addition, the transmittance of Rayleigh scattering is given by:

$$T_R = e^{(-0.0903 \cdot m^2 \cdot (1+m-m^{1.01}))} \quad (3.58)$$

Finally, aerosols scatter sunlight too. Considering the aerosol optical depth  $\tau_{A0.38}$  in a vertical path for a wavelength of 380 nm, for which there is no molecule absorption, and the aerosol optical depth  $\tau_{A0.50}$  at 500 nm wavelength, which corresponds to ozone absorption, the overall turbidity  $\tau_A$  is the following:

$$\tau_A = 0.2758 \cdot \tau_{A0.38} + 0.35 \cdot \tau_{A0.50} \quad (3.59)$$

The values of  $\tau_{A0.38}$  and  $\tau_{A0.50}$  vary significantly with the location and must be empirically measured. Their values may typically be between 0 and 1. Similarly, the overall atmospheric turbidity  $\tau_A$  may have values typically comprised between 0.02 and 0.5.

The transmittance of aerosol absorption and scattering  $T_A$  is function of the overall atmospheric turbidity  $\tau_A$  and the air mass ratio:

$$T_A = e^{(-\tau_A^{0.873} \cdot m^{0.9108} \cdot (1+\tau_A-\tau_A^{0.7088}))} \quad (3.60)$$

Considering an empirical transmittance of aerosol absorption  $K_1$ , with a recommended value of 0.1, up to 0.35 in urban areas, the transmittance of aerosol absorption  $T_{AA}$  is:

$$T_{AA} = 1 - K_1 \cdot (1 - m + m^{1.06}) \cdot (1 - T_A) \quad (3.61)$$

Considering the transmittance of empirical ratio of forward-scattered irradiance  $B_a$ , for which the recommended value is 0.84 according to Bird and Hulstrom (1981), the atmospheric albedo  $r_s$  is expressed as:

$$r_s = 0.0685 + (1 - B_a) \cdot \left(1 - \frac{T_A}{T_{AA}}\right) \quad (3.62)$$

The atmospheric albedo  $r_s$  depends on the transmittance of aerosol absorption and scattering and on the atmospheric turbidity.

Using the previous expressions from the equations (3.51), (3.54), (3.56), (3.57), (3.58), and (3.60), the direct horizontal beam radiation is expressed as:

$$I_{BH} = 0.9662 \cdot A \cdot T_A \cdot T_w \cdot T_{UM} \cdot T_0 \cdot T_R \quad (3.63)$$

Similarly, it can be inferred from equations (3.46), (3.51), (3.54), (3.57), (3.56), (3.60), and (3.58), (3.61), that the diffuse radiation from atmospheric scattering is:

$$I_{DH} = 0.79 \cdot A \cdot T_{AA} \cdot T_w \cdot T_{UM} \cdot T_0 \cdot \left( \frac{0.5 \cdot (1 - T_R) + B_a \cdot \left(1 - \frac{T_A}{T_{AA}}\right)}{1 - m + m^{1.02}} \right) \quad (3.64)$$

$B_a$  is the empirical ratio of forward-scattered irradiance which, as already mentioned above, has a recommended value of 0.84 according to Bird and Hulstrom (1981), although this value may vary between 0 and 1, according to the type of scattering.

Finally, the total horizontal irradiance  $I_{TH}$  is expressed as follows, on the basis of equations (3.63), (3.64), and (3.62).

$$I_{TH} = \frac{I_{BH} + I_{DH}}{(1 - \rho \cdot r_s)} \quad (3.65)$$

Where  $\rho$  is the ground reflectivity or albedo, and  $r_s$  is the atmospheric albedo.

### 3.4.1 Comparison of the some Relevant Empirical Models for Western Australia.

We are looking to model the GHI of the Perth area in Western Australia. Although most of empirical models are based on measurements made in the northern hemisphere, the direct beam component of Denashyar-Paltridge-Proctor (DPP) model is based on some data collected across all of Australia [135]. However, the diffuse radiation of the DPP model as referred by Sandia is based on measurements in Tehran, Iran, at a relatively high altitude (1191m). The diffuse part may thus not be applicable to a low altitude location in Australia [136]. However, in absence of a more appropriate diffusion model, this expression has been retained and also applied to the Laue and Meinel models.

Also, for optimal results, the zenith angle has not been calculated according to (3.7) but with the use of the Solar Position Algorithm (SPA) developed by NREL [113], which is used as a benchmark when it comes to the sun's position [97]. The chosen location is Curtin University's Green Energy Electric Park (GEEP) within the Bentley campus of Curtin University (32.01 S, 115.89 E, 18 m above sea level). In the absence of reliable and consistent measurements of the atmospheric pressure, the average value of 1011.11 hPa has been fed to the SPA algorithm. This value corresponds to the typical value of barometric pressure at the altitude  $h = 18$  m above sea level (ASL) of the chosen site [138]:



$$P = 100 \cdot \left[ \frac{44331.514 - h}{11880.516} \right]^{\frac{1}{0.1902632}} \quad (3.66)$$

However, the actual GHI and the ambient temperature have been both measured through a weather station Vantage Pro2™ that has been installed in the GEEP. A sampling period of 1 minute has been considered. For information, the resolution of the solar sensor is 1 W/m2. The weather station is mounted on a pole that shows a slight tilt comprised between 1 and 2 degrees. This tilt and the sensor delay has been taken into account in the models. The measurements have been made during 16 days of the period from June 2014 to October 2015 between 6 AM and 8 PM with a 1-minute sampling period. As per Table 3.2, each season is represented by 4 days. These measurements have been used as a benchmark against which all empirical models are compared.

**Table 3.2 – Measurement dates of the GHI and ambient temperature.**

<i>Measurement Dates (yyyy/mm/dd)</i>	<i>Season</i>
2015/03/05	<i>Summer</i>
2015/03/07	
2015/03/08	
2015/03/11	
2015/05/08	<i>Autumn</i>
2015/05/13	
2015/05/30	
2015/06/10	
2014/06/25	<i>Winter</i>
2015/06/27	
2015/07/11	
2015/08/14	
2015/09/21	<i>Spring</i>
2015/09/23	
2015/10/06	
2015/10/21	

Since all the models that are compared are fast, simple and empirical, speed is not expected to be an issue. After being given an array containing 960 zenith angle values, each model need less than 5ms to calculate 960 GHI values with a MATLAB® program running on a DELL Latitude E7440 laptop equipped with a 64-bit Intel processor i7-4600 2.10 GHz / 2.70 GHz. Therefore, accuracy prevails upon speed as a more relevant discriminator.

**Table 3.3 – Average values of the accuracy metrics for some existing models.**

<i>Pearson's Correlation Coefficient (<math>\rho</math>)</i>					
<b>Model</b>	<b>Summer</b>	<b>Autumn</b>	<b>Winter</b>	<b>Spring</b>	<b>Overall</b>
<i>Haurwitz</i>	<b>0.999586</b>	0.998877	0.998833	<b>0.999247</b>	<b>0.999136</b>
<i>DPP</i>	<b>0.999694</b>	<b>0.999519</b>	<b>0.999497</b>	<b>0.999653</b>	<b>0.999591</b>
<i>BD</i>	0.997681	0.995012	0.994903	0.996997	0.996148
<i>ABCG</i>	<b>0.999630</b>	0.998430	0.998386	<b>0.999432</b>	0.998970
<i>KC</i>	<b>0.999073</b>	0.997876	0.997816	0.998593	0.998340
<i>RS</i>	<b>0.999306</b>	0.998220	0.998174	0.998937	0.998659
<i>Meinel</i>	0.997597	0.997339	0.997282	0.998051	0.997567
<i>Laue</i>	0.997625	0.997381	0.997326	0.998076	0.997602
<i>TJ</i>	<b>0.999882</b>	<b>0.999804</b>	<b>0.999777</b>	<b>0.999733</b>	<b>0.999799</b>
<i>Dazhi</i>	<b>0.999228</b>	<b>0.999432</b>	<b>0.999477</b>	<b>0.999446</b>	<b>0.999396</b>
<i>nRMSE (%)</i>					
<b>Model</b>	<b>Summer</b>	<b>Autumn</b>	<b>Winter</b>	<b>Spring</b>	<b>Overall</b>
<i>Haurwitz</i>	12.86	25.18	24.97	14.39	19.35
<i>DPP</i>	<b>5.76</b>	12.87	12.52	<b>6.97</b>	<b>9.53</b>
<i>BD</i>	<b>9.93</b>	21.18	20.64	10.23	15.50
<i>ABCG</i>	<b>3.10</b>	<b>8.23</b>	<b>8.64</b>	<b>3.98</b>	<b>5.99</b>
<i>KC</i>	<b>8.54</b>	<b>8.12</b>	<b>8.37</b>	<b>8.48</b>	<b>8.38</b>
<i>RS</i>	11.32	24.53	24.38	12.94	18.29
<i>Meinel</i>	17.95	11.38	11.10	16.47	14.23
<i>Laue</i>	18.06	11.50	11.16	16.58	14.33
<i>TJ</i>	19.25	16.00	13.26	16.71	16.31
<i>Dazhi</i>	<b>7.52</b>	<b>3.59</b>	<b>5.59</b>	<b>6.22</b>	<b>5.73</b>
<i>MBE (%)</i>					
<b>Model</b>	<b>Summer</b>	<b>Autumn</b>	<b>Winter</b>	<b>Spring</b>	<b>Overall</b>
<i>Haurwitz</i>	10.61	18.82	18.55	11.85	14.96
<i>DPP</i>	4.76	11.12	10.74	5.90	8.13
<i>BD</i>	8.09	16.31	15.65	7.81	11.96
<i>ABCG</i>	<b>-0.44</b>	<b>5.56</b>	5.32	<b>0.58</b>	<b>2.75</b>
<i>KC</i>	<b>-4.56</b>	<b>4.10</b>	<b>3.97</b>	<b>-3.43</b>	<b>0.018</b>
<i>RS</i>	9.66	18.87	18.66	10.94	14.53
<i>Meinel</i>	11.41	6.47	<b>5.09</b>	10.64	8.40
<i>Laue</i>	11.53	6.64	5.27	10.76	8.55
<i>TJ</i>	14.98	11.22	9.04	12.99	12.06
<i>Dazhi</i>	<b>4.20</b>	<b>0.72</b>	<b>-0.49</b>	<b>3.43</b>	<b>1.97</b>
<i>MAE (%)</i>					
<b>Model</b>	<b>Summer</b>	<b>Autumn</b>	<b>Winter</b>	<b>Spring</b>	<b>Overall</b>
<i>Haurwitz</i>	10.75	19.11	18.84	12.00	15.18
<i>DPP</i>	5.05	11.13	10.82	6.11	8.28
<i>BD</i>	8.15	16.40	15.75	7.88	12.04
<i>ABCG</i>	<b>2.24</b>	5.70	6.10	<b>2.91</b>	<b>4.23</b>
<i>KC</i>	6.27	5.33	5.47	6.24	5.83
<i>RS</i>	9.73	18.99	18.78	11.01	14.63
<i>Meinel</i>	13.49	9.38	9.20	12.52	11.15
<i>Laue</i>	13.56	9.43	9.23	12.59	11.20
<i>TJ</i>	15.14	11.64	9.60	13.23	12.40
<i>Dazhi</i>	5.28	<b>2.45</b>	<b>3.85</b>	<b>4.56</b>	<b>4.03</b>

The Mean Bias Error (MBE), the Mean Absolute Error (MAE) and the normalised Root Mean Square Error (nRMSE) are three widely used metrics to assess the accuracy of an irradiance model or an irradiance forecast [24, 97]. The nRMSE quantifies the overall accuracy of a model while the MBE demonstrates if the model is too high or too low. The MAE shows how far off the model is. The accuracy

and offsets measured by these metrics may be easily changed by altering some coefficients in the model. Additionally, the Pearson's correlation coefficient ( $\rho$ ) quantifies how well a model matches the shape of the curve of practical data. A perfect match returns a result of 1. The respective expressions of the MBE, MAE, nRMSE and  $\rho$ , with respect to the modelled GHI and the measured GHI have been given in the previous chapter as (2.3), (2.4), (2.2), and (2.5).

The average values of the Pearson's correlation coefficients, the RMSE, the MBE, and the MAE for each season and for the entire period are shown in Table 3.3. Pearson's correlation coefficients greater than 0.9990, RMSE below 10%, and MBE and MAE below 5% appear in bold characters.

It can be observed that 7 models out of 10 (KC, RS, Haurwitz, DPP, ABCG, TJ, and Dazhi) demonstrate a Pearson's correlation coefficients greater than 0.9990 during summer time. The same models, except RS and RS also show a similarly good fit in spring time. However, only DPP, TJ and Dazhi demonstrate Pearson's correlation coefficients greater than 0.9990 for all seasons.

Only four models manage to show an RMSE below 10% for the whole period: DPP, TJ, ABCG and Dazhi. The lowest nRMSE comes however to ABCG and Dazhi with some results around 6%.

The KC, Dazhi and ABCG models demonstrate the best MBE results. In particular, the KC model is nearly as low in summer and in spring as it is high in winter and autumn.

As a result, the mean MBE of the KC model over the whole period is remarkably close to 0.

Dazhi and ABCG show the best MAE results, with mean values below 5% over the entire period.

As a first conclusion, empirical irradiance functions may be both practical and sufficiently accurate tools when used with a very accurate solar position algorithm like NREL's SPA. In particular, the DPP, TJ and Dazhi model demonstrate a close fit with the bell shape of the measured GHI over all seasons. The KC model is the best centred model while overall, Dazhi and ABCG models perform very well. Dazhi appears to be more accurate for Western Australia than the ABCG model it is derived from. It must also be noticed that the TJ model proves to have the closest shape with respect to the actual data. The GHI curves are shown in Figure 3.3.

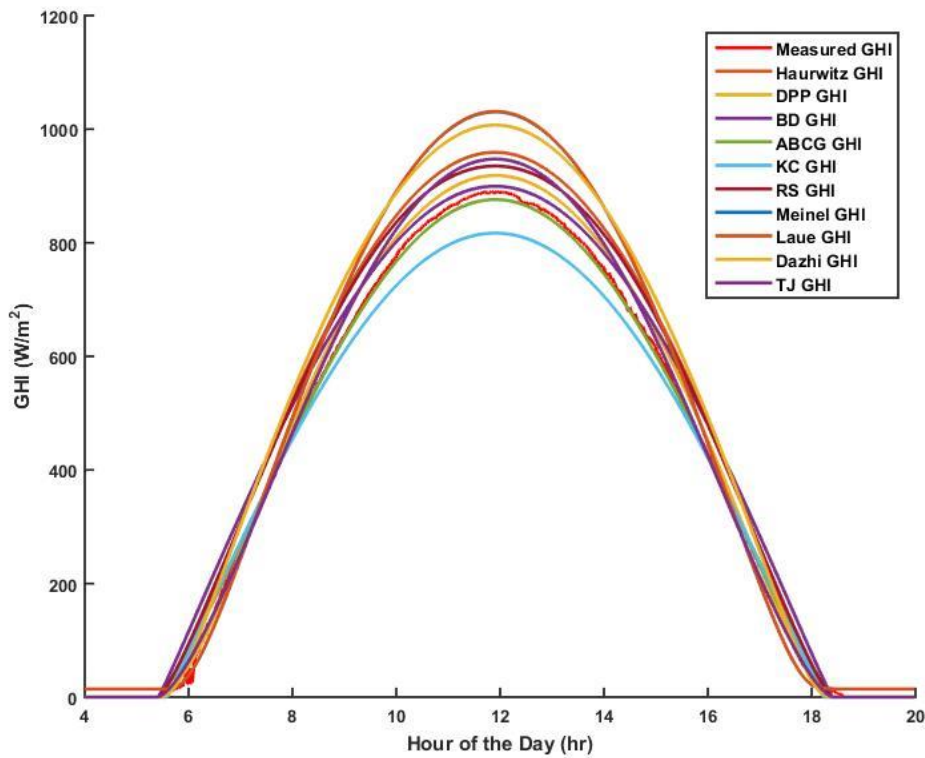


Figure 3.3 – Empirical GHI model curves compared with measured GHI every minute on 21/10/2015.

### 3.4.2 Initial Regression of Empirical Data.

After comparison of the empirical models, it appears that none is a perfect match for the measured data, despite some very good fit for some of the considered models. Thus, a regression can be a solution to ensure to have an optimal model for Western Australia. Such a regression can inspire from models that proved to be the most accurate as per the previously made comparison.

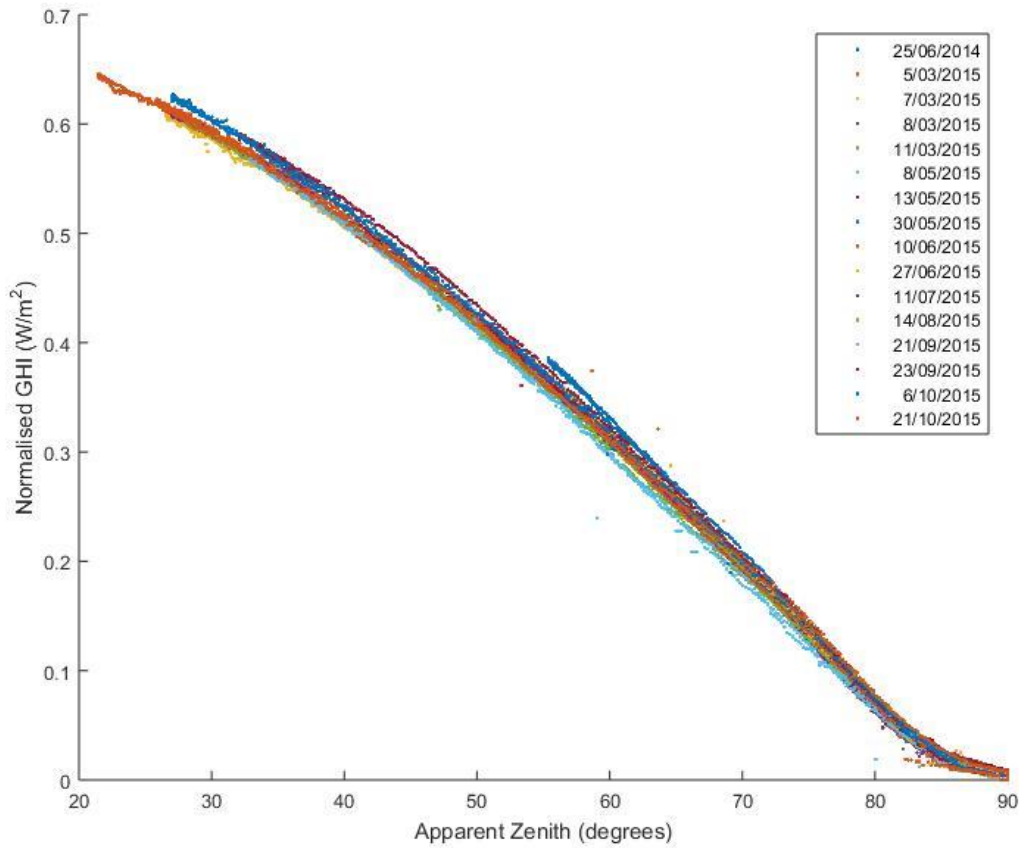
#### 3.4.2.1 Initial Dazhi-based Regression

As per the previous comparison of empirical clear-sky models, the smallest overall error over the whole period in terms of MAE and nRMSE proves to be from the Dazhi model. It has thus been hypothesised that a regression could be made on the basis of the Dazhi model, by using the measured data, to further improve this model. In particular, the Dazhi model does not have the most accurate Pearson's correlation coefficient with respect to the data. It has been hoped that this coefficient could be improved after regression. The new regression model under investigation is in the form:

$$GHI = a \cdot I_0 \cdot \cos(z)^b \cdot e^{-c \cdot (90^\circ - z)} \quad (3.67)$$

As per (3.16) and (3.11),  $I_0$  is constant for any given day number  $n$ . Hence, to simplify the problem, the measured GHI has been normalised with respect to the solar constant  $I_0$ :

$$GHI_{normalised} = \frac{GHI}{I_0} \quad (3.68)$$

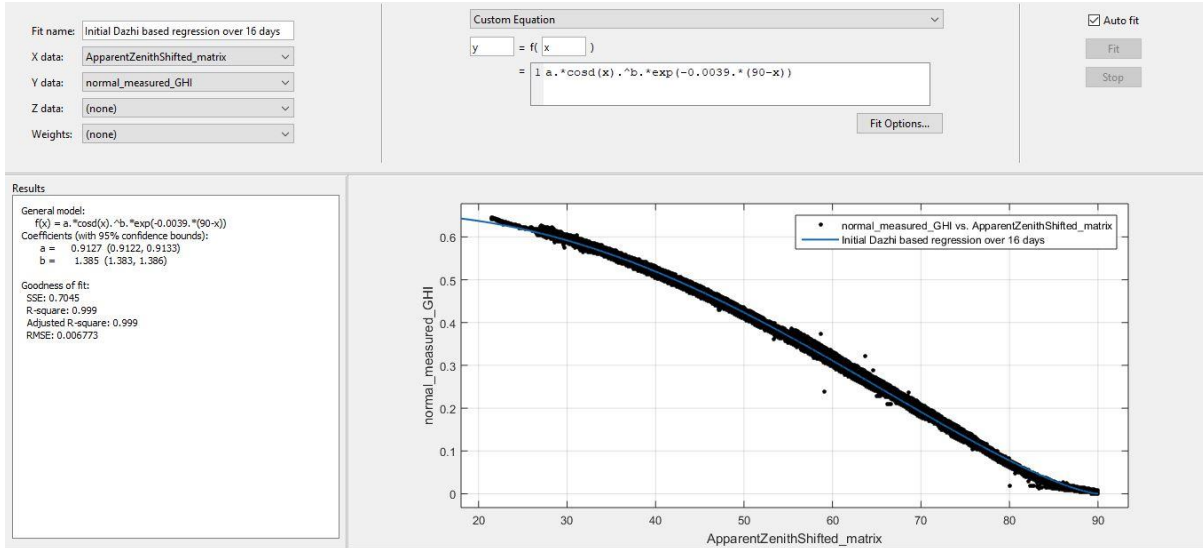


**Figure 3.4 – Normalised GHI in function of the zenith angle  $z$ , used for the initial Dazhi-based regression.**

Now, the problem is to determine  $a$ ,  $b$ , and  $c$ , given the following equation:

$$GHI_{normalised} = a \cdot [\cos(z)]^b \cdot e^{(c \cdot (90^\circ - z))} \quad (3.69)$$

MATLAB®'s curve-fitting toolbox has been used to determine the optimal curve fit with a custom equation. In particular, the data shown in Figure 3.4 has been fed to the regression tool from Figure 3.5. The output parameter 'y' is the normalised GHI whereas the input parameter 'x' is the apparent zenith angle.



**Figure 3.5 – Screenshot of the curve-fitting toolbox used for the initial Dazhi-based regression.**

Different values have been tested manually for the exponential factor  $c$  with the non-linear least square Levenberg-Macquardt method. By dichotomy, it has been found that the value with the best fit is  $c = -0.0039$ . The following custom equation is thus finally used to determine  $a$  and  $b$ :

$$GHI_{normalised} = a \cdot [\cos(z)]^b \cdot e^{(-0.0039 \cdot (90^\circ - z))} \quad (3.70)$$

Eventually, it has been found that values within 95% confidence bounds are  $a = 0.9127$  and  $b = 1.385$ .

The final equation is thus:

$$GHI = 0.9127 \cdot I_0 \cdot [\cos(z)]^{1.385} \cdot e^{(-0.0039 \cdot (90^\circ - z))} \quad (3.71)$$

### 3.4.2.2 Initial TJ-based Regression

With a Pearson's correlation coefficient of 0.999739 across all seasons and especially an astounding 0.999913 during summer days, the TJ model displays the closest shape with regards to actual data. However, the TJ model is characterized with one of the largest errors in terms of nRMSE, MBE and MAE. It has been nonetheless thought that the shape must be the most important property of a model and that altering the coefficients of the model through a regression could help to keep the advantage of a well-fitting shape while minimizing the error with respect to the actual data for any given apparent zenith angle.

Hence, the regression will be made on the basis of the following expression, where  $z$  is the apparent zenith angle,  $m$  is the air-mass ratio defined as per Equation (3.43),  $k$  is determined by equation (3.42), and  $A$  and  $C$  are functions of the day number  $n$ .

$$GHI = A \cdot e^{-k \cdot m} \cdot (C + \cos(z)) \quad (3.72)$$

$$A = \left[ a + b \cdot \sin\left(\frac{360}{365} \cdot (n - 275)\right) \right] \quad (3.73)$$

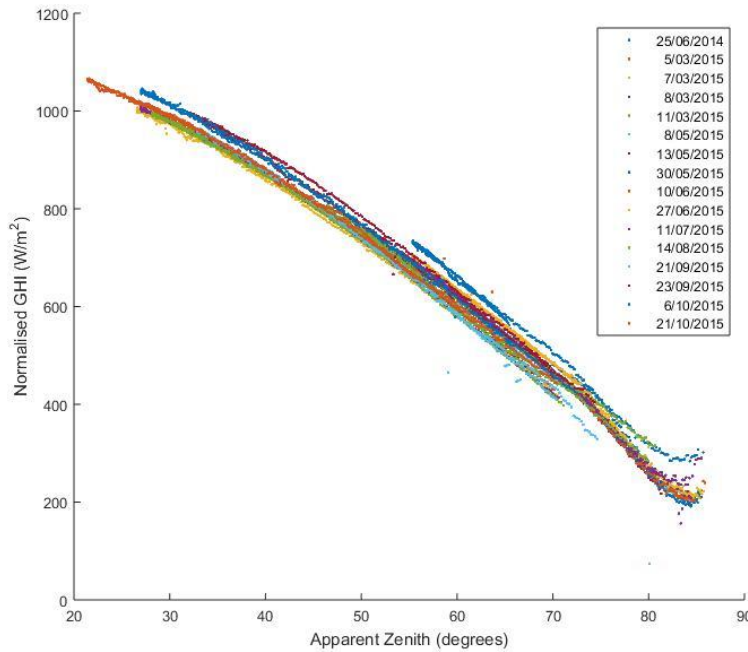
$$C = c + d \cdot \sin\left(\frac{360}{365} \cdot (n - 100)\right) \quad (3.74)$$

Like for the previous regression, the curve-fitting toolbox from MATLAB® is used to find out the best values for the new coefficients  $a$ ,  $b$ ,  $c$ , and  $d$ . However, to simplify the problem, the GHI can be normalised with respect to  $e^{-km}$ , which is an already known parameter. On the one hand, the optical depth of the atmosphere  $k$  is constant for a given day since it solely depends on the day number  $n$ . On the other hand, the air-mass ratio  $m$  is constant for any given altitude angle  $\beta$  or zenith angle  $z$ . The new normalised  $GHI_{normal}$  is:

$$GHI_{normal} = \frac{GHI}{e^{-k \cdot m}} = \frac{DNI \cdot \cos(z) + DHI}{e^{-k \cdot m}} = A \cdot (C + \cos(z)) \quad (3.75)$$

It must be reminded that  $GHI = DNI \cdot \cos(z) + DHI$ .

In (3.75), the apparent extra-terrestrial irradiance  $A$  also represents the normalised DNI or beam irradiance since  $DNI = I_B = A \cdot e^{-k \cdot m}$ , as expressed in (3.39) while the diffuse coefficient  $C$  also represents the normalised DHI since  $DHI = DNI \cdot C = A \cdot e^{-k \cdot m} \cdot C$  as per (3.44). However,  $C$  is a function of the day number  $n$ .



**Figure 3.6 – Normalised GHI in function of the zenith angle  $z$ , used for the Initial TJ-based regression.**

The normalised GHI is shown in Figure 3.6. The samples corresponding to a zenith value close to 90 degrees are ignored because and only values up to 86 degrees are considered. Indeed, the closer the zenith is to 90 degrees and the further the normalised GHI diverges to infinity, since it has been divided by  $m$ , which is also a function of  $z$ . Another aspect of the problem lies in the fact that the GHI was measured directly while the TJ expression considers both the DNI and the DHI into account. Finally, the problem consists in looking for the coefficients  $a$ ,  $b$ ,  $c$ , and  $d$  in a function of  $z$  and  $n$ :

$$\begin{aligned}
 GHI_{normal} &= A \cdot C_2 = f(z, n) \\
 &= \left[ a + b \cdot \sin\left(\frac{360}{365} \cdot (n - 275)\right) \right] \\
 &\quad \cdot \left[ c + d \cdot \sin\left(\frac{360}{365} \cdot (n - 100)\right) + \cos(z) \right]
 \end{aligned} \tag{3.76}$$

The expression (3.76) takes into account both the DNI and the DHI. The problem has been investigated by using the non-linear least square method and the Levenberg-Macquardt algorithm. In the curve-fitting toolbox, the output 'z' is  $GHI_{normalised}$ , the first input 'x' is the zenith angle  $z$  with values up to 86 degrees and the second input 'y' is the day number  $n$ .

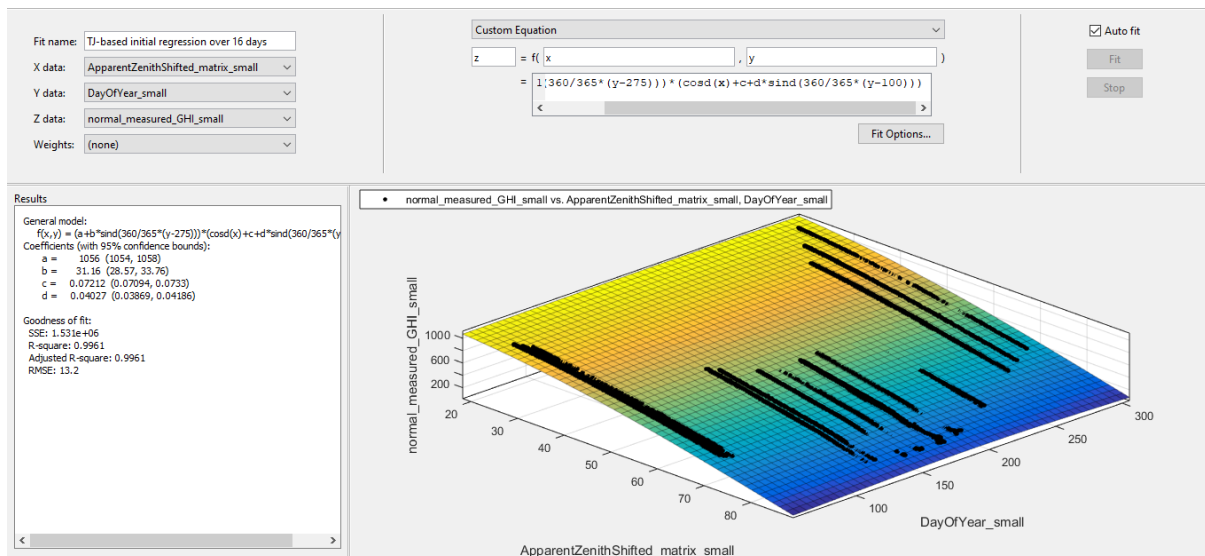


Figure 3.7 – Screenshot of the curve-fitting toolbox used for the initial TJ-based regression.

Finally, as shown in Figure 3.7, the best coefficients within 95% confidence bounds are:

$$\begin{aligned}
 GHI_{normal} &= 1056 + 31.16 \cdot \sin\left(\frac{360}{365} \cdot (n - 275)\right) \\
 &\quad \cdot \left[ 0.07212 + 0.04027 \cdot \sin\left(\frac{360}{365} \cdot (n - 100)\right) + \cos(z) \right]
 \end{aligned} \tag{3.77}$$

It can thus be inferred that the final expression of the DNI as per the TJ-based regression model is:

$$DNI = A \cdot e^{-k \cdot m} \tag{3.78}$$



$$A = 1056 + 31.16 \cdot \sin \left[ \frac{360}{365} \cdot (n - 275) \right] \quad (3.79)$$

The coefficient  $A$  is the apparent extra-terrestrial irradiance. The coefficients  $k$  and  $m$  are still the same as the ones used in the regular TJ model, given by (3.42) and (3.43). The DHI is:

$$DHI = C \cdot DNI \quad (3.80)$$

$$C = 0.07212 + 0.04027 \cdot \sin \left[ \frac{360}{365} \cdot (n - 100) \right] \quad (3.81)$$

Finally, the GHI is the following, with the DNI given by (3.78) and the DHI given by (3.80):

$$GHI = DHI \cdot \cos(z) + DNI \quad (3.82)$$

### 3.4.3 Refined Regression Made with an Extended Dataset

To enhance the performance of the regression models, the dataset has been extended to 88 days. An exhaustive table of all the dates has been documented in reverse chronological order in Table 3.4. The new dataset comprises 15 summer days, 16 winter days, 24 spring days, and 33 autumn days, within a period running from the 24<sup>th</sup> of June 2014 to the 1<sup>st</sup> of March 2018. Arguably, the days in this dataset have been selected to get a large number of days. Although this dataset may not be as balanced as the previous smaller dataset in terms of seasons, the dominant seasons are spring and autumn, which are characterised by a highly variable weather pattern.

**Table 3.4 – Second set of measurement dates for the GHI and ambient temperature.**

Date	Season	Date	Season	Date	Season	Date	Season
1/03/2018	Summer	13/09/2017	Winter	1/04/2017	Autumn	1/06/2016	Autumn
28/02/2018	Summer	25/08/2017	Winter	31/03/2017	Autumn	30/05/2016	Autumn
14/02/2018	Summer	29/06/2017	Winter	30/03/2017	Autumn	12/05/2016	Autumn
12/02/2018	Summer	28/06/2017	Winter	29/03/2017	Autumn	11/05/2016	Autumn
11/02/2018	Summer	18/06/2017	Autumn	25/03/2017	Autumn	10/05/2016	Autumn
4/01/2018	Summer	9/06/2017	Autumn	10/03/2017	Summer	21/10/2015	Spring
3/01/2018	Summer	8/06/2017	Autumn	17/02/2017	Summer	6/10/2015	Spring
1/01/2018	Summer	5/06/2017	Autumn	2/11/2016	Spring	23/09/2015	Spring
31/12/2017	Summer	4/06/2017	Autumn	1/11/2016	Spring	21/09/2015	Spring
21/12/2017	Spring	25/05/2017	Autumn	30/10/2016	Spring	14/08/2015	Winter
20/12/2017	Spring	7/05/2017	Autumn	23/10/2016	Spring	11/07/2015	Winter
14/02/2017	Spring	29/04/2017	Autumn	22/10/2016	Spring	27/06/2015	Winter
7/11/2017	Spring	28/04/2017	Autumn	18/10/2016	Spring	10/06/2015	Autumn
6/11/2017	Spring	27/04/2017	Autumn	12/10/2016	Spring	30/05/2015	Autumn
5/11/2017	Spring	25/04/2017	Autumn	29/09/2016	Spring	13/05/2015	Autumn
4/11/2017	Spring	20/04/2017	Autumn	21/09/2016	Winter	8/05/2015	Autumn
1/11/2017	Spring	11/04/2017	Autumn	5/09/2016	Winter	11/03/2015	Summer
30/10/2017	Spring	8/04/2017	Autumn	24/08/2016	Winter	8/03/2015	Summer
15/10/2017	Spring	5/04/2017	Autumn	14/08/2016	Winter	7/03/2015	Summer
9/10/2017	Spring	4/04/2017	Autumn	2/08/2016	Winter	5/03/2015	Summer
2/10/2017	Spring	3/04/2017	Autumn	1/07/2016	Winter	25/06/2014	Winter
18/09/2017	Winter	2/04/2017	Autumn	2/06/2016	Autumn	24/06/2014	Winter

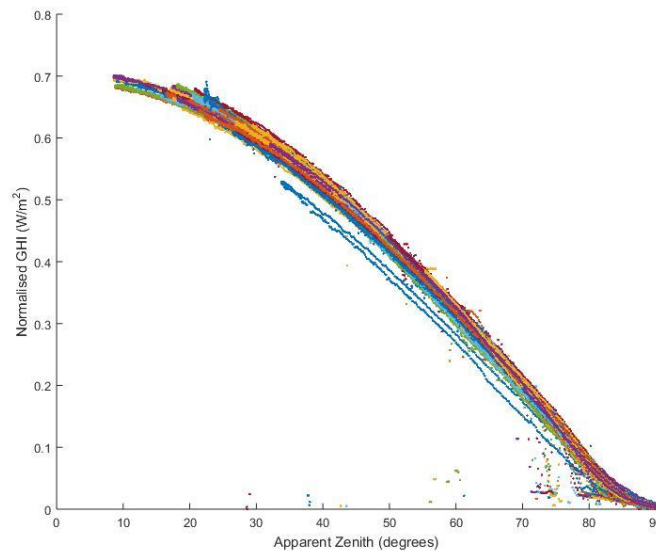
In this dataset, the GHI, the ambient temperature and the atmospheric pressure have been measured every minute. Although a measurement of the atmospheric pressure was available for most of the days, the average value of 1011.11 hPa has been taken for the days during which no measurement of the atmospheric pressure was available.

It must also be noted that over the four years that the dataset spans, the Davis weather station might have been turned on and off a few times. Although this might have caused some inconsistencies in terms of calibration, this problem can only affect the amplitude of the measured GHI and not its overall shape over a single day. Thus, the Pearson's correlation coefficient remains a very robust indicator to assess the accuracy of every model. Since the same measurements are used for each clear-sky model, this should not affect the relative accuracy of each model in terms of MAE, MBE, or NRMSE. Although, it may affect the absolute performance of the regression model, the relatively large number of samples within the dataset enables to find a good fit with the dominant calibration for the whole period.

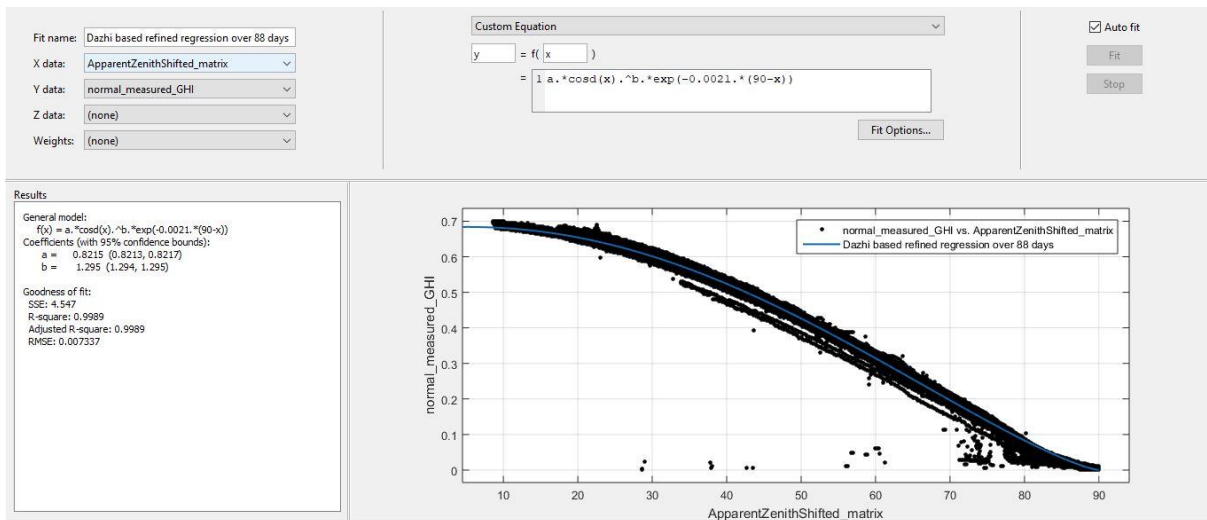
Since the Dazhi model and the Trelkeld-Jordan model proved to be the most accurate, these models can be reused as a basis for a new regression.

### ***3.4.3.1 Refined Dazhi model-based Regression***

The data shown in Figure 3.8 have been used to perform a bisquare Levenberg-Macquardt regression as per the screenshot in Figure 3.9, to find a more refined model inspired from the Dazhi model. A different colour has been chosen for each day in Figure 3.8. However, the excessively large number of colours present in the graph prevents from showing an exhaustive legend featuring every single day.



**Figure 3.8 – Normalised GHI in function of the apparent zenith angle  $z$ , used for the refined Dazhi-based regression.**



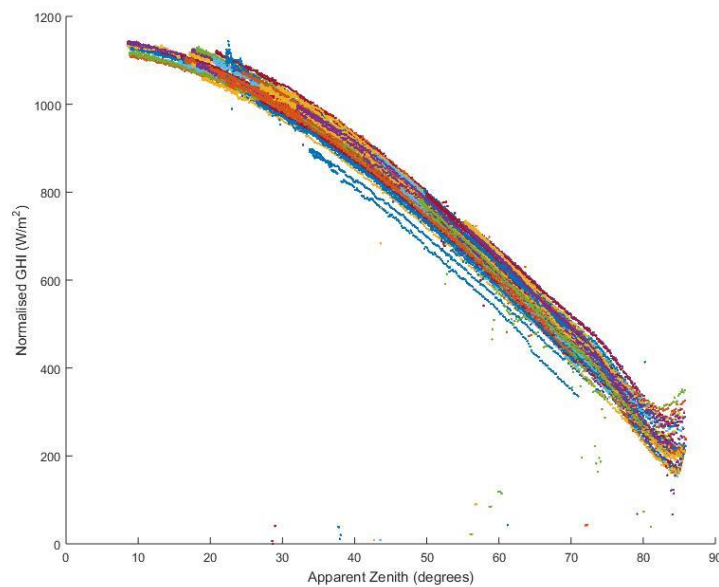
**Figure 3.9 – Screenshot of the curve-fitting toolbox used for the refined Dazhi-based regression.**

The best fit has been found to correspond to the equation (3.83):

$$GHI = 0.8215 \cdot I_0 \cdot [\cos(z)]^{1.295} \cdot e^{(-0.0021 \cdot (90^\circ - z))} \quad (3.83)$$

### 3.4.3.2 Refined Trelkeld-Jordan model-based Regression

Similarly to the refined regression made by inspiring from the Dazhi model, the same data, shown in Figure 3.10 has been used to perform a bisquare Levenberg-Macquardt linear regression with the results shown in Figure 3.11.



**Figure 3.10 – Normalised GHI in function of the apparent zenith angle z, used for the refined Trelkeld-Jordan model--based regression.**

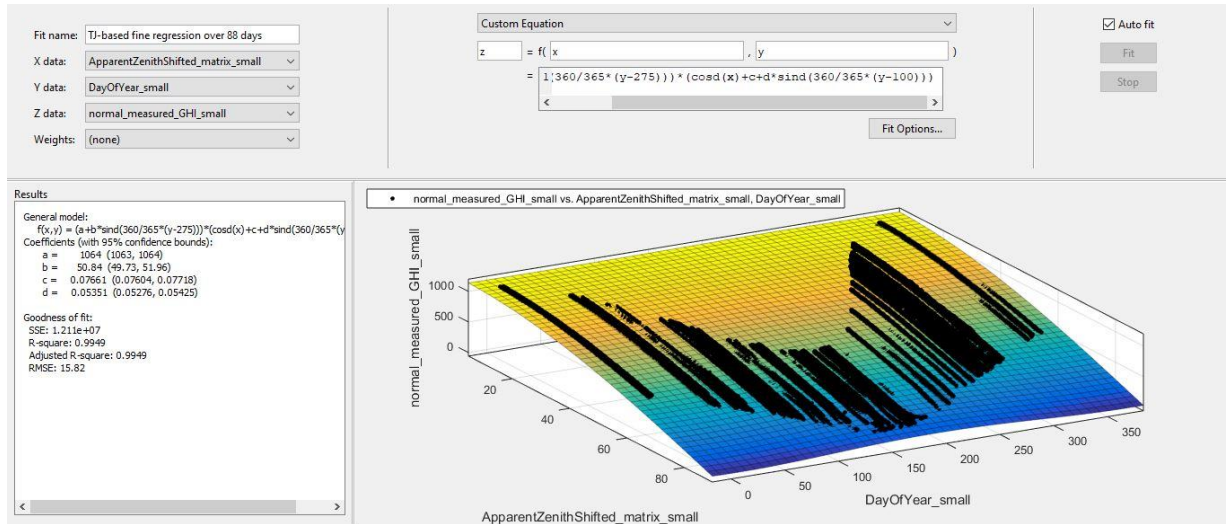


Figure 3.11 – Screenshot of the curve-fitting toolbox used for the refined TJ-based regression.

$$A = 1064 + 50.84 \cdot \sin \left[ \frac{360}{365} \cdot (n - 275) \right] \quad (3.84)$$

$$C = 0.07661 + 0.05351 \cdot \sin \left[ \frac{360}{365} \cdot (n - 100) \right] \quad (3.85)$$

The value of  $k$  remains unchanged compared with the original TJ model. It has been verified that the best fit was obtained with the expression previously introduced in Equation (3.42).

### 3.4.4 Accuracy of the Regression Models

#### 3.4.4.1 Original Dataset Regression

The accuracy of the two new models obtained by regression on the basis of the original 16-day dataset has been tested and compared against the previously mentioned models.

Table 3.5 – Average values of the accuracy metrics for the refined regression models.

<i>Pearson's Correlation Coefficient (<math>\rho</math>)</i>					
Model	Summer	Autumn	Winter	Spring	Overall
Dazhi Regres.	0.999785	0.999555	0.999569	0.999726	0.999659
TJ Regres.	0.999887	0.999840	0.999797	0.999804	<b>0.999832</b>
<i>nRMSE (%)</i>					
Model	Summer	Autumn	Winter	Spring	Overall
Dazhi Regres.	2.02	3.19	4.99	2.88	3.27
TJ Regres.	1.54	2.02	4.18	2.61	<b>2.59</b>
<i>MBE (%)</i>					
Model	Summer	Autumn	Winter	Spring	Overall
Dazhi Regres.	0.46	0.55	-0.57	-0.067	<b>0.095</b>
TJ Regres.	-0.21	0.12	-1.07	-0.32	-0.37
<i>MAE (%)</i>					
Model	Summer	Autumn	Winter	Spring	Overall
Dazhi Regres.	1.45	1.87	3.42	2.19	2.24
TJ Regres.	1.13	1.30	2.99	1.99	<b>1.85</b>

Table 3.5 shows the performance of both initial regressions according to the season and overall during the whole period. Both regressions fit best with the shape of the measured GHI curve during summer days. On the other hand, one of the poorest fit and highest error is achieved during winter days. The reason is most likely because the measured curves are the smoothest in the summer and the roughest in the winter. Table 3.6 compares the regression models with the original empirical models.

**Table 3.6 – Ranking of the top four most accurate empirical models overall over 16 days.**

<i>Rk.</i>	$\rho$		<i>nRMSE (%)</i>		<i>MBE (%)</i>		<i>MAE (%)</i>	
	<i>Model</i>	<i>Over. Value</i>	<i>Model</i>	<i>Over. Value</i>	<i>Model</i>	<i>Over. Value</i>	<i>Model</i>	<i>Over. Value</i>
1	<i>TJ Reg.</i>	<i>0.999832</i>	<i>TJ Reg.</i>	<i>2.59</i>	<i>KC</i>	<i>-0.018</i>	<i>TJ Reg.</i>	<i>1.85</i>
2	<i>TJ</i>	<i>0.999799</i>	<i>Dazhi Reg.</i>	<i>3.27</i>	<i>Dazhi Reg.</i>	<i>0.095</i>	<i>Dazhi Reg.</i>	<i>2.24</i>
3	<i>Dazhi Reg.</i>	<i>0.999659</i>	<i>Dazhi</i>	<i>5.73</i>	<i>TJ Reg.</i>	<i>-0.37</i>	<i>Dazhi</i>	<i>4.03</i>
4	<i>DPP</i>	<i>0.999591</i>	<i>ABCG</i>	<i>5.99</i>	<i>Dazhi</i>	<i>1.97</i>	<i>ABCG</i>	<i>4.23</i>

The initial TJ-based regression is the most accurate for all four considered metrics, with exception to the MBE. As shown in Table 3.6, the Dazhi-based model is the second most accurate model in terms of nRMSE, MBE, and MAE, but only comes in fourth position concerning Pearson’s correlation. If we look at the error of both regression in terms of MBE, it is visible that both achieve an error below 1%, which is much better than the regular Dazhi model, which happens to be the next accurate model. Thus, considering that all top 3 models in terms of MBE, namely KC, Dazhi-based and TJ-based regression models all display a less than 1% error, all three may be considered as equivalent and equally accurate. Following this reasoning, it can be concluded that the TJ-based model is the most accurate option. The Dazhi-based regression also has superior accuracy compared with the already established empirical models selected by Sandia. It is visible on Figure 3.12 that the shape of the regression models may not be a very good fit for dawn and dusk irradiance. However, this discrepancy can be considered as negligible since the difference is small in terms of absolute value. Moreover, the clear-sky irradiance at dawn and dusk is so low that the PV power output is negligible.

Although the accuracy of both the TJ and Dazhi models has been improved, the performance of both regression models could be even further improved if the regressions were based on a larger set of data. This is why a second regression based on both models has been carried out on new data set comprising a larger number of days, encompassing a longer period of time.

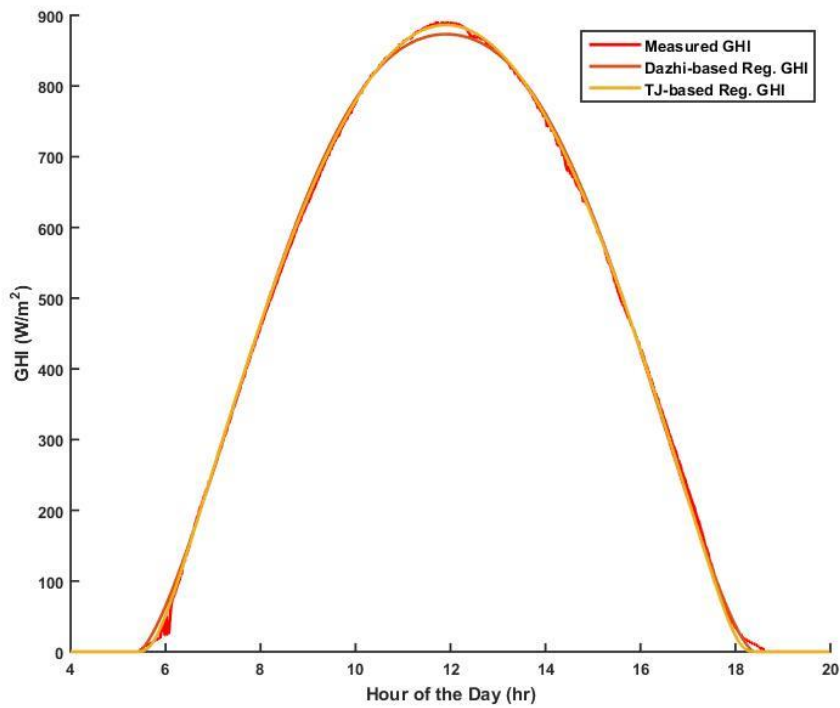


Figure 3.12 – First regression GHI model curves compared with measured GHI every minute on 21/10/2015.

### 3.4.4.2 Extended Dataset Refined Regression

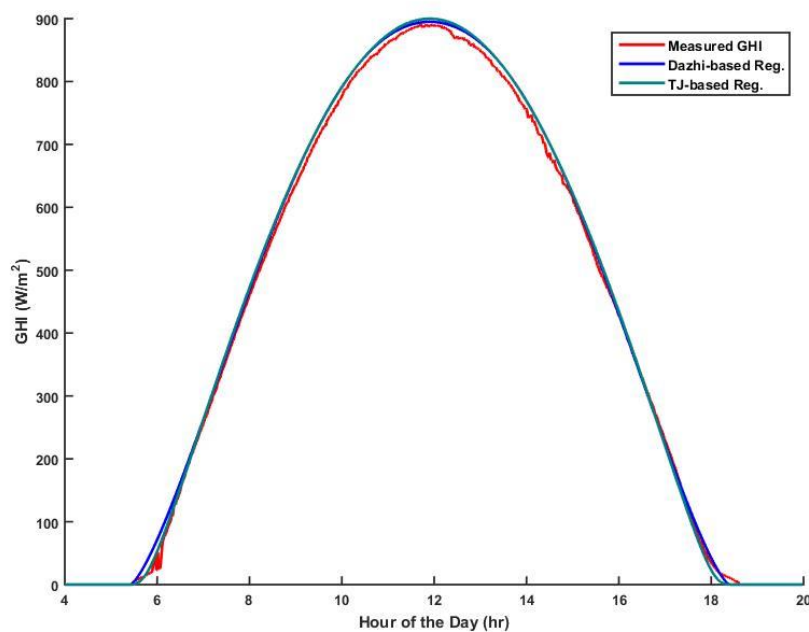
The average values of all four metrics over the three and a half year long period are summarised in Table 3.7. In that table, the initial regressions performed over 16 days are notes TJ-Reg1 and Dazhi\_Reg1 while the refined regressions obtained over 88 days are referred to as TJ-Reg2 and Dazhi-Reg2. All four regression models are noted in *italic*.

Table 3.7 – Ranking of the top four most accurate empirical models.

Rk.	$\rho$		<i>nRMSE</i> (%)		<i>MBE</i> (%)		<i>MAE</i> (%)	
	<i>Model</i>	<i>Over. Value</i>	<i>Model</i>	<i>Over. Value</i>	<i>Model</i>	<i>Over. Value</i>	<i>Model</i>	<i>Over. Value</i>
1	<i>TJ-Reg1</i>	<i>0.999145</i>	<i>TJ-Reg2</i>	<i>3.70</i>	<i>TJ-Reg2</i>	<i>-0.10</i>	<i>TJ-Reg2</i>	<i>2.29</i>
2	<i>TJ-Reg2</i>	<i>0.999142</i>	<i>TJ-Reg1</i>	<i>4.01</i>	<i>Dazhi-Reg2</i>	<i>0.58</i>	<i>TJ-Reg1</i>	<i>2.61</i>
3	TJ	<i>0.999119</i>	<i>Dazhi-Reg2</i>	<i>4.44</i>	ABCG	<i>0.70</i>	<i>Dazhi-Reg2</i>	<i>2.68</i>
4	DPP	<i>0.998934</i>	<i>Dazhi-Reg1</i>	<i>4.59</i>	<i>Dazhi-Reg1</i>	<i>-1.05</i>	<i>Dazhi-Reg1</i>	<i>2.96</i>
5	<i>Dazhi-Reg1</i>	<i>0.998932</i>	Dazhi	<i>6.16</i>	<i>TJ-Reg1</i>	<i>-1.41</i>	Dazhi	<i>4.19</i>
6	<i>Dazhi-Reg2</i>	<i>0.998901</i>	ABCG	<i>6.32</i>	Dazhi	<i>1.46</i>	ABCG	<i>4.20</i>

It clearly appears in Table 3.7 that the two TJ-based regressions are the most accurate empirical models with respect to the 88 days data, for all four metrics. In particular the refined TJ regression, TJ-Reg2

seems to outperform all of the other models, including the first regression model TJ-Reg1, in terms of nRMSE, MBE, and MAE. The Pearson's correlation coefficient of the refined TJ regression appears slightly lower than its counterpart from the first TJ regression. However, the difference is not extremely significant as it is only 0.000003. This difference is due to the fact that a certain number of curves from the large 88 day set that has been used to define TJ-Reg2 are not perfectly smooth, as opposed to the smoother curves from the 16 day set, which has been used as a basis to determine TJ-Reg1. This fact also explains the relatively low numbers for the overall Pearson's correlation coefficient of every model, compared to the results from Table 3.6. To illustrate how well the refined regression models fit with the measured data, the measured data as in Figure 3.12 has been compared with the two refined regression models in Figure 3.13.



**Figure 3.13 – Regression GHI model curves compared with measured GHI every minute on 21/10/2015, based on the extended 88-day dataset.**

### 3.4.5 Conclusions on the Regression Models

The refined Threlkeld and Jordan-based regression model TJ-Reg2 performs best overall, when considering all four criteria. The model could be even further improved if it could be built on the basis of a large number of perfectly smooth curves, which is difficult to measure over the years. Nonetheless, using an extended dataset spreading over a period of three and a half years has led to finding a new model that clearly outperforms all the original empirical models considered in this study, both in terms of shape fitness as well as error minimisation. In addition, none of the newly found regression models take more time to compute than their original counterparts. For all of these reasons, the refined TJ-based regression model TJ-Reg2 will be used further in this study as a reference for clear-sky irradiance.

### 3.5 CLOUD, SUN AND SHADOW PROJECTION MODELLING

As mentioned in a previous chapter, passing clouds can be monitored through several means of observations. In particular, satellite and ground cameras involve image processing, leading to potentially very accurate forecasts. In order to model the position and motion of a cloud, we will use a referential for the satellite view and a referential for the camera view.

#### 3.5.1 Single-point cloud model

##### 3.5.1.1 From sky Cartesian coordinates to camera sensor Cartesian coordinates

When modelling a cloud as it would be seen from an observer in the sky, e.g. an airplane or a satellite, spherical coordinates, also referred as sky coordinates, are the best-suited representation. Figure 3.14 shows a point represented with such sky coordinates.

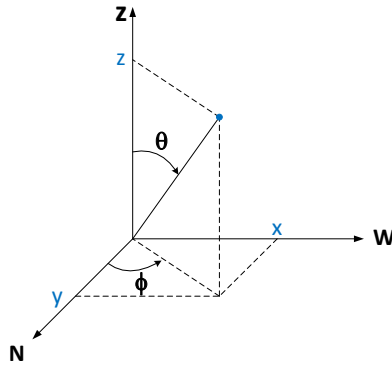


Figure 3.14 – Cartesian and spherical coordinates of a point object in the sky.

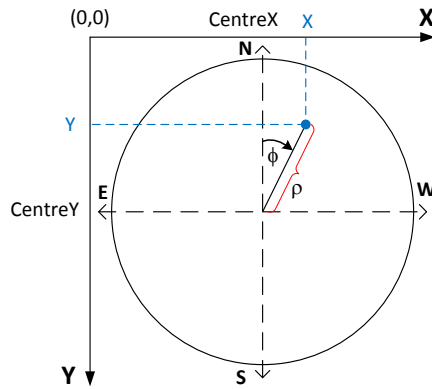
It is also important to define the equivalence between the Cartesian coordinates  $(x, y, z)$  and the equivalent spherical coordinates  $(r, \theta, \phi)$  of any given point object present in the sky:

$$r = \sqrt{x^2 + y^2 + z^2} \quad (3.86)$$

$$\theta = \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \quad (3.87)$$

$$\phi = \begin{cases} 0 & \text{if } x = 0 \\ \arctan\left(\frac{x}{y}\right) & \text{if } y \geq 0 \text{ \& } x \neq 0 \\ \arctan\left(\frac{x}{y}\right) + \frac{\pi}{2} & \text{if } y < 0 \text{ \& } x \geq 0 \\ \arctan\left(\frac{x}{y}\right) - \frac{\pi}{2} & \text{if } y < 0 \text{ \& } x < 0 \end{cases} \quad (3.88)$$



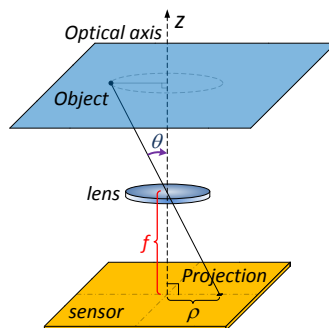


**Figure 3.15 – Cartesian and spherical coordinates of a point object in the sky.**

The same point seen through a ground camera located at the origin of the spherical frame shown in Figure 3.14 would have a 2-D projection with the polar coordinates  $(\rho, \phi)$  as shown in Figure 3.15.

It must be noted that the azimuth angle  $\phi$  is the same in both Figure 3.14 and Figure 3.15.

The radial distance  $\rho$  actually fully depends on the mapping function of the lens mounted on the camera, as illustrated in Figure 3.16. It is commonplace to use a fisheye lens on sky camera in order to have a 180 degree field of view, which enables to assess the total cloud sky cover.



**Figure 3.16 – Projection of a specific point object on the camera sensor.**

**Table 3.8 – Fisheye lens projection types and their respective mapping functions.**

<i>Projection type</i>	<i>Mapping function</i> $\rho = \text{radius (mm)}$ $f = \text{focal distance (mm)}$ $\theta = \text{angle (radians)}$
<i>Equidistant</i>	$\rho = f \cdot \theta$
<i>Stereographic</i>	$\rho = 2 \cdot f \cdot \tan(\theta/2)$
<i>Orthographic</i>	$\rho = f \cdot \sin(\theta)$
<i>Equisolid</i>	$\rho = 2 \cdot f \cdot \sin(\theta/2)$
<i>Rectilinear</i>	$\rho = f \cdot \tan(\theta)$

As shown in Table 3.8, fisheye lenses do not have a unique projection. Instead, they can have a relatively wide spectrum of mapping functions. In our case, we have considered an equisolid projection, which is the mapping function of the Sigma F2.8 EX DC lens we have used to shoot pictures of the sky. For that type of projection, the ideal radial distance, expressed in mm, is given by (3.83):

$$\rho = 2 \cdot f \cdot \sin\left(\frac{\theta}{2}\right) = 2 \cdot f \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)\right) \quad (3.83)$$

As already mentioned, the azimuth angle  $\phi$  given by (3.82) is the same in the sky and on the camera sensor. When converting the polar coordinates  $(\rho, \phi)$  of the projected punctual image to the Cartesian distance coordinates  $(d_X, d_Y)$  on the camera sensor, expressed in mm, we get the following expressions of the image coordinates in function of the object coordinates  $(x, y, z)$ :

$$\begin{aligned} d_X &= d_{CentreX} + \rho \cdot \sin(\phi) \\ &= d_{CentreX} + 2 \cdot f \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)\right) \cdot \sin(\phi) \end{aligned} \quad (3.89)$$

$$\begin{aligned} d_Y &= d_{CentreY} - \rho \cdot \sin(\phi) \\ &= d_{CentreY} - 2 \cdot f \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)\right) \cdot \cos(\phi) \end{aligned} \quad (3.90)$$

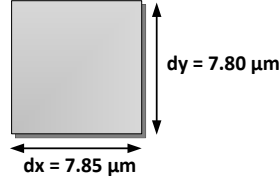
It must be noted that  $d_{CentreX}$  and  $d_{CentreY}$  are the horizontal and vertical position of the centre of the image, expressed in mm. It is however far more convenient to get to know the Cartesian coordinates  $(X, Y)$  in pixels rather than in mm. This is the purpose of the following section.

Firstly, it must be noted that lenses are imperfect. Their mapping functions may varies from one lens to the next, even with the same projection type. Hence, it is safer to write down the mapping function of an equisolid fisheye lens as follows, where  $cst$  and  $a$  are two coefficients to be determined empirically:

$$\rho = cst \cdot \sin\left(\frac{\theta}{a}\right) = cst \cdot \sin\left(\frac{1}{a} \cdot \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)\right) \quad (3.91)$$

However, for the sake of simplicity, we will assume for modelling purposes that the projection of our lens is perfectly equisolid.

Secondly, it must be noted that pixels of camera sensors are seldom perfect squares. For example, a Nikon D7100 has been used for verification purposes within this study. The size of its sensor is 23,500  $\mu\text{m}$  (horizontally)  $\times$  15,600  $\mu\text{m}$  (vertically) with a maximum resolution of 2992 horizontal pixels  $\times$  2000 vertical pixels. Thus, pixels have a size of  $dx = 7.85 \mu\text{m}$  (horizontally)  $\times$   $dy = 7.80 \mu\text{m}$  (vertically).



**Figure 3.17 – A sensor pixel of a Nikon D7100 DSLR camera.**

Given the rectangular shape of a pixel, as shown in Figure 3.17, the mapping function of an ideal equisolid lens mounted on the camera body is as follows:

$$\rho_x = cst_x \cdot \sin\left(\frac{\theta}{2}\right) \quad (3.92)$$

$$\rho_y = cst_y \cdot \sin\left(\frac{\theta}{2}\right) \quad (3.93)$$

$$cst_x = \frac{2 \cdot f}{dx} \quad (3.94)$$

$$cst_y = \frac{2 \cdot f}{dy} \quad (3.95)$$

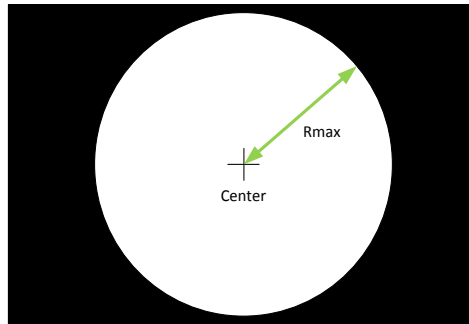
Therefore, the pixel Cartesian coordinates ( $X$ ,  $Y$ ) of the projected image of a sky object ( $x$ ,  $y$ ,  $z$ ) on the camera sensor, centred around  $CentreX$  and  $CentreY$ , expressed in pixels, are:

$$\left\{ \begin{array}{l} X = CentreX + \|\rho_x \cdot \sin(\phi)\| \\ = CentreX + \left\| cst_x \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)\right) \cdot \sin\left(\arctan\left(\frac{y}{x}\right)\right) \right\| \\ Y = CentreY - \|\rho_y \cdot \cos(\phi)\| \\ = CentreY - \left\| cst_y \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)\right) \cdot \cos\left(\arctan\left(\frac{y}{x}\right)\right) \right\| \end{array} \right. \quad (3.96)$$

It must be noted that the quantity added to  $CentreX$  or subtracted to  $CentreY$  must be rounded.

Since we have used a Sigma 4.5mm equisolid lens mounted on the body of our Nikon D7100 camera, the constant values should theoretically be  $cst_x = 1146$  and  $cst_y = 1154$ .

With an equisolid lens or any other mapping function, the image is a disk of radius  $Rmax$  as shown in Figure 3.18.



**Figure 3.18 – Typical projection of a circular fisheye lens on a sensor.**

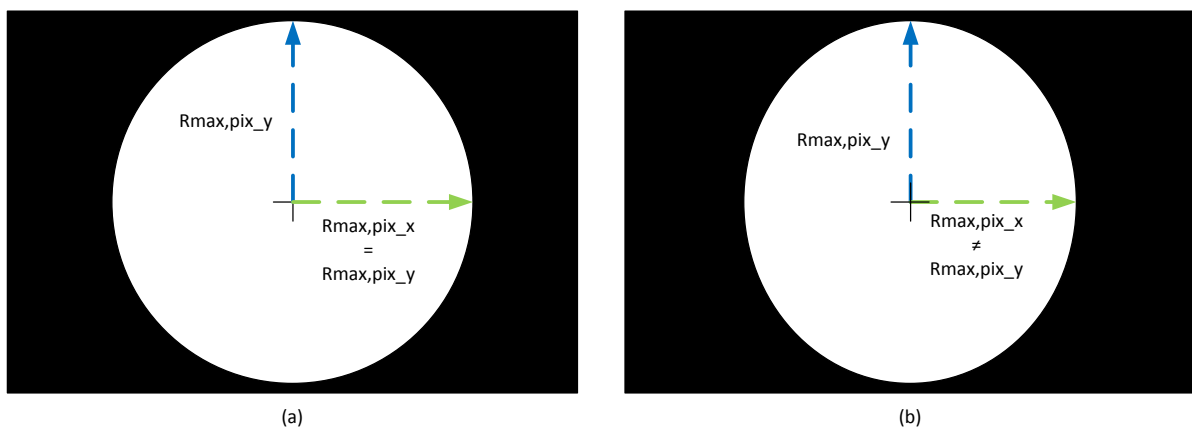
In the equisolid case, the radius  $R_{max}$  in mm of the projected image on the frame in radians is as formulated in Equation (3.97).

$$R_{max} = 2 \cdot f \cdot \sin\left(\frac{FoV}{4}\right) \quad (3.97)$$

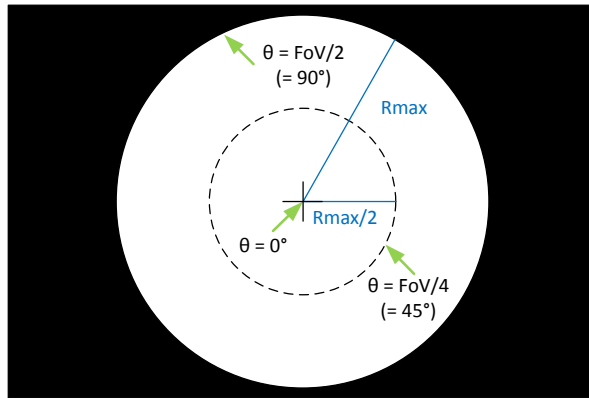
The pixel equivalent  $R_{max,pix}$  of  $R_{max}$  depends on the pixel unit size  $dx$  (horizontally) and  $dy$  (vertically) previously shown in Figure 3.17. The effective frame across the field of view, which is the projection of the image within the full frame, is a circle if pixels are exactly square shaped but it is an ellipse if pixels are rectangular in shape, as shown in Figure 3.19. In any of these cases, the relationship between the radius of the projection of the image and the pixel size on the horizontal and vertical axis are respectively given by Equations (3.98) and (3.99).

$$R_{max\_pix\_x} = \frac{R_{max}}{dx} \quad (3.98)$$

$$R_{max\_pix\_y} = \frac{R_{max}}{dy} \quad (3.99)$$



**Figure 3.19 – Projection with square shaped pixels (a) and with rectangular shaped pixels (b).**

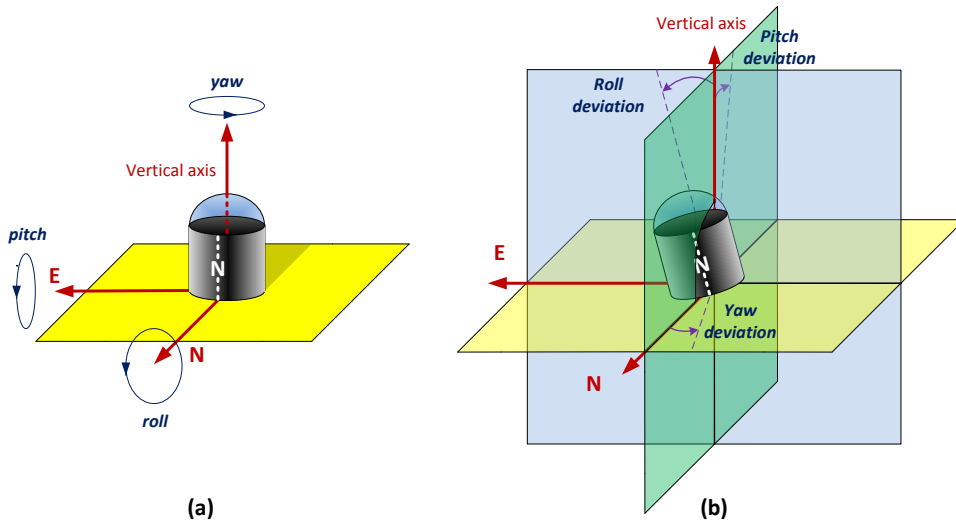


**Figure 3.20 – Position of the projections of point objects on the frame according to their angle  $\theta$  with respect to the optical axis.**

In ideal conditions, the circle or ellipse featuring the projection of the image is centred at the centre of the sensor. Also, the centre ideally features a remote point situated far away on a vertical axis, provided the camera is installed perfectly vertically. Therefore, the centre represents the projection of the normal axis with respect to the ground, with an angle of 0 degrees, while the edge of the complete disk-image of radius  $R_{max}$  represents an angle of half the field of view  $FoV/2$ . In the example shown in Figure 3.20, the field of view is assumed to be 180 degrees, or  $\pi$  radians. Thus,  $FoV/2$  corresponds to an angle of 90 degrees. The circles comprised between the centre and the rim of the circle-image represent angle values for  $\theta$  between 0 and  $FoV/2$ . For instance, the circle with a radius  $R_{max}/2$  represents an angle of  $FoV/4$ , which is in this case 45 degrees. However, the centre of the projection might be slightly systematically off-centre, due to manufacturing flaws. What has been explained above theoretically still holds when we consider the centre of the disk-image rather than necessarily the centre of the frame.

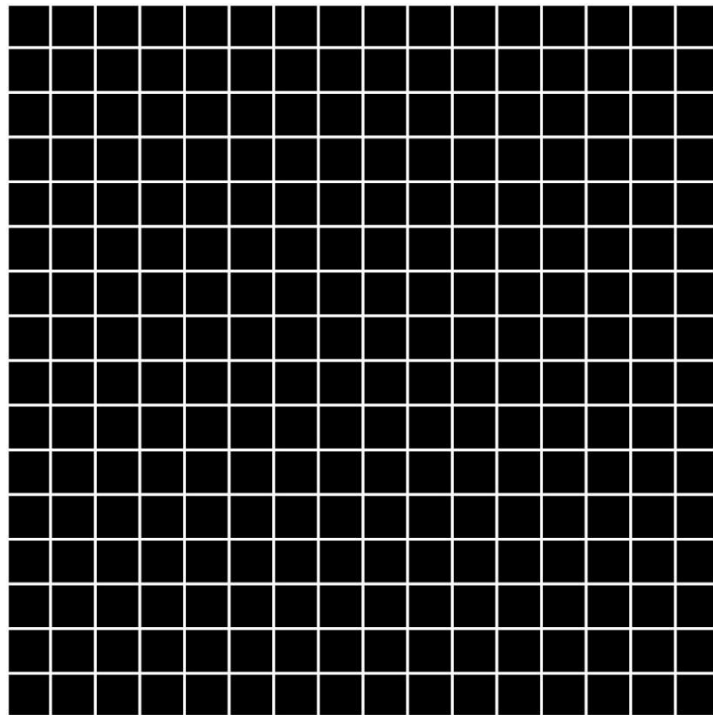
Besides some possible construction flaws, realistically, it is virtually impossible for the camera to be installed perfectly vertically or face exactly the true North, in which case the camera is oriented with a yaw, a pitch and a roll that all need to be considered. The yaw, pitch and roll of the camera are illustrated in Figure 3.21.

It is capital to be able to determine the deviation of the camera with respect to the vertical axis and with respect to the direction of true North. Neglecting those deviations within the 3-D referential leads to a non-linear 2-D projection error. For this reason, an object like the sun or stars must be observed over a reasonably long period of time in order to find a 3D correction matrix that corrects the effects of yaw, pitch and row. More details on the 3D rotation matrix will be given further in this chapter.



**Figure 3.21 – Illustration of the roll, pitch, and yaw of a sky camera (a) and their respective deviations on their respective planes (b).**

However, considering at first a simple and ideal model, if an ideal camera was installed on the ground perfectly vertical and perfectly facing North and if that camera was to shoot a 50 km wide mesh consisting of perpendicular and parallel lines as shown in Figure 3.22, located 3km above the camera, with each parallel line being 3.125 km apart from each other, the photo taken would look like the MATLAB ® model shown in Figure 3.23.



**Figure 3.22 – A 50-km wide mesh in the sky at 3000 m of altitude above the camera.**

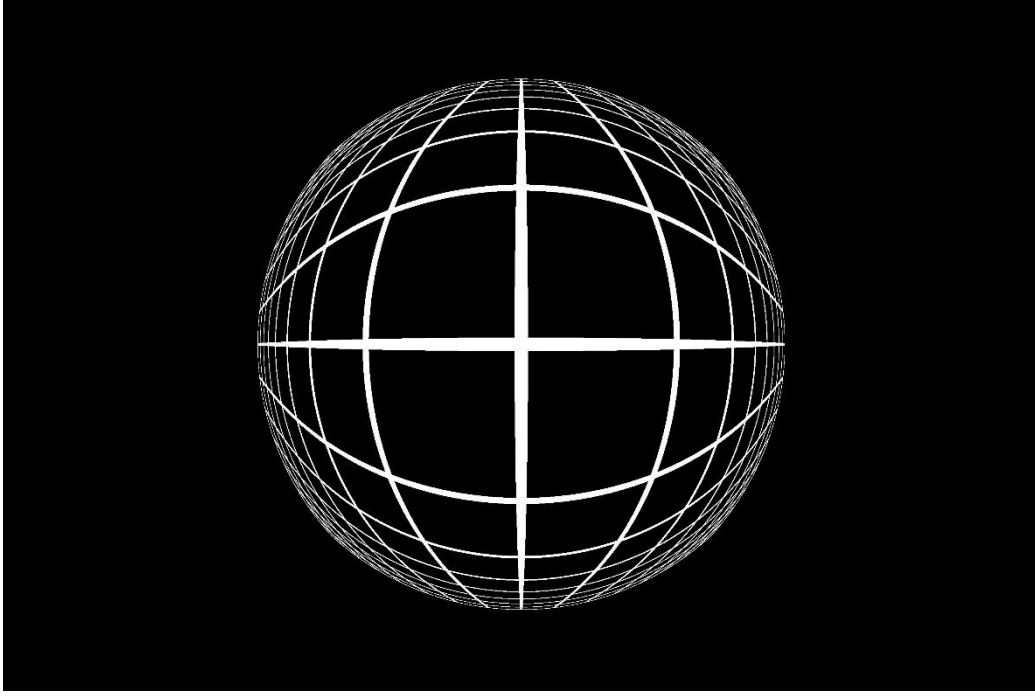


Figure 3.23 – Same mesh seen through a 4.5 mm equisolid circular fisheye lens.

### 3.5.1.2 From camera sensor Cartesian coordinates to sky Cartesian coordinates

Reciprocally, it may be useful to convert the Cartesian coordinates  $(X, Y)$  of the image on the camera sensor back to the Cartesian coordinates  $(x, y, z)$  of the corresponding object in the sky, provided the altitude  $z$  of the object is known.

The polar camera sensor  $(\rho, \phi)$  coordinates of the punctual projected image, when its Cartesian camera sensor  $(X, Y)$  coordinates are known, are:

$$\rho = \sqrt{(X - \text{Centre}X)^2 + (\text{Centre}Y - Y)^2} \quad (3.100)$$

$$\phi = \begin{cases} 0 & \text{if } X = \text{Centre}X \\ \arctan\left(\frac{(X - \text{Centre}X)}{(\text{Centre}Y - Y)}\right) & \text{if } Y \leq \text{Centre}Y \quad \& \quad X \neq \text{Centre}X \\ \arctan\left(\frac{(X - \text{Centre}X)}{(\text{Centre}Y - Y)}\right) + \frac{\pi}{2} & \text{if } Y > \text{Centre}Y \quad \& \quad X \geq \text{Centre}X \\ \arctan\left(\frac{(X - \text{Centre}X)}{(\text{Centre}Y - Y)}\right) - \frac{\pi}{2} & \text{if } Y > \text{Centre}Y \quad \& \quad X < \text{Centre}X \end{cases} \quad (3.101)$$

The zenith angle  $\theta$  of the sky object can be inferred from (3.102), with  $\rho$  and  $f$  expressed in mm:

$$\theta = 2 \cdot \arcsin\left(\frac{\rho}{2 \cdot f}\right) \quad (3.102)$$

Since the value of  $\rho$  given by (2.5) is expressed in pixels, the expression of  $\theta$  depends on whether a projection is considered on the x-axis or on the y-axis:

$$\theta_x = 2 \cdot \arcsin \left( \frac{\sqrt{(X - CentreX)^2 + (CentreY - Y)^2}}{cst_x} \right) \quad (3.103)$$

$$\theta_y = 2 \cdot \arcsin \left( \frac{\sqrt{(X - CentreX)^2 + (CentreY - Y)^2}}{cst_y} \right) \quad (3.104)$$

The altitude  $z$  is supposed to be known. Indeed, all information related to the altitude has been lost during the projection of the 3-D object into the 2-D sensor. The altitude should thus be retrieved from another source than the photos. Finally, the sky Cartesian coordinates  $(x, y, z)$  of the original object for a given altitude  $z$  are expressed on the basis of (3.101), (3.103), and (3.104) according to pixel coordinates  $(X, Y)$  of the image on the camera sensor:

$$\begin{cases} x = z \cdot \tan(\theta_x) \cdot \sin(\phi) \\ y = z \cdot \tan(\theta_y) \cdot \cos(\phi) \\ z \end{cases} \quad (3.105)$$

### 3.5.2 Four-point cloud model

Although clouds are 3-dimensional objects, it is the 2-dimensional projection of the cast shadow that may create light disruption on a PV system. It matters thus to model clouds as a surface corresponding to that projection.

#### 3.5.2.1 From sky Cartesian coordinates to camera sensor Cartesian coordinates

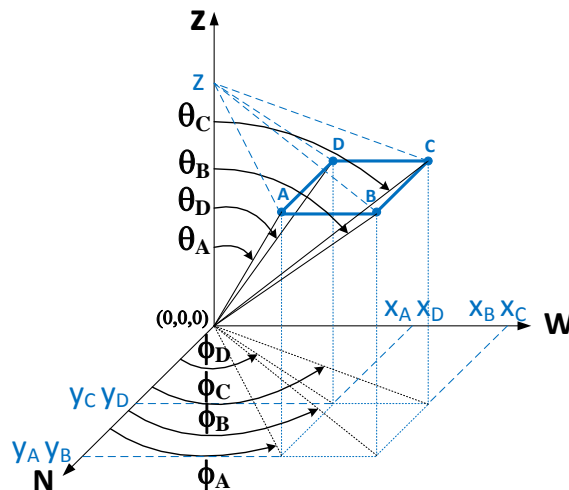
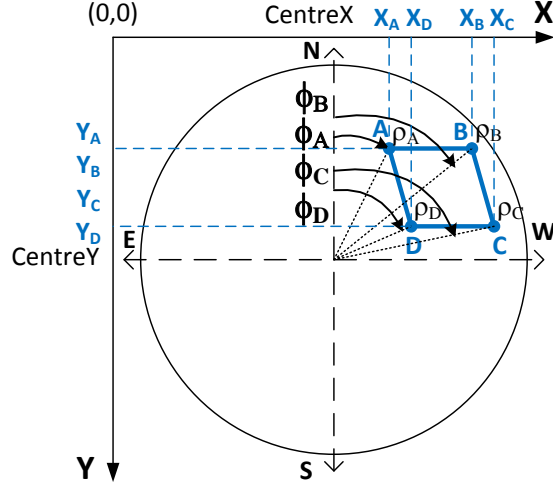


Figure 3.24 – Cartesian and spherical coordinates of a parallelogram object ABCD in the sky.





**Figure 3.25 – Polar coordinates of the image of the parallelogram object on a ground camera sensor.**

The coordinates of each 4 points of the parallelogram can be inferred from the previous calculation (3.105) made for a single-point object. For the sake of simplicity, we assume that all 4 points are located at the same altitude  $z$ . This is assumption is particularly true when considering the base of cumulus clouds, which is a flat horizontal surface. The Cartesian coordinates  $(X_A, Y_A)$ ,  $(X_B, Y_B)$ ,  $(X_C, Y_C)$  and  $(X_D, Y_D)$  of the projected single-point images of a sky single-point object  $(x_A, y_A, z)$  on the camera sensor, expressed in pixels, are:

$$\left\{ \begin{array}{l} X_A = CentreX + \|\rho_x \cdot \sin(\phi)\| \\ = CentreX + \left\| cst_x \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x_A^2 + y_A^2 + z^2}}\right)\right) \cdot \sin\left(\arctan\left(\frac{y_A}{x_A}\right)\right) \right\| \\ Y_A = CentreY - \|\rho_y \cdot \cos(\phi)\| \\ = CentreY - \left\| cst_y \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x_A^2 + y_A^2 + z^2}}\right)\right) \cdot \cos\left(\arctan\left(\frac{y_A}{x_A}\right)\right) \right\| \end{array} \right. \quad (3.106)$$

$$\left\{ \begin{array}{l} X_B = CentreX + \|\rho_x \cdot \sin(\phi)\| \\ = CentreX + \left\| cst_x \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x_B^2 + y_B^2 + z^2}}\right)\right) \cdot \sin\left(\arctan\left(\frac{y_B}{x_B}\right)\right) \right\| \\ Y_B = CentreY - \|\rho_y \cdot \cos(\phi)\| \\ = CentreY - \left\| cst_y \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x_B^2 + y_B^2 + z^2}}\right)\right) \cdot \cos\left(\arctan\left(\frac{y_B}{x_B}\right)\right) \right\| \end{array} \right. \quad (3.107)$$

$$\left\{ \begin{array}{l} X_C = CentreX + \|\rho_x \cdot \sin(\phi)\| \\ = CentreX + \left\| cst_x \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x_C^2 + y_C^2 + z^2}}\right)\right) \cdot \sin\left(\arctan\left(\frac{y_C}{x_C}\right)\right) \right\| \\ Y_C = CentreY - \|\rho_y \cdot \cos(\phi)\| \\ = CentreY - \left\| cst_y \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x_C^2 + y_C^2 + z^2}}\right)\right) \cdot \cos\left(\arctan\left(\frac{y_C}{x_C}\right)\right) \right\| \end{array} \right. \quad (3.108)$$

$$\left\{ \begin{array}{l} X_D = CentreX + \|\rho_x \cdot \sin(\phi)\| \\ = CentreX + \left\| cst_x \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x_D^2 + y_D^2 + z^2}}\right)\right) \cdot \sin\left(\arctan\left(\frac{y_D}{x_D}\right)\right) \right\| \\ Y_D = CentreY - \|\rho_y \cdot \cos(\phi)\| \\ = CentreY - \left\| cst_y \cdot \sin\left(\frac{1}{2} \cdot \arccos\left(\frac{z}{\sqrt{x_D^2 + y_D^2 + z^2}}\right)\right) \cdot \cos\left(\arctan\left(\frac{y_D}{x_D}\right)\right) \right\| \end{array} \right. \quad (3.109)$$

The quantities  $\rho_x$ ,  $\rho_y$ ,  $cst_x$ , and  $cst_y$ , are respectively given by Equations (3.92), (3.93), (3.94), and (3.95).

### 3.5.2.2 From camera sensor Cartesian coordinates to sky Cartesian coordinates

As calculated previously for a single-point object, for a fixed value of  $z$  for all 4 points of sky object, the Cartesian sky coordinates  $(x_A, y_A, z)$  respectively in function of the Cartesian sensor coordinates  $(X_A, Y_A)$  are:

$$\begin{cases} x_A = z \cdot \tan(\theta_{A,x}) \cdot \sin(\phi_A) \\ y_A = z \cdot \tan(\theta_{A,y}) \cdot \cos(\phi_A) \\ z \end{cases} \quad (3.110)$$

### 3.5.3 A possibly more Realistic Cloud Model

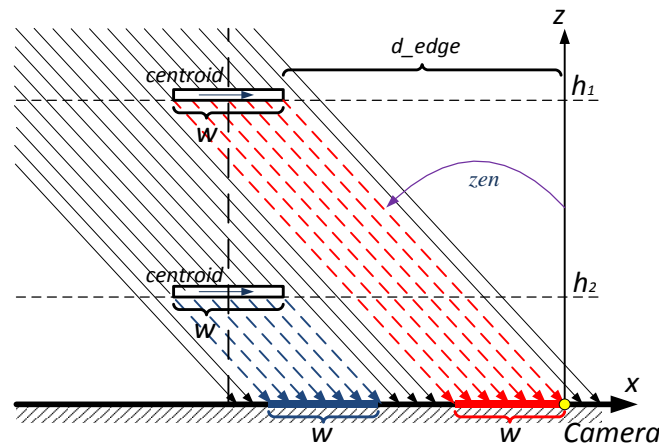
Since the contour of cumulus clouds is composed of fractals, it is possible to model it with a realistic random pattern [33]. Such a feature can be added to the four-point cloud model previously presented. Such a feature helps to take the dynamics of clouds into account, with respect to a simple warped rectangular shape.

### 3.5.4 Simple shadow model

#### 3.5.4.1 In the sky domain

Since the Sun is about 109 times bigger than the Earth in diameter, sunrays seemingly strike the surface of the Earth with the same incidence angle, which corresponds to the zenith angle  $zen$  of the Sun. It

must be noted that both the zenith angle of the Sun and the altitude of the cloud influences the position of the cast shadow. However, the altitude of a cloud does not determine the size of its cast shadow. In Figure 3.26 , an example is given showing two clouds of identical shape and size located in a 2-D plane  $(x, z)$ . Assuming this plane corresponds to a projection of the 3-D sky referential  $(x, y, z)$ , the centroids of both clouds have the same coordinates  $(x, y)$  but two different altitudes  $h_1$  and  $h_2$ . As one can see, the width  $w$  of the cast shadow is the same for both clouds.



**Figure 3.26 – Shadow model of a couple of clouds in a sky frame.**

Also, if the small yellow circle on the ground in Figure 3.26 represents the camera, provided the wind pushes clouds to the right hand side of the picture, towards the positive side of the x-axis, the top cloud starts casting a shadow over the camera as soon as the closest edge of the cloud gets located with an angle that equals the zenith angle  $zen$ .

In other words, this means the site of the camera will start to be shaded as soon as the distance between the vertical axis crossing the camera site and the edge of a cloud with a base at an altitude  $h_1$  becomes  $d\_edge = h_1 \times \tan(zen)$ .

In sum, in a sky referential, it is *the position of the cast shadow with respect to the position of the camera/solar module* that must be used to determine if the sunlight is dimmed.

### **3.5.4.2 Projected on the camera sensor**

If the camera shown in Figure 3.26 was to take a shot of the moment that immediately precedes a shading episode, triggered by a lone cloud of rectangular shape, the photo taken would look like Figure 3.27. The projection of the cloud gets distorted by the mapping function of the lens. If the cloud moves downwards within the image, i.e. towards the positive side of the y-axis, the cloud basically starts masking the sun as soon as the edge of the cloud overlaps the solar disk. Also, it suffices to consider

only clouds located within a region of interest around the sun. This region of interest is represented by an orange circle centred on the Sun in Figure 3.27.

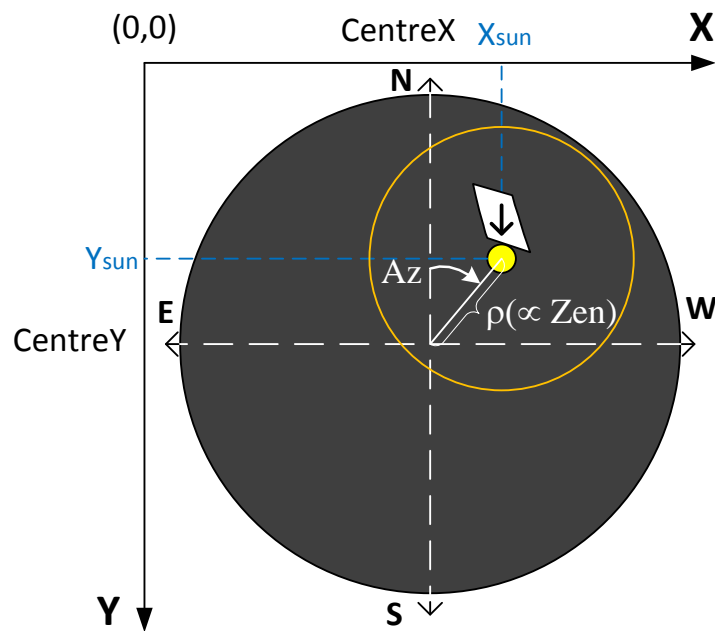
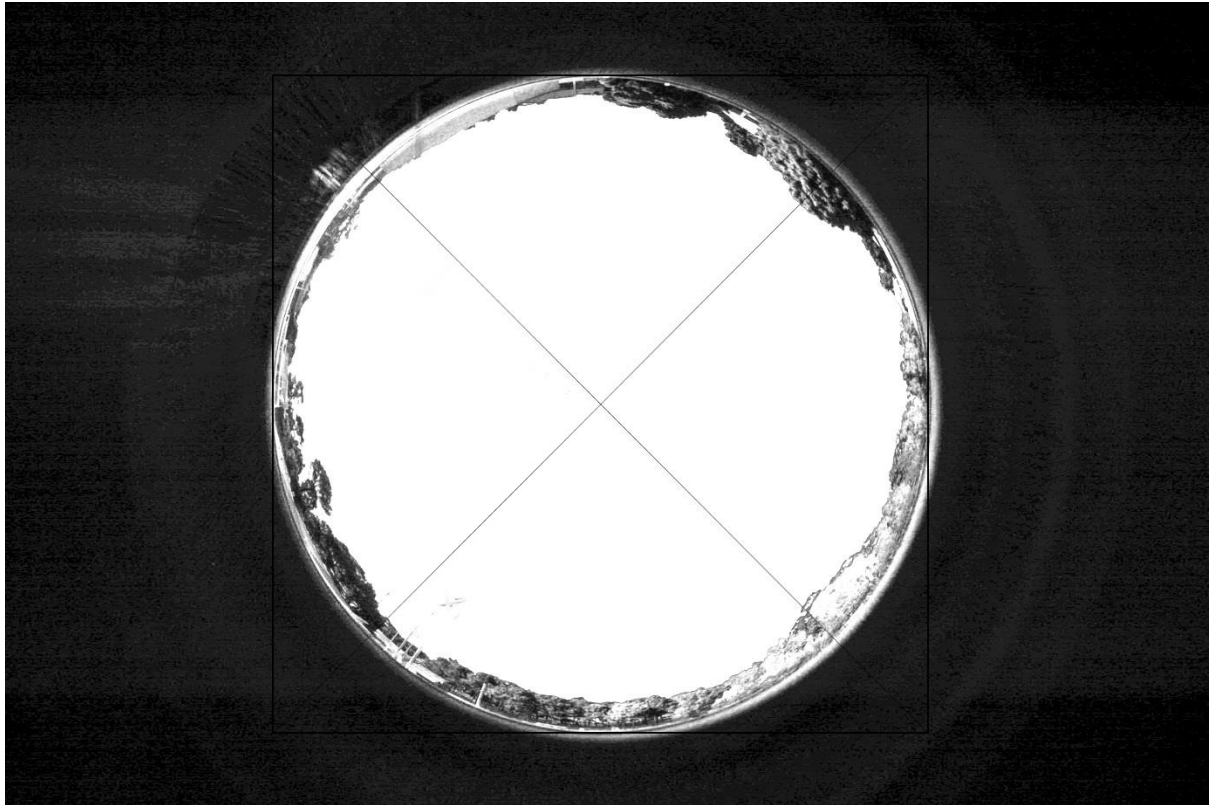


Figure 3.27 – Image referential frame.

To summarize, in a camera sensor referential, it is the *position of the image of a cloud with respect to the position of the solar disk* that must be used to define whether a shading episode occurs. However, only the actual size of the solar disk matters to define whether the Sun gets covered or not. This means that an object obstructing the glare around the solar disk does not cause any interruption of direct sunlight.

### 3.5.5 Camera Calibration

As discussed in literature [139], calibrating sky cameras is primordial prior to processing images. It is in fact necessary to characterize all the angular offset of the camera as they significantly affect the 2D projection of sky objects onto the sensor. In particular, it matters to find the exact location of the centre of the camera sensor, which corresponds to the location of the projection of the optical axis onto the sensor. The centre can be approximately found geometrically by fitting the circular image within a rectangle and by drawing diagonals from each corner. It helps to greatly enhance the contrast and luminosity of the image as shown in Figure 3.28. The crossing intersection of the diagonals is the estimated centre. This centre is the projection of the optical axis and does not necessarily match with the centre of the image. However, using a polynomial approximation of the projection by using a known pattern leads to more accurate results.



**Figure 3.28 – Geometrical approximate method to locate the centre of the projected image.**

With that latter approach, the OcamCalib Calibration Toolbox [140] developed by Davide Scaramuzza for MATLAB ® on the basis of the Caltech Calibration Toolbox by Jean-Yves Bouguet [141] is an interesting tool to serve the purpose of characterizing a sky camera equipped with a fisheye lens. To fit with most of cameras, the distortion mapping function of the lens is approximated by a polynomial function. In order to do so, the toolbox needs to be fed with a certain number of input images showing various orientations of a given checkerboard pattern. The toolbox is able to detect corners. Then, based on the information of the size of each square of the checkerboard specified by the user, the position of detected corners are used to model the mapping function.

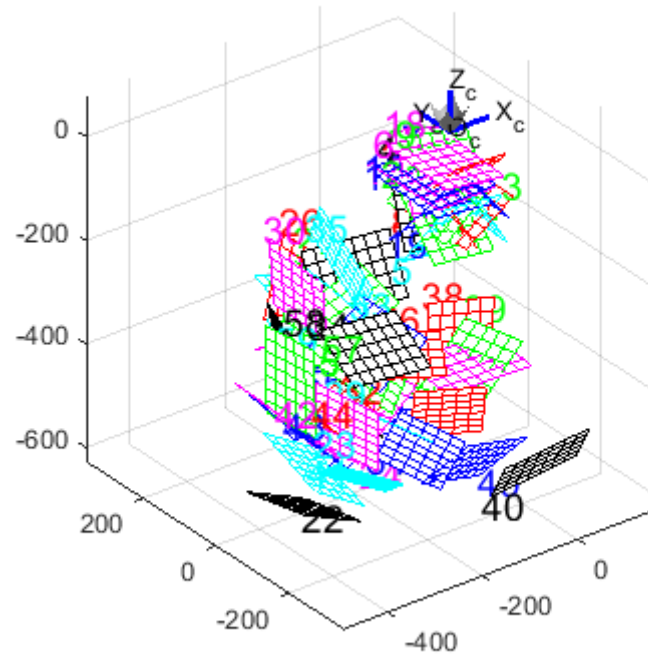


Figure 3.29 – Orientation of the reference checkerboard with respect to the camera.

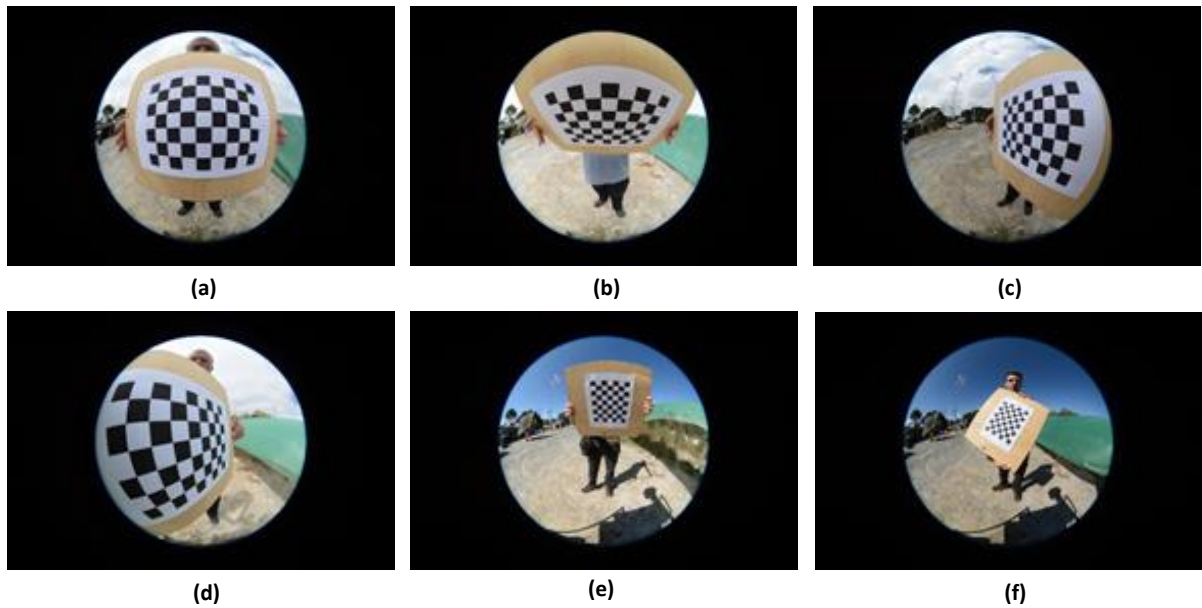


Figure 3.30 – Camera calibration procedure.

In our case, we know that the mapping function should be in the form of  $2 \cdot f \sin(\theta/2)$  since we chose to use an equisolid Sigma 4.5mm circular fish-eye lens.

A set of 61 24-JPEG images with a resolution of 2992×2000 has been fed to the toolbox to approximate the position of the centre along with the mapping function. The orientation of the checkerboard with respect to the camera is shown in Figure 3.29. As an illustration, some of the positions as seen from the camera during the calibration process are shown in Figure 3.30.

Assuming that both the camera and its lens are well aligned, the centre of the image has been found to be [1004.935883, 1484.065429]. In an ideal case, where the camera sensor is compound of perfectly square pixels and where the centre of the optical axis perfectly matches with the centre of the sensor, the 2D projection  $p$  of a vector  $P$  in the 3D world  $f$  is approximated with a 4<sup>th</sup> degree polynomial function  $F$  such that:

$$\text{if } p = \begin{bmatrix} u \\ v \end{bmatrix} \text{ then } P = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} u \\ v \\ f(u, v) \end{bmatrix} = \begin{bmatrix} u \\ v \\ f(\rho) \end{bmatrix} \text{ where } \rho = \sqrt{u^2 + v^2} \quad (3.111)$$

The distance  $\rho$  corresponds to the length of the vector  $p$ .

In a realistic case where pixels are not squares and where the optical axis is off-centre, we have the additional following relationship between the coordinates of the actual vector  $[u', v']$ , the ideal vector  $[u, v]$  and the actual centre  $[x_c', y_c']$ .

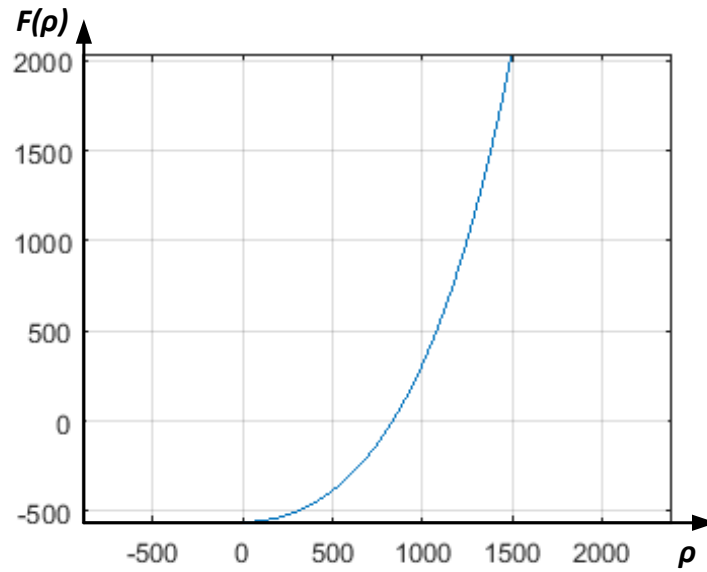
$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} c & d \\ e & 1 \end{bmatrix} \cdot \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} x_c' \\ y_c' \end{bmatrix} \quad (3.112)$$

Any given point  $p(u, v)$  within the circular region of the camera sensor is the 2D projection of a unit vector  $P(x, y, z)$  in the 3D world. In terms of length  $\rho$ , the retroprojection function  $F$  is in the following form:

$$F(\rho) = a_0 + a_1 \cdot \rho + a_2 \cdot \rho^2 + a_3 \cdot \rho^3 + a_4 \cdot \rho^4 \quad (3.113)$$

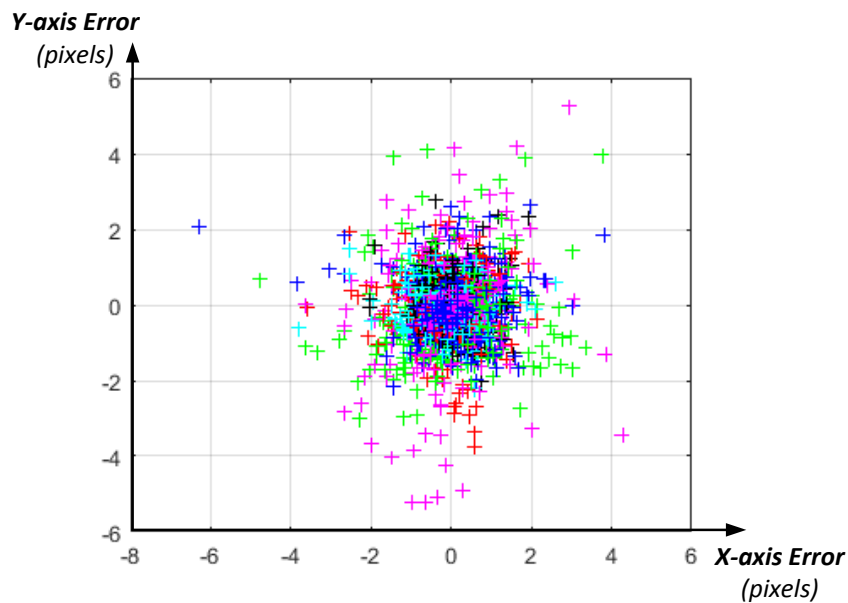
The retroprojection has been found to be:

$$F(\rho) = -561.61 + 6.49 \cdot 10^{-2} \cdot \rho^2 - 3.18 \cdot 10^{-9} \cdot \rho^3 + 2.29 \cdot 10^{-10} \cdot \rho^4 \quad (3.114)$$



**Figure 3.31 – Retroprojection function of the lens.**

That retroprojection function is illustrated in the graph in Figure 3.31 and every point of a retroprojected image is characterised by an error along the X-axis and the Y-axis expressed in pixels as shown in Figure 3.32:



**Figure 3.32 – Error of the retroprojection function along the X and Y axes.**

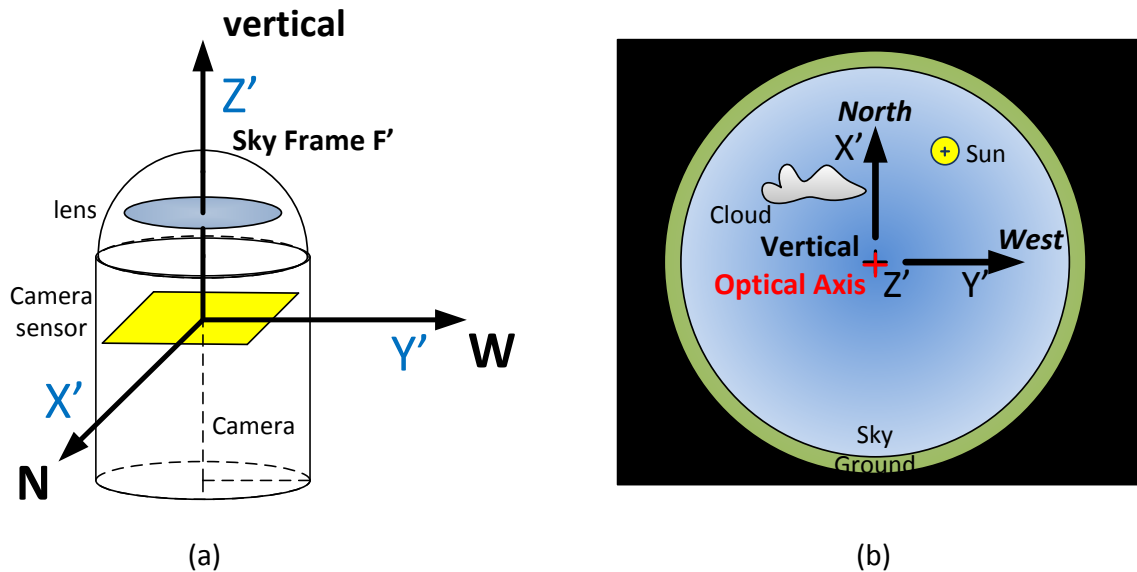
A sine wave can be represented per a Taylor series as shown in Equation (3.115):

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} \quad (3.115)$$



### 3.5.6 3D rotation matrix

As mentioned in the Single-point cloud model discussed in Section 3.5.1, an ideal camera should be installed so that its optical axis perfectly matches with the normal vector to the ground, which is the vertical direction, and so that the top edge of the sensor perfectly faces North, as shown in Figure 3.33. In particular, Figure 3.33 (a) illustrates the camera and the sky frame  $F$  in 3D space while Figure 3.33 (b) represents the corresponding 2D projection on the camera sensor.



**Figure 3.33 – Configuration of an ideal camera and the projected image on the sensor.**

In reality, the optical axis of the camera may be characterized by a tilt with respect to the normal to the horizon, which is in the vertical direction, on a vertical plane. Also, on a horizontal plane, the orientation of the camera may be twisted with respect to the direction indicating true North, as shown on Figure 3.34. These angles in the 3D frame have the effect of introducing a non-linear error in the 2D projection. As a result, it is difficult to correlate the expected position of the solar disk and clouds with the position of their projections on the camera sensor. A methodology is thus necessary to correct such a projection error whenever needed.

Like in the previous figure, Figure 3.34 (a) illustrates the camera and the sky frame  $F$  and the camera  $F'$  in 3D space while Figure 3.34 (b) represents the corresponding 2D projection on the camera sensor. However, unlike previously, it must be noted that the projected optical axis do match with the vertical direction. This affects the position, size and shape of the solar disk as well as any passing cloud.

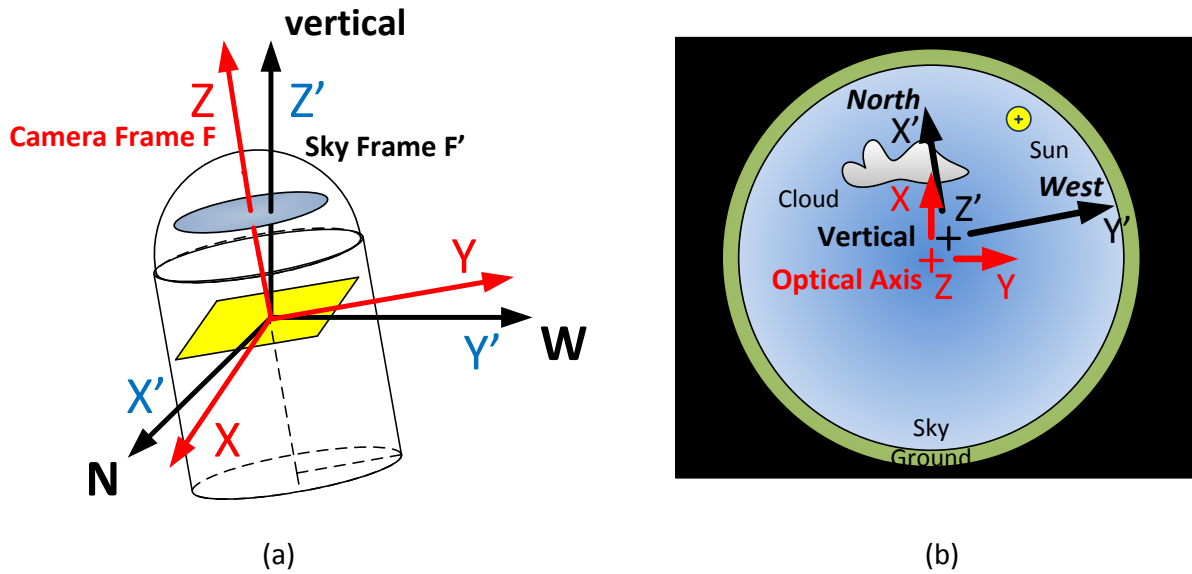


Figure 3.34 – Configuration of a real camera and the projected image on the sensor.

Therefore, in the 3D space, from the point of view of a frame of reference  $F$  composed of the true North for the  $X'$  axis, the true West for the  $Y'$  axis, and the true vertical for the  $Z'$  axis, following the left hand rule, the camera possesses its own frame of reference  $[X, Y, Z]$ . A 3D rotation can be applied to translate coordinates from one frame to the other, through the use of a  $3 \times 3$  rotation matrix  $R$ .

Let us consider the unit vector  $M'$  of a given direction in the 3D sky frame of reference  $F'$  composed of the unit axis  $[u', v', z']$  as shown in Figure 3.35 with  $u'$  and  $v'$  respectively aligned with the North and West directions while  $z'$  is aligned with the vertical axis. The vector  $M'$ , which could feature the sun position in the sky for example, can either be represented by a zenith angle  $\theta$ , an azimuth angle  $\phi$  and a radius  $r$  of 1 in spherical coordinates or by the Cartesian coordinates  $[x', y', z']$ :

$$\begin{aligned}
 M' &= \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \text{ in unit frame } F' = \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} \\
 \left\{ \begin{array}{l} r = \sqrt{x'^2 + y'^2 + z'^2} = 1 \\ \theta = \cos^{-1}(z') \\ \phi = \tan^{-1}\left(\frac{y'}{x'}\right) \end{array} \right. & \quad (3.116)
 \end{aligned}$$

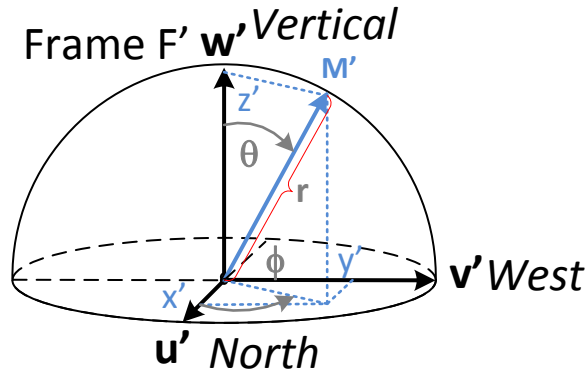


Figure 3.35 –Cartesian and spherical coordinates of a given point  $M'$  within the 3D unit frame  $F'$ .

The same vector is seen as the unit vector  $M$  rotated by  $R$  in the 3D camera frame of reference  $F'$ , so that:  $R \cdot M' = M$ , as illustrated in Figure 3.36.

$$M = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ in unit frame } F = \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

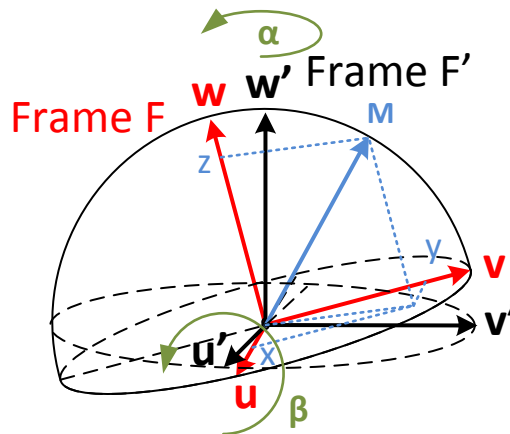


Figure 3.36 – Cartesian and spherical coordinates of a given point  $M$  within the 3D unit frames  $F$  and  $F'$ .

The frame  $F$  corresponds the frame  $F'$  after an anticlockwise rotation  $\alpha$  about the  $w'$  axis, while staring down at it, which corresponds to an azimuth twist, followed by an anticlockwise rotation  $\beta$  about the  $u'$  axis while staring down at it, which is in fact a zenith twist. The convention of staring down the axes will be adopted for every rotation angle. Also, it matters to twist the azimuth first and then the zenith.

Similarly, the single anticlockwise rotation by an angle  $\alpha$  about the  $w'$  axis as shown in Figure 3.37 is performed with the elementary azimuth rotation matrix  $R_z$ :

$$R_z = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

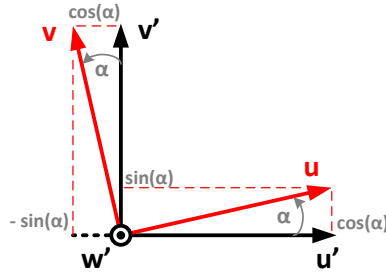


Figure 3.37 – Simple rotation of the unit frame around the  $w'$  axis by an angle  $\alpha$  as seen in the  $u'$ - $v'$  plane.

Similarly, the single anticlockwise rotation by an angle  $\beta$  about the  $u'$  axis as shown in Figure 3.38 is performed with the elementary zenith rotation matrix  $R_x$ :

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta) & -\sin(\beta) \\ 0 & \sin(\beta) & \cos(\beta) \end{bmatrix}$$

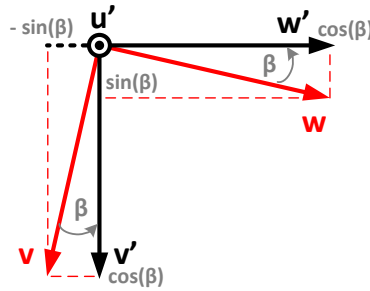


Figure 3.38 – Simple rotation of the unit frame around the  $u'$  axis by an angle  $\beta$  as seen in the  $v'$ - $w'$  plane.

The complete 3D transformation of the frame of reference after two rotations can thus be modelled by the following rotation matrix  $R$  that consists in the product of the two elementary rotation matrices:

$$R = R_z \cdot R_x = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \cdot \cos(\beta) & \sin(\alpha) \cdot \sin(\beta) \\ \sin(\alpha) & \cos(\alpha) \cdot \cos(\beta) & -\cos(\alpha) \cdot \sin(\beta) \\ 0 & \sin(\beta) & \cos(\beta) \end{bmatrix}$$

In order to find the position of the corresponding 2D projection  $m$  on the camera sensor starting from  $m'$ , it is necessary to convert  $m'$  in 3D. This can be done by using the projection function found during the calibration phase of the camera. The rotation  $R$  must then be applied to the resulting 3D vector  $M$ , giving the transformed vector  $M$  so that  $R \cdot M' = M$ . Likewise, the reprojection function can be used to find the 2D projected position  $m$  of the 3D rotated vector  $M$ .

### 3.5.7 Solar disk and cloud model

#### 3.5.7.1 Solar disk in the sky domain

The solar disk as it can be observed in the sky is characterized by a couple of parameters, as shown in Figure 3.39.

- The position of its centre, characterized by its azimuth and zenith ( $az, zen$ ), expressed in degrees
- Its radius  $R_{sun,angular}$  also expressed in degrees.

#### 3.5.7.2 Simplified solar disk projection on the camera sensor

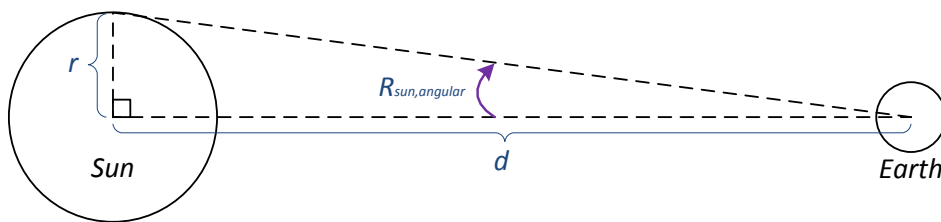


Figure 3.39 – Angular radius of the sun as seen from the earth.

The contour of the sun, also called solar disk, gets projected as an ellipsoid on the camera sensor. Also, the projected image of the solar disk on the camera sensor is characterized by the following parameters, all expressed in pixels:

- The position of its centre ( $X_{sun}, Y_{sun}$ )
- Its horizontal and vertical radius ( $R_{image\_sun\_x}, R_{image\_sun\_y}$ )

The relationship between the position of the centre solar disk ( $az, zen$ ) in the sky and the position of the centre of its projection on the camera sensor is exactly identical to the single-point object projection previously mentioned.

Therefore, such a relationship is defined by a couple of equations that are similar to (3.117).

$$\begin{cases} X_{sun} = CentreX - \left\| cst_x \cdot \sin\left(\frac{zen}{2}\right) \cdot \sin(az) \right\| \\ Y_{sun} = CentreY - \left\| cst_y \cdot \sin\left(\frac{zen}{2}\right) \cdot \cos(az) \right\| \end{cases} \quad (3.118)$$

The quantities  $cst_x$  and  $cst_y$  are respectively given by (3.94) and (3.95).

Given that the mean distance between the sun and the earth is  $d = 149.6$  million km and given that the mean equatorial radius of the sun is  $r = 696,342$  km, an elementary geometry calculation gives the angular radius of the Sun seen from the Earth in degrees:

$$R_{sun,angular} = \tan\left(\frac{d}{r}\right) = \tan\left(\frac{149.6}{0.696342}\right) \approx 0.7^\circ \quad (3.119)$$

Since the pixels of the camera sensor are not necessarily perfect square, the projection of the sun is characterised by a horizontal radius and a vertical radius, both expressed as follows in pixels:

$$R_{image_{sun_x}} = CentreX - \left\| \frac{cst_x}{2} \cdot \left( \sin\left(\frac{az + R_{sun,angular}}{2}\right) - \sin\left(\frac{az - R_{sun,angular}}{2}\right) \right) \right\| \quad (3.120)$$

$$R_{image_{sun_y}} = CentreY - \left\| \frac{cst_y}{2} \cdot \left( \sin\left(\frac{zen + R_{sun,angular}}{2}\right) - \sin\left(\frac{zen - R_{sun,angular}}{2}\right) \right) \right\| \quad (3.121)$$

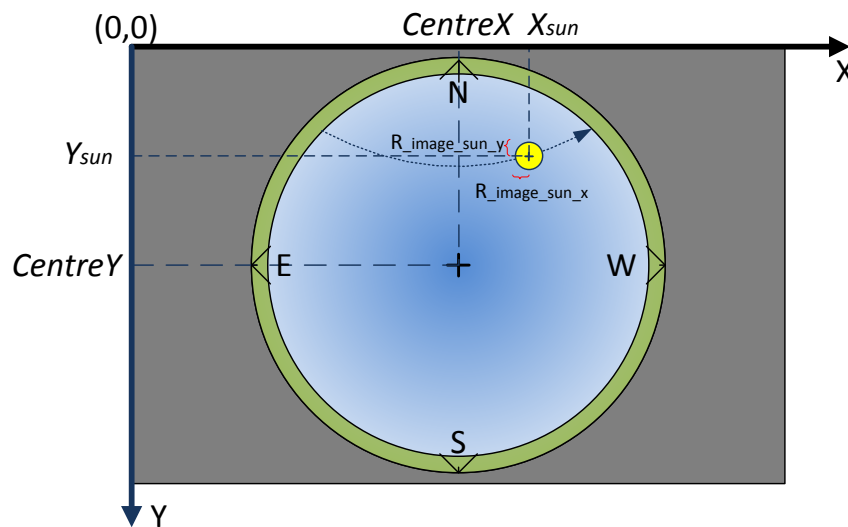
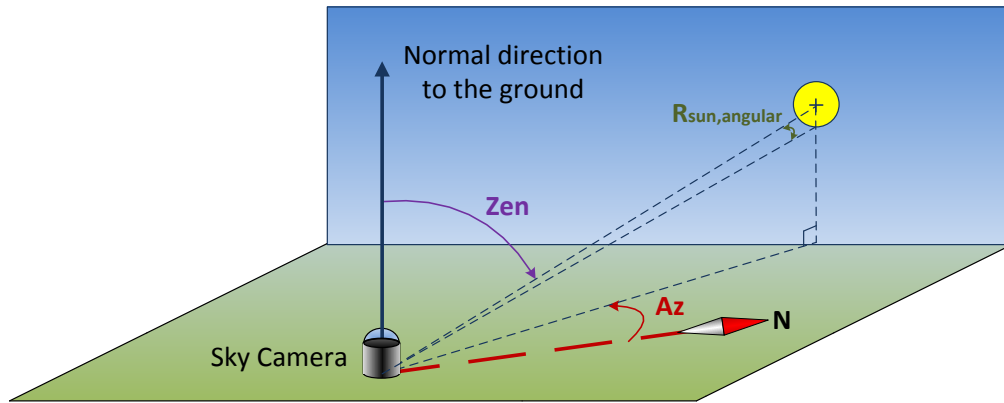


Figure 3.40 – Angular position and angular radius of the sun in the sky.



**Figure 3.41 – Position of the projected solar disk on the camera sensor.**

Figure 3.40 and Figure 3.41 respectively illustrate the position of the Sun in the sky and the position and size of the projected image of the solar disk on the camera sensor.

### **3.5.7.3 Actual solar disk projection on the camera sensor**

If the camera was ideally set up, it should be sufficient to calculate the position of the sun in the sky in terms of zenith and azimuth angle by using NREL’s Solar Position Algorithm (SPA) that convert the circular coordinates to a Cartesian unit vector  $M' = [x', y', z']$  and 2D-project that sun position onto the camera sensor with the reciprocal of the projection function to find the position  $m = [X, Y]$ . However, as previously mentioned, cameras are actually characterised by a tilt and twist angle that induce a mismatch between the position of the sun as it should be and as it actually appears on the image. As an illustration, we considered a set of 56 images shot every 30 minutes to constitute a clear-sky images covering all 4 seasons. Over that sample, the error on the sun position can be as much as 45 pixels in the images of resolution  $2992 \times 2000$  and 21 pixels in average, which is quite significant. This is visible in Figure 3.42, which shows a clear sky picture taken on 28/11/2016 at 11:05 AM. The cross in red in that picture indicates the predicted position of the sun whereas the cross in black shows the location of the actual centre of the solar disk. A tool inspired from Son Lam Phung’s MATLAB® function from 2004 [142] and similar to the “magic wand” feature in Adobe Photoshop has been used to identify the actual centre of the solar disk featured by that cross in black.



**Figure 3.42 – Clear sky picture showing the position mismatch between predicted and actual projection of the centre of the solar disk.**

The error of the sun position can be mitigated with a 3D rotation matrix. However, since the orientation of the camera can be affected by maintenance operation, it has been necessary to recalculate the rotation matrix every day. The matrix is determined by the location of the sun itself, as performed by Gauchet et al. [143] and Savoy et al. [144] rather than the stars like in astronomical studies. It is indeed more convenient to work on daytime conditions given the fact that the camera has been set up for direct sunlight conditions in the middle of the day.

The rotation matrix is found by solving a system so that  $R \cdot M' = M$ , where  $M'$  is the theoretical 3D unit vector of the sun position and  $M$  is the actual 3D unit vector of the sun position, both in given in Cartesian sky coordinates. The rotation matrix  $R$  is in a  $3 \times 3$  matrix with 9 coefficients to find:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

In order to solve the system, both  $M'$  and  $M$  actually need to be  $n \times 3$  matrices containing  $n$  vectors:

$$M' = \begin{bmatrix} M'_{11} & M'_{12} & M'_{13} \\ \dots & \dots & \dots \\ M'_{n1} & M'_{n2} & M'_{n3} \end{bmatrix}, \quad M = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ \dots & \dots & \dots \\ M_{n1} & M_{n2} & M_{n3} \end{bmatrix}$$



Each vector is drawn from each individual image shot throughout the day. Nevertheless, we have only considered zenith angles up to 70 degrees.

For example, the rotation matrix  $R$  has been identified as the following respectively as of the 1<sup>st</sup> of July 2016, as of the 28<sup>th</sup> of November 2016, as the 1<sup>st</sup> of January 2017, and as of the 4<sup>th</sup> April 2017:

$$R_{01_07_2016} = \begin{bmatrix} 1.0174 & 0.0184 & 0.0022 \\ 0.0172 & 1.0055 & 0.0277 \\ -0.0011 & 0.0159 & 0.9652 \end{bmatrix}$$

$$R_{28_11_2016} = \begin{bmatrix} 0.9539 & 0.0491 & -0.0717 \\ -0.0839 & 1.0088 & -0.0400 \\ 0.0536 & 0.0346 & 0.9954 \end{bmatrix}$$

$$R_{25_01_2017} = \begin{bmatrix} 1.0089 & 0.0331 & -0.0356 \\ -0.0261 & 1.0000 & -0.0221 \\ 0.0554 & 0.0225 & 1.0011 \end{bmatrix}$$

$$R_{04_04_2017} = \begin{bmatrix} 0.9833 & 0.0408 & -0.0449 \\ -0.0463 & 1.0012 & -0.0237 \\ 0.0338 & 0.0238 & 0.9987 \end{bmatrix}$$

It can be observed that all four matrices slightly vary from one day to another. However, it is essential to calculate an individual rotation matrix for each day to achieve an optimal correction and a minimum error. Since the camera in use has been uninstalled and reinstalled for maintenance purposes over the months, it is unfortunately impossible to find a unique rotation matrix for every day of the year that would minimize the error as effectively as a daily matrix.

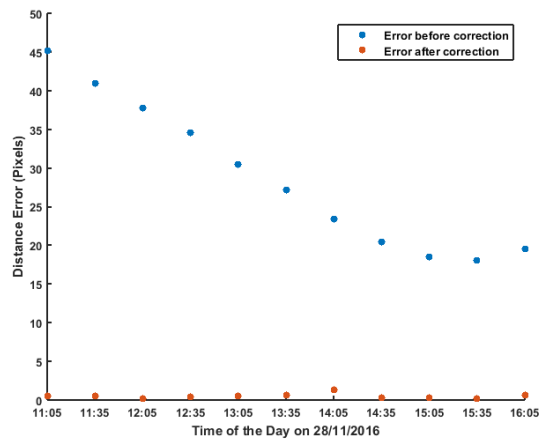
Also, it is worth noticing that the matrix does not look exactly like the rotation matrix as previously introduced, possibly due to the non-linear nature of the projection function through the fisheye lens as well as the occasional swing of the camera pole on windy days.

The dataset features a day from each season and the time resolution has been set to 30 minutes.

**Table 3.9 – Distance error between the predicted and the actual position of the sun on four different days, expressed in pixels, before and after application of each respective 3D rotation matrix.**

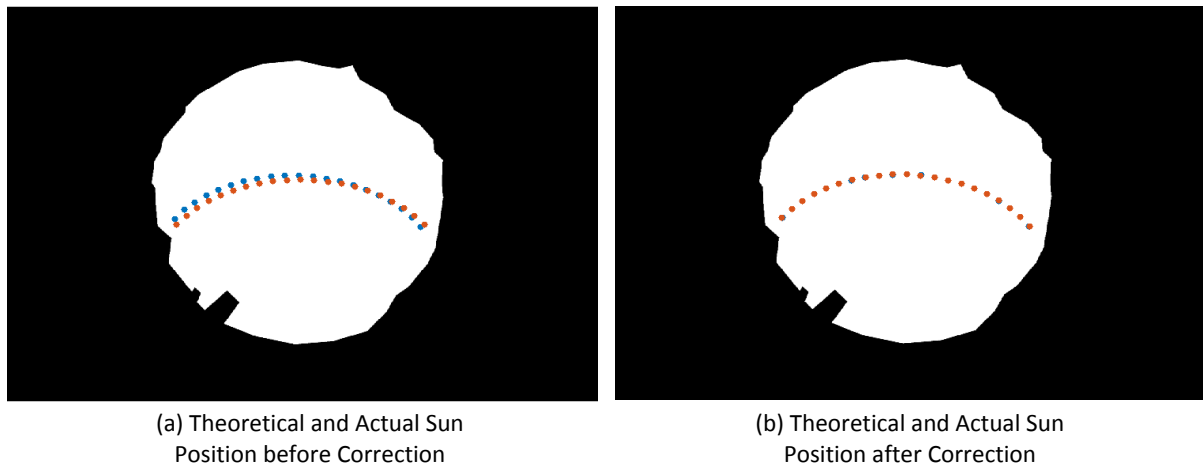
Dates	Season	Start Time	End Time	No. of images	Before Correction			After Correction		
					RMSE (pix.)	Min Error (pix.)	Max Error (pix.)	RMSE (pix.)	Min Error (pix.)	Max Error (pix.)
1/07/2016	Winter	12:00	15:00	7	8.71	1.67	12.58	1.43	0.37	2.97
28/11/2016	Spring	11:05	16:05	11	30.17	18.01	45.17	0.57	0.13	1.35
25/01/2017	Summer	7:29	17:29	21	24.13	9.09	33.30	1.69	0.20	3.85
4/04/2017	Autumn	8:30	16:00	16	22.64	7.02	40.12	2.34	0.26	6.86

The distance error for the specific day of 28<sup>th</sup> of November 2016 expressed in pixels between the predicted and the actual position of the sun is plotted in Figure 3.43, before and after correction.



**Figure 3.43 – Distance error between the expected projected position of the centre of the solar disk on the camera sensor and its actual position, expressed in pixels, before and after 3D matrix correction.**

To further illustrate the effectiveness of the 3D rotation to minimize the error, the actual positions in are shown in red and the predicted positions are shown in blue in Figure 3.44. The part on the left noted (a) shows the actual and predicted trajectories of the sun over the course of the day on the 25<sup>th</sup> of January 2016 with the background masked, without any correction while the part on the right noted (b) shows the same trajectories after correction with an adequately calculated rotation matrix.



**Figure 3.44 – Expected projected position of the centre of the solar disk and its actual position on the camera sensor on 28/11/2016 for all the studied hours before (a) and after correction (b).**

Although Savoy et al. [144] used a 2-minute time resolution, a high temporal resolution may not be needed to generate an effective 3D correction matrix. Especially considering the fact that the 3D matrix analysis can be time consuming, it matters to define the optimal minimum time resolution and number

of sky images required for an effective 3D correction. In our application, it takes an average time of 0.8 seconds per image to generate an effective 3D rotation correction matrix.

**Table 3.10 – Investigating the optimal time resolution to calculate the 3D rotation matrix on 25/01/2017.**

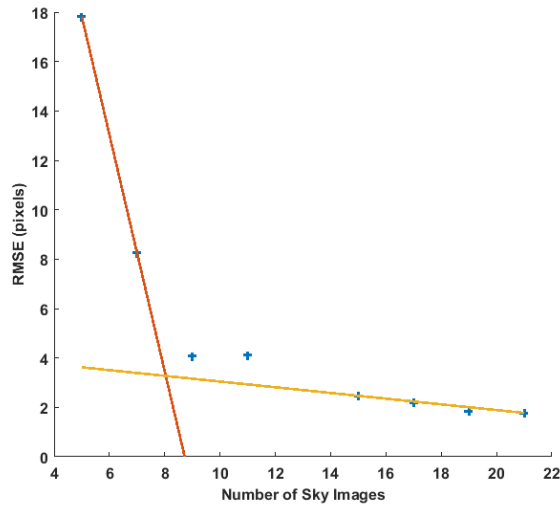
			Before Correction			After Correction		
Time Res. (min)	Number of Images	Analysis Time (seconds)	RMSE (pix.)	Min Error (pix.)	Max Error (pix.)	RMSE (pix.)	Min Error (pix.)	Max Error (pix.)
2	358	241.18	24.16	8.34	35.58	1.73	0.09	6.73
4	155	130.95				1.73	0.10	6.61
10	62	46.42				1.73	0.16	6.71
20	31	25.38				1.73	0.10	6.65
30	21	16.87				1.73	0.15	6.48
60	10	8.54				1.77	0.11	7.29

As shown in Table 3.10, the RMSE remains completely unchanged between 2 minutes and 30 minutes while the processing time varies by a factor 14 and the number of images varies by a factor 17. The lowest maximum distance error is even achieved for a resolution of 30 minutes. Thus, the optimal time resolution that guarantees the most accurate result with a minimal computation effort is 30 minutes on that day.

**Table 3.11 – Investigating the optimal number of sky images to calculate the 3D rotation matrix on 25/01/2017 with a 30 minutes time resolution.**

		Before Correction			After Correction		
Number of Images	Analysis Time (seconds)	RMSE (pix.)	Min Error (pix.)	Max Error (pix.)	RMSE (pix.)	Min Error (pix.)	Max Error (pix.)
5	5.82	24.16	8.34	35.58	17.83	0.13	47.35
7	7.59				8.25	0.12	26.69
9	8.67				4.08	0.06	15.72
11	10.24				4.11	0.10	15.06
15	12.02				2.47	0.05	9.90
17	13.48				2.17	0.20	9.20
19	16.31				1.84	0.07	7.68
21	16.63				1.78	0.0089	6.95

A further optimisation would consist in determining the minimum number of sky images required for an effective 3D correction. An increasing number of sky images centred on and around solar noon has been selected to answer that question. A rotation matrix has been calculated for each set of images. Table 3.11 shows the RMSE obtained after correction on 25/01/2017 when a 30 minutes time resolution is applied to generate a rotation matrix.

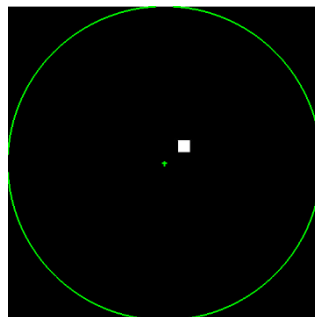


**Figure 3.45 – Scatter plot of the distance error between the expected and actual centre of the solar disk in function of the number of sky images, with asymptotes.**

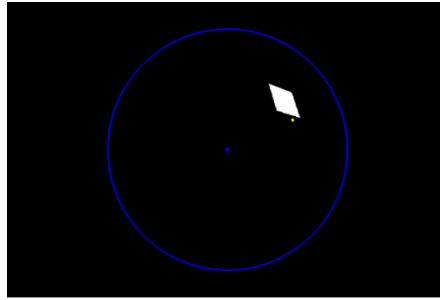
The RMSE has been scatter-plotted with blue crosses in Figure 3.45. The asymptotes corresponding to the lowest and the highest RMSE intersect around a number of images of 8. Since the RMSE does not increase significantly passed this value, we can conclude that the optimal setting to calculate the 3D rotation matrix is at least 8 sky images centred around solar noon, with a time resolution of 30 minutes.

#### ***3.5.7.4 Simplified cloud model***

When the projection function is known, the position of clouds can be modelled in a similar manner to what has been previously cased in the simplified solar disk projection model. Figure 3.46 shows the example of a cloud located above the camera as it could be seen by a satellite. The cloud has been modelled as a square with a side 1.8 km. The entire image in picture 10 shows clouds within a 50 km radius, featured by a green circle, around the location of the camera is indicated by a green cross in the centre.



**Figure 3.46 – Example of a position for a cloud in satellite view.**



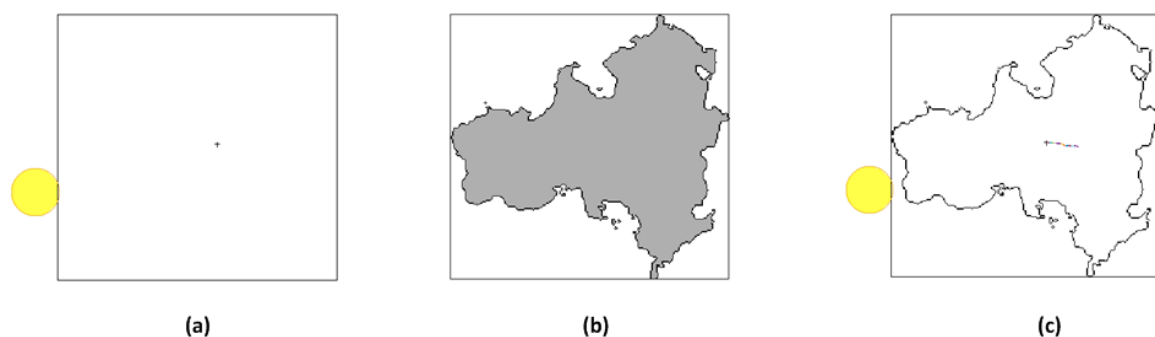
**Figure 3.47 – Same cloud seen by the camera with the solar disk shown in yellow.**

Figure 3.47 shows the exact same cloud as in Figure 3.46, as seen by the previously mentioned fisheye ground camera. The solar disk is also represented by a yellow circle. This circle may look quite small though. Indeed, as previously demonstrated, the actual solar disk is  $1.4^\circ$  wide in the sky. The brightness of Sun induces a glaring effect that makes it look bigger that it actually is.

### **3.5.7.5 Realistic cloud model**

In the simplified model presented above, clouds have been modelled as rectangular in shape. However, the shape of clouds may be very complex. If we focus on the specific case of cumulus clouds, their contour can be modelled as fractals [33]. Thus, in order to forecast as accurately as possible the moments when the sun gets covered or uncovered by the sun, it matters to take into account the actual contour of clouds, especially for those located within the region of interest.

With that said, tracking every single point of a cloud may be excessively time-consuming and even not absolutely necessary. Instead, each target identified as a cloud of significant size within the region of interest can be fitted inside a bounding box (BB). It is then sufficient to track the centroid of that BB. Also, the contour of the cloud may be saved as a binary image and may be called to check if an overlap will occur between the solar disk and a predicted future position of the cloud.



**Figure 3.48 – First example of cloud tracking and sun cover measurement.**

As an example, a cumulus cloud could be entirely enclosed inside a BB as shown in Figure 3.49 (a). The centroid of this BB is marked by a cross inside it. Also, since one of the edges of the BB is in

contact with the rim of the solar disk, the sun could possibly be covered by the cloud enclosed in the BB. For that reason, it is useful to save the shape of the cloud at the current instant  $k$  in order to determine whether the sun is indeed being covered or not. A light-weight binary image containing the contour of the cloud like in Figure 3.49 (b) superposed to its bounding box leads to the conclusion that the sun is actually not being covered as it can be seen in Figure 3.49 (c). Also, in Figure 3.49 (c), the trajectory followed by cloud has been drawn to show the direction that has been taken by the cloud since the beginning of the sequence. The size of each segment shows the distance that has been travelled by the cloud between two steps, which corresponds to the speed of the clouds per step.



**Figure 3.49 – Second example of cloud tracking and sun cover measurement.**

The actual moment step when the solar disk gets covered may happen in a later stage within the same sequence, as shown in the second example in Figure 3.49. Although the BB covers a fair part of the solar disk, given the shape of the cloud, the sun just starts getting covered. It must also be noted that the shape of the cloud has been updated and has changed since the time of the first example. An advantage of the approach is to take into account the dynamics of the cloud with minimum calculations.

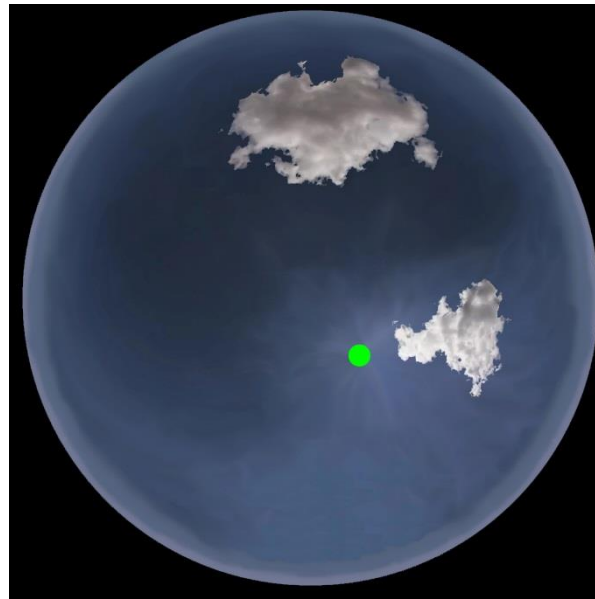


**Figure 3.50 – Third example of cloud tracking and sun cover measurement.**

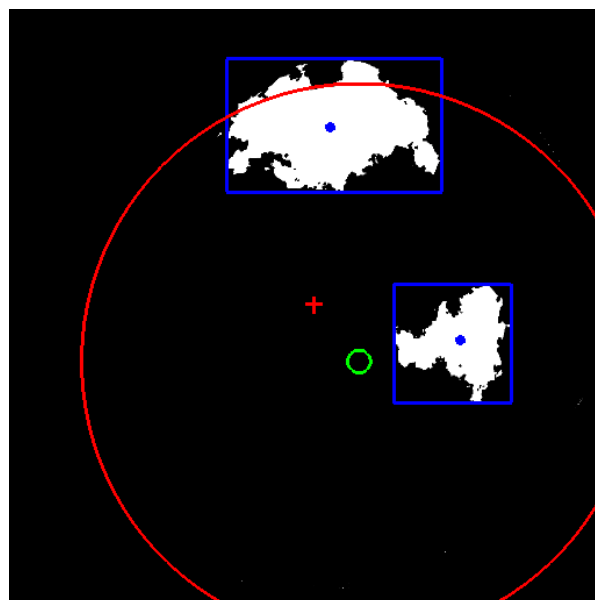
Another case that may happen in an even later stage is depicted in Figure 3.50. In that case, the BB entirely encapsulates the solar disk. One could thus assume that the sun might be entirely covered. However, superposing an updated version of the shape of the cloud on top of its bounding box shows

that sun is actually only partially covered. The cloud covered pixels inside the solar disk are shown in orange in Figure 3.50 (c). The trajectory of the centroid is also featured in the same figure.

When a prediction filter tracks the BB that encloses a cloud, two quantities can be tracked: the position of the centroid of the cloud and the speed. It makes sense to track the speed because clouds follow a continuous movement, even though they may be characterized by some dynamics such as shape and size changes over time.



**Figure 3.51 – Example of an image of the sky simplified with Photoshop.**



**Figure 3.52 – Bonding boxes enclosing binary shapes.**

The image on Figure 3.51 features a sky configuration with only two sufficiently large clouds. The solar disk has been masked in green to be easily recognized. As it can be seen in Figure 3.52, it is the bounding boxes and centroids of sufficiently large clouds that are tracked rather than their entire contour. The BB and the centroid of each cloud is represented in blue.

The binary shape of the cloud is saved at every sample but only used to detect shading rather than for tracking purposes. A shading occurs when as soon as there is an overlap between the solar disk and the actual shape of a cloud occurs, like shown in Figure 3.53.

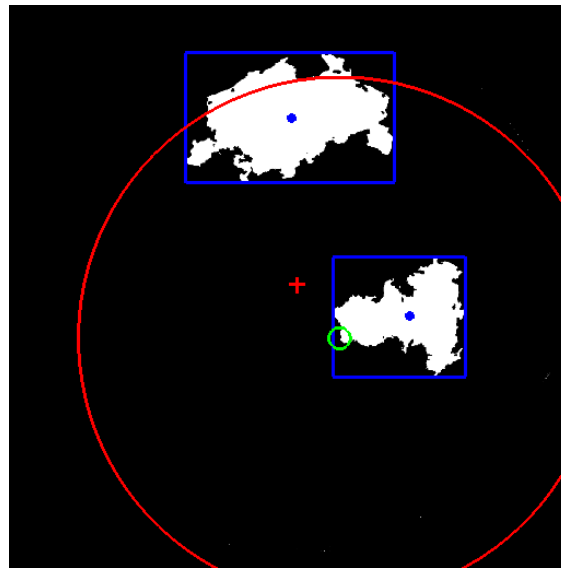
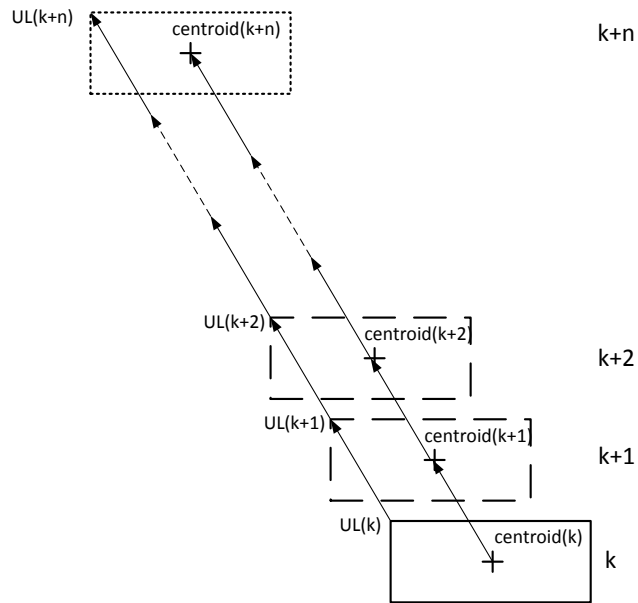


Figure 3.53 – Shading episode as seen through a sky camera.

### 3.5.8 Estimation of the shading time and duration

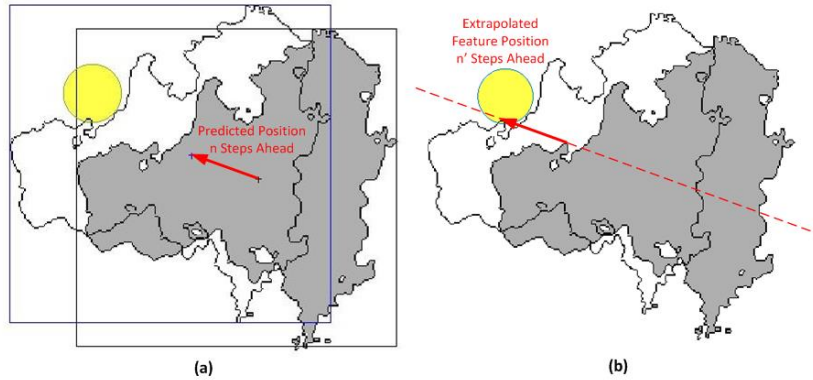
As previously mentioned, the position of the centroid of the bounding boxes encapsulating relevant clouds is tracked. To be considered as relevant, clouds must be large enough and their centroid must be within the region of interest around the sun. Using a predictive filter that will be described in the further chapters, the future position of the centroid is forecast a few runs ahead. The current position of the upper-left (UL) corner of the BB is saved at each run. By combining both pieces of information, it becomes relatively easy to forecast the next positions of the upper-left corner of the BB in a sky referential, assuming the BB will not change in shape nor size. Based on the previous positions of the centroid of a specific relevant cloud, a recursive prediction filter can estimate a transformation function  $F$ . Provided this function  $F$  is known at time  $k$ , it can be used to predict the future positions of the centroid of the same cloud one step ahead:  $X_{centroid}(k+1) = F \times X_{centroid}(k)$ . The same function can also be used iteratively  $n$  times to predict the future state of the cloud  $n$  steps ahead:  $X_{centroid}(k+n) = F^{(n)} \times X_{centroid}(k)$ . By applying the same transformation to the upper left corner of the bounding box UL:  $X_{UL}(k+n) = F^{(n)} \times X_{UL}(k)$  as shown in Figure 3.54.





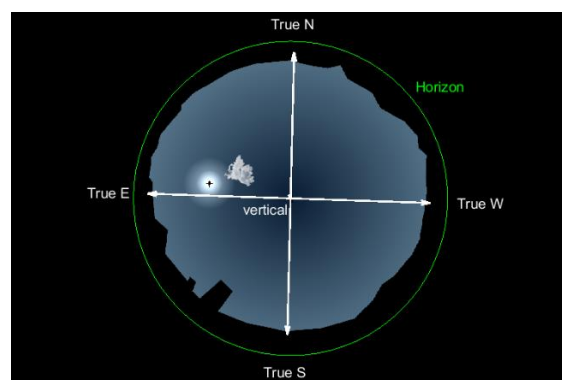
**Figure 3.54 – Current position and predicted positions of a single cloud bounding box.**

Such an approach is particularly relevant for very short-term forecasts, up to a few steps ahead, for which the persistence assumption has proven to be very accurate [53]. Up to a few steps ahead, this method can be applied in both the sky referential and the camera sensor referential. As an example, when the sun is just merely a few steps away from beginning to be shaded, the position of the solar disk as seen on the camera sensor can be assumed to be virtually fixed and the shape of the cloud can be assumed to be fixed as well, both in terms of actual shape and size but also in terms of distortion. It may thus be sufficient to apply this method directly in the camera sensor domain, without retro-projecting the image. As shown in Figure 3.55 (a), by using bounding box-centroid method, the position of the bounding box that encloses the cloud can be forecast  $n$  steps in the future through an extrapolation based on the last measurement of the motion of the centroid of the cloud. The latest updated shape of the cloud can thus be superposed on the position of the forecast the bounding box to assess the possible overlap with the solar disk. Alternatively, a state of the art method would consist in extrapolating the position of several features located around the contour of the cloud shape. By drawing linear approximations of the trajectory of such features, it can be found that at least one feature may cross the solar disk  $n'$  steps ahead, where  $n'$  is not necessarily found to be the same number as  $n$ .

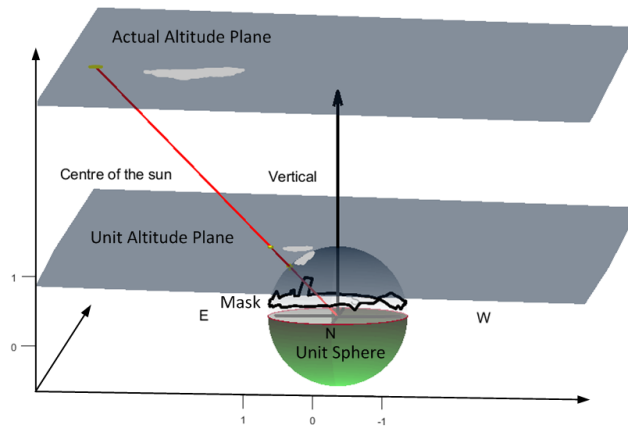


**Figure 3.55 – Direct application of the bonding box and centroid forecasting approach and the feature approach in the camera sensor referential.**

Besides the possible variations in shape, size and altitude of clouds, a Euclidian space referential like the sky does not imply any distortion, unlike the projection on the camera sensor. Hence, for forecasts several dozens or hundreds of steps ahead, it may be greatly preferable to apply this method to a sky Euclidian referential, which means converting camera images to sky coordinate images. An example is illustrated in Figure 3.56 and Figure 3.57. The image as seen on the camera sensor that is shown Figure 3.56 actually corresponds to the real three-dimensional world situation shown in Figure 3.57. The mask is shown on both figures and the axes corresponding to the true cardinal directions is shown in Figure 3.56. The cloud that is shown with a warped shape and at a warped distance for the centre of the sun in Figure 3.56 can be retro-projected onto the surface of the unit sphere as shown in Figure 3.57. The actual shape, size and motion of the cloud is comprised onto the actual altitude plane. However, for tracking purposes, the actual altitude of clouds is not required and it sufficient to track clouds within the unit altitude plane only.

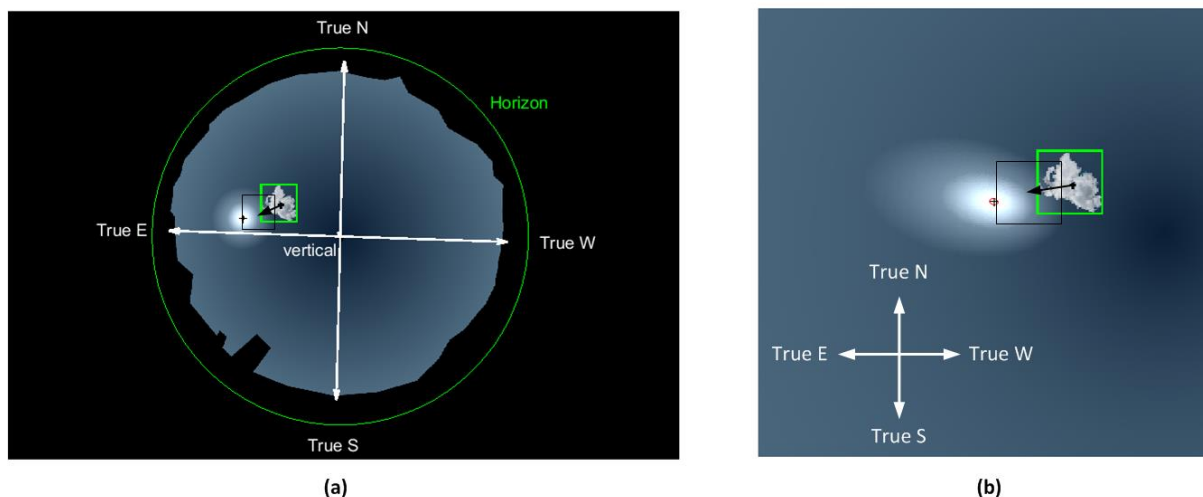


**Figure 3.56 – Example of a reprocessed sky image taken on 28/01/2017 at the GEEP lab at 9:23 am.**



**Figure 3.57 – Retro-projection of the reprocessed sky image taken on 28/01/2017 at the GEEP lab at 9:23 am onto the unit sphere, unit altitude plane and actual altitude plane, within the actual 3D world.**

The only difference between tracking cloud at the actual altitude they move and at the unit altitude is the value of the speed. Thus, it is necessary and also sufficient to retro-project clouds onto the unit altitude plane. Figure 3.58 (a) features the bounding box and a forecast position of the bounding box several steps ahead within the sky image sensed by the camera. The projected bounding box gets more and more warped as it gets closer to the rim of the image, or conversely less and less warped as it moves towards the centre of the image. This contrasts with the situation of the equivalent bounding box that fits the same cloud when it is retro-projected onto the unit altitude plane. On this plane, as shown in Figure 3.58 (b), the bounding box can be safely assumed to retain the same size as it is not affected by any optical projection function.



**Figure 3.58 – Comparison the bounding boxes within the camera sensor domain (a) and within the unit altitude plane.**

Only the dynamic changes of the shape of the cloud due to the effect of the wind may affect the bounding box size. However, such an effect may be slow and may only have a limited impact on the size of the

bounding box. It must be noted that for optimum results, it is additionally absolutely necessary to update the position of the centre of the solar disk and the size of the solar disk itself at every forecast step.

The exact same reasoning can also be applied to forecasts of cloud clearing times. The time when a cloud clears out after covering the sun can thus be forecast through the bounding box approach or through the feature approach that has just been described. However, due to the significant effect of distortion due to the use of a fisheye lens, it is strongly recommended to track the cloud on the unit altitude plane rather than directly within the domain of the camera sensor. Although the dynamic change of shape of the cloud can hardly be forecast and needs to be assumed to remain constant, the

Considering some cloud images in sky coordinates, the projection of the cast shadows matters a lot more than the position of the clouds themselves. Yet, as shown in Figure 3.54, the position of the cast shadow on the ground depends on both the altitude of the base of the cloud and on zenith angle of the sun. A cumulus clouds with a typical flat base located at an altitude  $h$  while the zenith of the sun is  $zen$  casts a shadow on the ground with the same size and shape. Indeed, the shadow is nothing but a 2-D projection of the base of the cloud. The shadow may be enclosed within a bonding box the same way as the cloud itself can be. However, at the instant  $k_s$ , the upper-left corner of the bonding box of the shadow is shifted by a distance of  $d_{cs}(k_s) = h \times \tan(zen)$  with respect to the upper-left corner of the bonding box of the cloud. Figure 3.59 shows a side view of the possible movement of a single cloud alongside the  $x$ -axis on the ground.

Figure 3.60 represents the same cloud and the same movement seen from above, with a top view. In such a situation, with the cloud moving towards the camera, the exact spot occupied by the camera starts to be shaded when the edge of the shade crosses that spot.

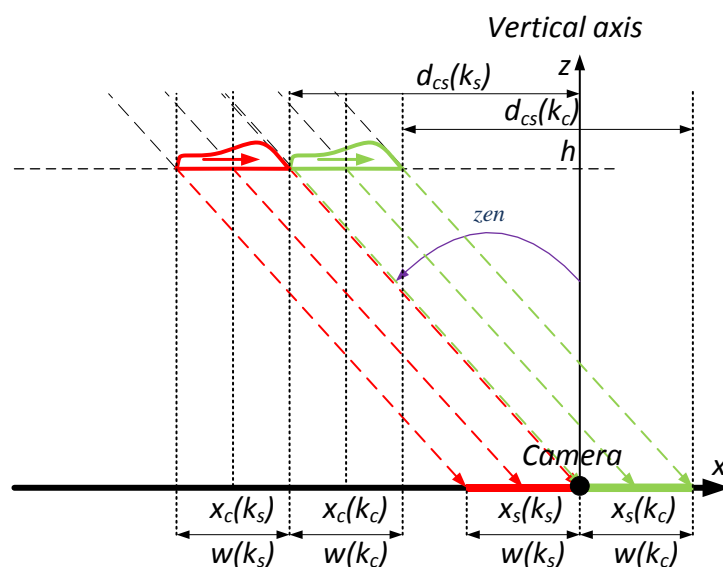


Figure 3.59 – Shading Model in  $(y, z)$  two-dimensional sky plane or side view of a shading episode.

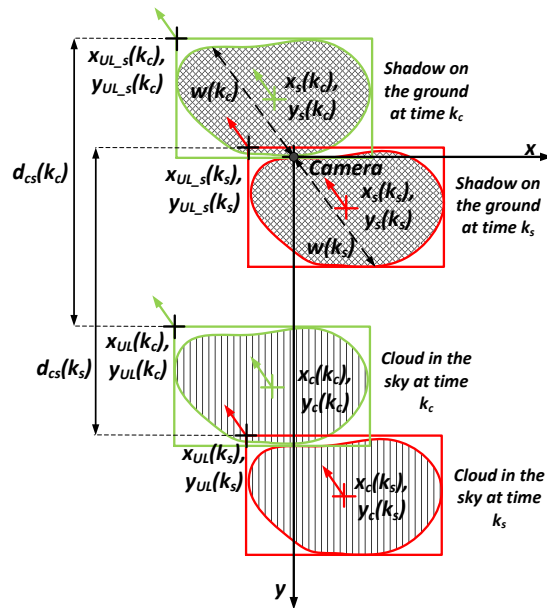


Figure 3.60 – Shading model in  $(x, y)$  two-dimensional sky plane or top view of a shading episode.

That spot remains shaded until the opposite edge crosses it. That means the shading lasts for the time it takes for the cloud to move over a distance corresponding to its diagonal dimension according to the direction of movement, noted  $w$  in Figure 3.60. After getting pushed by the wind, the centroid of the cloud moves from  $(x_c(k_s), y_c(k_s))$  to  $(x_c(k_c), y_c(k_c))$ . Assuming its diagonal dimension  $w$  remains unchanged during the period  $[k_s, k_c]$ , the cloud then moves across its own diagonal  $w$  during that time. The duration of the shade thus corresponds to the time the cloud needs to move across its own diagonal  $w$ .

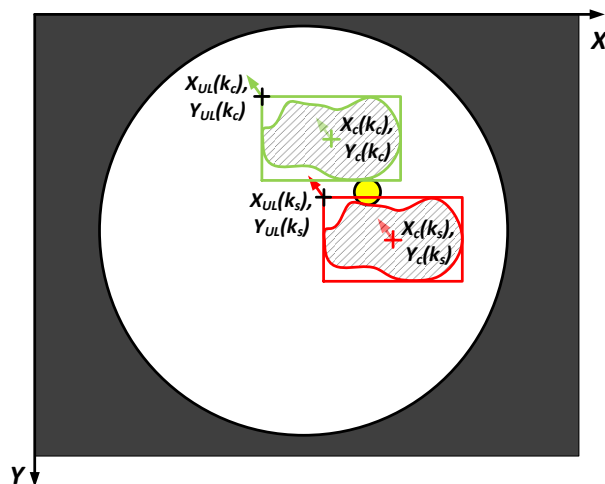


Figure 3.61 – Shading model as seen through a ground camera of a shading episode.

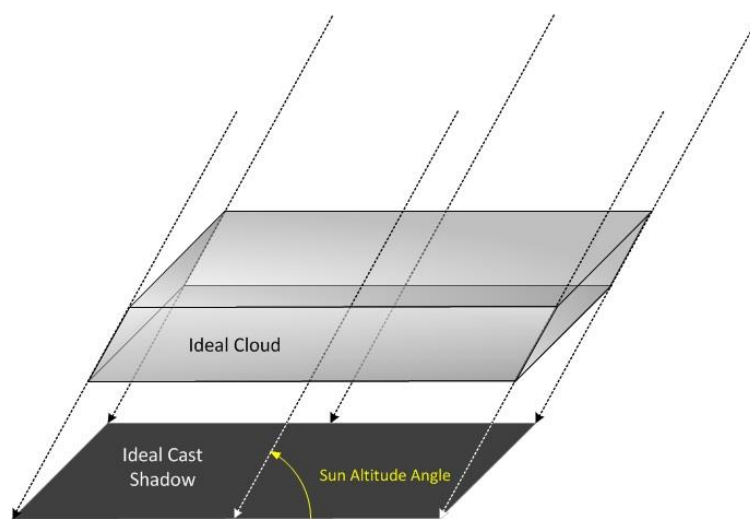
The same shading episode can be observed through the ground camera. As seen in Figure 3.61, as soon as the edge of a cloud crosses the contour of the solar disk, it induces some shade. In the example of a

single cloud moving from the bottom to the top of the image, the shade starts at the instant  $k_s$ , when the upper edge of the actual shape of the cloud gets in contact with the solar disk. The cloud has to stop covering entirely the solar disk for the shade to cease. This happens at time  $k_c$ , when the lower edge of the actual shape of the cloud crosses the solar disk and when the solar disk gets completely uncovered.

The predicted time of sun cover and uncover are exactly the same as in a sky referential. Therefore, it is always possible to convert the position of clouds from a camera sensor referential into a sky referential, by applying the transformation equations mentioned in the previous paragraphs. This method is valid for any forecast horizon. If however, the forecast is only intended to be a few steps ahead, the sun cover and uncover times may be worked out directly on the camera images, without any conversion. Indeed, a distortion of shapes and distances is still involved due to the mapping function of the lens but it can be considered as relatively negligible and accounted for by the prediction filter, especially from one step to the next.

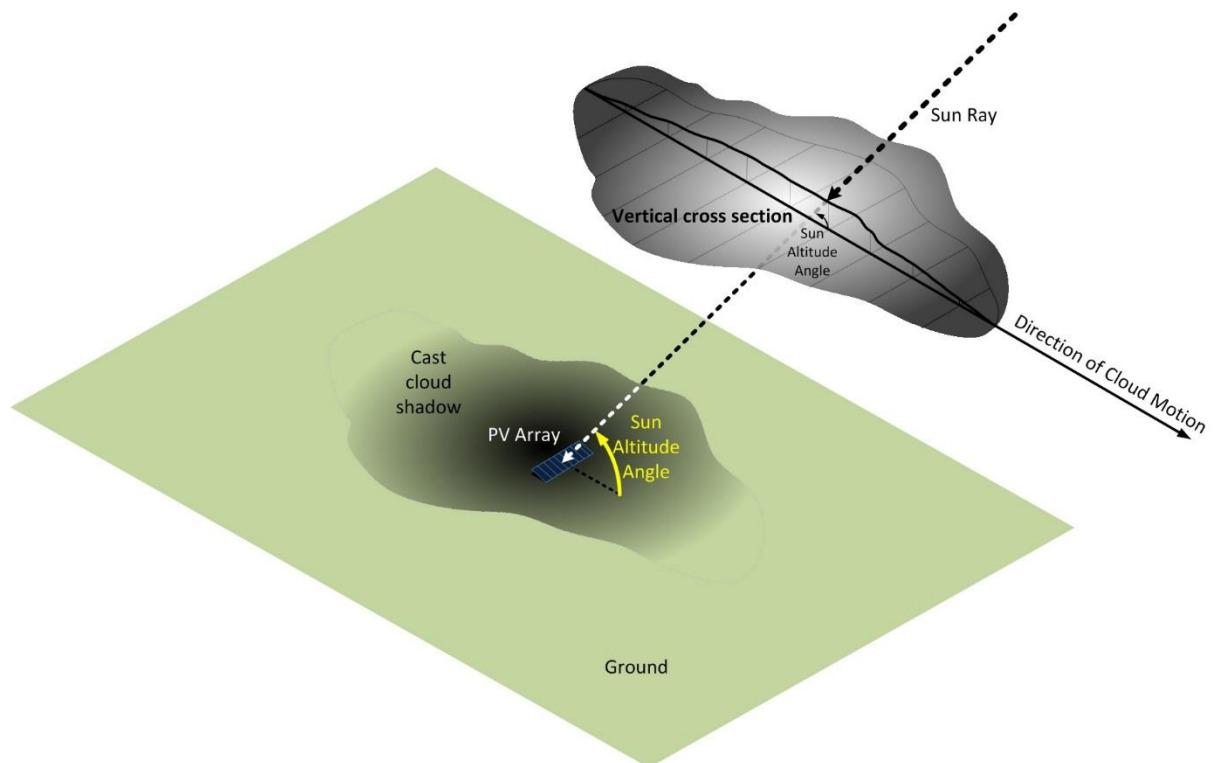
### 3.5.9 Shading intensity

It must be duly noted that the thickness of clouds of all types is not uniform. This thus results in a non-uniform shading on the ground. The intensity of the shading depends on the optical thickness of the cross section of the cloud that obstructs a given point at a given time. For clouds to be characterised with a perfectly even and consistent optical thickness, their density would need to be perfectly homogeneous and they would need to be shaped as perfect parallelepipeds with an angles that fit with the sun azimuth at all time as shown in Figure 3.62. This situation is obviously impossible since the density of clouds may vary with the altitude, since the shape of clouds changes dynamically and since the sun azimuth angle varies continuously throughout the day.



**Figure 3.62 – Model of an ideal cloud with a perfectly uniform optical thickness and sides.**

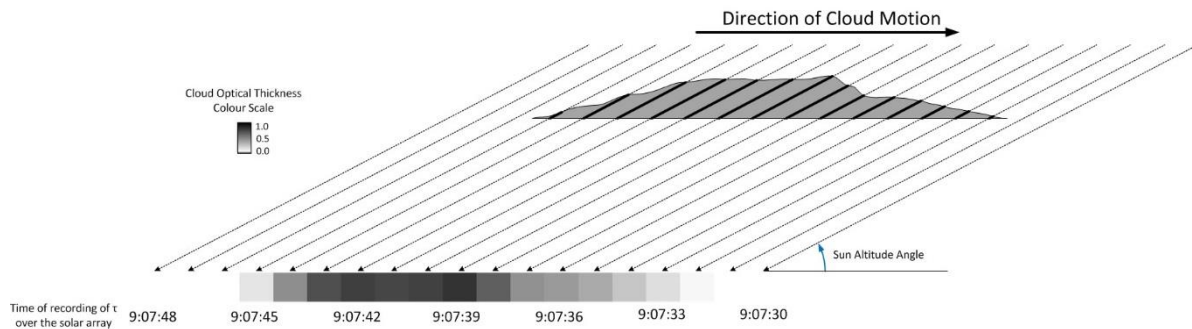
The intensity of the cast shadow received at a given point on the ground actually depends on the optical thickness of the entire path crossed by the sun ray that illuminates that given point. It is noticeable that the diffraction angle of the sun ray through the cross section of the cloud is dependent on the sun altitude angle. For the sake of simplicity, we shall assume that the sunlight passing through a cloud, in particular a low altitude cloud such as a cumulus cloud, is not bent. This means that we assume that direction of a sunray remains linear even through a cloud. As a result, the intensity of the shading that is perceived on the ground due a cloud depends on the density of the cloud, the thickness of the cloud and the sun altitude angle. For a cloud in motion, the instantaneous shading is determined by the path followed by sunlight through the section of the cloud overhead the shaded point that intersects with the plane of direction of the cloud motion. Compared with that vertical section of the cloud, the global shape of the cloud is irrelevant and does not play a role in the shading intensity of the shaded point, which happens to be a PV array in our application. That situation is illustrated in Figure 3.63, in which the vertical cross section is shown with a thick contour.



**Figure 3.63 – A model corresponding to realistic shading conditions.**

The specific example of the estimated vertical cross section of the cloud shown in Figure 3.63 has been obtained based on the calculated optical thickness of a small cumulus cloud on 19<sup>th</sup> of April 2016 passing over our monocrystalline solar array at the Green Electric Energy Park at Curtin University between 9:07:32am and 9:07:45am. The estimated sun altitude angle at each instant has also been used to reconstruct the vertical section of the cloud. Both the reconstructed shape of the cloud cross-section and the calculated optical thickness of the section of the cloud that hovered over the PV array at a

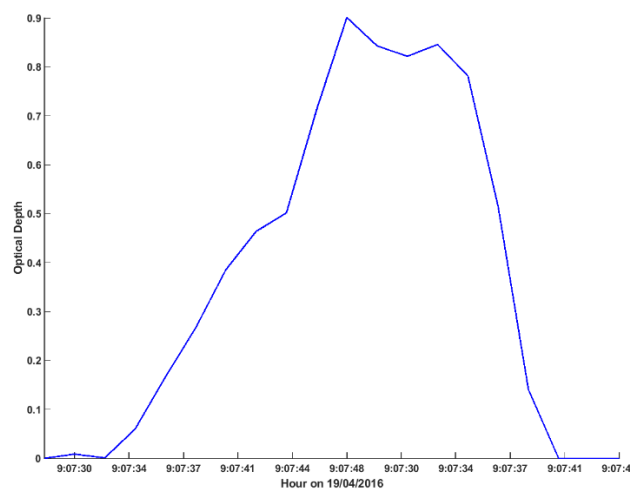
specific time are shown in Figure 3.64. The section of the cloud passing over the array and the corresponding optical thickness is shown for every second of the period.



**Figure 3.64 – Reconstructed vertical cross-section of the cloud passing over the monocrystalline PV array, the times when each section of the clouds passed over the array and the corresponding optical thicknesses.**

Since cumulus clouds grow vertically, the top of the cloud has been reconstructed from the bottom up, starting from the flat base that is a typical characteristic of cumulus clouds.

It is noticeable that the calculated optical thickness varies significantly. The smallest value corresponding to the edge of the cloud and the highest value to the centre. Thus, understandably, the shading is minimal when only the edge of the cloud passes over and the shading climaxes when the thickest centre passes over. The sun altitude angle was estimated to be  $27.7^\circ$  for that day and time. For the same vertical profile, a different sun altitude angle would result in a different length for the path followed by the sunlight through the cloud, resulting in a different optical depth for each instant, compared with the values featured in Figure 3.64. Those optical depths values are shown in Figure 3.65 with a one-second time resolution between 9:07:33am and 9:07:45am.

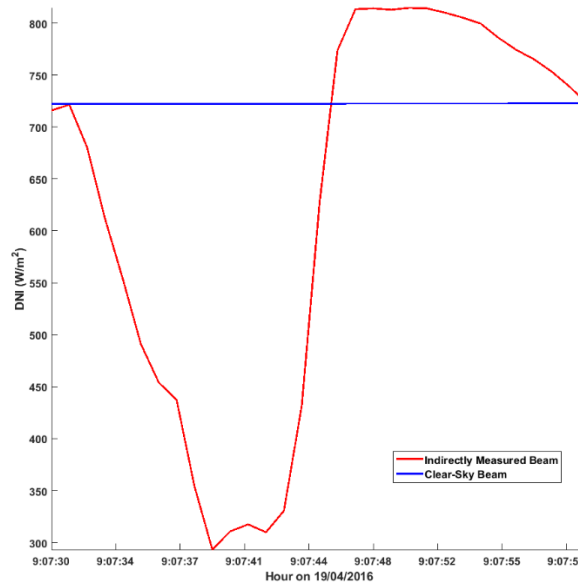


**Figure 3.65 – Optical depth of the detected small passing cloud.**



The optical thickness shown in Figure 3.65 has been calculated based on the beam intensity that was inferred from the measured current output from the monocrystalline array compared with the modelled clear-sky beam intensity, as shown in Figure 3.66. The refined TJ-based regression model has been used to estimate the clear-sky beam intensity while the optical depth  $\tau$  has been calculated by using the expression (3.122) which is derived from an expression introduced in the previous chapter.

$$\tau = \log\left(\frac{I_0}{I}\right) \quad (3.122)$$



**Figure 3.66 – Measured and simulated beam intensities.**

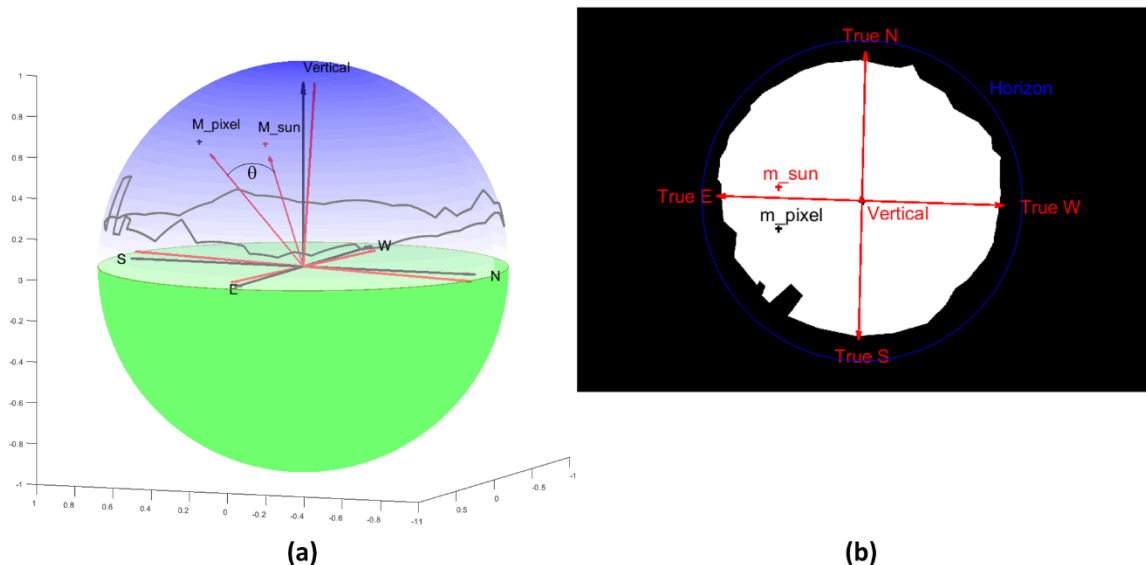
In the expression (3.122),  $I$  represents the instantaneous actual beam intensity whereas  $I_0$  is the estimated clear-sky beam intensity. However, it must be noted that the exact same value of optical depth can be found with a ratio of measured and modelled clear-sky GHIs.

It is also noticeable in the indirectly measured value of beam irradiance that not only the light intensity is dimmed when a cloud blocks the sunlight but also that the close proximity of a cloud that does not obstruct the sun or no longer obstructs the sun induces an amount of sunlight received by the PV array that exceeds the clear-sky light intensity. This can be due to the large reflection of sun light on the surface of cloud, effectively transforming the cloud in a secondary source of light. The quantity of reflected sunlight can indeed increase as the edge of a cloud moves closer to the target and it can decrease as the edge of the cloud moves away. This latter effect is visible in Figure 3.66, as the small cloud that was completely obstructing the sun clears away from 9:07:45.

### 3.5.10 Sun-Pixel Angle

It can be useful to know how to define the angle between centre of the sun and the actual 3D position of an object corresponding to a specific projection on the sensor. This angle is called sun-pixel angle. In particular, some cloud tracking methods may require to know whether a cloud is located within a close angular perimeter from the centre of the solar disk. Also, it can be sometimes necessary to carry out a separate chromatic analysis and clear sky pixel discrimination between the circumsolar region, which is prone to strong scattering and the extrasolar region. In those examples, it may be required to know the sun-pixel angle for every pixel within the sky area. When a single point  $M_{pixel}$  is considered on the camera sensor, a simple dot product suffices to calculate the sun-pixel angle, as per Equation (3.123). In this equation,  $\theta$  is the sun pixel angle between  $\vec{M}$ , the vector in the 3D world within the unit sphere centred on the position of the camera that points towards the spherical position of the centre of the sun and  $\vec{M}_{pixel}$ , the equivalent vector of  $M_{pixel}$  within same the unit sphere.

$$\cos(\theta) = \frac{\vec{M} \cdot \vec{M}_{pixel}}{\|\vec{M}\| \cdot \|\vec{M}_{pixel}\|} \quad (3.123)$$



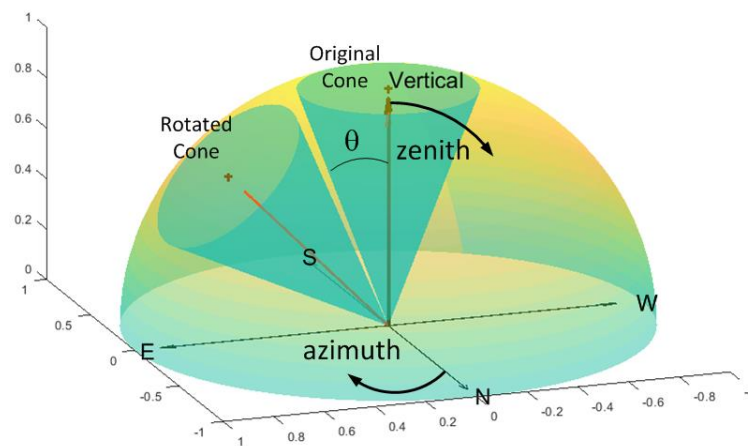
**Figure 3.67 – 3D unit sphere (a) and 2D projection (b) featuring the sun and a pixel located in a 20 degrees periphery of the sun on 25/01/2017 at 9:19:41 am.**

This is illustrated in Figure 3.67 (a) where the unit sphere is featured with the vector  $M_{sun}$  pointing at the sun on 25/01/2017 at 9:19:41 am and the vector  $M_{pixel}$  pointing at a pixel located at an angle  $\theta$  of 20 degrees from  $M_{sun}$ . In this figure, the ground is represented in green and the sky in blue. The actual cardinal directions are shown with black axes while the cardinal directions seen by the camera is shown as red frame. Also, the horizon is shown with a blue gradient [145] and the outline of the masked background is shown with a grey line The equivalent 2D projection on the camera sensor is shown

Figure 3.67 (b) where the projected position of the both the sun and the same pixel with an angle of 20 degrees are shown for the exact same moment. Also, the horizon is shown in blue and the true cardinal directions are shown in red.

However, the process of calculating the dot product can be excessively time consuming when browsing every pixel of a high-definition large sky image. For example, when processing the  $2992 \times 2000$  pixel images of our application, even after excluding masked pixels, 1,612,194 pixels remains to be processed. This means that it takes 95s to identify the pixels that are in the periphery of an angle with the sun in only one single image.

Instead, it is about 600 times faster to draw a vertical cone with the desired opening angle  $\theta$  and then rotate it by the zenith angle first and then by the azimuth angle of the centre solar disk, as shown in Figure 3.68. With this method, it only takes 0.16 s to identify the pixels that are within a given angle  $\theta$  around the sun.



**Figure 3.68 – 3D unit sphere showing the original cone and the rotated cone.**

After applying this method, the 3D position of the vector pointing at the sun on 25/01/2017 at 7:19:41 AM looks like what is shown in Figure 3.69.

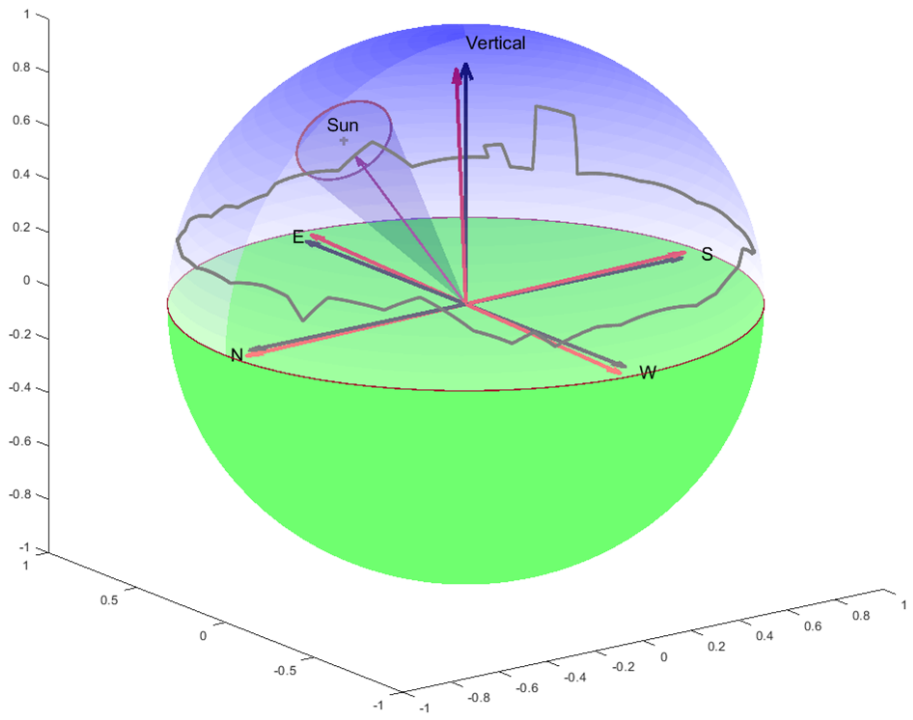


Figure 3.69 – 3D unit sphere showing the position of the sun and all of the pixels within 10 degrees around it on 25/01/2017 at 7:19:41 am.

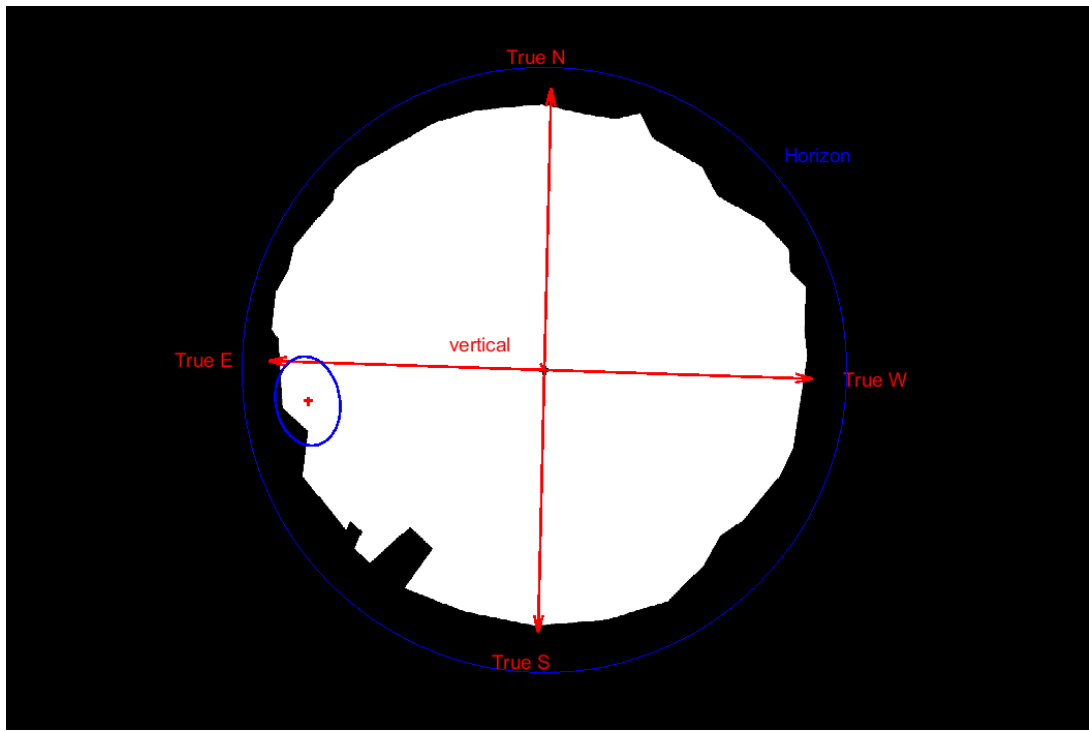


Figure 3.70 – 3D unit sphere showing the position of the sun and all of the pixels within 10 degrees around it on 25/01/2017 at 7:19:41 am.

The corresponding 2D projection on the camera sensor is shown in Figure 3.70. It can be noticed that the circle that represents the pixels within  $10^\circ$  around the centre of the solar disk is distorted, because of the projection function of the fisheye lens. To further illustrate that fact, the position of the sun every couple of hours has been plotted in Figure 3.71 throughout the day on 25/01/2017.

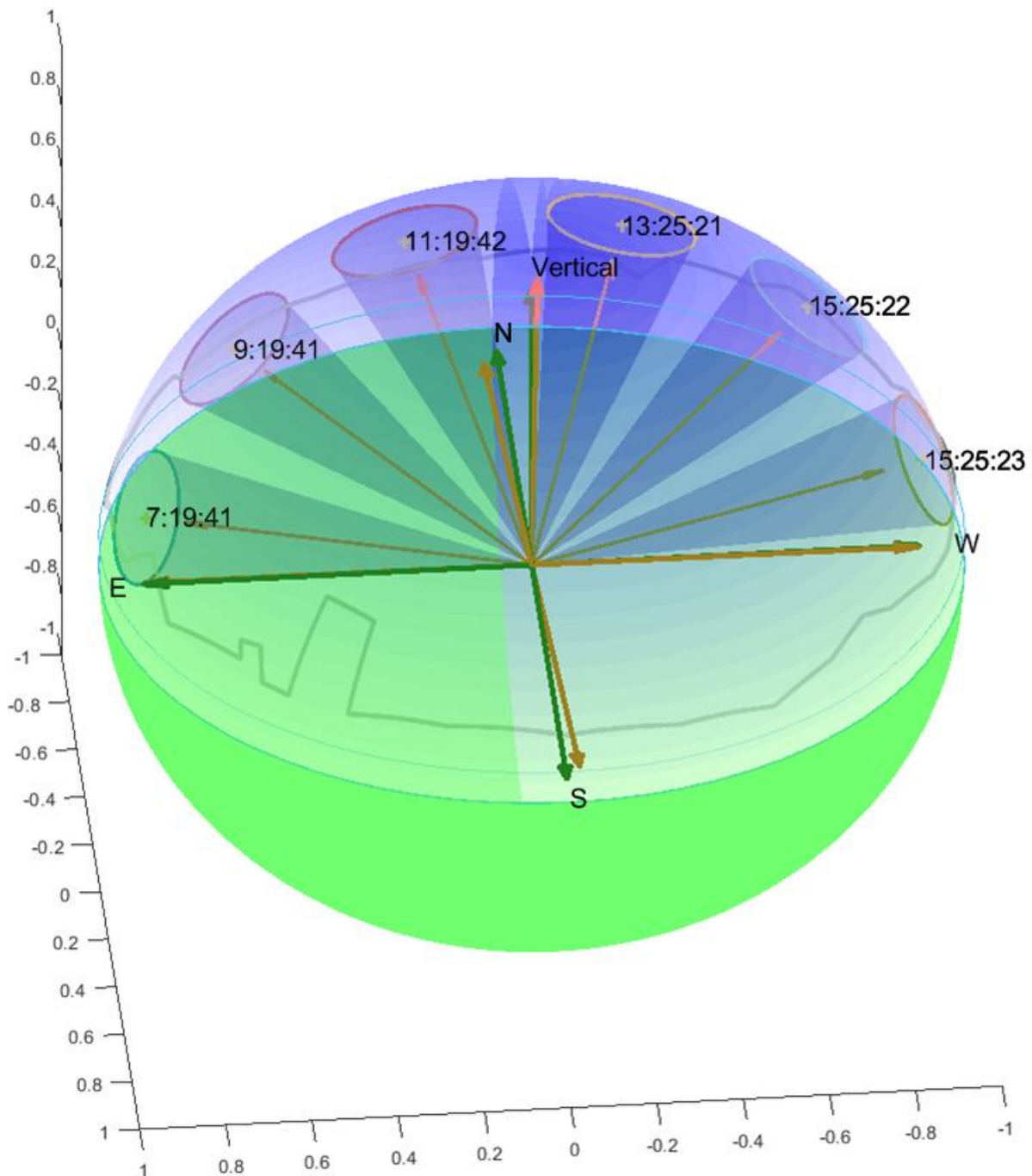


Figure 3.71 – 3D positions of the sun every two hours during daytime on 25/01/2017.

The corresponding 2D projections of the position of the sun and of the 10° perimeters around it can be seen in Figure 3.72. It can be noticed that the higher the zenith angle, the lower the sun and thus the closer it appears to the horizon and the more distorted the angular perimeter around it.

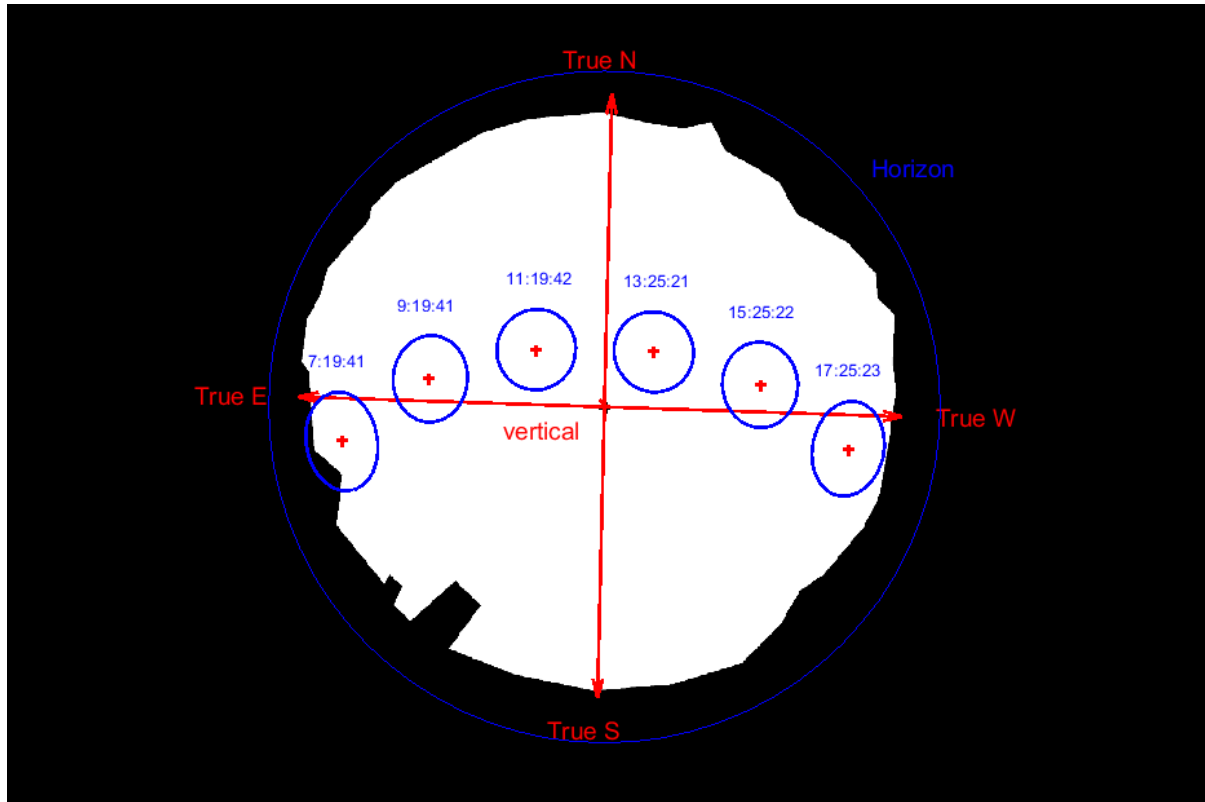


Figure 3.72 – Projected 2D positions of the sun every two hours during daytime on 25/01/2017.

### 3.6 CONCLUSIONS ON IRRADIANCE, CLOUD, SUN AND SHADOW MODELLING

Some astronomical considerations are sufficient to define the extra-terrestrial solar irradiance that would be received on the surface of the earth at a given place and time if there was no atmosphere. However, gases, aerosols and moisture in the atmosphere diffracts the extra-terrestrial beam radiance differently from one region of the world to another. This why empirical and physical irradiance models have been developed. For our application, we considered empirical models only and compared them with our own empirical measurements. It resulted from comparisons that the KC and Dazhi models were closest fits with the measurements. Based on this conclusion, we made a couple of linear regression inspired from each model and found that the lowest error was based on KC. This is thus the model we will consider using throughout this study. We also studied the optics of fisheye lenses that are used in ground sky-cameras like ours. It is important to calibrate properly a sky camera in order to translate accurately the 2D projection of the solar disk and clouds into their actual 3D positions in the sky.

## 4 CLOUD/SKY DISCRIMINATION AND CLOUD TRACKING

---

### 4.1 INTRODUCTION AND PREREQUISITES

Tracking clouds cannot be possible without two crucial preliminary steps. The first one consists in making sure of the quality of the pictures by choosing a proper sky camera and setting up the camera adequately. In addition, a regular maintenance routine is also needed to make sure that the shots can actually be processed. Once this first step is complete, the next step consists in defining a method to distinguish in every individual sky image the pixels corresponding to cloud and the pixels corresponding to clear sky.

Thus, throughout this chapter, we will exhaustively review the hardware considerations that should be taken when installing and using a sky camera. An example of application of the advisory tool for cloud/clear sky pixel discrimination criterion will be shown and used on a dataset of sky images that have been adequately prepared.

Introducing the cloud tracking part of the overall methodology, the Bayes theory will be presented and a couple of the possible practical implementations will be shown. The practical implementations will be tested on the pre-processed sky image dataset and their performance will be assessed. The implementations will be compared to each other and compared to the state-of-the-art methods in the field of cloud tracking. The results of the comparisons will then be discussed in depth.

This chapter is strongly based on an article published in the peer reviewed IEEE journal Transactions on Sustainable Energy [3].

### 4.2 HARDWARE CONSIDERATION, CAMERA SETTINGS AND MAINTENANCE

It must be stressed that it has been found that any hardware-related loss of information in the input pictures can unfortunately not be compensated by any software-based processing method at a later stage of the workflow. In other words, like many other systems, the quality of the hardware setup along with proper settings determine the quality of the input data, guaranteeing that this input data, which consists in sky images, are usable. This means that the hardware should be carefully chosen and maintained to avoid any loss of information that could compromise the entire system.

#### 4.2.1 Preparation of a Hardware Setup

In order to ensure the quality of the entire sky imaging system, prior to installing an effective sky camera, four important points should be taken into account:

- Choosing an adequate camera can be a very tough since most of cameras that are available on the market have not been designed to take direct photographs of the sun. For most cameras, direct sun exposure does not correspond to normal lighting conditions and may result in overly overexposed pictures, unfit for the purpose of cloud tracking. In particular, the very large majority of automatic cameras such as compact cameras and surveillance cameras may be designed with some fixed settings that the user cannot change. Such fixed setting may unfortunately not be incompatible with sun tracking. Moreover, such settings cannot be modified on automatic cameras. Thus, the ability to set up light exposure on the photographs must be considered first and foremost. In addition to the availability of light exposure settings, the resolution of the camera sensor is another factor that can affect both the quality of images and the period of the time lapse.
- The choice of a lens is also very important since it determines the field of view of the camera.
- The camera must be enclosed in a proper weatherproof casing to be protected from the rain, hail, snow wind, dust, UV radiation, birds and insects. That casing must be reasonably easy to reach to be wiped on a regular basis or whenever needed.
- The location of the camera is another extremely important factor. The wider the field of view of the camera, the more objects become visible on the pictures. In particular, with a fisheye lens, undesirable objects do not only constitute clutter but can also obstructs the view. Therefore, a sky camera should be ideally installed a few meters away from the solar panels for which the forecast is planned but away from light poles, fences and tall buildings. For that purpose, it is desirable to mount the sky camera on a rooftop or up a high pole.

#### **4.2.2 Recommended Choices and Recommended Settings**

We have endeavoured to cover the above mentioned problems. Any type of leisure camera (i.e. either a compact, bridge or digital single-lens reflex camera) equipped with a relatively narrow angle lens could be easily mounted on a tracking device that follows the course of the sun. However, this considerably limits the field of view of the camera and it does not enable to estimate the total cloud cover (TCC), which is a necessary parameter to consider when identifying cloud types. Thus, we chose to use a camera equipped with a fisheye lens, which has the widest available angle. Also, in order to be able to mount a fisheye lens on the camera and take photos with a reasonably high definition, change the exposure settings manually and take photos with a programmed time lapse, we chose a digital single-lens reflex (DSLR) camera. In particular, we chose a Sigma 4.5 fisheye lens and used images from two DSLR cameras that enabled exposure settings: a Canon 500D and a Nikon D7100. The Nikon D7100 was preferred in the end since it is a cheaper option and since the sensor features some very good chromatic properties, which is important for clear-sky and cloudy pixels discrimination. In addition, the D7100 features a larger range of exposure settings than its competitors of the same price range.



Empirically, the best settings for the Nikon D7100 have been defined as a centred-weighted area corresponding to a circle with a 10 mm diameter, along with a sensitivity of 100 ISO and a white balance set as direct sunlight 5200K. In terms of size of images, the size of  $2992 \times 2000$  has been determined as the best compromise between high resolution and ease to process.

However, we will mostly focus on the images we got from the Canon 500D to expose the principles of our proposed method. The images from the Nikon D7100 will be used in the next chapter to implement the method in actual online conditions and to correlate cloud tracking with PV power variations.

It must also be noted that the solar disk itself on the images constitutes the most overexposed part of the picture. For this reason and because the solar disk needs to be told apart and isolated from the rest of the picture, some systems dispose of a tracker that follows and masks the sun. We decided for our part not to mask the sun, like some other studies before, and rather used a custom software post-processing method to track the solar disk in the images. This method will be presented in the next chapter.

### **4.2.3 Notes on Maintenance Routine and Camera Limitations**

For protection purposes, we enclosed our D7100 camera under a dome, in a weather-proof plastic box, at the top of a 4 meter high pole. The pole has been altered for the camera to be lowered down for maintenance on a weekly basis.

However, using a DSLR under a dome implies that a large area becomes exposed and susceptible to get dirty quickly. It then becomes a requirement to wipe the outside of dome of on a weekly basis to remove dust, streak from the rain and bird droppings in order to keep good quality imagery. Failing to carry out that maintenance routine leads to a build-up of smudges and streaks, which are out-of-focus objects that make the images blurry and overexposed, effectively causing dramatic losses of information.

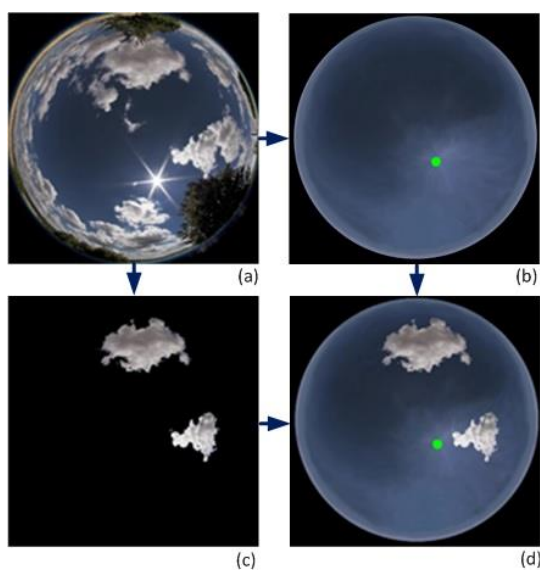
An additional drawback to using DSLR camera is the lifespan of the system. DSLR cameras require the use of a physical shutter, which can only be triggered somewhere between a few 100,000 and a few 1,000,000 times. This limitation constrains the user to restrain the time lapse to times when taking pictures is relevant, for example during day time only.

It must be said that for research purposes, despite all of the above mentioned drawbacks, DSLR cameras remain affordable and very relevant to provide very high quality images that are usable for the purpose of cloud tracking.

## 4.3 CLEAR SKY PIXELS / CLOUDY PIXELS DISCRIMINATION

### 4.3.1 Preparation of the Input Data

Once an adequate set of pictures has become available, pre-processing these images is another important and possibly difficult task. However, this step is the next essential prerequisite prior to tracking clouds. The pre-processing step starts with the discrimination between pixels corresponding to clouds and pixels corresponding to clear-sky. For that purpose, for the sake of simplicity, we have decided to illustrate the method by using pictures from the Canon 500D that we have edited with Photoshop. A video of the time lapse is visible online [146]. Some frames have been extracted from the time frame and edited in order to make a new set of images by following the process in Figure 4.1.

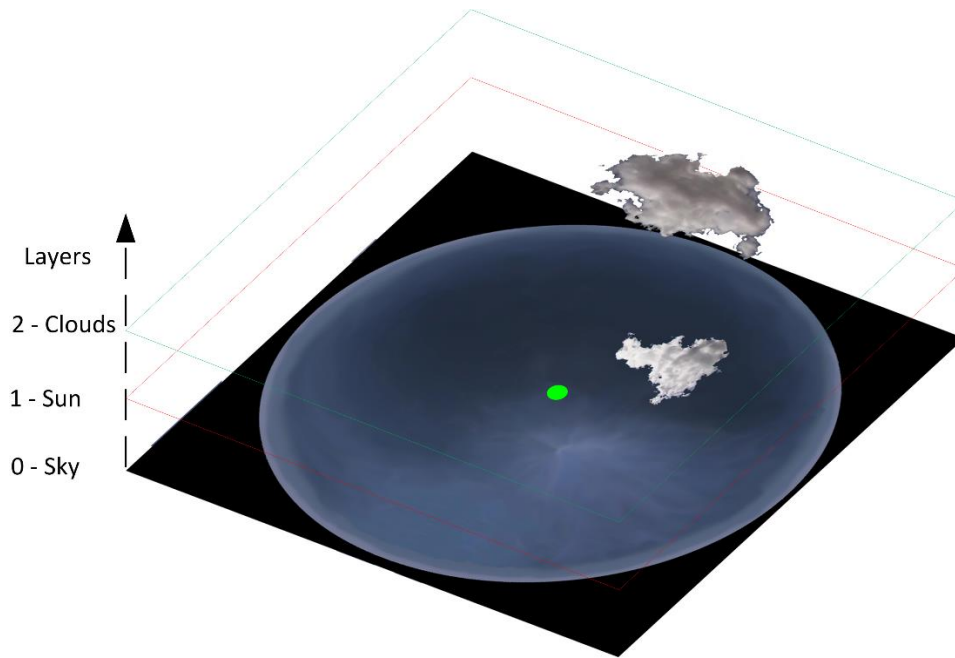


**Figure 4.1 – Edition of a sky image with Photoshop.**

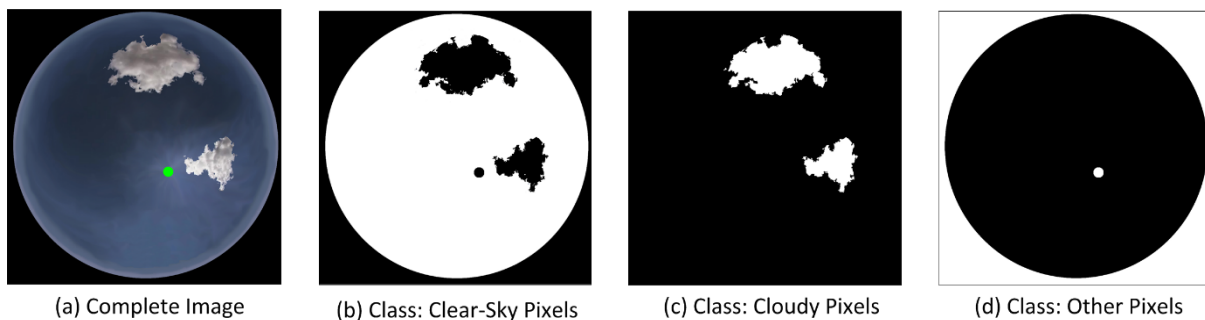
To illustrate this, from an original frame (a) as shown in the top left corner of, both background and clouds have been removed and the sun has been repositioned and coloured in green. The idea is to get a pure clear sky, as per the edited image (b) shown in the top right corner of Figure 4.1. In such an image, the only elements left from the original picture are:

- The black corners of the fisheye lens, which have been edited to be perfectly black.
- The clear sky, which has been extended to the entire panoramic field of view.
- The sun, which has been deprived from its rays and limited to a simple solar disk. Due to its extreme brightness, the solar disk naturally appears both oversized compared to its actual size and extremely overexposed. For this reason, the sun can be very easily confused with clouds. This why colouring the sun in green helps to differentiate it from both clear sky and clouds.

In the meantime, two clouds from the original frame (a) have been retained with their original shape, position and colours in order to constitute a separate layer that can be added on top of the clear sky. The cloud layer on its own is shown superposed on a black background is visible in (c) at the bottom left corner of Figure 4.1. The reconstructed image is obtained by superimposing all the layers of both (b) and (c). Finally, a complete sky image is reconstructed as per frame (d) in the bottom right corner. A reconstructed image only consists in the superposition of Photoshop layers as shown in Figure 4.2.



**Figure 4.2 – View of the layers stacked to reconstruct an image.**



**Figure 4.3 – Various pixel classes.**

To summarise, each complete reconstructed image contains pixels that can be classified into 3 classes:

- Clear-Sky Pixels
- Cloudy Pixels
- Other Pixels

A manual classification is illustrated in Figure 4.3 in the instance of a specific given sky image. The relevant pixels are shown in white.

A set of 30 sky images has been prepared according to the preparation process explained above. As it can be seen in the few samples shown in Figure 4.4 below, in the sequence of images, the top cloud is static while the cloud on the left moves from right to left. The sun is initially not covered. It only starts getting covered at the instant  $k = 17$  and remains covered from that instant. Our objective will be to automate an accurate pixel classification that could be later on implemented online.

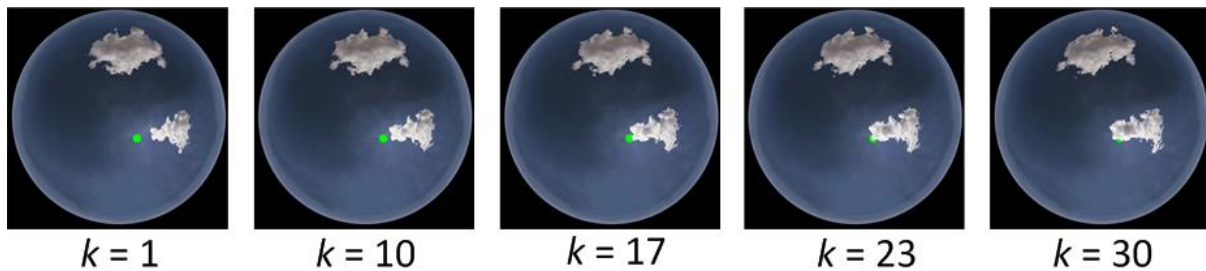
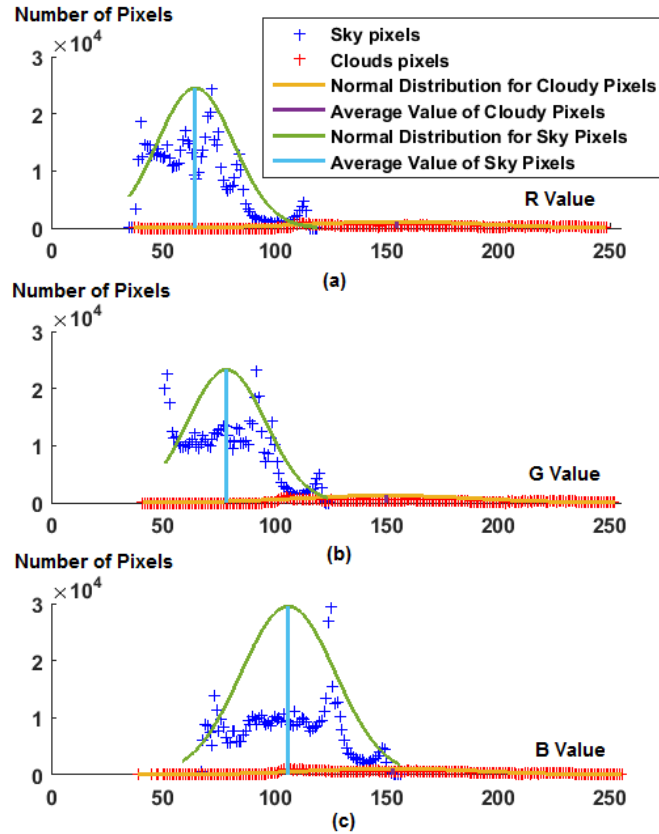


Figure 4.4 – Example of a few images from the dataset at different instants  $k$ .

### 4.3.2 Chromatic Analysis of the Input Data

Cloud discrimination methods generally use the chromatic properties of pixels based on the way that visible sunlight is scattered according to the molecules that are present in the atmosphere [59]. When the atmosphere is clear and does not contain aerosols or clouds, the sky appears blue because the molecules present in the dry air scatter more blue light than red light. In contrast, clouds look white or grey because blue and red light are scattered in similar proportion [58].

Based on their chromatic properties, pixels can be discriminated by either using hue saturation value (HSV) or hue saturation level (HSL) colour models. However, red, green, and blue (RGB) colour models are more commonly used [58, 59, 147] and have thus been chosen in this study. It must be noted that the red (R), green (G) and blue (B) features of the class ‘Clear-Sky Pixels’ and the class ‘Cloudy Pixels’ are independent from one another. It is assumed that the number of occurrences of the R, G and B values for each pixel of each class follows a normal distribution, as it often occurs in nature.



**Figure 4.5 – Number of occurrences of each red value (a), green value (b), and blue value (c) for the ‘clear-sky pixels’ class (in blue) and the ‘cloudy pixels’ class (in red) of a given sky image.**

The number of occurrences of the red (R), green (G), and blue (B) values for each pixel corresponding to the classes ‘Clear-Sky Pixels’ and ‘Cloudy Pixels’ of the photo shown in Figure 4.3 (a). Only the pixels of each of these two classes (respectively shown in Figure 4.3 (b) and (c)) that have been sorted manually.

It can be observed that there is a significant overlap between both classes, with many cloudy pixels (in red) that are closer to the mean value of the clear-sky class than they are from the mean value of their own class. This may lead to a significant confusion between both classes since a large number of cloudy pixels can be mistaken for clear-sky pixels. This is why a robust and accurate discrimination criterion must be used to classify pixels.

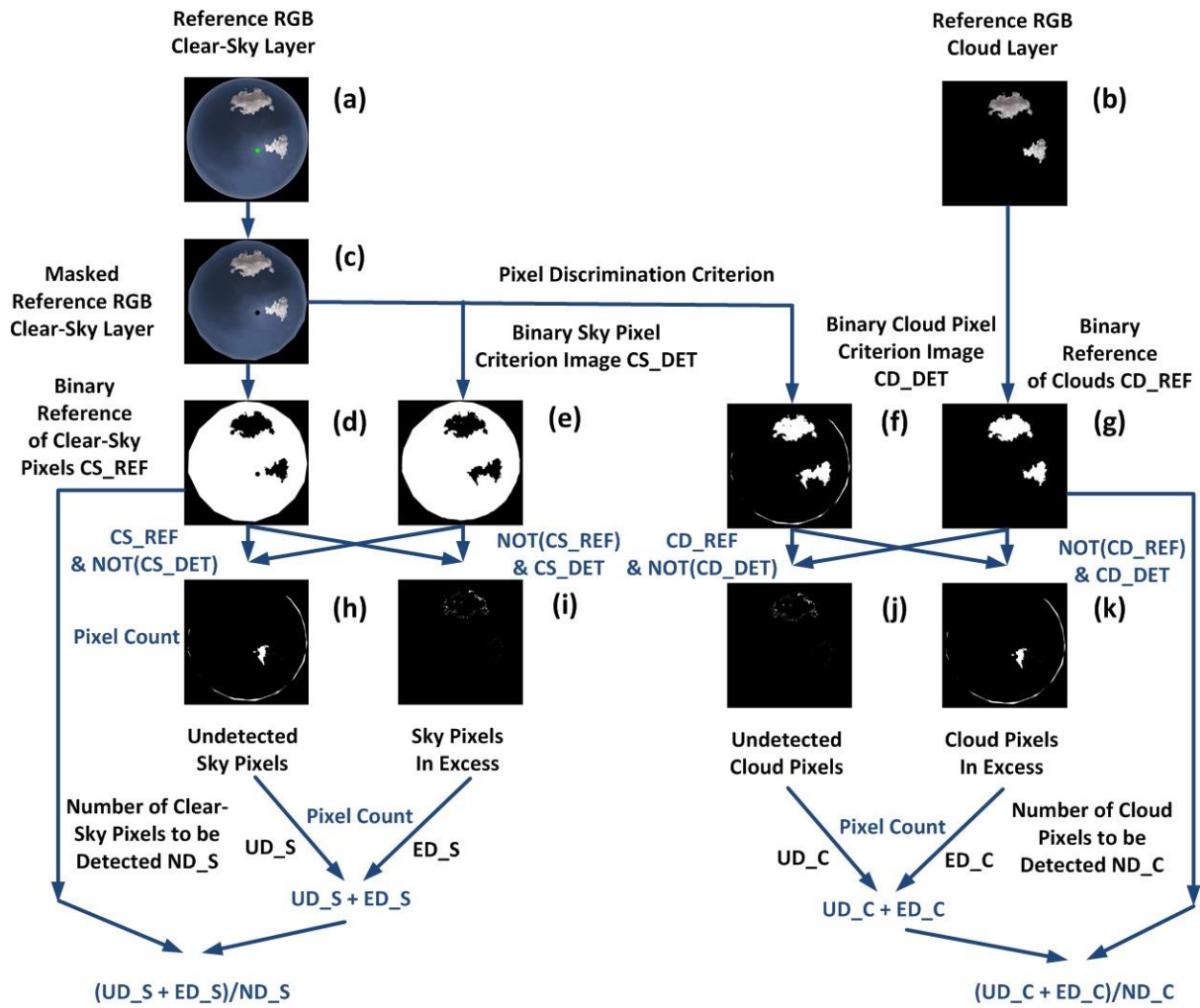


Figure 4.6 – Flow chart of the pixel discrimination assessment method.

### 4.3.3 Pixel Discrimination Criterion Assessment Method

A method has been designed to assess the accuracy of a given criterion. It consists in automatically classifying pixels in a reconstituted image like the one in Figure 4.3 (a) according to a given criterion and measure the errors in classification with respect to the count of pixels in references where all pixels belong to the same class. The complete flow to assess a pixel discrimination criterion is shown in Figure 4.6. In all of the binary images shown in Figure 4.6, white pixels correspond to the state 1 while black pixels correspond to the state 0.

The starting point is the RGB cloud layer of each image, shown in Figure 4.6 (b) and previously introduced in Figure 4.1 (c). It is used as a basis to obtain a reference binary image CD\_REF of the cloudy pixels as shown in Figure 4.6 (g). The reference number of cloud pixels counted in the image CD\_REF is noted ND\_C.

Meanwhile, the corresponding complete RGB frame (Figure 4.6 (a)) is masked (Figure 4.6 (c)) and used to obtain a reference binary image of the clear sky CS\_REF (Figure 4.6 (d)). Clear-sky pixels are

detected in the masked complete RGB frame (Figure 4.6 (c)) with the pixel discrimination criterion that is under test. The criterion is used to identify clear-sky pixels, which are shown in white in the binary sky pixel criterion image CS\_DET shown in Figure 4.6 (e). Also, the complement image is subsequently generated to identify the pixels of the original frame that correspond to clouds, as shown in the binary cloud pixel criterion image CD\_DET shown in Figure 4.6 (f).

It is possible to assess the accuracy of the discrimination criterion by comparing the binary sky pixel criterion image CS\_DET (Figure 4.6 (e)) with its reference clear-sky image CS\_REF (Figure 4.6 (d)). The reference number of clear-sky pixels that are counted in the image CS\_REF is noted ND\_S.

There are two different errors to measure between the extracted clear-sky pixel image and its reference image. One error is the number UD\_S of clear-sky pixels that have not been detected while another error is the number ED\_S of pixels that have been wrongly detected clear-sky pixels whereas there were not. The error of undetected clear-sky pixels is obtained by combining the complement of effectively detected clear-sky pixels CS\_DET with the reference CS\_REF, which can be noted  $CS\_REF \ \& \ \overline{CS\_DET}$ . The position of the undetected pixels is shown in the resulting image, which is Figure 4.6 (h). The number of undetected pixels UD\_S is the count of pixels in white in Figure 4.6 (h). On the other hand, the error of pixels erroneously detected as clear-sky results from the combination of CS\_DET with the complement of CS\_REF, which can be noted  $CS\_DET \ \& \ \overline{CS\_REF}$ . The position of the pixels in excess can be seen in Figure 4.6 (i). The number of pixels wrongly detected as clear-sky is noted ED\_S and is the number pixels in white in Figure 4.6 (i).

The total error in absolute value corresponding to the detection of clear-sky pixels is the sum of the undetected pixels and of the pixels that were erroneously detected  $UD\_S + ED\_S$ . The equivalent relative error value, expressed as a percentage is then the absolute value divided by the reference number of expected clear-sky pixels, which means  $(UD\_S + ED\_S) / ND\_S$ .

A similar process can be conducted to assess the error in the count of cloudy pixels. The number of cloudy pixels CD\_DET detected by the criterion under test can be compared with the reference number of cloudy pixels CD\_REF. On the one hand, the position of undetected cloudy pixels as seen in Figure 4.6 (j) results from the combination of the reference cloudy pixels CD\_REF and the complement of detected cloudy pixels CD\_RET, which is  $CD\_REF \ \& \ \overline{CD\_DET}$ . The number of undetected cloudy pixels is noted UD\_C and corresponds to the number of pixels in white in Figure 4.6 (j). On the other hand, the position of incorrectly detected cloudy pixels as seen in Figure 4.6 (k) results from the combination of detected cloudy pixels CD\_RET and the complement of the reference cloudy pixels

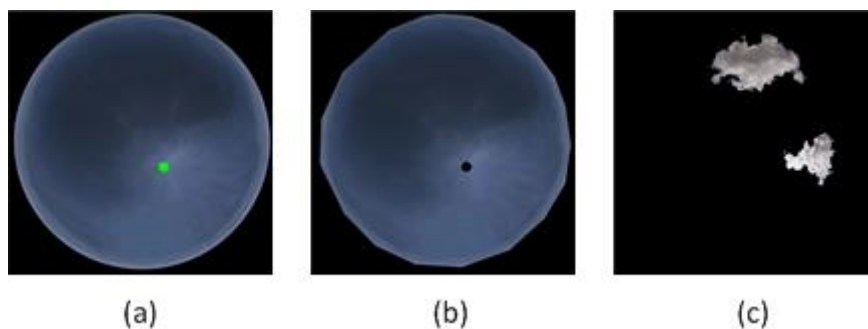
CD\_REF and, which is  $CD\_DET \& \overline{CD\_REF}$  . The number of erroneously detected cloudy pixels is noted ED\_C and corresponds to the number of pixels in white in Figure 4.6 (k).

Similarly to what was presented for the clear-sky pixels, the total error in absolute value for the detection of cloudy pixels is the sum of the undetected pixels and of the pixels that were incorrectly detected UD\_C + ED\_C. The corresponding relative error value is  $(UD\_C + ED\_C) / ND\_C$ .

#### 4.3.4 Discrimination Criteria

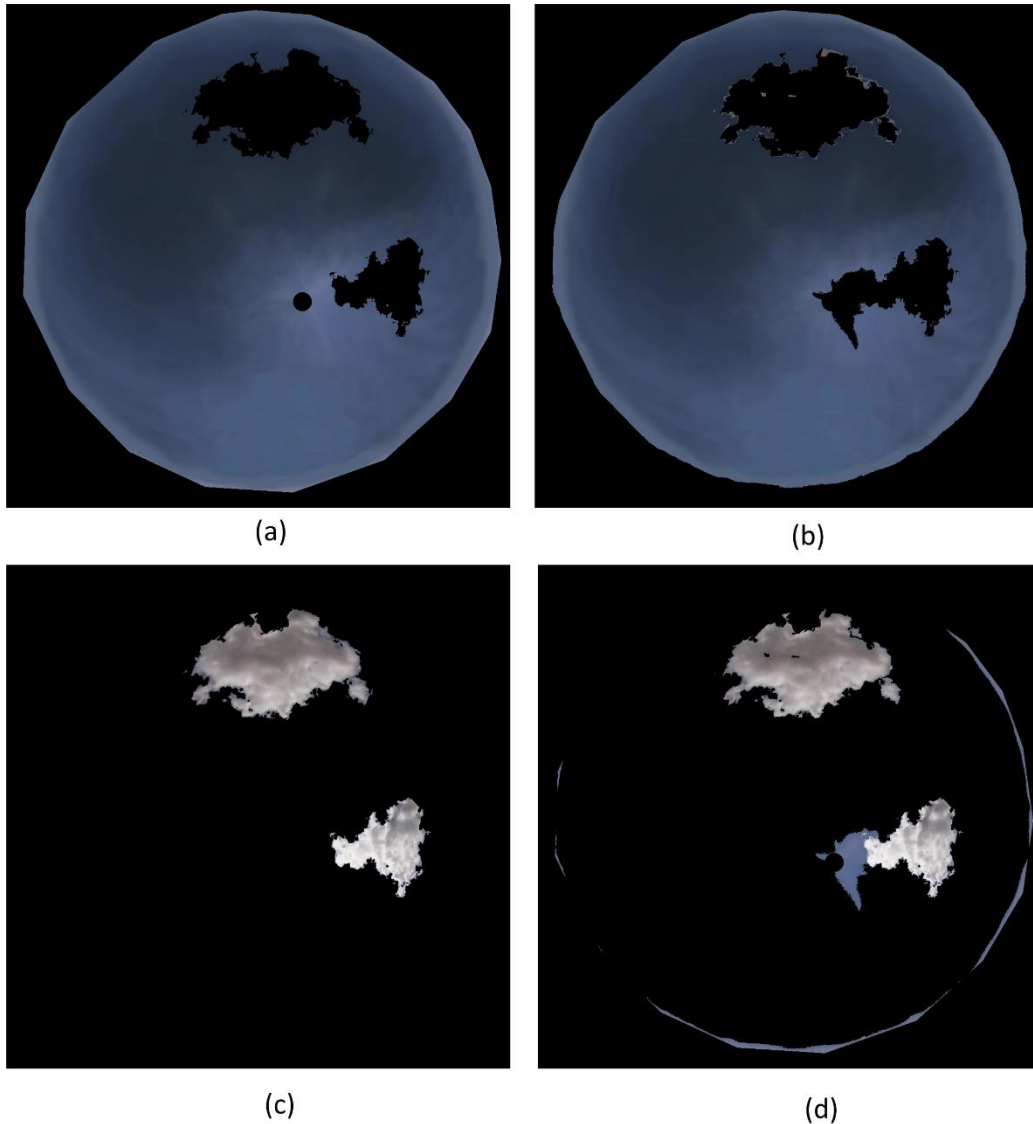
##### 4.3.4.1 Simple Bayes Filter

We made the hypothesis that a simple Bayes classifier [148] could be trained to classify pixels since the R, G and B features of the each class of pixels are independent from one another. A simple Bayes classifier is also sometimes referred to as naïve Bayes classifier and is a classifier based on Bayes' theorem. Its purpose is to classify the data in various classes. In the case of continuous data, it is assumed that the distribution of data in each class is Gaussian. For our application, we have trained a simple Bayes with the chromatic values of the single layer of clear-sky pixels so that it could recognize clear-sky pixels. We used for that purpose a reference clear-sky image in which the sun and the background have been masked as per Figure 4.7 (b). We have also trained it with the chromatic values of in the single layer of cloudy pixels as shown in Figure 4.7 (c). Overall, the training may exceed a full minute when run on MATLAB (80 seconds).



**Figure 4.7 – Reference clear-sky image (a), masked reference clear-sky image (b), and reference cloud layer (c).**





**Figure 4.8 – Comparison between the reference sky (a) and the simple Bayes pixel criterion for sky pixels (b) and between the reference cloud (c) and the simple Bayes pixel criterion for clouds (d).**

When assessed with the method explained in the previous section, a simple Bayes filter approach proves however to be relatively inaccurate and quite slow. While the average error across all 30 images happen to be about 20%, it also takes in average around 10 seconds to run the pixel discrimination with a trained simple Bayes pixel discrimination method. More detailed results can be found in Appendix A.

Basically, a confusion between both classes can be noticed around the solar disk and in the dark shaded parts of the clouds, as one can see in Figure 4.8. The reference clear-sky pixels are shown on the top left-hand side of Figure 4.8 (a) while the pixels that are identified as clear-sky by the simple Bayes filter are shown on the top right-hand side of Figure 4.8 (b). Similarly, the reference cloud layer is shown on the bottom left-hand side of Figure 4.8 (c). That reference cloud image should be compared with the pixels that have been identified as clouds shown in the bottom right-hand side of Figure 4.8

(d). A significant confusion between both classes can be noticed around the solar disk and near the horizon as well as on the edge of clouds.

Since we would like to discriminate pixels in a less than 1 second with the highest accuracy possible, alternative pixel discrimination criteria should be investigated.

#### 4.3.4.2 Threshold-based Criteria

A few reliable RGB-based criteria are readily available in literature as shown in Table 4.1, despite the fact that automatic cloud detection is a relatively new topic that is still under investigation. Generally speaking, these criteria consists in combining the values of two features (mostly R and B) with the purpose of separating the mean values of each class apart. In particular, Long et al. have used the ratio R/B as a discrimination criterion [147] while Heinle, Macke, and Srivastav have observed an improved accuracy with the difference B-R with respect to the ratio [58]. On the other hand, Kazantzidis et al., have enhanced the B-R criterion for even more accuracy [59]. Nonetheless, when we plot the average number of occurrences of each criterion across all pixels of the entire dataset, Long et al.'s R/B ratio criterion, shown in Figure 4.9 (a), spreads each class's mean values furthest apart, compared with the other difference criteria, which can be seen in Figure 4.9 (b) and Figure 4.9 (c). The blue component used in Kazantzidis et al.'s approach is shown in Figure 4.5 (c).

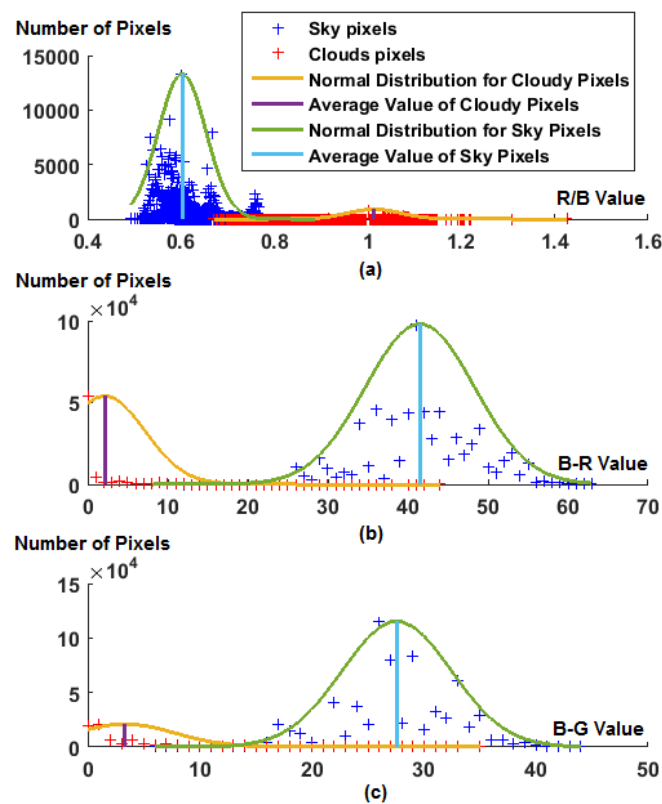


Figure 4.9 – Number of occurrences of each R/B value (a), R-B value (b), and B-G value (c).

It must be noted that the three above mentioned criteria are intended for directly detecting clear-sky pixels and not cloudy pixels. Cloudy pixels are actually inferred indirectly. The criteria are presented in with some indicative thresholds  $Th_i$ , where  $1 \leq i \leq 4$ , in Table 4.1. The threshold values suitable for times during the middle of the day are listed in Table 4.2, when applied to our dataset.

**Table 4.1 – Criteria used for pixel discrimination.**

<i>Criterion #</i>	<i>Criterion Formula</i>	<i>Average Minimal Error (%)</i>
<b>1 [147]</b>	$R / B \leq Th1$	0.68
<b>2 [58]</b>	$B - R > Th2$	2.08
<b>3 [59]</b>	$((B - R > Th2) \& (B - G > Th3) \& (B > Th4))$	2.08

**Table 4.2 – Indicative optimal threshold values for a minimal error.**

<i>Threshold</i>	<i>Th1</i>	<i>Th2</i>	<i>Th3</i>	<i>Th4</i>
<i>Optimal Value</i>	0.77	20-22	0-12	55-57

Please note that the threshold values for each criterion depend heavily on the camera sensor and may significantly vary from a camera to another.

The performance of all three criteria has been assessed according to the method explained in Section 4.3.3. The average value of the total error for all 30 images expressed as a percentage in the rightmost column of Table 4.1.

Although the multicolour criterion has been identified in literature as the most accurate, in our specific case, Long et al.'s R/B criterion proves to be the most accurate. Also, the R/B criterion demonstrates to be extremely robust over all 30 images, as the optimal threshold value for a minimal error is the same. In comparison, the two other criteria require variable optimal threshold values from one image to the next. For all these reasons, the ratio criterion has been chosen as the preferred criterion for this study.

#### **4.4 AUTOMATIC CLOUD CLASSIFICATION AND OPTICAL DEPTH**

It might be necessary to classify clouds automatically in the case of an online hardware implementation of the forecasting method presented in this thesis. Hence, a set of spectral and textural features needs to be used to identify cloud types. Such a set can be the ones used by Heinle et al. [58] on the basis of statistical metrics introduced by Haralick et al [149] computed from Grey Level Co-occurrence Matrices (GLCM), as already mentioned in Chapter 2 and detailed in Heinle et al (2010) [58] have automated the classification of clouds according to their type as a further step following the clear/cloudy classification. On the basis of their Fisher Distances, a technique used in satellite imagery (Pankiewicz, 1995), 12 features have been selected to classify clouds. Seven of these features are spectral while 4 are textural and 1 is simply the total cloud cover. Indeed, the spectral features of an image do not provide any information about the spatial distribution of clouds.

As a matter of fact, solely using colours proves to be insufficient to distinguish altocumulus from stratocumulus clouds. This is why textural features must also be taken into consideration. In addition, the cloud cover is also used as a feature to tackle the issue of spatial distribution. In particular, this helps to distinguish stratiform clouds from the rest.

The colour information of each cloud image is partitioned into a set of three monochromatic (also referred to as Grey Level) red (R), green (G) and blue (B) data sets. Each monochromatic partial image is treated separately as levels of R or G or even B. The average colour and the tonal variation of the image both provide some useful information to distinguish thick and thin clouds, or even high altitude bright clouds from the others. The colour component B proved to have the highest separation power because it is the dominant colour of the sky.

The aspect of automatic cloud classification based on their type will be further investigated in the next chapter, based on physical observations. However, an automatic cloud classification is not vital for this study since only cumulus clouds have been tracked. For that purpose, a manual selection has been made to retain cumulus clouds only in the sky images that have been processed.

#### **4.5 CLOUD TRACKING AND RANDOM FINITE SET BASED FORECASTING APPROACHES**

A few cloud tracking methods have been developed in recent years [30, 38, 125]. In particular, Jayadevan and Cronin, aided by Brooks, Rodriguez, Koch, Leuthold, and Lonij have processed the sky images from a wide-angle camera mounted on a dual-axis sun-tracker [38, 150]. They have extracted cloud motion vectors from the processed pictures on the basis of a block-motion estimation algorithm, determining the cloud velocity in the images, in pixels/second. Assuming cloud velocity to be constant, they have managed to forecast cloud position up to 10 minutes ahead. Similarly, Chow, Urquhart, Lave, Dominguez, Kleissl, Shields, and Washom have determined cloud velocity and direction up to 5 minutes ahead through a cross-correlation method applied to two consecutive images [30]. Both methods require to project images in sky coordinates to remove the distortion due to the distance of objects from the optical axis of the camera. However, these methods do not account for the cloud dynamics and may become relatively inaccurate in the case of clouds with fast changing shapes. In contrast, Wood-Bradly, Zapata, and Pye have used an implementation of the Lucas-Kanade optical flow method with OpenCV functions to track clouds for up to 40 seconds [125]. Although very accurate for very-short term forecasts, such an approach is unfortunately too computationally expensive as too many features are tracked. Thus, such a method does not permit an online implementation. Also, all of the above mentioned techniques used images from automatic cameras that saturated around the sun, preventing proper tracking of the clouds that could mask the solar disk.

In the method discussed in this thesis, tracking clouds from sky pictures can be seen as a typical multi-target image processing problem in the random finite set framework.

Firstly, the number of targets is not known in advance and needs to be assessed. During the time between two consecutive shots, some new targets may be born or some known targets may not survive. Spawning, e.g. the birth of new targets from existing ones, may also occur [151].

Secondly, the identification of individual clouds is not required. The main problem consists in forecasting when any cloud whatsoever masks the sun.

Thirdly, the data captured by sensors is never perfect and may include some clutter mingled with actual data. Clutter can be defined as some irrelevant information that does not correspond to any actual target.

Random Finite Set (RFS) modelling is the most fitted tool to deal with all of these characteristics. The Probability Hypothesis Density (PHD) filter and Cardinalised Probability Hypothesis Density (CPHD) filter are two practical implementation [152] of the RFS approach. They are actually approximations of the RFS model involving a few assumptions that will be described in the upcoming sections.

#### 4.5.1 Probability Hypothesis Filtering (PHD)

##### 4.5.1.1 Assumptions made by the PHD Filter

A PHD filter assumes that all targets evolve and generate measurements independently from one another. It also assumes that the predicted targets and clutter are Poisson RFSs. An RFS  $\psi$  is Poisson if the objects of  $\psi$  are spatially distributed according to the distribution  $S_\psi(y)$ , and in which the number  $|\psi|$  of objects in  $\psi$  is Poisson-distributed. The number of targets can be expressed as in Equation 4.1.

$$N_\psi = \int D_\psi(y) dy \quad (4.1)$$

The Poisson distribution is given by Equation 4.2.

$$P_\psi(n) = e^{-N_\psi} \cdot \frac{N_\psi^n}{n!} \quad (4.2)$$

This implies in particular that the predicted multi-target distribution  $f_{k+1|k}(X | Z(k))$  is approximated by the expression (1) where  $D_{k+1|k}(X | Z(k))$  is the probability hypothesis density a time  $k+1$  knowing measurements  $Z(k)$  at time  $k$ :

$$f_{k+1|k}(X|Z^{(k)}) \cong e^{N_{k+1|k}} \prod_{x \in X} D_{k+1|k}(X|Z^{(k)}) \quad (4.3)$$

Additionally, the birth RFSs and the surviving RFSs are assumed independent from one another. Finally, the clutter RFS is assumed independent from measurement RFSs. The PHD filter can either be implemented with a particle filter as a Sequential Monte-Carlo (SMC) filter [153, 154] or with a Kalman filter as a Gaussian-Mixture (GM) filter [151]. While a SMC-PHD has been successfully used for the purpose of cloud tracking with satellite images by Adesso et al. [124], the GM approach has been preferred in this study. Indeed, it is better-suited for linear systems and has also been identified as superior [155].

PHD filters are meant to propagate the intensity. However, in our practical cloud tracking application, the filter actually assesses and propagates the information related to the centroid of each relevant cloud. If we consider the  $i^{th}$  cloud among the total number of  $J_k$  at time  $k$ , the propagated information is in the form of the triplet  $\{w^{(i)}_k, m^{(i)}_k, P^{(i)}_k\}$ . The quantity  $w$  represents the weight of the triplet,  $m$  the mean value of the cloud's centroid coordinates and  $P$  is the covariance. In the case of a GM-PHD, the Kalman filter is characterized at any given time  $k$  by its transformation matrix  $F_k$ , the noise related to transformation matrix  $Q_k$ , the observation matrix  $H_k$  and the noise related matrix  $R_k$ . With that said, the following three conditions must be fulfilled for linear GM-PHD implementations:

Each target follows a linear Gaussian dynamic model and is detected with a sensor that has linear Gaussian measurement model. In other words, the transition function  $f_{k|k-1}(x|\zeta)$  and the likelihood function  $g_k(z|x)$  are both assumed to follow a normal distribution, as shown in the equations 4.4 and 4.5. The target  $x$  is characterized by a previous state  $\zeta$  while  $z$  is a measurement.

$$f_{k|k-1}(x|\zeta) = N(x; F_{k-1}\zeta, Q_{k-1}) \quad (4.4)$$

$$g_k(z|x) = N(z; H_k x, R_k) \quad (4.5)$$

The probabilities for targets to be detected  $p_{D,k}$  or to survive  $p_{S,k}$  at time  $k$  are state independent, which means  $p_{S,k}(x) = p_{S,k}$  and  $p_{D,k}(x) = p_{D,k}$ .

The intensities of birth  $\gamma_k$ , spawn  $\nu_\beta$ , and survival  $\nu_S$  RFSs are the following respective Gaussian mixtures:

$$\gamma_k(x) = \sum_{i=1}^{J_{\gamma,k}} w_{\gamma,k}^{(i)} N(x; m_{\gamma,k}^{(i)}, P_{\gamma,k}^{(i)}) \quad (4.6)$$

$$\nu_{\beta,k|k-1}(x) = \sum_{j=1}^{J_{k-1}} \sum_{l=1}^{J_{\beta,k}} w_{k-1}^{(j)} w_{\beta,k}^{(l)} * N(x; m_{\beta,k|k-1}^{(j,l)}, P_{\beta,k|k-1}^{(j,l)}) \quad (4.7)$$

$$v_{S,k|k-1}(x) = P_{S,k} \sum w_{k-1}^j N(x; m_{S,k|k-1}^{(j)}, P_{S,k|k-1}^{(j)}) \quad (4.8)$$

The total predicted intensity for time  $k$  is a Gaussian mixture given by Equation 4.9.

$$v_{k|k-1}(x) = v_{S,k|k-1}(x) + v_{\beta,k|k-1}(x) + \gamma_k(x) \quad (4.9)$$

The quantities  $J_{\gamma,k}$ ,  $w_{\gamma,k}^{(i)}$ ,  $m_{\gamma,k}^{(i)}$ , and  $P_{\gamma,k}^{(i)}$ , with  $i=1, \dots, J_{\gamma,k}$ , are given model parameters upon which depends the shape of the birth intensity. In a similar manner,  $J_{\beta,k}$ ,  $w_{\beta,k}^{(i)}$ ,  $F_{\beta,k-1}^{(i)}$ ,  $d_{\beta,k-1}^{(i)}$ , and  $Q_{\beta,k-1}^{(i)}$ , with  $j = 1, \dots, J_{\beta,k}$ , are responsible for the shape of a target with a previous state  $\zeta$ .

If all the conditions previously mentioned are fulfilled, the intensity at time  $k-1$  is the following Gaussian mixture, as per Equation 4.10:

$$v_{k-1}(x) = \sum_{i=1}^{J_{k-1}} w_{k-1}^i N(x; w_{k-1}^{(i)}, P_{k-1}^{(i)}) \quad (4.10)$$

Assuming all the conditions above mentioned are fulfilled, the predicted intensity for time  $k$  is in the Gaussian mixture form given by Equation 4.11.

$$v_{k|k-1}(x) = \sum_{i=1}^{J_{k|k-1}} w_{k|k-1}^{(i)} N(x; m_{k|k-1}^{(i)}, P_{k|k-1}^{(i)}) \quad (4.11)$$

Then, the posterior intensity at time  $k$  is also in Gaussian mixture form and is shown in Equation 4.12. It models the intensity of the sum of all missed terms and all possible combinations between targets and measurements.

$$v_k(x) = (1 - P_D)v_{k|k-1}(x) + \sum_{z \in Z_k} v_{D,k}(x; z) \quad (4.12)$$

In Equation 4.12, the various parameters are expressed as follows:

$$v_{D,k}(x; z) = \sum_{j=1}^{J_{k|k-1}} w_k^{(j)}(z) N(x; m_{k|k}^{(j)} \cdot P_{k|k}^{(j)}) \quad (4.13)$$

The implementation of the overall PHD filter recursion can be summarized in Figure 4.10. The implementation starts with an initialization stage, with the initial weight  $w_0$ , initial mean  $m_0$ , initial covariance  $P_0$ , and the cardinality of the filter. Also, the number of targets is initialized to 0.

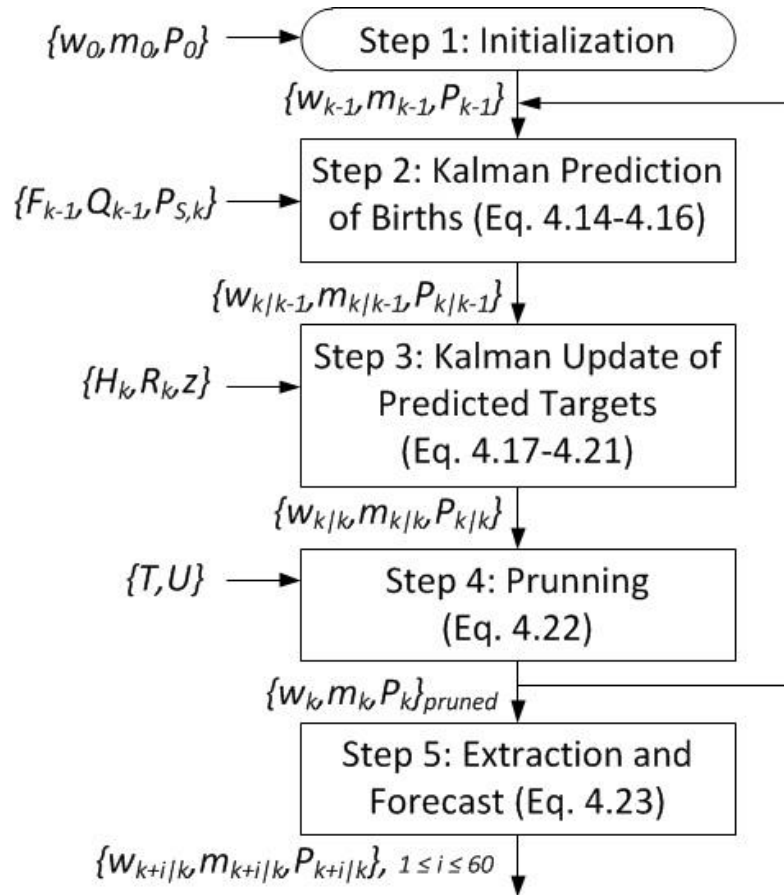


Figure 4.10 – Flowchart of the implementation of a PHD filter.

The total predicted intensity for time  $k$  is a Gaussian mixture given by Equation 4.9. If all the conditions previously mentioned are fulfilled, the intensity at time  $k-1$  is a Gaussian mixture given by Equation 4.10. Assuming all the conditions above mentioned are fulfilled, the predicted intensity for time  $k$  is in Gaussian mixture form as per Equation 4.11. Then, the posterior intensity at time  $k$  is also in Gaussian mixture form and is shown in Equation 4.12. It models the intensity of the sum of all missed terms and all possible combinations between targets and measurements.

#### 4.5.1.2 Initialization of the Filter

In each image, only the objects with an area greater than 600 square pixels are considered as targets since smaller targets cannot create any significant shade. In our application, only two clouds of significant size are present in the images. They are identified as targets and their shape are enclosed within bounding boxes. It is drastically simpler and faster to track bounding boxes instead of entire clouds. The main information is the centroid of each target. The precise shape of the cloud is also saved at every run.



#### 4.5.2 Kalman Prediction of Births

The next step of the filter is the prediction of the triplet  $\{w_{k|k-1}, m_{k|k-1}, P_{k|k-1}\}$  based on the triplet  $\{w_{k-1}, m_{k-1}, P_{k-1}\}$  from the previous run. The weight  $w_{k|k-1}$  is predicted on the basis of the probability rate  $p_{S,k}$  at time  $k$  that a target survives until the next step, as per Equation 4.14. In our implementation, the survival rate is fixed at a constant value of 99%, at any time, as it is very unlikely for a genuine target to vanish from one step to the next. On the other hand, the mean  $m_{k|k-1}$  and the covariance  $P_{k|k-1}$  are predicted through a Kalman filtering method, as per equations 4.15 and 4.16 respectively.

$$w_{k|k-1}^{(j)} = P_{S,k} * w_{k-1}^{(j)} \quad (4.14)$$

$$m_{k|k-1}^{(j)} = F_k * m_{k-1}^{(j)} \quad (4.15)$$

$$P_{k|k-1}^{(j)} = Q_k + F_k * P_{k-1}^{(j)} * F_k^T \quad (4.16)$$

##### 4.5.2.1 Kalman Update of Predicted Targets

The step that follows consists in updating the state  $\{w_{k|k}, m_{k|k}, P_{k|k}\}$  based on the measurement  $z$  made at time  $k$ . However, in the absence of measurements for this run, the updated mean value and covariance remain the predicted ones. This means  $m_{k|k} = m_{k|k-1}$  and  $P_{k|k} = P_{k|k-1}$ . Also, the updated weight becomes the predicted weight multiplied by the probability of non-detection  $p_{nD} = 2\%$ ,  $w_{k|k} = p_{nD} * w_{k|k-1}$ .

$$w_{k|k}^{(j)}(z) = \frac{P_{D,k} * w_{k|k-1}^{(j)} * q_k^{(j)}(z)}{K_k(z) + P_{D,k} \sum_{l=1}^{J_{k-1}} w_{k|k-1}^{(l)} * q_k^{(l)}(z)} \quad (4.17)$$

$$m_{k|k}^{(j)}(z) = m_{k|k-1}^{(j)} + K_k^{(j)} * (z - H_k m_{k|k-1}^{(j)}) \quad (4.18)$$

$$P_{k|k}^{(j)} = P_{k|k-1}^{(j)} * (I - K_k^{(j)} H_k) \quad (4.19)$$

The quantities  $q_k$  and  $K_k$  are defined as follows:

$$q_k^{(j)}(z) = N(z; H_k * m_{k|k-1}^{(j)}, R_k + H_k * P_{k|k-1}^{(j)} * H_k^T) \quad (4.20)$$

$$K_k^{(j)} = P_{k|k-1}^{(j)} * H_k^T * (H_k * P_{k|k-1}^{(j)} * H_k^T + R_k)^{-1} \quad (4.21)$$

##### 4.5.2.2 Pruning

The increasing number of Gaussian components over time causes a computational problem. The number of components in the posterior intensity increases without any limit. Therefore, a simple pruning method can be used to reduce the number of components that are propagated to the next step

through the recursion. This can be done by truncating components that are characterized by a weight  $w_k$  below a given threshold. In our application, this threshold is set to  $T=10^{-5}$ . Also targets verifying a Mahalanobis distance below  $U=20$  are labelled. The Mahalanobis distance is a measure of the distance between a point  $P$  and a distribution  $D$ . It is widely used in clustering. For any given observation  $x$  from a set  $X$  defined by a mean value  $m$  and a covariance matrix  $S$ , the Mahalanobis distance is defined as:

$$d_M(x) = \sqrt{(x - m) \cdot S \cdot (x - m)^T} \quad (4.22)$$

Targets are also merged if their number exceeds a maximum of  $J_{max}=100$ , which is not the case in our example where only two clouds are featured.

#### 4.5.2.3 Extraction and Forecast

To compute the forecasted position of the centre of mass of each target  $j$  at time  $k+1$ , the extracted mean value  $m_{ex,k}$  at time  $k$  is multiplied by the transformation matrix  $F_k$  as per the equation (21). In our application, the position of a target  $j$  is forecasted 60 steps ahead, based on the same fixed  $F_k$ , as per (21) where  $i$  represents the number of steps ahead. As a downside, the estimate of the number of targets tends to lose accuracy in the presence of false alarms and missed detections. As an improvement of the PHD filter, Mahler [156] has designed the CPHD as a generalization of the recursion.

$$m_{k+i|k}^{(j)} = F_k^i * m_{k|k}^{(j)} \quad i = 1, \dots, 60 \quad (4.23)$$

#### 4.5.3 Cardinalised Probability Hypothesis Filtering (CPHD)

The overall design of the CPHD filter is similar to the PHD with the exception that the CPHD assumes generalized Poisson RFSs. Another difference is the nature of the prior, predicted and clutter RFSs, which are assumed to be independent and identically distributed with a probability density  $v/N$ , where  $N = \int v(x) dx$ . Also, the Newton-Girard formula is used to compute the elementary symmetric function of order  $j$  [157] as per Equation 4.24), where conventionally,  $e_0(Z) = 1$ .

$$e_j(Z) = \sum_{S \subseteq Z, |S|=j} \left( \prod_{\zeta \in S} \zeta \right) \quad (4.24)$$

Moreover, for a posterior intensity  $v_{k-1}$  and a posterior cardinal distribution  $p_{k-1}$  given at time  $t-1$ , the intensity  $v_{k|k-1}$  and the cardinality distribution associated with the predicted multi-target state  $p_{k|k-1}$  are given by equations 4.25 and 4.26.

$$v_{k|k-1}(x) = \int P_{S,k}(\zeta) f_{k|k}(x|\zeta) v_{k-1}(\zeta) d\zeta + \gamma_k(x) \quad (4.25)$$

$$p_{k|k-1}(n) = \sum_{j=0}^n p_{\Gamma,k}(n-j) \prod_{k|k-1} [v_{k-1}, p_{k-1}](j) \quad (4.26)$$

Where:

$$\prod_{k|k-1} [v, p](j) = \sum_{l=j}^{\infty} C_j^l \frac{\langle p_{S,k}, v \rangle^j \langle 1 - p_{S,k}, v \rangle^{l-j}}{\langle 1, v \rangle^l} p(l) \quad (4.27)$$

Similarly, if the posterior intensity  $v_{k|k-1}$  and the posterior cardinal distribution  $p_{k|k-1}$  are given at time  $t$ , then the updated cardinality distribution state  $p_k$  and the updated intensity  $v_k$  are given by the equations 4.28 and 4.29.

$$p_k(n) = \frac{Y_k^0[v_{k|k-1}, Z_k](n) p_{k|k-1}(n)}{\langle Y_k^0[v_{k|k-1}, Z_k], p_{k|k-1} \rangle} \quad (4.28)$$

$$v_k(n) = \frac{\langle Y_k^1[v_{k|k-1}, Z_k], p_{k|k-1} \rangle}{\langle Y_k^0[v_{k|k-1}, Z_k], p_{k|k-1} \rangle} [1 - P_{D,k}(x)] v_{k|k-1} \\ + \prod_{z \in Z_k} \frac{\langle Y_k^1[v_{k|k-1}, Z_k \setminus \{z\}], p_{k|k-1} \rangle}{\langle Y_k^0[v_{k|k-1}, Z_k], p_{k|k-1} \rangle} \psi_{k,z}(x) v_{k|k-1}(x) \quad (4.29)$$

Where:

$$Y_k^u[v, Z](n) = \sum_{j=0}^{\min(|Z|, n)} (|Z| - j)! P_{K,k}(|Z| - j) P_j^n + u \\ * \frac{\langle 1 - P_{D,k}, v \rangle^n}{\langle 1, v \rangle^n} e_j(\Xi_k(v, Z)) \quad (4.30)$$

$$\psi_{k,z}(x) = \frac{\langle 1 - K_k \rangle}{K_k(z)} g_k(z|x) P_{D,k}(x) \quad (4.31)$$

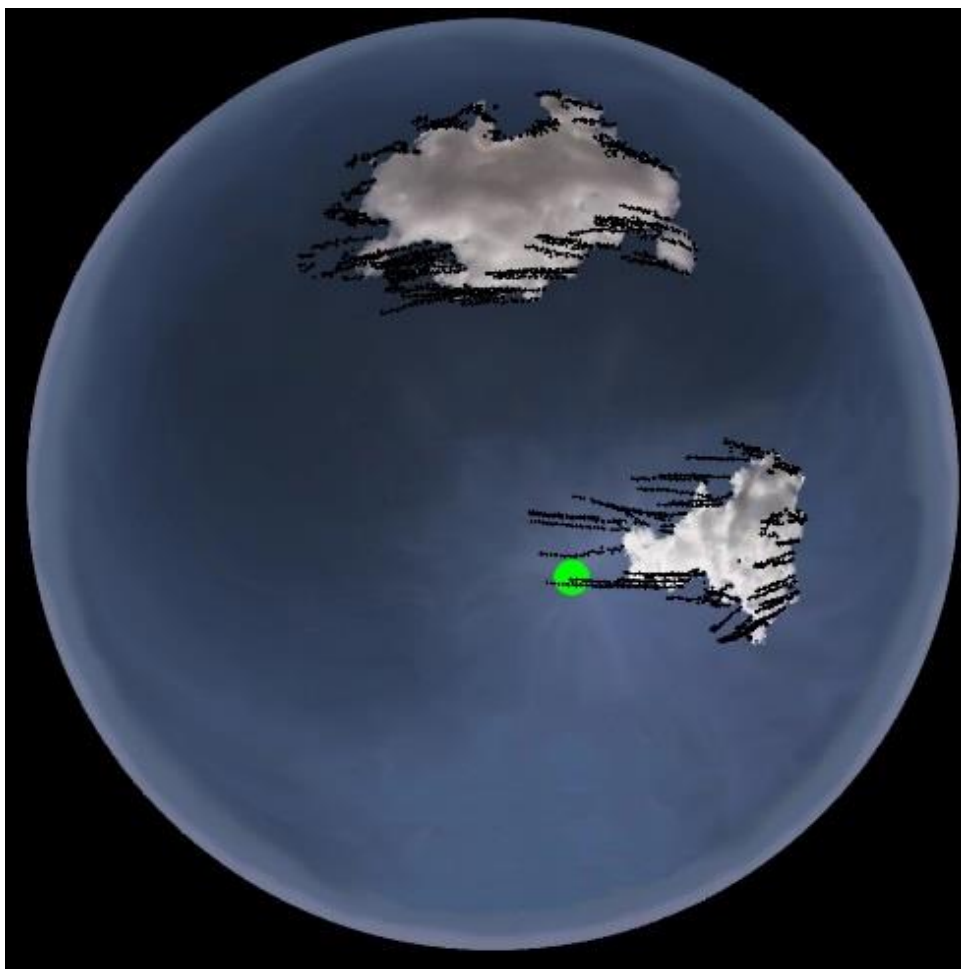
$$\Xi_k(v, Z) = \{ \langle v, \Psi_{k,z} \rangle : z \in Z \} \quad (4.32)$$

The number of targets can be estimated through a maximum a posteriori (MAP) estimator. The number of components to propagate is pruned by the minimum between the MAP estimator and the weight threshold. Further details regarding the CPHD filter and its implementation can be found in Mahler's [155] and Vo et al.'s [157] analyses.

#### 4.5.4 State of the Art Methods

In contrast with the proposed Bayes filter, a few state-of-the-art approaches have been used in literature for the purpose of cloud tracking for very-short term solar power forecasting. An extensive literature review has been conducted to depict the state-of-the-art of very short-term PV power forecasting [1]. Among recent projects, Jayadevan, Cronin, Brooks, Rodriguez, Koch, Leuthold, and Lonij have extracted cloud motion vectors (CMV) from the pictures of a wide-angle camera mounted on a dual-axis sun-tracker [38, 150]. They utilized a block-motion estimation algorithm to determine the cloud velocity in pixels/second in each image. With a persistence approach, they have managed to

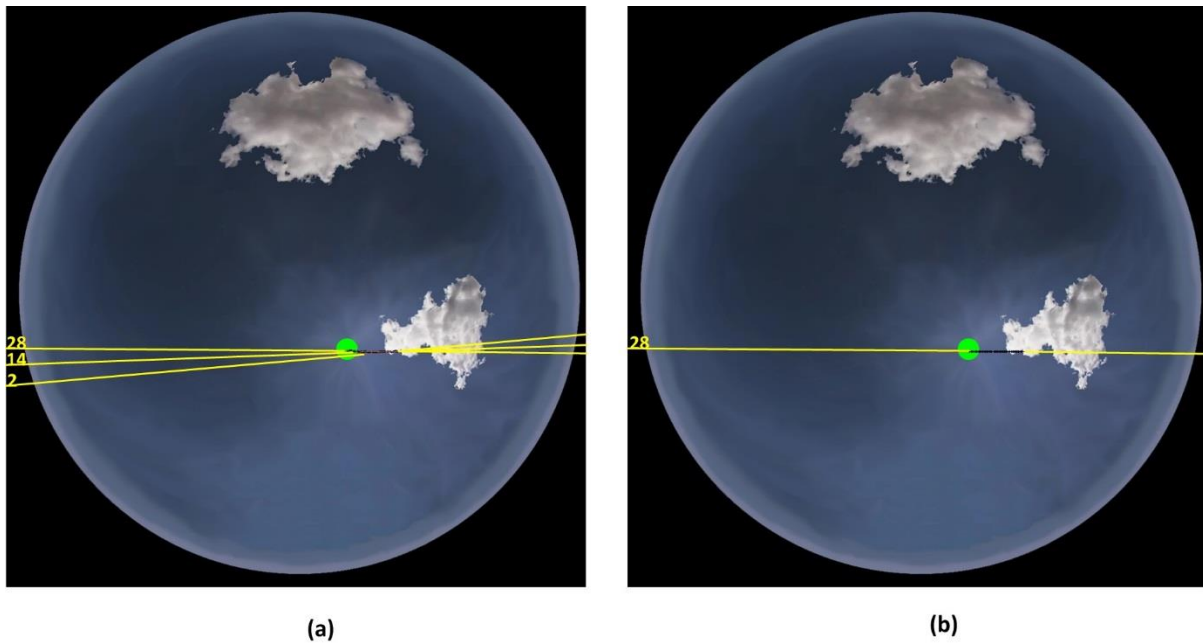
forecast cloud positions up to 10 minutes ahead. Similarly, Chow, Urquhart, Lave, Dominguez, Kleissl, Shields, and Washom have determined cloud velocity and direction up to 5 minutes ahead through a cross-correlation method (CCM) applied to two consecutive images [30]. However, these methods do not account for the cloud dynamics and may become relatively inaccurate when dealing with clouds that are characterized by fast changing shapes. Another team composed of Wood-Bradley, Zapata, and Pye have used the Lucas-Kanade optical flow (OF) method to track clouds for up to 40 seconds [125]. With a 5.3% to 21.2% error in sun-shading time (also called ‘time to collision’) estimates, it is well suited for very-short term forecasts. However, such an approach is unfortunately computationally too expensive and thus may not be suitable since it requires to extract a large number of features to track clouds. Thus, such a method may not permit an online implementation.



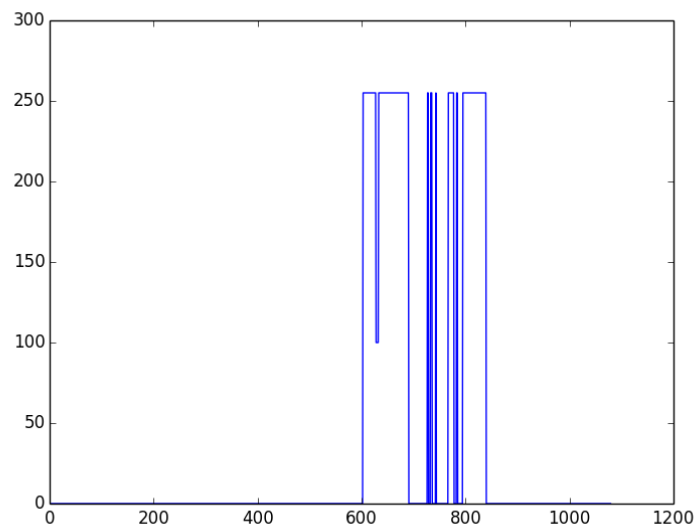
**Figure 4.11 – Movement of all extracted features over time.**

The Lucas-Kanade algorithm for Optical Flow (LKOF) has been implemented in Python with the mathematical toolbox Numpy for comparison purposes. The implementation has been based on Wood-Bradley, Zapata, and Pye’s work [125]. The method consists in discriminating the cloudy pixels from the clear-sky pixels in each RGB sky image on the basis of the R/B ratio criterion, as previously described. The resulting binary cloud image is then smoothed with a 5×5 Gaussian kernel by using an

OpenCV function. Some relevant features have been extracted where high gradients of colours are detected, which effectively, enabled to identify the edges of clouds. Across all pictures, a total of 119 features have been identified. A view of the movement of these features over time is visible in Figure 4.11. Among those features, the feature identified as #35 has been identified as very relevant since its trajectory covers the sun. The trajectory of this feature over time can be seen in Figure 4.12 (a). One can see that the direction of the lines joining two consecutive positions of the same feature may change quite a lot over time. The trajectory lines at step  $k=2$ ,  $k=14$ , and  $k=28$  are featured in Figure 4.12 (a).



**Figure 4.12 – Position of the feature 35 over time (a) and after linear regression (b).**



**Figure 4.13 – Cloud Signal along regression line for feature #35 at step 28.**

At each step, the features are tracked using a Lucas-Kanade Optical Flow function in OpenCV. The locations of the features across all images is stored in a list. However, since the locations of the features do not form a straight line, a linear regression of all the locations is undertaken to obtain the average direction of the movement of each feature. As an example, the linear regression for the feature number 35 is shown in Figure 4.12 (b). Such a regression line can be read as 1-dimensional representation of the trajectory of the feature as seen in Figure 4.13. Across each line, the pixels encountered corresponding to clear-sky pixels are given the value 0 while pixels corresponding to a cloudy pixel is given the value 255. If some pixels along the line belong to the sun, there are given the value 100 and the corresponding feature is then considered as risky in terms of sun cover. The time to sun cover is calculated by dividing the distance from the current feature position to the nearest sun pixel by the average speed of the feature along the regression line.

## **Comparison and Discussion**

### **4.5.4.1 Performance Metrics**

To evaluate the prediction performance of each filter, the input images themselves are used as a benchmark and two different metrics are assessed against that benchmark.

Firstly, we look at how accurately we can forecast in advance the time when the sun will start being covered. In other words, we assess during each step how many more steps are incorrectly forecast until the sun gets covered by any cloud. This is the Time to Sun Collision (TSCE). Although, that metric is assessed at every step, it takes 5 steps for the filter to model properly the dynamics of the system.

Secondly, for a finer assessment, from  $k = 15$  onward, when the sun starts being covered, we compare the predicted portion of the sun that gets covered with the reference ratio of the solar disk that actually gets covered as per our reference images. The error in prediction of the coverage of the solar disk for the next step is the Solar Disk Cover Ratio Error (SDCRE). It can only be assessed after the sun gets covered, which is after  $k = 15$ .

Both errors are relative and thus expressed as a percentage.

**Table 4.3 – Relative prediction errors against the benchmark.**

<i>Step</i>	<i>TSCE (%)</i>				<i>SDRCE (%)</i>	
	<i>LKOF</i>		<i>PHD</i>	<i>CPHD</i>	<i>PHD</i>	<i>CPHD</i>
	<i>Max.</i>	<i>Min.</i>				
5	45.5	45.5	100.0	100.0	N/A	N/A
6	50.0	50.0	33.3	33.3	N/A	N/A
7	288.9	44.4	125.0	125.0	N/A	N/A
8	312.5	50.0	114.3	114.3	N/A	N/A
9	314.3	14.3	-33.33	-33.3	N/A	N/A
10	350.0	0.0	0.00	0.0	N/A	N/A
11	380.0	20.0	50.0	50.0	N/A	N/A
12	600.0	0.0	66.7	66.7	N/A	N/A
13	666.7	0.0	0.0	0.0	N/A	N/A
14	7300.0	0.0	0.0	0.0	N/A	N/A
<b>15</b>	<b>1900.0</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>	<b>0.15</b>	<b>0.15</b>
16	1500.0	0.0	0.	0.0	0.70	0.70
17	1500.0	0.0	0.0	0.0	0.23	0.23
18	1300.0	0.0	0.0	0.0	0.54	0.54
19	1100.0	0.0	0.0	0.0	0.54	0.54
20	1700.0	0.0	0.0	0.0	1.16	1.16
21	800.0	0.0	0.0	0.0	3.03	3.03
22	500.0	0.0	0.0	0.0	2.80	2.80
23	400.0	0.0	0.0	0.0	2.40	2.40
24	300.0	0.0	0.0	0.0	0.31	0.31
25	0.0	0.0	0.0	0.0	3.03	3.03
26	0.0	0.0	0.0	0.0	0.23	0.23
27	0.0	0.0	0.0	0.0	1.40	1.40
28	0.0	0.0	0.0	0.0	1.86	1.86
29	0.0	0.0	0.0	0.0	0.85	0.85
30	0.0	0.0	0.0	0.0	0.15	0.15
<i>ATE</i>	<i>1030.8</i>	<i>22.4</i>	<i>45.6</i>	<i>45.6</i>	<i>N/A</i>	<i>N/A</i>
<i>ACE</i>	<i>N/A</i>	<i>N/A</i>	<i>N/A</i>	<i>N/A</i>	<i>1.21</i>	<i>1.21</i>
<b><i>Average Total Error</i></b>	<b><i>819.5</i></b>	<b><i>8.6</i></b>	<b><i>15.2</i></b>	<b><i>15.2</i></b>	<b><i>1.21</i></b>	<b><i>1.21</i></b>

#### 4.5.5 Comparison Results

Table 4.3 shows the TSCE and SDCRE. It can be observed that both filters perform with almost identical accuracies regarding both the cover time and the cover ratio. The precision of the sun-cloud collision time is absolute up to 2 steps ( $k = 13$ ) before it actually happens ( $k = 15$ ). That means that both filters can forecast shading times up to 6 seconds ahead with an absolute accuracy. Also, the precision varies but remains within an acceptable range, up to 6 steps ahead ( $k = 9$ ).

This means the RFS filters can be safely used to forecast 18 seconds ahead. The forecast 9 steps ahead (at  $k = 6$ ) even suggest that a forecast 27 seconds is achievable with some improvement. When there are some errors of sun-cloud collision times, like at  $k = 5-9$  and  $k = 11-12$ , these errors can be large but also quite inconsistent. For example, the collision time is largely underestimated at time  $k = 8$  and then overestimated at the next time  $k = 9$ . The problem may come from an imperfect time lapse or the fact that the filter was given a too short time to adapt before a sun collision happens. Thus, more datasets would need to be tested in the future.

The three bottom rows of the table correspond to average errors. For each method, the relative Average Time Error (ATE) is the mean error across all Time to Sun Collisions prior to  $k = 15$ . The Average Cover Error (ACE), presented in second last position, is the mean error across all Solar Disk Cover Ratios, which means, after  $k = 15$ . The last row is the average total error, which is the global mean value of all the errors for every  $k$  value, before and after the sun gets covered

The Lucas-Kanade Optical Flow (LKOF) prediction results can be seen in the left most column of Table 4.3. The second leftmost column, right to the step numbers is the maximum TSCE while the next column on the right is the minimum TSCE. As one can see, the minimum TSCE of the LKOF is only half of that of the PHD/CHPD and hence better. However, the maximum TSCE is considerably much higher. The reason is that only one single feature has been identified as relevant and tracked in the case of the PHD/CPHD. This feature is the centroid of the cloud that eventually covered the sun. That means that for one object, only one feature was tracked, leading to minimal calculations and to one single forecast result. In contrast, the LKOF identified and tracked all along 115 features among which only between 1 and 8 features were actually relevant. That means that the LKOF resulted in a number of forecasts between 1 and 8 at each step, after some extra computation for the 114 to 107 irrelevant features. As previously mentioned, one single feature, tracked as feature #35 has been identified as very accurate and led to the minimal TSCE. However, the features #27 and #39 led the maximum TSCE 19 times over 25. To summarize, the LKOF can be very accurate as well as it can be very inaccurate, depending on the proper selection of features. The PHD and CPHD are more consistent in that regard.



The error in terms of SDRCE for the PHD and CPHD fluctuates between 0.15% and 3%, which is very low. For reference, with cross-correlation method (CCM), Chi et al. managed to forecast collisions at 6% for 30 seconds [30].

Also, it has been successfully tested that all types of false alarms could be effectively filtered out by both PHD and CPHD.

Indicatively, the runtime for 60 forecasts ahead with the PHD is 0.0021 s whereas it may take 0.0157 s, i.e. 7.5 times longer, for the CPHD to run with MATLAB® on the same laptop computer equipped with an Intel 7-4600U dual core 2.1 GHz / 2.7 GHz. However, both filters are definitely fast enough for online applications. To iterate over 29 images, the PHD and CPHD may only require respectively 0.06111 s and 0.45 s, leaving the PHD with a very comfortable margin for heavy image processing and intense calculations. In comparison, the optical-flow based method LKOF requires 0.6724 s per image on the same machine. This corresponds to a total execution time of 19.50 s over 29 images. This number may even further increase in the case of real image with a large number of small clouds and hence a large number of features. This would then make the LKOF a less suitable for real-time forecasts in the second-range [125]. Unlike optical flow-based approaches, the number of features tracked by the RFS-based is very limited and the number of targets could easily be increased without affecting too much the computation time.

Pre-processing is also simpler and faster with the RFS approach. OF and CCM-based methods may require a sky coordinate transformation of the images since the fisheye distortion may affect their ability to recognize a given individual cloud. These methods may also require to compute the angle of each sky pixel with respect to the centre of the solar disk, which might be challenging for a real-time application with high-resolution images.

#### **4.5.6 Conclusion of the comparison**

In this simple application, both the PHD and CPHD filters prove to be adequate tools for cloud cover motion forecasting using a ground camera. As expected in such a simplified case, no significant difference in the accuracy of the estimate is discernible between both filters.

## 5 DATA COLLECTION AND ANALYSIS

---

### 5.1 INTRODUCTION

The effectiveness of the RFS based approach that is used in this study has been demonstrated in the previous chapter on a simplified dataset. Although the dataset was representative of the sunlight dimming conditions that may cause dramatic photovoltaic power disruption, it did not necessarily include all of the clutter that an online hardware implementation would have to deal with.

This chapter presents the experimental setup that has been designed, installed and used for shade forecasting. The setup consists of some hardware equipment, more precisely a camera body and a lens enclosed under a sealed dome that is installed at a certain height as well as a piece of software to control the camera.

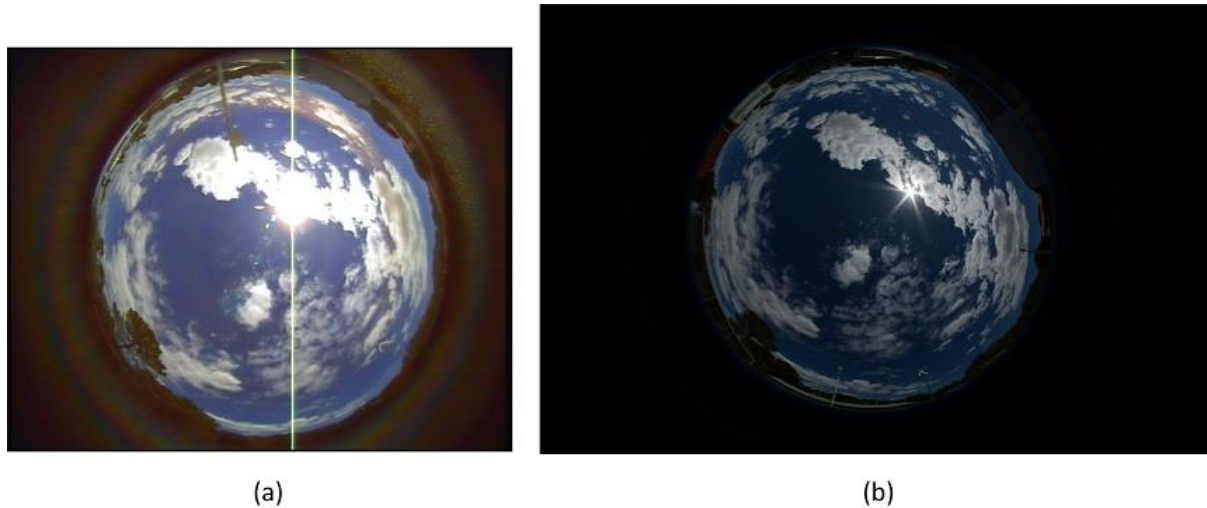
A chromatic analysis of the sky images is performed in order to determine the optimal criterion and its optimal thresholds. This step is necessary since the camera in use in this chapter is different from the one used in the previous chapter. As part of this analysis, a clear sky library has been constituted and the chromatic properties of each cloud class has been analysed for the purpose of automating cloud classification.

Finally, clouds have been tracked on several datasets that are representative of the most extreme corner cases the system can encounter in summer and in winter.

### 5.2 EXPERIMENTAL SETUP

From its installation on the 26<sup>th</sup> of February 2015 onwards, a Moonlight AllSkyCam video camera has been used to get a first set of images. Although the pictures obtained when the sun is completely hidden by a cloud are somewhat passable, the pictures shot when the sun was visible were excessively overexposed, due to some fixed internal exposure settings. This leads to an irremediable loss of important information for the purpose of sun cover time forecasting. Unfortunately it is not possible to amend those settings. Therefore, the images from a Canon 500D have been used first to showcase the forecasting method, providing a single dataset. However, in order to obtain more datasets that could be properly used to experiment with the method, a Nikon D7100 DSLR camera has been installed on the site of Green Electric Energy Park (GEEP) within the Bentley campus of Curtin University (32.01 S, 115.89 E, 18 m ASL). To compare the performance of AllSkyCam with that of Nikon D7100, a photo of the sky has been taken just a few metres apart at the exact same moment on the 24<sup>th</sup> of February 2016 at 1:22:33 pm with the already installed AllSkyCam Figure 5.1 (a) and with the newly installed Nikon

7100 Figure 5.1 (b). The picture from the Nikon D7100, when optimal settings have been applied manually, does not only feature a different frame format and a higher resolution, but, first and foremost, also enables to distinctively tell the solar disk apart from the nearby clouds, which is not the case on the picture obtained with the AllSkyCam.



**Figure 5.1 – Comparison between pictures of the sky obtained with the AllSkyCam (a) and with the Nikon D7100 (b) at the exact same instant (24/02/2016, 1:22:33 pm) a few metres apart.**

The sky images collected from the camera have been cross-correlated with the global horizontal irradiance, the ambient temperature and the atmospheric pressure measured by a Davis Vantage Pro 2 weather station [158] located only about 10 metres away from the camera. The images from the camera have also been compared with the current, voltage and power output of a 1.5kW photovoltaic array composed of 8 monocrystalline SUNTECH STP190S – 24/Ad+ modules [159] connected in series, mounted on an East-West single axis tracker and located about 6 metres away from the camera. The array has been connected to the grid through an SMA SunnyBoy 1700 rated for 1.7kW [160], the DC link voltage of the PV system being controlled via an MPPT. In addition to this, the direct normal irradiance received by the PV array and the measured temperature of its modules have also been collected to be compared with the images.

The above mentioned weather station, photovoltaic array and sky camera are featured in Figure 5.2 among some other pieces of equipment at the GEEP lab. A close-up view of the camera setup mounted on its bracket can be seen details in Figure 5.3. It can be noted that a computer and a portable hard-drive have been fitted in a weather proof box at the bottom of the pole. More details about the installation of this camera can also be found in the Appendices B and C.



**Figure 5.2 – View of the camera setup among other installations at the GEEP lab.**



**Figure 5.3 – View of the camera setup mounted on its pole.**

The data coming from the weather station and from the solar array have been sampled every second and saved on the GEEP lab local server. Whenever required, the data has been queried via the LabView custom GEEP client software shown in Figure 5.4. It must be noted that the GEEP lab and GEEP software have been managed since 2012 by Assoc. Prof. Sumedha Rajakaruna.

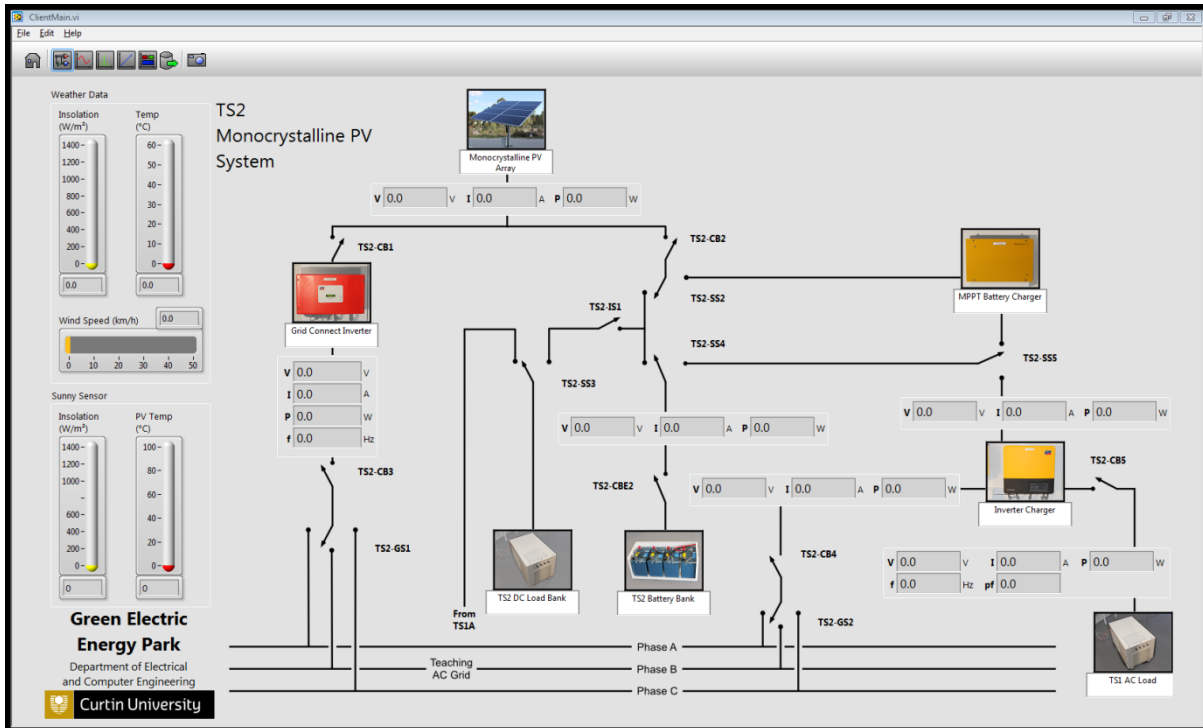


Figure 5.4 – Labview-based GEEP client software.

The capture and storage of the sky images has been managed with the free open-source software DigiCamControl, which has been originally developed in C# in a .NET framework by Duka Istvan [161][162]. A screenshot of the main window of the program is visible in Figure 5.5. The version 2.0.21.0 of the software has been further developed and customised to match up with our needs [163]. In particular, a 3-second time lapse feature has been developed with the ability to save every picture directly onto an external hard-drive with a name corresponding to the exact date and time of the instant that photo was shot. The time lapse has been programmed to run every day during sunlight hours only. The start and stop times were set manually every month to match with the length of the day.

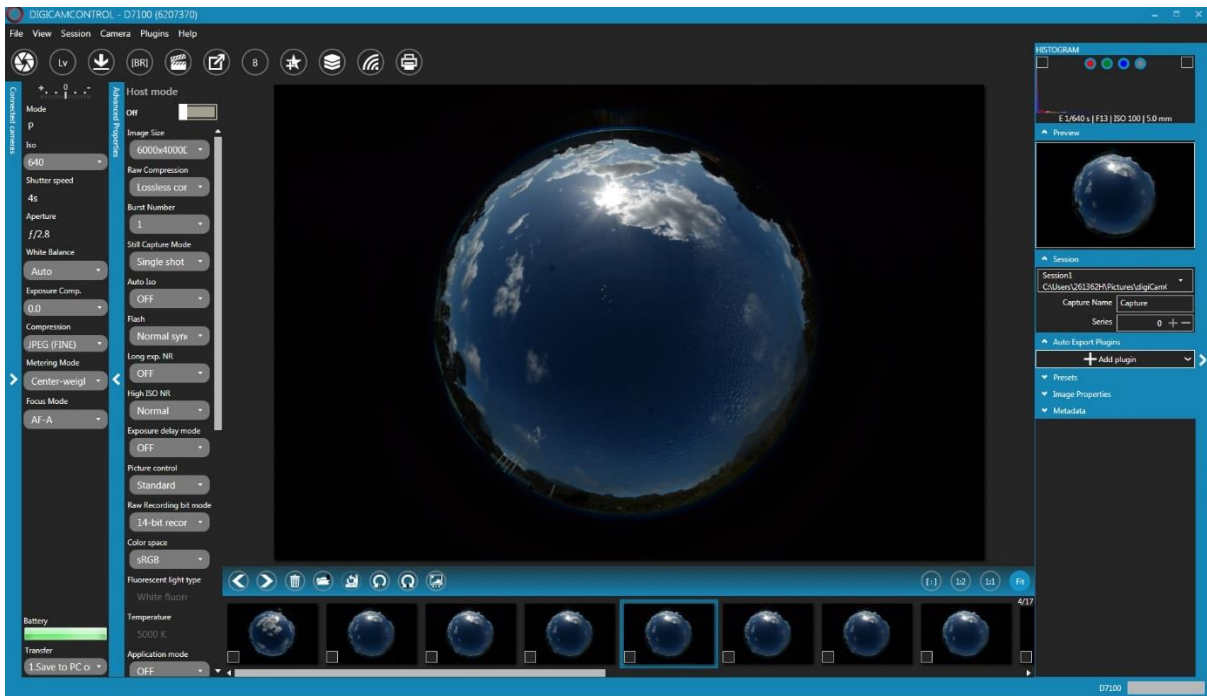


Figure 5.5 – Overview of the main window of the enhanced DigiCamControl software version 2.0.21.0.

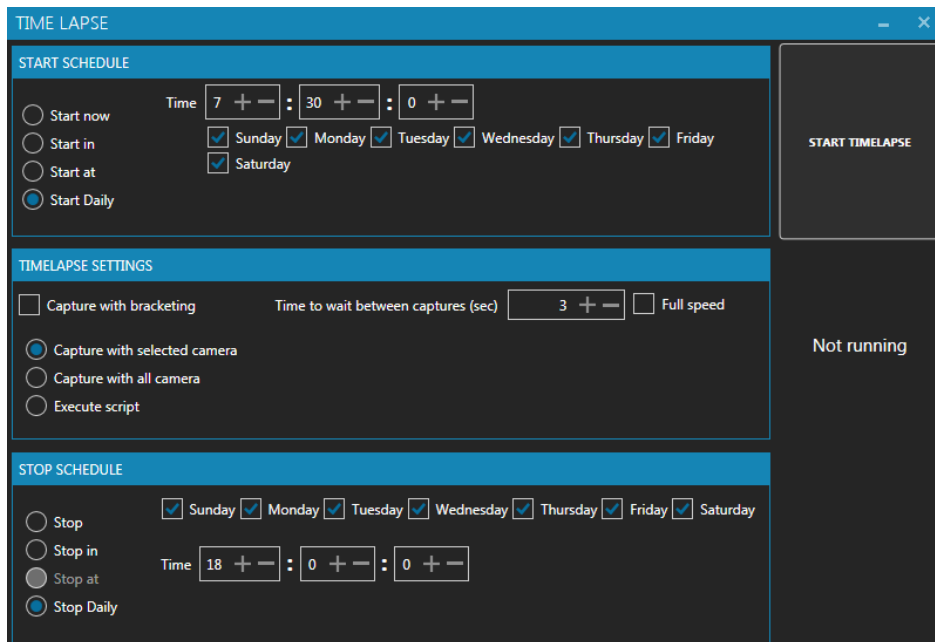


Figure 5.6 – Time lapse window of the DigiCamControl software.

The time lapse is controlled from the window that is visible in Figure 5.6. The following pseudo-code illustrates the strategy used in the file *CameraControl\ViewModel\TimelapseViewModel.cs* to achieve a tested average absolute error in the shot time of 26 ms over a sample of 1000 shots [164]. A timer has been used for optimal results. A comparison of consistency in shooting times has led to the recommendation to use a timer and an event handler rather than comparing times to trigger a shot.

**Given that** (the START TIME LAPSE button has been pressed) **&** (a Time Lapse has been scheduled) **then:**

**step 1.** (initialisation)

*TotalNumberOfShots=0*

Get *TimeBetweenShots* as defined by the user in seconds in the graphical interface

Create a local timer *Timer* with a period *per = TimeBetweenShots\*1000*

Start *Timer*

On each incrementation of *Timer*, run the event handler

**step 2.** (checking whether the time lapse should start)

**If** (no Time Lapse is already running) **then:**

**If** (a Time Lapse is scheduled on a week day like today) **&** (the scheduled start time is passed) **then:**

**step 3.** (starting a new Time Lapse)

*StartTime = Now*

**If** (*Timer* is active) **then**

*TimeOfLastShot = Now*

*TotalNumberOfShots++*

**step 4.** (shooting a picture)

Execute the adequate command to shoot a picture

**step 5.** (checking if the timer needs to be stopped)

**If** (the Time Lapse is to be stopped on a week day like today) **&** (the scheduled stop time is passed) **then:**

*ShouldStop = true*

**If** (*Timer* is active) **&** (*ShouldStop*) **&** (*Timer* has been scheduled to start daily) **then:**

*Timer* is set to inactive

**Step 6.** (stopping the local timer and re-initialisation of the status flags)

Reset the flags to indicate that no Time Lapse is running anymore

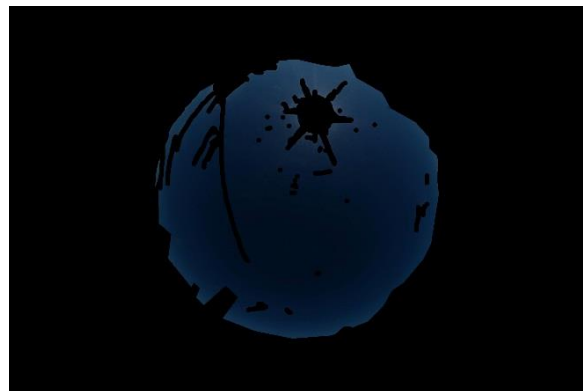
Stop *Timer* and dispose of *Timer*

### 5.3 QUICK CHROMATIC ANALYSIS

An initial step towards tracking clouds in actual sky images is to discriminate clear-sky pixels on one side and cloudy pixels on the other side. For that purpose, a couple of images have been selected. One image consists in picture of the clear sky taken on 21<sup>st</sup> of April 2016 at 12:48:47 PM and the other one is a picture of cumulus clouds. The colour of the pictures have not been touched. However, the clear sky picture has been masked so that only clear-sky pixels remained. Similarly, a selection of cloudy pixels has been made in the image featuring clouds so that only cloudy pixels remained. An effort has also been made to make sure that the contour of clouds was preserved as much as possible. This is important since the correct detection of complex contours can be a discriminating factor. The process is shown throughout the Figures 5.7 – 5.11.



**Figure 5.7 – Original clear-sky image.**



**Figure 5.8 – Fully masked clear-sky image.**





**Figure 5.9 – Original image containing cumulus clouds.**



**Figure 5.10 – Masked cumulus cloud image.**



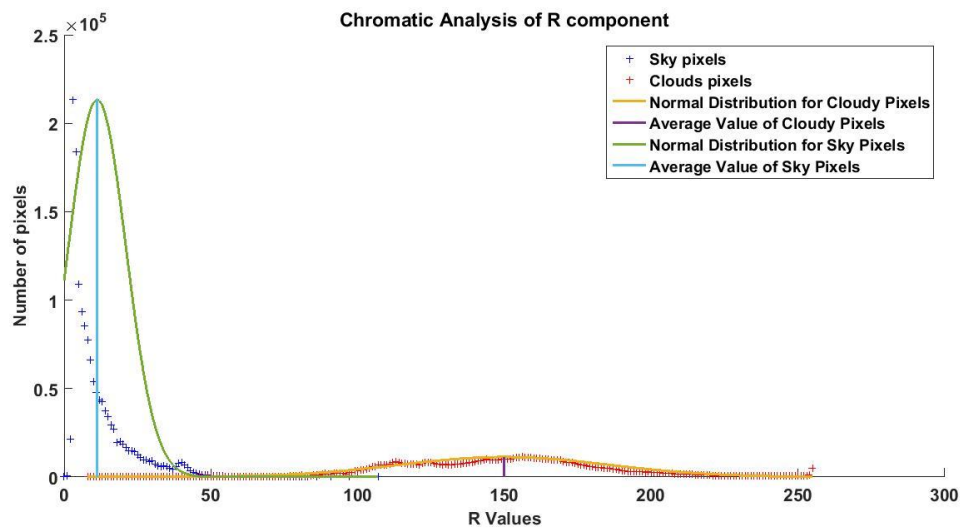
**Figure 5.11 – Recomposed complete masked image.**

The original clear-sky image in Figure 5.7 has been masked manually by using Photoshop in order to get the masked clear-sky only image in Figure 5.8. Similarly, the cloudy image shown in Figure 5.9 has been masked manually to obtain the masked cloud only image in Figure 5.10. A chromatic analysis of the masked clear-sky image, shown in Figure 5.8, and the masked cloud image shown in Figure 5.10 has been made in order to determine the range of R, G and B values of the class of clear sky pixels and the values of cloudy pixels. Based on these two images, the chromatic range is shown in Table 5.1.

**Table 5.1 – Chromatic range of clear-sky and cloudy Images.**

<i>Class</i>	<i>Clear-Sky Pixels</i>		
<i>Primary Colour</i>	<i>Red</i>	<i>Green</i>	<i>Blue</i>
<i>Minimal Value</i>	0	12	24
<i>Maximal Value</i>	107	117	144
<i>Rounded Mean Value</i>	11	31	54
<i>Class</i>	<i>Cloudy Pixels</i>		
<i>Primary Colour</i>	<i>Red</i>	<i>Green</i>	<i>Blue</i>
<i>Minimal Value</i>	8	15	18
<i>Maximal Value</i>	255	255	255
<i>Rounded Mean Value</i>	150	159	166

By superposing Figure 5.8 and Figure 5.10 as two distinct layers in Photoshop, it is possible to make a recomposed image shown in Figure 5.11 that contains both clear sky pixels and clouds. The advantage is that the exact number of clear sky pixels and clouds is known. This way, it is possible to compare each pixel discrimination method and to determine which one is the best. A first qualitative guess can nonetheless be made by looking at the number of occurrences of each chromatic value for each class. For the two images mentioned, the complete chromatic analysis of the R, G, and B value distribution is shown in Figure 5.12, Figure 5.13, and Figure 5.14, respectively. A Gaussian distribution has been assumed as it most often occurs in the nature. This is why normal distribution curves of each class are shown based on the mean values and variance values.



**Figure 5.12 – Chromatic analysis of the R component of clear-sky and cloudy pixels.**

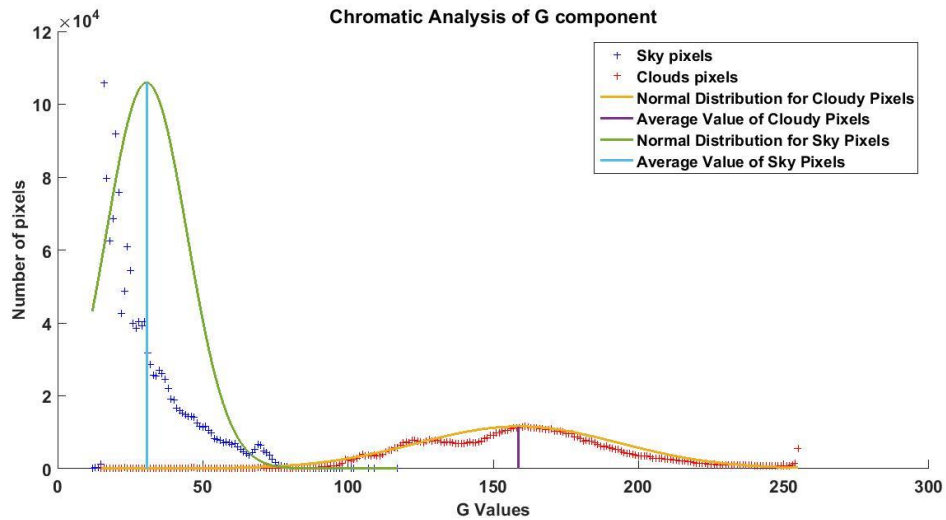


Figure 5.13 – Chromatic analysis of the G component of clear-sky and cloudy pixels.

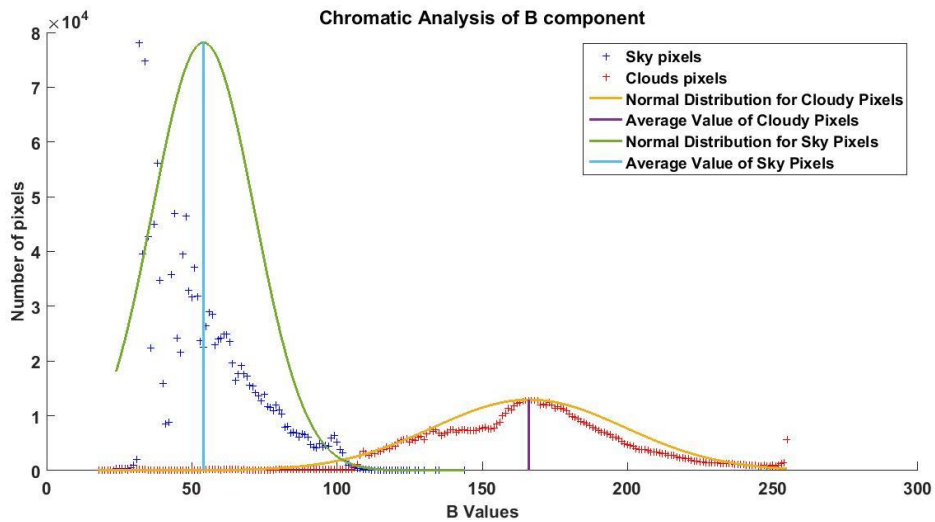


Figure 5.14 – Chromatic analysis of the B Component of clear-sky and cloudy pixels.

It is clearly visible in the distribution of R, G, and B that clouds diffuse each component of sunlight in a similar manner. However, as already mentioned in reference [122], the molecule present in the air in the absence of cloud diffuse more blue light than red light. This fact is also visible in the chromatic range shown in Table 5.1. For this reason, it is far easier to identify clear-sky pixels than cloudy pixel in a sky image that has been taken with a proper exposure. Also, it makes perfect sense to detect sky pixels by comparing the blue values with respect to red values, since there is a ratio of 5 between them.

However, another feature that is visible is the overlap of both classes. The minimal R, G, and B values of pixels belonging to the class Clouds is lower than the average value of pixels belonging to the class Clear-Sky. This results in many cloudy pixels having a closer R, G, or B value to the average value of clear-sky pixels than the average value of the class they belong to. Theoretically, a simple probabilistic classifier like a simple Bayes filter, or naïve predictor, could theoretically be trained to classify pixels

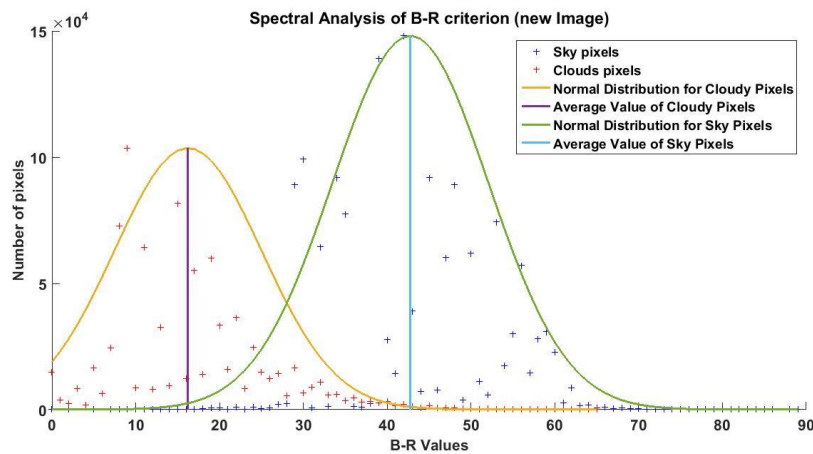
as either cloudy or clear sky pixels. However, the mean values of the features of each class are so close that there is a significant overlap between both classes and a lot of confusion can be expected between both classes. In particular, after training and naïve Bayesian classification, an error of about 44% was found. In order to tell both classes apart more easily, an operation must be made in order to move away the average value of each class. In particular, combining two colours together is a common way to proceed. The figures below show various criteria used in the literature. The criteria in question have been mentioned by Heinle et al. [58] and are described in Table 5.2 below.

**Table 5.2 – Clear-sky pixel discrimination criteria under consideration.**

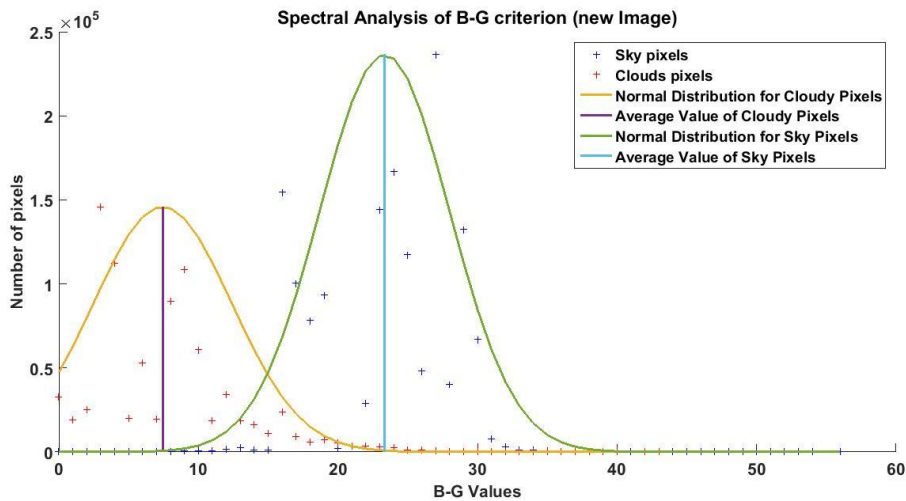
<i>Criterion</i>	<i>Criterion Formula</i>
1 [147]	$R / B \leq T_1$
2 [58]	$B - R > T_2$
3 [59]	$((B - R > T_2) \& (B - G > T_3) \& (B > T_4))$

Those criteria have already been mentioned in Chapter 3. However, it must be reminded that their threshold values vary greatly according to the camera sensor that is used. Thus, a completely new chromatic analysis is required for the Nikon D7100 camera.

For the sake of simplicity, the chromatic analysis is first made on the entire sky. The second criterion in Table 5.2, also often referred to as difference criterion,  $R - B$ , is shown in Figure 5.15. Although the average values of both sky pixels and cloud pixels class seem to be relatively far apart, there is a significant overlap of both classes.



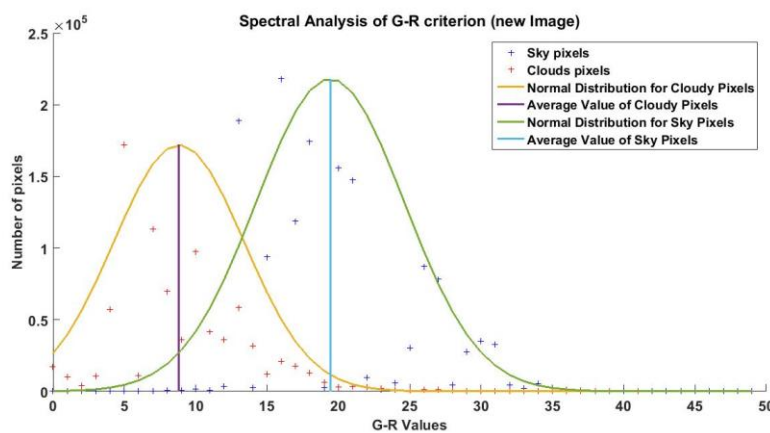
**Figure 5.15 – B-R chromatic analysis.**



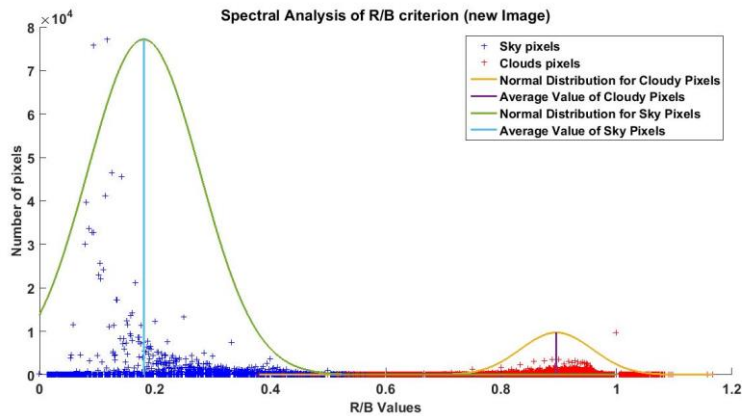
**Figure 5.16 – B-G chromatic analysis.**

To compensate for this overlap and the class confusion that may result from it, Kazantzidis et al [122] suggested to consider another difference term, namely the B – G difference shown in Figure 5.16. Also, it has been suggested to consider the B criterion, previously featured in Figure 5.14. The combination of the two difference criterion with the blue criterion is presented as the third criterion in Table 5.3. This criterion can also be called multicolour. Although, it has proven to be effective and reliable for some camera sensors, there is still too much overlap with every different criterion, when using the Nikon D7100. As an alternative, the difference criterion G – R seem to be characterised by the biggest overlap among all difference criteria, as shown in Figure 5.17.

Visibly, when the entire sky is analysed, the ratio criterion R/B displays the largest gap between both classes, as seen in Figure 5.18. Thus, a simple visual analysis of the chromatic analysis can already lead to suspect that the ratio criterion, or first criterion as per Table 5.3, is likely to be the most suitable for the Nikon D7100.

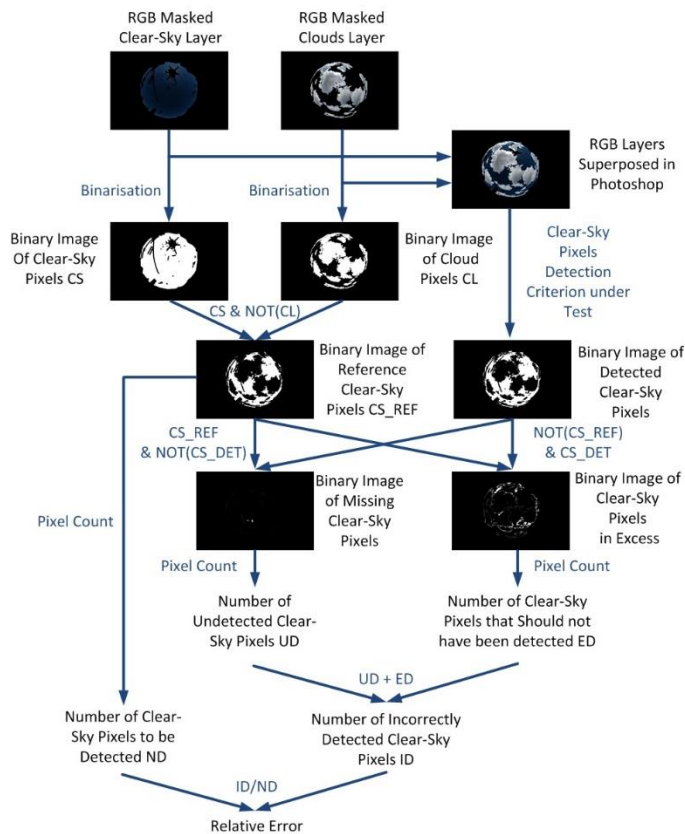


**Figure 5.17 – G-R chromatic analysis.**



**Figure 5.18 – R/B chromatic analysis.**

For each clear-sky pixel criterion shown in Table 5.2, thresholds are sought to minimise the error due to incorrectly detected pixels. Either too many or too few clear-sky pixels may be identified by an imperfect criterion. Figure 5.19 illustrates the modus operandi that is applied to estimate the best thresholds for each criterion, which generates the lowest error. For a given criterion, the same algorithm is run for various thresholds and the corresponding relative error is calculated. The threshold value generating the lowest error is kept as the best. In further details, black pixels are filtered out of the clear-sky image. The result is the binary image CS that only features clear-sky pixels. The same filter is also applied to the cloud-only image. The result is the binary image CL that only features cloudy pixels.



**Figure 5.19 – Algorithm applied to calculate the error in clear-sky pixel detection.**

**Table 5.3 – Optimal threshold values.**

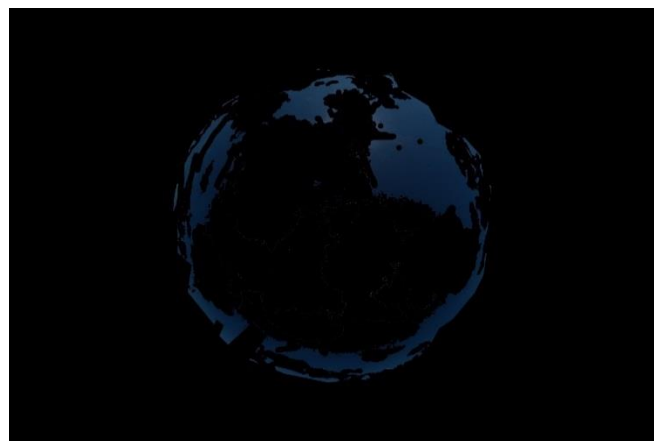
<i>Thresholds</i>			
$T_1$	$T_2$	$T_3$	$T_4$
<b>0.55</b>	28	15	30

The thresholds of each criterion obtained by using the above mentioned method is shown in Table 5.4

The results of the process are shown in further details throughout the sky images shown in Figures 5.20–5.24. The white pixels in Figure 5.20 correspond to the position of the pixels that have been manually identified as clear-sky pixels in a given sky image. This image can thus be used as a reference. With respect to this reference, a simple naïve Bayes filter as shown in Figure 5.21 confuses a large number of clear-sky pixels for cloudy pixels, mostly in the central region of the sky image. In Figures 5.21-5.24, the pixels of the original image are shown where each respective criterion has identified a clear-sky pixel.



**Figure 5.20 – Reference binary sky image featuring clear-sky pixels only.**



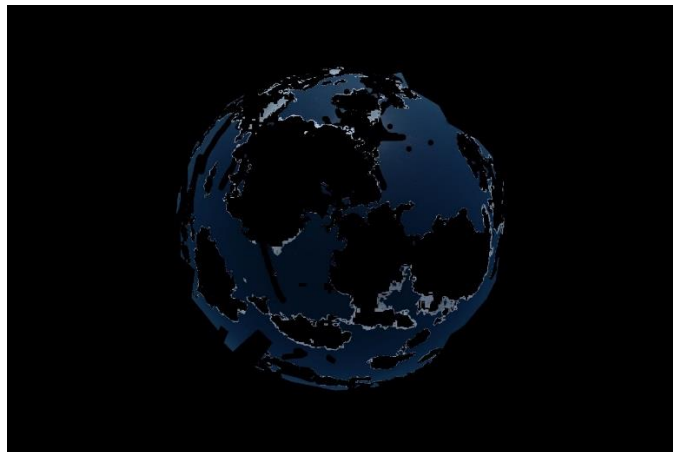
**Figure 5.21 – Clear-sky pixels detected with a Naive Bayes classifier.**



**Figure 5.22 – Clear-sky pixels detected with the R/B criterion.**



**Figure 5.23 – Clear-sky pixels detected with the B-R criterion.**



**Figure 5.24 – Clear-sky pixels detected with the multicolour criterion.**

Similarly to what has been done with Figure 5.21, a simple visual examination of the pixels identified as clear-sky pixels by the ratio R/B criterion in Figure 5.22 proves this criterion to be quite accurate, due to the resemblance of the reference sky image from Figure 5.20. In comparison, a relatively large number of cloudy pixels get misclassified as clear-sky pixels when using the difference criterion  $R - B$  in Figure 5.23. Although complementing the difference criterion with a couple of more criterion to



practically implement the multicolour criterion does reduce the number of cloud pixels incorrectly detected as clear-sky pixels, too many cloud pixels remain misclassified.

As a result, the ratio criterion proves once again to be the most accurate of all four criteria through a simple quick visual evaluation of the accuracy of each criterion.

**Table 5.4 – Best relative errors corresponding to the optimal threshold values.**

<i>Criterion</i>	<i>R/B</i>	<i>B-R</i>	<i>Multicolour</i>	<i>Simple Bayes</i>
<i>Rel. Error (%)</i>	4.35	15.62	12.08	44.46

Table 5.4 summarises the quantitative relative pixel classification errors. Arguably, this finally confirms that the R/B ratio to be the most accurate approach to detect clear-sky pixels with the sensor of Nikon D7100 DSLR camera. However, it must be reminded that only a special case has been considered, in which the scattering effect around the solar disk and at the horizon are minimal. The analysis of a larger number of sky configurations is still required at this stage to draw some more relevant conclusions.

## 5.4 DETAILED CHROMATIC ANALYSIS OF A CLEAR-SKY LIBRARY

For a more detailed analysis of the sky, selection of 54 clear-sky images has been made to be used as a reference clear-sky library. Within this clear-sky library, a separate analysis of chromatic features has been performed for the clear sky pixels located within the circumsolar region and outside of it. Following a recommendation given in [165], only times of the days corresponding to an azimuth angle lower or equal to 70 degrees have been considered. The underlying reason for that selection is because most of the solar power generation occurs whenever the azimuth angle of the sun remains lower than 70 degrees. The dataset corresponding to the clear-sky library is described in Table 5.5 and covers a wide range of times across all four seasons.

**Table 5.5 – Description of the clear-sky dataset.**

<b>Date</b>	<b>Frequency</b>	<b>Frequency</b>	<b>Number of Images</b>	<b>Season</b>
01/07/2016	12:00:01 - 15:00:00	30 minutes	7	Winter
28/11/2016	11:05:01 - 16:05:02	30 minutes	11	Spring
25/01/2017	07:29:41 - 17:29:22	30 minutes	21	Summer
04/04/2017	08:30:54 - 16:00:56	30 minutes	16	Autumn

**Table 5.6 – Average values of features for the circumsolar region and for the extrasolar region within the dataset.**

<i>Feature</i>	<i>Circumsolar Clear Sky</i>	<i>Extrasolar Clear Sky</i>
<i>Number of pictures</i>	54	54
<i>Mean (R)</i>	57.6250	18.7972
<i>Mean (G)</i>	78.6470	41.5237
<i>Mean (B)</i>	105.1245	70.1827
<i>Standard Deviation (B)</i>	32.7548	18.7942
<i>Skewness (B)</i>	0.3326	0.3907
<i>Difference (R-G)</i>	-21.0220	-22.7264
<i>Difference (R-B)</i>	-47.4995	-51.3855
<i>Difference (G-B)</i>	-26.4775	-28.6591
<i>Ratio R/B</i>	<b>0.5340</b>	<b>0.2639</b>
<i>Energy (B)</i>	0.0065	0.0165
<i>Entropy (B)</i>	0.5324	0.4443
<i>Contrast (B)</i>	68.2780	7.0902
<i>Homogeneity (B)</i>	0.7078	0.8665

When analysing the features of every clear-sky pixels across the entire clear-sky library, a distinction has been made between the circumsolar region, which corresponds to a radius of 35 degrees around the centre of the solar disk and the extrasolar region, which corresponds to the rest of the sky. After such an analysis, the average values of each feature have been compiled in Table 5.6.

It appears that the average threshold values for the R/B criterion varies quite significantly between the circumsolar region and the extrasolar region. Thus, a different threshold value should be considered for each region. The difference of threshold seems less dramatic for some alternative criteria, such as the difference criterion R – B and even less so for G – B and R – G. However, the ratio criterion R/G has proven to be the most accurate for the totality of the sky in the previous section. Thus, despite this inconsistency, it has been decided to use the ratio criterion to discrimination clear-sky pixels from cloud pixels.

## **5.5 CLOUD FEATURE ANALYSIS AND AUTOMATED CLOUD CLASSIFICATION**

The previous step has consisted in analysis a clear-sky library to determine an accurate criterion and some thresholds to distinguish clear-sky pixels from clouds, effectively to identify some clouds, broadly speaking. However, as already mentioned in Chapter 2, there are 6 cloud types on top of a clear-sky situation and each has a different effect on a photovoltaic system. Hence, a further step would consist in analysing the features of each cloud type to identify some characteristic features that are particular to each cloud type, as perceived by the Nikon D7100 sensor. Once the features of each type of clouds are known, it is then relatively easy to automate the cloud classification and effectively activate the forecast when cumulus clouds are detected. Also, some alerts may be given when a dramatic long-term sunlight disruption is to be expected, like what may happen with rain clouds.

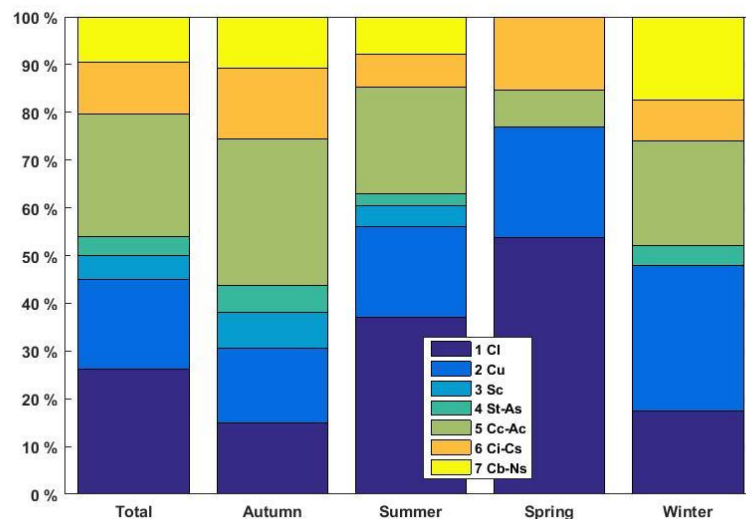
### 5.5.1 Clear-sky and Cloud Data Collection

To carry out an analysis of the features of each cloud type, a manual cloud classification has been done across a collection of 860,579 sky images. Those images were taken during daytime over a period of 159 days spanning from 15/04/2016 to 17/03/2017 and covering all four seasons.

Although the camera was not in operation every day during that period, all cloud types have been recorded. After a manual classification, the frequency of each cloud type has been documented in Table 5.7 and represented graphically in Figure 5.25. It must be noted that only a relative small number of days could be recorded in the spring and in the winter.

**Table 5.7 – Cloud type frequency for the period under consideration.**

Cloud Type	Total Frequency (%)	Autumn Frequency (%)	Summer Frequency (%)	Spring Frequency (%)	Winter Frequency (%)
1	26.3	14.9	37.1	53.8	17.4
2	18.6	15.7	19.0	23.1	30.4
3	5.1	7.4	4.3	0.0	0.0
4	4.0	5.8	2.6	0.0	4.3
5	25.5	30.6	22.4	7.7	21.7
6	10.9	14.9	6.9	15.4	8.7
7	9.5	10.7	7.8	0.0	17.4
<b>Number of days</b>	<b>159</b>	<b>64</b>	<b>75</b>	<b>9</b>	<b>11</b>



**Figure 5.25 – Graphical representation of the cloud type frequency distribution for the period under consideration.**

It can be noticed that overall for the whole period, the type Cirrocumulus-Alto cumulus dominates and was recorded as occurring a quarter of the time. A large proportion of those clouds have been detected during all seasons. This type of clouds correspond to high altitude clouds that often cover a large portion of the sky as it will be shown in the next section. Also, being located at a high altitude, such clouds

have a relative slow movement when seen from the surface, which does not result in swift irradiance variations.

Cumulus clouds come next and occur overall 18% of the period while this type of clouds has been identified as causing most of fast stochastic sunlight disruption [31, 32]. Such clouds were present during all four seasons but largely dominant in the winter.

Some other cloud type that were also significantly present were cirrus and cirrostratus that are thin clouds located at a very altitude. However, such clouds are not expected to cause swift sunlight variations. Also, rain clouds occurred quite often in winter and autumn. Considering their thickness, it is expected for those clouds to dim sunlight very significantly. However, due to the size of those clouds and the portion of the sky they usually cover, such clouds are very likely to obstruct sunlight for a period of time that may largely exceed 30 minutes, which is a timeframe beyond the scope of very short-term forecasting.

### 5.5.2 Clear-Sky and Cloud Feature Characterisation

The Fisher Distance based automatic cloud classification method developed by Pankiewicz in 1995, and used by Heinle et al. in 2010 [58], has been applied to the sky and cloud dataset. As already presented in details in Section 2.8.3, this method consists in analysing 12 different features. The results have been compiled within Table 5.8 – 5.10. The features have been calculated for the pixels that have been identified as cloudy pixels for all classes except for the clear-sky class, for which clear-sky pixels only have been considered. Some additional features such as the normalised (B-R)/(B+R) difference, as well as the saturation and the Euclidean geometric distance (EGD) have been also been added. The analysis includes the average as well as the extreme minimum and maximum values of every feature.

**Table 5.8 – Cloud feature characterisation of sky images – red (R) and green (G) features.**

<i>Number of pictures</i>		26,318	29	18	26	20	46	23
<b>Feature</b>	<b>Type</b>	<b>Cl</b>	<b>Cu</b>	<b>Sc</b>	<b>St-As</b>	<b>Cc-Ac</b>	<b>Ci-Cs</b>	<b>Cb-Ns</b>
<i>Min (R)</i>	<i>Min</i>	1	1	1	1	1	1	1
	<i>Avg</i>	1	1	1	9.845	1	1	9.391
	<i>Max</i>	1	1	1	68	1	1	43
<b>Mean (R)</b>	<i>Min</i>	10.293	58.342	136.468	95.290	92.447	71.171	120.584
	<b>Avg</b>	<b>34.212</b>	<b>153.535</b>	<b>152.153</b>	<b>136.065</b>	<b>122.415</b>	<b>113.434</b>	<b>142.120</b>
	<i>Max</i>	144.282	216.024	178.680	155.670	159.988	168.890	150.407
<i>Max (R)</i>	<i>Min</i>	251	188	243	181	226	220	182
	<i>Avg</i>	254.997	249.276	252.500	226.077	246.650	243.217	227.870
	<i>Max</i>	255	255	255	255	255	255	255
<i>Min (G)</i>	<i>Min</i>	1	1	1	1	1	1	1
	<i>Avg</i>	1	1	1	10.654	1	1	10.826
	<i>Max</i>	1	1	1	73	1	1	47
<b>Mean (G)</b>	<i>Min</i>	25.362	63.172	143.310	102.268	103.193	82.028	126.506
	<b>Avg</b>	<b>58.934</b>	<b>161.223</b>	<b>160.534</b>	<b>142.814</b>	<b>131.833</b>	<b>122.621</b>	<b>151.255</b>
	<i>Max</i>	176.818	219.034	189.681	162.932	170.550	180.157	162.853
<i>Max (G)</i>	<i>Min</i>	251	194	244	185	229	225	189
	<i>Avg</i>	254.997	249.379	252.889	228.270	246.850	244.674	232.522
	<i>Max</i>	255	255	255	255	255	255	255

**Table 5.9 – Continuation of the cloud feature characterisation of sky images from the Nikon D7100 camera – blue (B) feature.**

<i>Number of pictures</i>		26,318	25	18	26	32	46	23
<b>Feature</b>	<b>Type</b>	<b>Cl</b>	<b>Cu</b>	<b>Sc</b>	<b>St-As</b>	<b>Cc-Ac</b>	<b>Ci-Cs</b>	<b>Cb-Ns</b>
<i>Min (B)</i>	<i>Min</i>	1	1	1	1	1	1	1
	<i>Avg</i>	1	1	1	11	1	1	10.696
	<i>Max</i>	1	1	1	76	1	1	50
<i>Mean (B)</i>	<i>Min</i>	39.383	67.485	153.647	108.031	117.048	95.737	129.478
	<i>Avg</i>	<b>86.895</b>	<b>169.110</b>	<b>170.045</b>	<b>149.137</b>	<b>143.441</b>	<b>132.824</b>	<b>158.463</b>
	<i>Max</i>	207.104	221.180	198.220	169.330	179.399	191.152	173.357
<i>Max (B)</i>	<i>Min</i>	246	198	251	188	234	223	198
	<i>Avg</i>	254.935	249.655	170.045	229.500	247.250	245.717	236.435
	<i>Max</i>	255	255	255	255	255	255	255
<i>Standard Deviation (B)</i>	<i>Min</i>	11.523	33.730	35.962	8.564	25.6494	23.070	13.227
	<i>Avg</i>	<b>28.150</b>	<b>61.66</b>	<b>49.185</b>	<b>22.875</b>	<b>38.9614</b>	<b>44.034</b>	<b>24.885</b>
	<i>Max</i>	46.061	93.417	62.453	42.194	49.307	71.869	43.230
<i>Skewness (B)</i>	<i>Min</i>	-4.178	-2.815	-2.517	-1.407	-2.760	-1.870	-1.666
	<i>Avg</i>	<b>1.627</b>	<b>-1.476</b>	<b>-1.849</b>	<b>-0.345</b>	<b>-1.149</b>	<b>-0.654</b>	<b>-0.746</b>
	<i>Max</i>	6.949	0.692	-1.187	1.076	0.0459	0.811	0.502
<i>Energy (B)</i>	<i>Min</i>	0.0030	4.88·10 <sup>-4</sup>	7.31·10 <sup>-4</sup>	0.0014	0.0010	6.45·10 <sup>-4</sup>	0.0021
	<i>Avg</i>	<b>0.0119</b>	<b>0.0099</b>	<b>0.0016</b>	<b>0.0092</b>	<b>0.0019</b>	<b>0.0023</b>	<b>0.0062</b>
	<i>Max</i>	0.0392	0.0810	0.0026	0.0274	0.0044	0.0078	0.0185
<i>Entropy (B)</i>	<i>Min</i>	0.418	0.312	0.648	0.295	0.606	0.373	0.337
	<i>Avg</i>	<b>0.551</b>	<b>0.774</b>	<b>0.884</b>	<b>0.457</b>	<b>0.772</b>	<b>0.686</b>	<b>0.528</b>
	<i>Max</i>	0.753	0.961	0.970	0.812	0.919	0.930	0.653
<i>Contrast (B)</i>	<i>Min</i>	4.0064	116.252	183.422	1.904	34.586	11.784	2.069
	<i>Avg</i>	<b>27.759</b>	<b>1030.0</b>	<b>448.232</b>	<b>20.337</b>	<b>173.217</b>	<b>246.929</b>	<b>16.599</b>
	<i>Max</i>	922.456	2512.1	652.713	151.504	537.156	1024	106.738
<i>Homogeneity (B)</i>	<i>Min</i>	0.739	0.275	0.396	0.552	0.417	0.326	0.567
	<i>Avg</i>	<b>0.836</b>	<b>0.462</b>	<b>0.497</b>	<b>0.756</b>	<b>0.543</b>	<b>0.543</b>	<b>0.743</b>
	<i>Max</i>	0.908	0.568	0.591	0.877	0.644	0.776	0.867

**Table 5.10 – Cloud feature characterisation of sky images from the Nikon D7100 camera – other features.**

<i>Number of pictures</i>		26,318	25	18	26	32	46	23
<i>Feature</i>	<i>Type</i>	<i>Cl</i>	<i>Cu</i>	<i>Sc</i>	<i>St-As</i>	<i>Cc-Ac</i>	<i>Ci-Cs</i>	<i>Cb-Ns</i>
<i>Difference (R-G)</i>	<i>Min</i>	-155	-69	-76	-55	-65	-103	-51
	<i>Avg</i>	<b>-24.721</b>	<b>-7.689</b>	<b>-8.387</b>	<b>-6.749</b>	<b>-9.418</b>	<b>-9.187</b>	<b>-9.135</b>
	<i>Max</i>	142	18	26	24	30	50	21
<i>Difference (R-B)</i>	<i>Min</i>	-167	-137	-135	-98	-119	-130	-88
	<i>Avg</i>	<b>-52.683</b>	<b>-15.575</b>	<b>-17.892</b>	<b>-13.073</b>	<b>-21.026</b>	<b>-19.390</b>	<b>-16.342</b>
	<i>Max</i>	179	26	64	61	50	89	65
<i>Difference (G-B)</i>	<i>Min</i>	-135	-73	-100	-57	-60	-110	-82
	<i>Avg</i>	<b>-27.961</b>	<b>-7.887</b>	<b>-9.505</b>	<b>-6.324</b>	<b>-11.608</b>	<b>-10.202</b>	<b>-7.207</b>
	<i>Max</i>	155	31	46	56	40	87	61
<i>Ratio R/B</i>	<i>Min</i>	0.0066	0.0625	0.0526	0.0588	0.0200	0.0102	0.0714
	<i>Avg</i>	<b>0.320</b>	<b>0.893</b>	<b>0.885</b>	<b>0.906</b>	<b>0.833</b>	<b>0.833</b>	<b>0.898</b>
	<i>Max</i>	96	8	9	6	20	60	23
<i>Normalised (R-B)/(R+B)</i>	<i>Min</i>	-1	-1	-1	-1	-1	-1	-1
	<i>Avg</i>	<b>-0.535</b>	<b>-0.0683</b>	<b>-0.0671</b>	<b>-0.0507</b>	<b>-0.0996</b>	<b>-0.102</b>	<b>-0.0554</b>
	<i>Max</i>	1	1	1	1	1	1	1
<i>Saturation</i>	<i>Min</i>	0.0039	0.0039	0.0039	0.0039	0.0039	0.0039	0.0039
	<i>Avg</i>	<b>0.678</b>	<b>0.192</b>	<b>0.163</b>	<b>0.0968</b>	<b>0.196</b>	<b>0.213</b>	<b>0.105</b>
	<i>Max</i>	1	1	1	1	1	1	1
<i>Euclidian Geometric Distance</i>	<i>Min</i>	1.155	1.155	1.155	1.155	1.155	1.154	1.154
	<i>Avg</i>	<b>155.648</b>	<b>379.073</b>	<b>383.713</b>	<b>349.195</b>	<b>320.468</b>	<b>294.208</b>	<b>368.955</b>
	<i>Max</i>	624.620	624.620	624.620	624.620	624.620	624.620	624.620
<i>Cloud Cover (R,G,B) (%)</i>	<i>Min</i>	0.0445	5.085	28.897	69.846	38.718	7.455	94.193
	<i>Avg</i>	<b>0.499</b>	<b>26.533</b>	<b>49.526</b>	<b>96.667</b>	<b>69.261</b>	<b>54.084</b>	<b>99.423</b>
	<i>Max</i>	9.997	54.787	75.533	100	96.3414	99.058	100

Excluding clear-sky conditions, the cloud type that covers the least the sky happens to be cumulus clouds with a maximum total cloud cover slightly over 50%. Following this cloud type, the closely related stratocumulus cover the sky up to about 75%. As a comparison, stratus and altostratus clouds cover at least about 70% of the sky and over 95% in average. There is wide variation of the sky covered by cirrocumulus/altostratus and an even wider range covered by cirrus/cirrostratus clouds. In contrast, rain clouds always cover a very large portion of the sky, with average total cloud cover close to 100%.

Although the ratio R/B criterion is effective to tell apart clear-sky pixels from any sort of cloud pixel, due to the homogeneity of the average values for all clouds and their contrast with respect to the sky value. However, it is not sufficient to tell apart cloud types due to the lack of difference of average value across all types of cloud types regarding this feature.

Across all the explored features, besides the total cloud cover, contrast, standard deviation and skewness of the blue component seem to be the most discriminant features to distinguish clearly each cloud type. Thus, it can be recommended to use the latter as priority to distinguish cloud types, when using a Nikon D7100.

## **5.6 MEASUREMENT OF THE OPTICAL THICKNESS**

An even further analysis of clouds would be the observation of their direct and indirect effect on the sunlight. The most obvious direct effect is how much clouds have dimmed the light of the sun, which is in relation with their optical thickness, and how long for.

Several days' worth of recording between 4am and 8pm have been considered for each cloud type, with 1 second of time resolution. However, for practical reasons, as already discussed in Section 5.4 of this chapter, only azimuth angles smaller than 70 degrees have been considered. One motivation for doing that is of course the fact that the irradiance is particularly weak when the sun is lower than an elevation of 30 degrees, limiting the effect of clouds on sunlight dimming. However, the same decision is also motivated by the fact that it is quite difficult to model the low level of irradiance that occurs when the sun is particularly low. As a result, for zenith angles above 70 degrees, the inaccuracy of the model rather than the effect of clouds would cause most of the discrepancy between modelled clear-sky irradiance and measured irradiance.

The images of the Nikon D7100 have been used to select the days when only one single cloud has been detected.

### 5.6.1 Cloud type 2: Cumulus Clouds

Cumulus clouds have been identified as the sole cloud type during 9 different days within the dataset. Those days are noted at the top of Table 5.11. The time mentioned as total cloud time effectively corresponds to the total duration, in seconds, during which some cumulus clouds has been identified throughout the day, when the azimuth was below 70 degrees. The time resolution in use is 1 second.

**Table 5.11 – Observations made on cumulus clouds.**

Days	19/04/2016	2/05/2016	25/05/2016	27/05/2016	10/07/2016	27/12/2016	12/01/2017	19/01/2017	5/02/2017
Total cloud time (s)	5949	18940	9522	7431	3869	12566	10174	6861	2820
Number of clouds	19	11	20	21	14	18	28	7	12
Min cloud duration (s)	39	49	17	50	50	49	101	255	48
Mean cloud duration (s)	312.1052632	<b>1720.818182</b>	475.1	352.8571429	275.3571429	697.1111111	362.3571429	979.1428571	<b>234</b>
Max cloud duration (s)	1025	3996	1793	1178	1383	3844	2265	2664	1178
Min cloud $\tau$	0.070000109	0.073783758	0.070031344	0.073580276	0.070009868	0.070390772	0.072536438	0.127726961	0.070002461
Mean cloud $\tau$	<b>1.044621473</b>	0.817862478	0.907447985	<b>0.521725511</b>	0.735419933	1.163301755	0.861775301	0.853993206	0.793624763
Max cloud $\tau$	1.805447147	2.3555839	1.746411466	1.86864493	1.991883105	2.001238301	1.553876941	1.408658465	2.194907803
$\sigma$ cloud $\tau$	0.400800013	0.403553516	0.404483384	0.345088463	0.478118976	0.376406302	0.356528783	0.283366398	0.636569322
Total peak time (s)	12098	5072	8876	6627	14584	25116	26753	8678	11238
Number of peaks	30	13	22	42	24	31	35	15	23
Min peak duration (s)	7	100	9	0	18	29	24	12	26
Mean peak duration (s)	402.2666667	389.1538462	402.4545455	<b>156.7857143</b>	606.6666667	<b>809.1935484</b>	763.3714286	577.5333333	487.6086957
Max peak duration (s)	1891	1228	1792	664	2869	6860	2919	1946	3484
Min peak $\tau$	-0.362145087	-0.351956192	-0.448637771	-0.459790748	-0.480305556	-0.324205268	-0.364156733	-0.382051957	-0.252412032
Mean peak $\tau$	-0.100033621	-0.148088712	-0.155252168	-0.119635915	-0.133260615	-0.104543549	-0.113785958	-0.113252257	-0.064761781
Max peak $\tau$	-0.02000154	-0.024759552	-0.020014036	-0.020010096	-0.020037968	-0.020006903	-0.020077914	-0.020000919	-0.020006843
$\sigma$ peak $\tau$	0.072776942	0.079337471	0.101576708	0.117812901	0.087308681	0.063448801	0.069677562	0.088751574	0.054416615

From the results compiled in Table 5.11, it appears that cumulus clouds created some shading that lasted continuously in average between 4 and 29 minutes, accompanied by some peak amplitude episodes that lasted in average between 2 minutes 36 seconds and 13 minutes 29 seconds. It must also be noted that irradiance peaks do not occur systematically in the presence of clouds. The average optical thickness of the clouds varies significantly from day to day between approximately 0.52 and 1.04.

Figure 5.26 is a good illustration of the effect of passing cumulus clouds on the global horizontal irradiance (GHI) when clouds pass by all day long, as they did on 2/05/2016. The clear-sky irradiance modelled with the refined TJ regression is shown in blue whereas the measured GHI is shown in red. Some peaks can be occasionally observed between a few consecutive shadings. The maximum shade around 13:17 corresponded to an irradiation reduction of 85.7%. The corresponding optical thickness values throughout the same day, for zenith angles smaller than 70 degrees, are shown in Figure 5.27.



The green and blue horizontal lines in that figure represent the thresholds of tolerance. As a comparison, the power output from the PV array is shown in Figure 5.28.

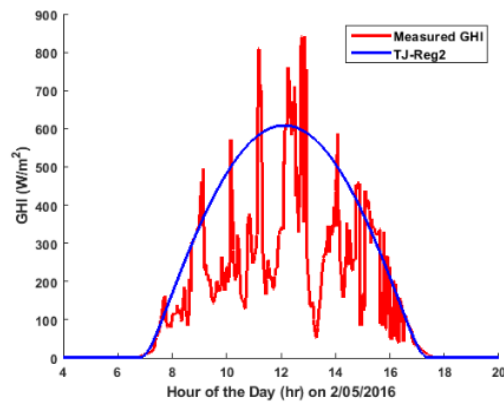


Figure 5.26 – Measured and clear-sky GHI on 2/05/2016.

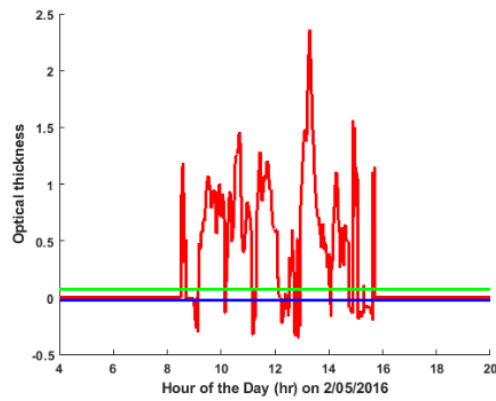


Figure 5.27 – Optical thickness on 2/05/2016.

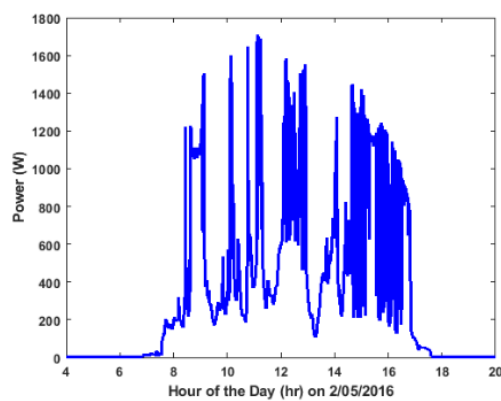
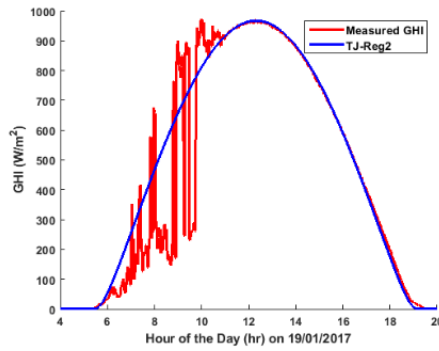
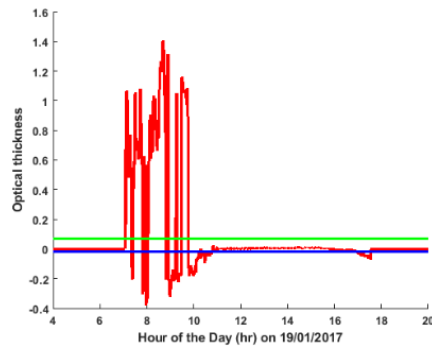


Figure 5.28 – Power output on 2/05/2016.

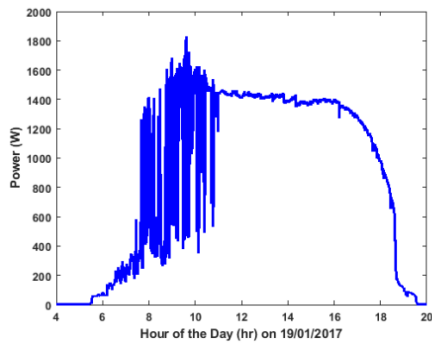
It can be observed in Figure 5.28 that the peaks in irradiance create power surges up to 1.7kW, which exceeds the rated power of the solar array, which is 1.5kW. Some similar quantities are represented in Figures 5.29-5.31 based on some measurements performed on 19/01/2017.



**Figure 5.29 – Measured and clear-sky GHI on 19/01/2017.**



**Figure 5.30 – Optical thickness on 19/01/2017.**



**Figure 5.31 – PV Power output on 19/01/2017.**

An illustration of a possible cause of irradiance peaks is shown in Figure 5.32. A cloud that is close enough from the sun but that does not obstruct the sun may act as a reflector if the incidence angle of the sunlight enables a reflection. In this case, the cloud will act as a secondary source of light. It is for example the case for the illuminated cloud that is visible on the left of the solar disk in Figure 5.32.



Figure 5.32 – View of the sky on 19/01/2017 at 9:59 am.

### 5.6.2 Cloud type 3 : Stratocumulus Clouds

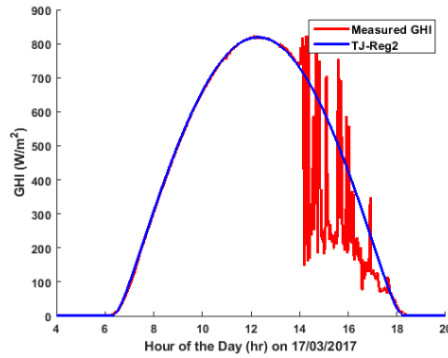
Some stratocumulus have been identified during 8 days of the dataset. Those days are mentioned on the top of Table 5.12. Similarly to the classification of cumulus clouds, the zenith angle of the selected samples has been filtered to remain below 70 degrees and the time resolution of the samples is 1 second.

Table 5.12 – Observations made on stratocumulus clouds.

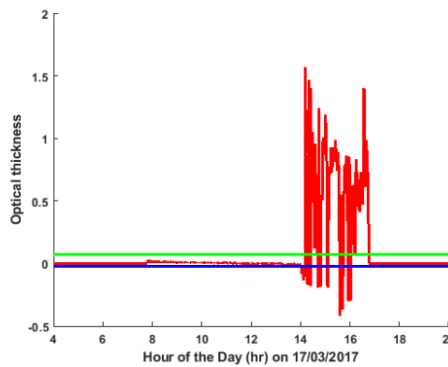
Days	7/05/2016	25/05/2016	28/05/2016	6/01/2017	5/02/2017	17/03/2017	24/03/2017	26/03/2017
Total cloud time (s)	13688	9522	4722	19979	2820	7821	5078	10677
Number of clouds	24	20	24	24	12	12	19	31
Min cloud duration (s)	49	17	3	102	48	69	6	50
Mean cloud duration (s)	569.333333	475.1	<b>195.75</b>	<b>831.45833</b>	234	650.75	266.263158	343.419355
Max cloud duration (s)	2510	1793	923	6406	1178	2010	819	1075
Min cloud $\tau$	0.0703764	0.07003134	0.07001551	0.0714931	0.0700025	0.07002348	0.07017843	0.07003247
Mean cloud $\tau$	<b>0.9278182</b>	0.90744798	0.84663228	1.0085809	0.7936248	<b>0.75475442</b>	0.92185024	0.85208469
Max cloud $\tau$	2.2165052	1.74641147	1.55082513	2.0712714	2.1949078	1.56735083	1.60996877	1.70340863
$\sigma$ cloud $\tau$	0.4631798	0.40448338	0.43102746	0.3743414	0.6365693	0.24339903	0.47158736	0.35032563
Total peak time (s)	8426	8876	11561	18308	11238	1940	18498	16852
Number of peaks	24	22	32	25	23	11	22	44
Min peak duration (s)	40	9	12	100	26	43	6	4
Mean peak duration (s)	350.08333	402.454545	360.28125	731.32	487.6087	<b>175.363636</b>	<b>839.818182</b>	382
Max peak duration (s)	1332	1792	1794	3151	3484	511	2407	1998
Min peak $\tau$	-0.519597	-0.4486378	-0.3616931	-0.301512	-0.252412	-0.4169338	-0.2474739	-0.3497863
Mean peak $\tau$	-0.167384	-0.1552522	-0.0992819	-0.132999	-0.064762	-0.1763909	-0.1004807	-0.1504563
Max peak $\tau$	-0.02001	-0.020014	-0.0200211	-0.020022	-0.020007	-0.0200286	-0.0200599	-0.0200007
$\sigma$ peak $\tau$	0.1031319	0.10157671	0.07129267	0.0695594	0.0544166	0.10347209	0.05498564	0.08793759

The shading caused by individual stratocumulus clouds lasted between 3 minutes 15 second and 14 minutes, accompanied by some peak amplitude episodes that lasted in average between 3 minutes and 14 minutes. The average optical thickness of the clouds is more consistent than for cumulus clouds as it varies between approximately 0.75 and 0.92.

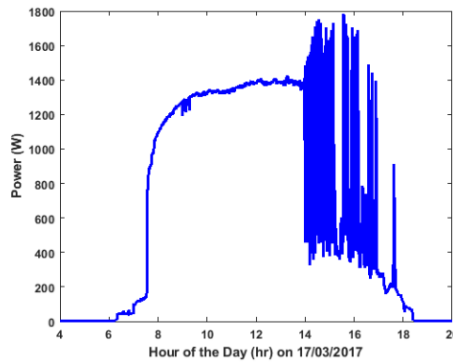
Figure 5.33 features the clear-sky and measured GHI on the day of 17/03/2017 during which shading only occurred in the afternoon. It can be observed that the effect of stratocumulus clouds looks very similar to the effect of cumulus clouds, as previously observed, in terms of duration and intensity.



**Figure 5.33 – Measured and clear-sky GHI on 17/03/2017.**



**Figure 5.34 – Optical thickness on 17/03/2017.**



**Figure 5.35 – PV Power output on 17/03/2017.**

The corresponding optical thickness and the resulting PV power output on the same day can be seen respectively in Figure 5.34 and Figure 5.35. Some intense irradiance peaks are present in Figure 5.33, leading to some very strong reverse optical thickness values in Figure 5.34 and some equally important power surges in Figure 5.35 that largely exceed the rated power of the PV array.

### 5.6.3 Cloud type 4 : Stratus/Altostratus Clouds

A total of 11 days have been identified as being characterised by the passage of stratus and altostratus clouds. As shown in Tables 5.13 and 5.14, a single cloud can shade for up to 6 hours 33 minutes. Some broken pieces of a cloud can alternatively create some shading for at least 9 minutes and 26 seconds.

**Table 5.13 – Observations made on stratus and altostratus clouds – First part.**

Days	17/04/2016	18/04/2016	24/04/2016	27/04/2016	29/04/2016	19/05/2016
Total cloud time (s)	27940	27132	21832	18151	21701	23609
Number of clouds	1	5	21	10	14	1
Min cloud duration (s)	27939	103	51	6	32	23608
Mean cloud duration (s)	27939	5425.4	1038.61905	1814.1	1549.07143	<b>23608</b>
Max cloud duration (s)	27939	18956	3639	7891	7225	23608
Min cloud $\tau$	0.13844912	0.0883357	0.0700986	0.07007278	0.07002974	0.69049541
Mean cloud $\tau$	1.05894744	0.95422079	0.56480456	0.83688376	0.7963897	1.02043845
Max cloud $\tau$	2.84217499	1.87742377	1.68360657	2.08838741	1.99688668	1.35219421
$\sigma$ cloud $\tau$	0.44757282	0.38788681	0.32583347	0.38061275	0.39198905	0.15965802
Total peak time (s)	0	460	3144	7969	4304	0
Number of peaks	0	4	15	17	13	0
Min peak duration (s)	0	49	16	0	101	0
Mean peak duration (s)	<b>0</b>	114	208.6	<b>467.764706</b>	330.076923	<b>0</b>
Max peak duration (s)	0	204	359	2046	768	0
Min peak $\tau$	0	-0.1645708	-0.3730512	-0.3958562	-0.3629718	0
Mean peak $\tau$	0	-0.0863214	-0.1266405	-0.1378324	-0.1498863	0
Max peak $\tau$	0	-0.0218016	-0.0200395	-0.0200256	-0.0231545	0
$\sigma$ peak $\tau$	0	0.04506961	0.08378251	0.10121516	0.08688845	0

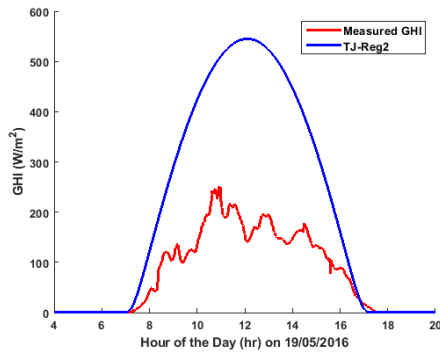
**Table 5.14 – Observations made on stratus and altostratus clouds – Second part.**

Days	26/05/2016	7/07/2016	18/01/2017	2/02/2017	11/02/2017
Total cloud time (s)	17405	22062	14190	16566	32169
Number of clouds	29	1	25	10	24
Min cloud duration (s)	31	22061	11	4	49
Mean cloud duration (s)	599.172414	22061	<b>566.6</b>	1655.6	1339.375
Max cloud duration (s)	3279	22061	4334	9232	7788
Min cloud $\tau$	0.07000637	0.3515243	0.07005727	0.0700546	0.07002586
Mean cloud $\tau$	0.47642489	0.7732448	0.60246503	0.5717536	0.5392232
Max cloud $\tau$	1.998643	1.1292582	1.47808697	1.1818053	1.52792562
$\sigma$ cloud $\tau$	0.33385158	0.1738666	0.31445128	0.2166575	0.28820771
Total peak time (s)	3910	0	13385	7458	2403
Number of peaks	21	0	36	17	14
Min peak duration (s)	17	0	12	3	45
Mean peak duration (s)	185.190476	0	370.805556	437.70588	170.642857
Max peak duration (s)	460	0	1434	1639	358
Min peak $\tau$	-0.3432187	0	-0.324853	-0.395484	-0.4476265
Mean peak $\tau$	-0.1300833	0	-0.1379224	-0.148877	-0.2147724
Max peak $\tau$	-0.0200203	0	-0.0200041	-0.020013	-0.0200341
$\sigma$ peak $\tau$	0.08207285	0	0.07651672	0.128823	0.13717958

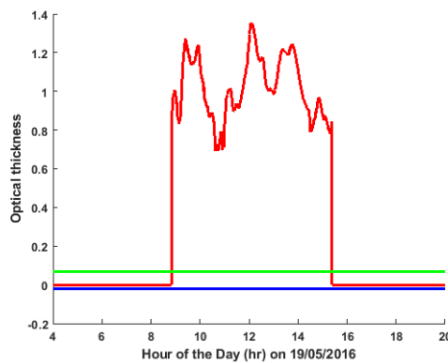
It can be observed that irradiance peaks occur less often than for cumulus and stratocumulus clouds. The reason is the overextended size of this class of clouds. As seen in a previous section of this chapter, stratus clouds can spread over such a large horizontal distance that they can cover the sky entirely over an extended amount of time. Thus, as shown in Tables 5.13 and 5.14, no irradiance intensity peaks were recorded during days when only one single cloud was present.

A typical shape of the measured GHI drawn in red, and to be compared with its corresponding clear-sky irradiance shown in blue, can be seen in Figure 5.36. On that day, only one single cloud was present all day long and no irradiance intensity peaks were observed.

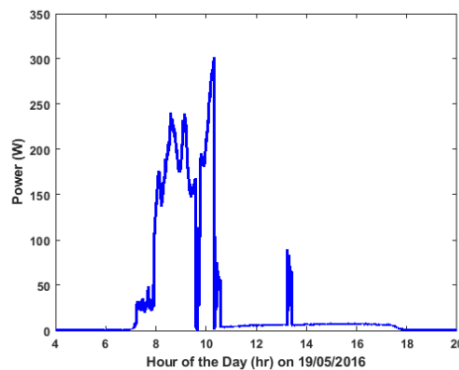
The effect of such a single cloud resulted in a consistently strong opacity throughout the entire day, as shown in Figure 5.37.



**Figure 5.36 – Measured and clear-sky GHI on 19/05/2016.**



**Figure 5.37 – Optical thickness on 19/05/2016.**



**Figure 5.38 – PV Power output on 19/05/2016.**

The thickness of the cloud slowly varied over time and oscillated around the average value of about 1, varying between 0.7 as minimum and 1.35 as a maximum. As a result, the PV power generation for the entire day has been very low. The occasional peak generations never exceeded 300W, which corresponds to only 20% of the installed capacity. For several hours in a row, the effective operating power was even close to 0W.

It can be concluded from those observations that stratus and altostratus clouds may cause a very consistent continuous power cut, which places them out of the scope of very-short term PV power forecasting.

### 5.6.4 Cloud type 5 : Cirrocumulus/Alto cumulus

Cirrocumulus and alto cumulus clouds have been observed during 14 days of the dataset as shown in Tables 5.15 and 5.16. Those clouds may cause very fast variations of irradiance, with shading occurring between 3 minutes and 33 minutes in average, with a minimum of 3 seconds only on 16/01/2017. Also, it must be noted that the optical thickness of the clouds varied greatly between 0.35 and 1.05 in average.

From the observations that have been made, irradiance peaks occur very often in presence of these clouds, with shading being virtually systematically accompanied by peaks. In average, for the investigated days, peaks lasted between 2 minutes 42 seconds and 8 minutes 7 seconds. An extreme variability of the peak optical thickness has been observed, between -0.047 and -0.21.

**Table 5.15 – Observations made on cirrocumulus and alto cumulus clouds – First part.**

Days	21/06/2016	26/06/2016	16/01/2017	18/01/2017	27/01/2017	28/01/2017	2/02/2017
Total cloud time (s)	6047	512	1779	14190	1384	5572	16566
Number of clouds	4	1	10	25	5	10	10
Min cloud duration (s)	153	511	3	11	51	79	4
Mean cloud duration (s)	1510.75	511	176.9	566.6	275.8	556.2	1655.6
Max cloud duration (s)	3843	511	358	4334	716	2920	9232
Min cloud $\tau$	0.08834771	0.34656324	0.07004569	0.07005727	0.07134518	0.07002575	0.0700546
Mean cloud $\tau$	0.64289459	0.55657672	0.51936905	0.60246503	0.52013592	1.05077472	0.5717536
Max cloud $\tau$	1.2774113	0.97660881	1.68046954	1.47808697	1.08540087	1.98376429	1.1818053
$\sigma$ cloud $\tau$	0.30515362	0.25190008	0.4715895	0.31445128	0.29952941	0.51498426	0.2166575
Total peak time (s)	5948	2489	6928	13385	4693	5762	7458
Number of peaks	14	11	29	36	16	25	17
Min peak duration (s)	7	14	1	12	2	6	3
Mean peak duration (s)	423.857143	225.272727	237.896552	370.805556	292.3125	229.48	437.70588
Max peak duration (s)	1655	770	3482	1434	1127	922	1639
Min peak $\tau$	-0.4415364	-0.0963885	-0.1537618	-0.324853	-0.312408	-0.174301	-0.395484
Mean peak $\tau$	-0.1229816	-0.0468114	-0.0607316	-0.1379224	-0.0933142	-0.0544276	-0.148877
Max peak $\tau$	-0.0200014	-0.0200111	-0.0200071	-0.0200041	-0.0200234	-0.0200071	-0.020013
$\sigma$ peak $\tau$	0.11465041	0.02128357	0.03446247	0.07651672	0.07909435	0.03177483	0.128823

**Table 5.16 – Observations made on cirrocumulus and alto cumulus clouds – Second part.**

Days	5/02/2017	7/02/2017	8/02/2017	11/02/2017	27/03/2017	28/03/2017	6/04/2017
Total cloud time (s)	2820	30556	6324	32169	3017	21372	10846
Number of clouds	12	15	12	24	8	19	28
Min cloud duration (s)	48	51	27	49	46	50	38
Mean cloud duration (s)	234	2036.0667	526	1339.375	376.125	1123.84211	386.35714
Max cloud duration (s)	1178	13118	3177	7788	1844	11018	5893
Min cloud $\tau$	0.0700025	0.0712886	0.0700088	0.07002586	0.07002236	0.07054546	0.0700056
Mean cloud $\tau$	0.7936248	0.8061889	0.7628048	0.5392232	0.34497655	0.77143827	0.5139924
Max cloud $\tau$	2.1949078	1.5689519	1.4876315	1.52792562	0.81206294	1.77992752	1.5128825
$\sigma$ cloud $\tau$	0.6365693	0.3201265	0.2792991	0.28820771	0.22964578	0.35405062	0.3287414
Total peak time (s)	11238	4930	22372	2403	12175	4709	3928
Number of peaks	23	17	58	14	57	21	24
Min peak duration (s)	26	16	10	45	1	10	11
Mean peak duration (s)	487.6087	289	384.72414	170.642857	212.596491	223.238095	162.66667
Max peak duration (s)	3484	768	6337	358	1638	1946	563
Min peak $\tau$	-0.252412	-0.394791	-0.177641	-0.4476265	-0.3573679	-0.4792665	-0.264155
Mean peak $\tau$	-0.064762	-0.144943	-0.048884	-0.2147724	-0.0567133	-0.1740411	-0.091857
Max peak $\tau$	-0.020007	-0.020048	-0.020001	-0.0200341	-0.0200004	-0.0200117	-0.020005
$\sigma$ peak $\tau$	0.0544166	0.0966714	0.0363307	0.13717958	0.06068089	0.11979943	0.05543

A comparison of the measured GHI and the clear-sky GHI has been made in Figure 5.39. The large intensity of both shading and peaks can be observed. Also, both shading and peaks look very brief.

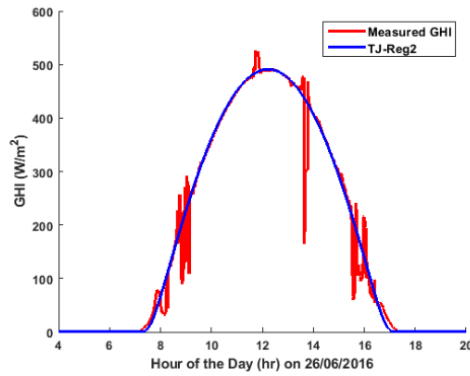


Figure 5.39 – Measured and clear-sky GHI on 26/06/2016.

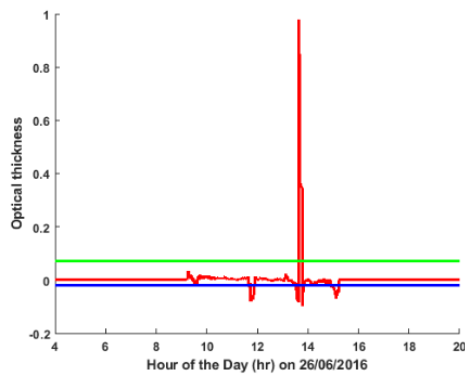


Figure 5.40 – Optical thickness on 26/06/2016.

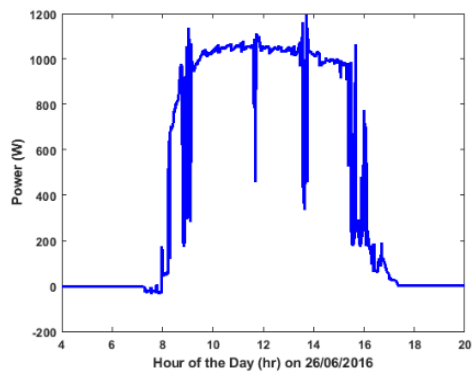


Figure 5.41 – PV power output on 26/06/2016.

The resulting optical thickness through the day, as seen in Figure 5.40, is characterised by very sharp, swift variations. The PV power output on the same day that is plotted in Figure 5.41 looks overall consistent and almost like it could be expected on a clear-sky day, with the exception of some very brief and large power drops preceded and followed by some equally brief power surges.

### 5.6.5 Cloud type 6 : Cirrus/Cirrostratus

A total of 15 days have been classified as days during which either cirrus or cirrostratus clouds passed over the sky camera and the PV array.



**Table 5.17 – Observations made on cirrus and cirrostratus clouds – First part.**

Days	21/04/2016	22/04/2016	3/05/2016	20/05/2016	26/05/2016	31/05/2016	16/06/2016
Total cloud time (s)	1095	13936	22902	6111	17405	5251	7915
Number of clouds	9	16	6	11	29	30	17
Min cloud duration (s)	18	6	203	41	31	49	32
Mean cloud duration (s)	120.666667	870	3816	554.545455	599.172414	174.033333	464.588235
Max cloud duration (s)	306	2973	12663	1434	3279	614	3022
Min cloud $\tau$	0.07000191	0.07001943	0.082042	0.07002098	0.07000637	0.07005331	0.07373782
Mean cloud $\tau$	<b>0.15527231</b>	0.42889334	0.7200895	0.86243795	0.47642489	0.29027485	0.55786287
Max cloud $\tau$	0.37858025	0.96333886	2.7475918	1.48086132	1.998643	1.46703482	1.70696004
$\sigma$ cloud $\tau$	0.09287793	0.2406574	0.2990863	0.33107214	0.33385158	0.2436032	0.37200535
Total peak time (s)	6784	10077	1485	2661	3910	8634	3347
Number of peaks	18	30	7	12	21	34	17
Min peak duration (s)	10	16	49	101	17	49	3
Mean peak duration (s)	<b>375.888889</b>	334.9	211.14286	220.75	185.190476	252.941176	195.882353
Max peak duration (s)	2407	2153	511	461	460	768	563
Min peak $\tau$	-0.0752579	-0.3226735	-0.14243	-0.368144	-0.3432187	-0.180011	-0.2588884
Mean peak $\tau$	-0.0423849	-0.09926	-0.066978	-0.1433612	-0.1300833	-0.0759811	-0.0773964
Max peak $\tau$	-0.0200056	-0.0200648	-0.020786	-0.0224158	-0.0200203	-0.0200129	-0.0200122
$\sigma$ peak $\tau$	0.01429603	0.07320902	0.0331685	0.10798153	0.08207285	0.04106444	0.05977279

**Table 5.18 – Observations made on cirrus and cirrostratus clouds – Second part.**

Days	7/07/2016	29/11/2016	15/12/2016	1/02/2017	3/02/2017	7/03/2017	8/03/2017	27/03/2017
Total cloud time (s)	22062	570	22086	7012	1794	3275	12668	3019
Number of clouds	1	6	26	17	7	22	45	8
Min cloud duration (s)	22061	1	46	50	51	49	14	47
Mean cloud duration (s)	<b>22061</b>	<b>94</b>	848.461538	411.47059	255.28571	147.86364	280.51111	376.375
Max cloud duration (s)	22061	204	5278	2509	768	512	1332	1844
Min cloud $\tau$	0.3515243	0.07000021	0.0700556	0.0713703	0.0966666	0.0702194	0.0700099	0.07002236
Mean cloud $\tau$	0.7732448	0.64513778	0.45119303	<b>0.9254451</b>	0.720678	0.2152958	0.2599684	0.34486928
Max cloud $\tau$	1.1292582	2.99425572	2.41750935	2.4263691	1.5228519	0.8762939	3.6111322	0.81206294
$\sigma$ cloud $\tau$	0.1738666	1.0895309	0.3296937	0.9675387	0.4037574	0.1781556	0.3340134	0.22977835
Total peak time (s)	0	5514	12498	5114	5527	3823	4987	12174
Number of peaks	0	35	64	42	22	19	27	57
Min peak duration (s)	0	1	5	0	14	46	28	1
Mean peak duration (s)	<b>0</b>	156.542857	194.28125	120.7619	250.22727	200.21053	183.7037	212.578947
Max peak duration (s)	0	1793	1382	1162	1197	819	870	1639
Min peak $\tau$	0	-0.1509303	-0.2842887	-0.225167	-0.310316	-0.192571	-0.220567	-0.3573679
Mean peak $\tau$	0	-0.0571672	-0.0750734	-0.038168	-0.081997	-0.074565	-0.091756	-0.0567236
Max peak $\tau$	0	-0.02	-0.0200031	-0.020003	-0.02003	-0.020004	-0.020002	-0.0200004
$\sigma$ peak $\tau$	0	0.03652224	0.05484804	0.0242874	0.062818	0.0514348	0.0536395	0.06068712

Figure 5.42 shows an interesting example of measured GHI superposed to its corresponding clear-sky irradiance curve. It can be seen that the periods of shading are very short, not unlike for cirrocumulus and altocumulus. However, when it occurs, the shading is not as intense as for the previous class.

In contrast, the irradiance peaks are quite intense and prolonged in time. The optical thickness shown in Figure 5.43 demonstrates swift variations of low intensity optical thickness when a shading occurs.

In comparison, peaks are sustained over time and may be intense, although the intensity of peaks varies greatly.

The curve of the PV array output power shown in Figure 5.44 is quite close to what could be expected during a clear-sky day. It is not totally unlike the output power curve from Figure 5.41, with nonetheless a major difference in terms of intensity. The sunlight attenuation is more moderate than with cirrocumulus altocumulus since the maximum optical depth of clouds did not exceed 0.38 on 21/04/2016.

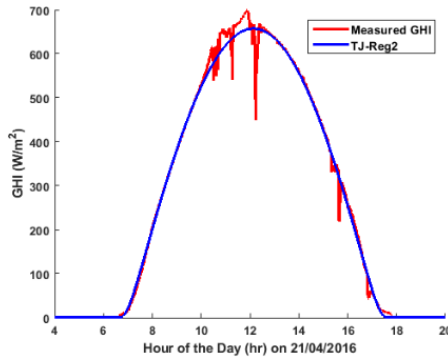


Figure 5.42 – Measured and clear-sky GHI on 21/04/2016.

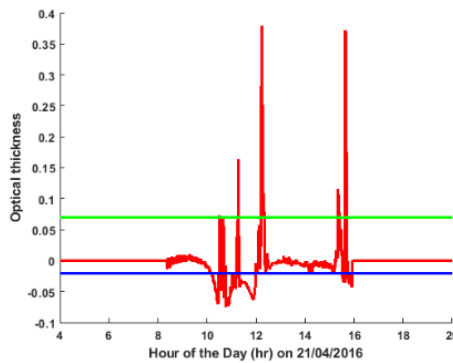


Figure 5.43 – Optical thickness on 21/04/2016.

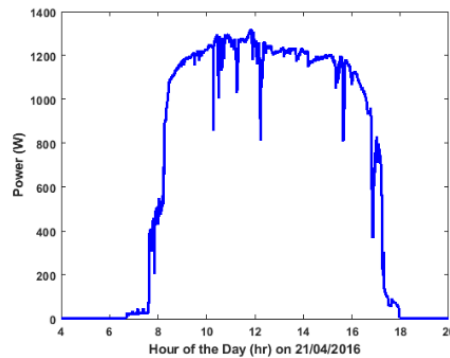


Figure 5.44 – PV power output on 21/04/2016.

### 5.6.6 Cloud type 7: Cumulonimbus/Nimbostratus

Finally, 11 days have been selected because they were characterised by the presence of rain clouds, namely cumulonimbus or nimbostratus. The results of the observations during those days has been compiled in Tables 5.19 and 5.20. Although some broken pieces of such rain clouds may appear on the edge of a large cloud, rain clouds tend to be very large and usually cover the sky entirely for long periods of time. The longest period of shading time among the selected days is 10 hours 16 minutes for a single cloud. In contrast, the shortest average time is about 15 minutes. However, rain clouds are a mature version of cumulus or stratus clouds and some younger clouds can sometimes surround the mature one.

**Table 5.19 – Observations made on cumulonimbus and nimbostratus clouds – First part.**

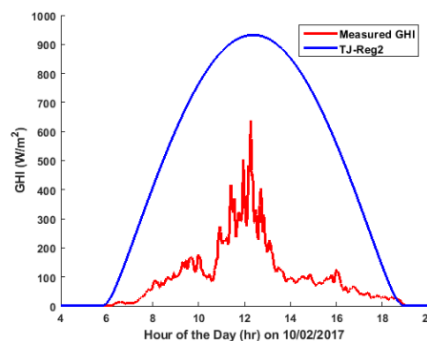
Days	17/04/2016	29/04/2016	2/05/2016	15/05/2016	21/05/2016	19/06/2016
Total cloud time (s)	27940	21701	18941	13240	18157	20536
Number of clouds	1	14	11	15	12	4
Min cloud duration (s)	27939	32	50	101	102	1605
Mean cloud duration (s)	27939	1549.07143	1720.9091	<b>881.666667</b>	1512.08333	5133
Max cloud duration (s)	27939	7225	3996	3074	14732	10352
Min cloud $\tau$	0.13844912	0.07002974	0.0737838	0.07011774	0.07716377	0.10750353
Mean cloud $\tau$	1.05893937	<b>0.7963897</b>	0.8179338	0.80504437	1.57048075	0.86622403
Max cloud $\tau$	2.84217499	1.99688668	2.3555839	1.82083447	3.97427145	1.75584946
$\sigma$ cloud $\tau$	0.44756399	0.39198905	0.4035452	0.38012107	0.77730016	0.30868756
Total peak time (s)	0	4304	5074	6310	4072	615
Number of peaks	0	13	13	20	13	3
Min peak duration (s)	0	101	101	50	101	101
Mean peak duration (s)	0	330.076923	389.30769	314.5	312.230769	204
Max peak duration (s)	0	768	1229	1434	819	358
Min peak $\tau$	0	-0.3629718	-0.351956	-0.438827	-0.3229329	-0.3709165
Mean peak $\tau$	0	-0.1498863	-0.148243	-0.1513001	-0.1537336	-0.2856986
Max peak $\tau$	0	-0.0231545	-0.022453	-0.0205882	-0.0200134	-0.1395024
$\sigma$ peak $\tau$	0	0.08688845	0.0793206	0.10730688	0.08576833	0.07777129

**Table 5.20 – Observations made on cumulonimbus and nimbostratus clouds – Second part.**

Days	22/06/2016	30/01/2016	7/02/2017	9/02/2017	10/02/2017
Total cloud time (s)	14384	36994	30556	36168	36077
Number of clouds	9	1	15	1	1
Min cloud duration (s)	153	36993	51	36167	36076
Mean cloud duration (s)	1597.22222	<b>36993</b>	2036.0667	36167	36076
Max cloud duration (s)	5945	36993	13118	36167	36076
Min cloud $\tau$	0.07271929	0.8348859	0.0712886	0.7290805	0.37917291
Mean cloud $\tau$	0.6870043	<b>1.83043166</b>	0.8061889	1.5778994	1.75433701
Max cloud $\tau$	1.78607164	2.81139471	1.5689519	2.8909436	2.65741139
$\sigma$ cloud $\tau$	0.31628096	0.43147137	0.3201265	0.4350542	0.41790509
Total peak time (s)	4824	0	4930	0	0
Number of peaks	11	0	17	0	0
Min peak duration (s)	168	0	16	0	0
Mean peak duration (s)	437.545455	0	289	0	0
Max peak duration (s)	1126	0	768	0	0
Min peak $\tau$	-0.3409986	0	-0.394791	0	0
Mean peak $\tau$	-0.129662	0	-0.144943	0	0
Max peak $\tau$	-0.0200075	0	-0.020048	0	0
$\sigma$ peak $\tau$	0.09873647	0	0.0966714	0	0

The optical thickness of those rain clouds varies greatly between 0.8 and 1.83. It can be nonetheless said that, overall, the optical thickness is high. This causes some significant, extended periods of time during which the sunlight is very strongly dimmed, without any irradiance peaks, as it can be seen in Figure 5.45.

The optical thickness remained remarkably consistent and high throughout the same day, as shown in Figure 5.46. As a direct consequence, the power generated by the PV array on the same day is low overall, with some occasional peaks around noon. It can be noted that unlike what could be observed with some other clouds, the power curve closely follows the GHI curve.



**Figure 5.45 – Measured and clear-sky GHI on 10/02/2017.**

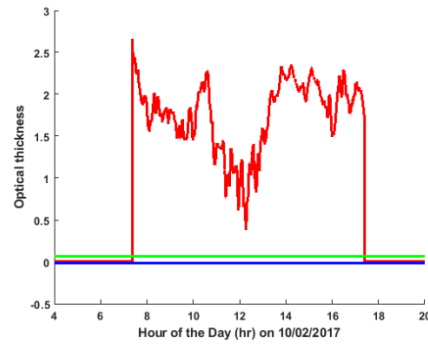


Figure 5.46 – Optical thickness on 10/02/2017.

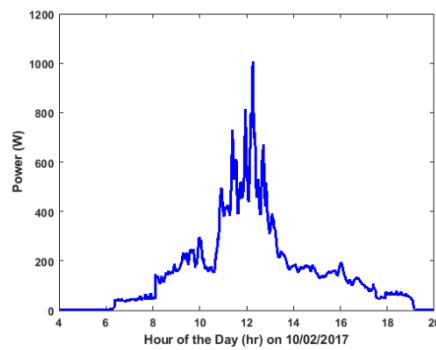


Figure 5.47 – PV power output on 10/02/2017.

### 5.6.7 Comparison of All Types of Cloud

The average values and standard deviation values for all cloud types have been regrouped in Table 5.21.

Table 5.21 – Observations made on all cloud types across the whole dataset.

Cloud Type	2.Cu	3.Sc	4.St-As	5.Cc-Ac	6.Ci-Cs	7.Cb-Ns
Total cloud time (s)	78132	74307	242757	153154	147101	274694
Number of clouds	150	166	141	183	250	84
Min cloud duration (s)	17	3	4	3	1	32
Mean cloud duration (s)	<b>519.88</b>	<b>446.6325</b>	<b>1720.681</b>	<b>835.9071</b>	<b>587.404</b>	<b>3269.167</b>
Max cloud duration (s)	3996	6406	27939	13118	22061	36993
Min cloud $\tau$	0.07	0.070002	0.070006	0.070002	0.07	0.07003
Mean cloud $\tau$	<b>0.877372</b>	<b>0.907166</b>	<b>0.762451</b>	<b>0.664931</b>	<b>0.559851</b>	<b>1.244716</b>
Max cloud $\tau$	2.355584	2.216505	2.842175	2.194908	3.611132	3.974271
$\sigma$ cloud $\tau$	0.429999	0.41335	0.385201	0.355449	0.413451	0.615569
Total peak time (s)	119042	95699	43033	108418	86535	30129
Number of peaks	235	203	137	362	405	90
Min peak duration (s)	0	4	0	1	0	0
Mean peak duration (s)	<b>505.5617</b>	<b>470.4236</b>	<b>306.4</b>	<b>298.4972</b>	<b>212.1429</b>	<b>319.5638</b>
Max peak duration (s)	6860	3484	2046	6337	2407	1434
Min peak $\tau$	<b>-0.48031</b>	<b>-0.5196</b>	<b>-0.44763</b>	<b>-0.47927</b>	<b>-0.36814</b>	<b>-0.43883</b>
Mean peak $\tau$	-0.11304	-0.12367	-0.1432	-0.09131	-0.07543	-0.14915
Max peak $\tau$	-0.02	-0.02	0	-0.02	0	0
$\sigma$ peak $\tau$	0.081033	0.083588	0.099628	0.0874	0.061933	0.095938

Unsurprisingly, rain clouds prove to cause the longest and most dramatic shading. In contrast, cumulus and stratus clouds also have a relatively high optical thickness while they cause the shortest disruptions. However, the duration and intensity of peaks is fairly consistent regardless of the cloud type. It can possibly be said that cirrus and cirrostratus clouds cause shorter and less dramatic irradiance peaks than

all of the other types. This result makes sense considering the thin, elongated shape of such clouds. It also seems that the longest and most intense irradiance peaks are caused by the densest and most compact clouds, which are cumulus and stratocumulus.

## **5.7 CONCLUSIONS**

A comprehensive analysis of the chromatic features and properties of the clear sky and of each type of cloud has been carried out. It was found that the ratio criterion was the most accurate to discriminate clear-sky pixels from cloud pixels. However, when the scattering effect is strong, two separate thresholds should be applied in the circumsolar region and in the extrasolar region. Also, an analysis of the distribution of clouds over a period covering nearly a year has demonstrated that cumulus clouds can be present all year long during all seasons. A study of the features of each cloud type, as recommended in the literature has been done, so that the resulting properties of each cloud type could be used for a fully automated online cloud classification. Finally, the effect of clouds on sunlight has been investigated. A large data set has highlighted the fact rain clouds were blocking sunlight the most and the longest. It was also found that cumulus and stratocumulus clouds may cause a significant shading as well, making them some of the most light-blocking blocks. However, those two cloud types also cause the short sunlight disruption. As a result, both cumulus and stratocumulus clouds are likely to make the DC voltage of grid-connected PV systems fairly unstable. A solution to that problem will be investigated in the next chapter.

## 6 PV POWER CONTROL

---

As seen in the previous chapters, clouds may affect the continuity of sunlight in various ways, based on their type. In particular, fast, the significant sunlight variations of irradiance caused by compact clouds such cumulus clouds may induce some undesirable power output variations. When such problems occurs, a grid-connected PV system may turn into a load and may consume some power from the grid. One simple possibility would consist in disconnecting the PV system from the grid by curtailing the active power through the MPPT [166]. Alternatively, a more elegant although more costly solution would consist in taking advantage of the potential past excess power generated by the system to regulate the power on the short-term [11]. Effectively, this latter solution could be applied to smoothen the power output of the PV system for quality purposes, prior to injecting it to the grid. This solution has been already extensively proposed to smoothen wind and solar power output [16, 17, 167]. A case study is carried out in this chapter to illustrate this strategy when applied for short-term PV power smoothening.

### 6.1 BATTERY STORAGE MANAGEMENT

Batteries have long been thought as a solution to mitigate the effects that sunlight fluctuations may have on PV systems [11]. However, on a relatively short-term horizon, proper control is required to manage the state of charge of a battery bank. For that purpose, it has been decided to observe the effects of sunlight disruptions of various amplitudes on a verified PV system model.

#### 6.1.1 PV Array Model

##### 6.1.1.1 *Initial Simulink Model*

The grid-connected 1.5kW monocrystalline PV array ‘Teaching Station 2’ from the GEEP lab, which has already been described in the previous chapter, can be modelled as shown in Figure 6.1. A model has been developed in MATLAB-SIMULINK to feature the PV array, with 8 SunTech STP190S-24/Ad+ monocrystalline modules in series, the maximum power point tracking (MPPT), the 1.7kW SMA SunnyBoy inverter, the DC link regulator and the grid as shown in Figure 6.2. The simulation results of the SIMULINK model have been compared with some measurements taken at the GEEP lab.

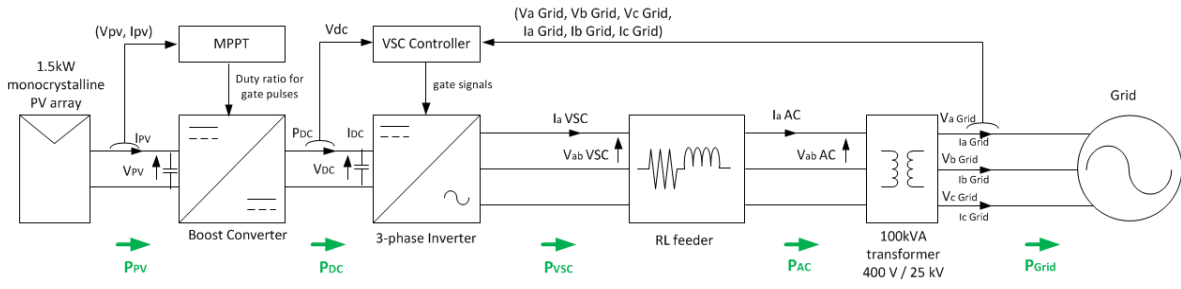


Figure 6.1 – Diagram of the GEEP lab 1.5kW monocrystalline PV system.

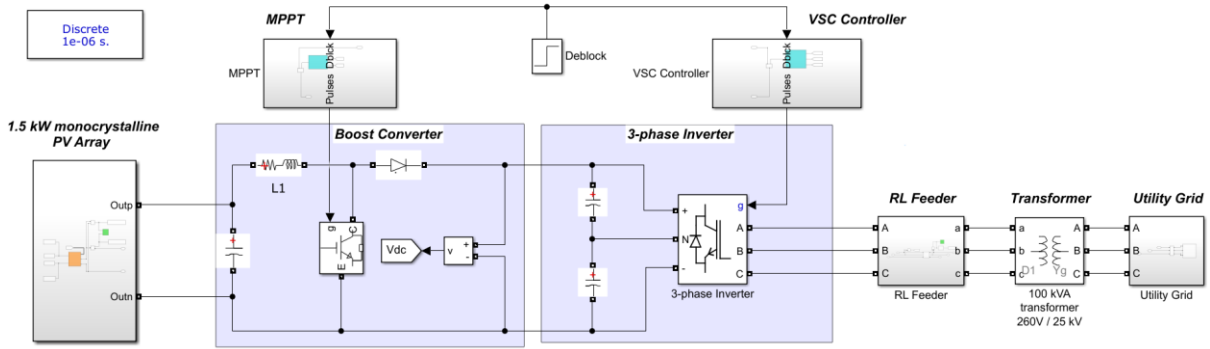


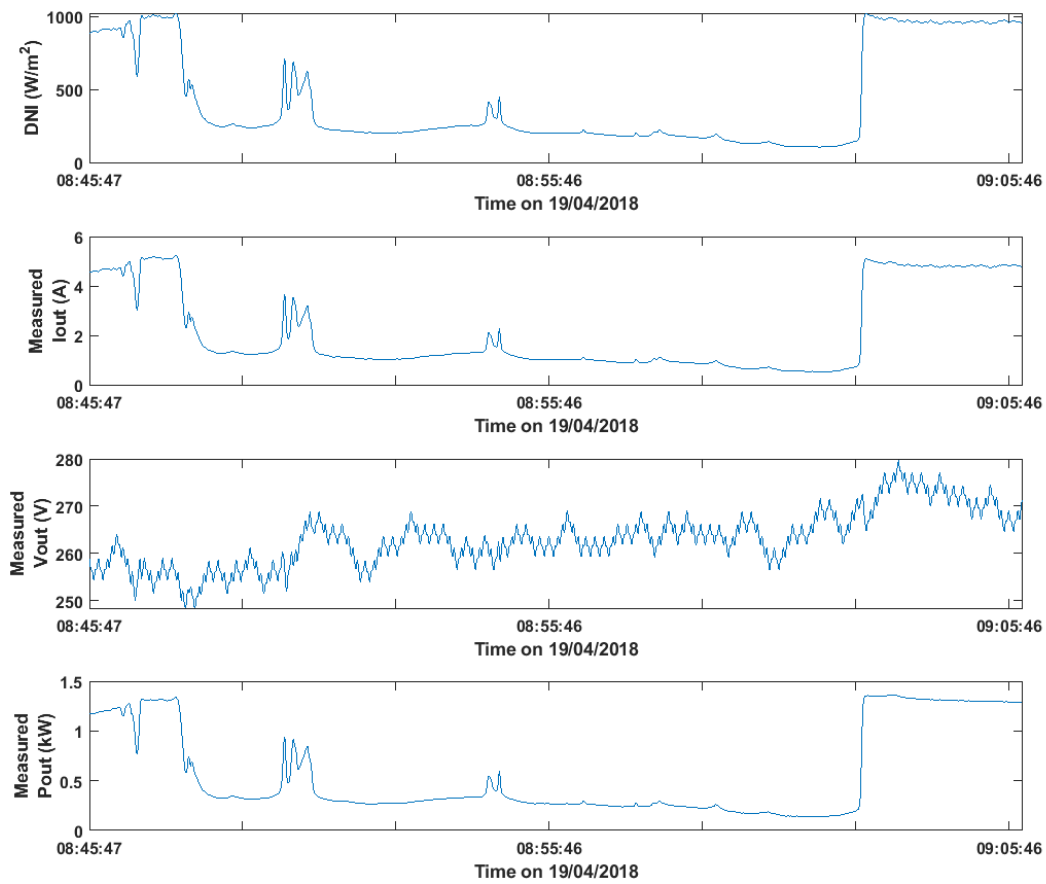
Figure 6.2 – MATLAB SIMULINK model of the GEEP lab 1.5kW Monocrystalline PV system.

Since the MATLAB SIMULINK model that has been designed is discrete with a sample time of  $1\mu\text{s}$ , simulations cannot be run for longer than about 20 seconds. In order to verify the accuracy of the model, it is thus necessary to choose a representative sample of measured data to be fed to the model. The data measured between 8:45:47 and 9:05:46 am has been identified as a particularly typical situation of dramatic transient shading. The measured direct normal irradiance (DNI) received by the array as well as the PV output current, voltage and power during that period of time are visible in Figure 6.3. It must be noted how close the variations of the DNI and the output current  $I_{out}$  since they are linked by the relationship given in Equation 6.1, in which the short-circuit current  $I_{sc}$  is proportional to the DNI  $G$ .

$$I_{out} = I_{sc} - I_0 \cdot \left( e^{\frac{q(V_{out} + R_s \cdot I_{out})}{k \cdot T_{cell}}} - 1 \right) - \frac{V_{out} + R_s \cdot I_{out}}{R_p} \quad (6.1)$$

Assuming that the second and third terms much smaller than the first, a simplification can be made as in Equation 6.2 where  $I_{ref}$  is the output current measured at  $G_{ref} = 1000 \text{ W/m}^2$  and  $T_{cell} = 25^\circ\text{C}$ .

$$I_{out} \approx I_{sc} = I_{ref} \cdot \left( \frac{G}{G_{ref}} \right) \times [1 + \alpha \cdot (T_{cell} - T_{cell,ref})] \quad (6.2)$$

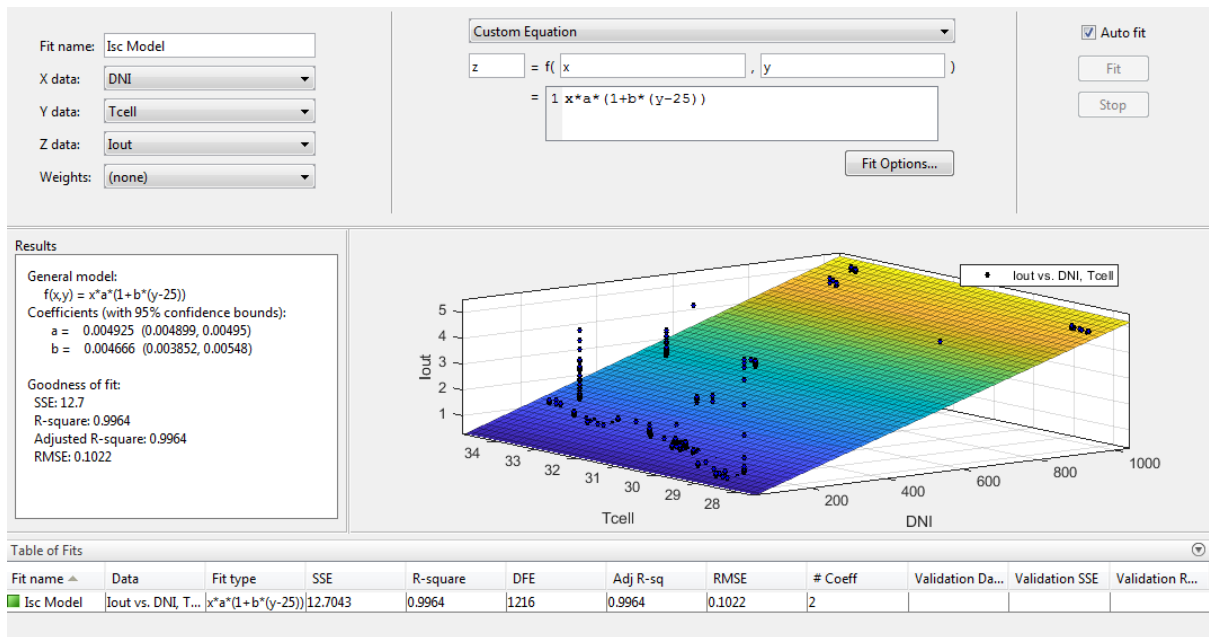


**Figure 6.3 – DNI and output signals of the 1.5kW monocrystalline array from the GEEP measured on 19/04/2018.**

By using MATLAB's linear toolbox, a linear regression between the output current, the cell temperature and a coarse measurement of the DNI for the same period of time lead to the relationship from Equation 6.3 that fits the measurements at 99.64% and only gives an RMSE of 0.1022 A in shown in Figure 6.4.

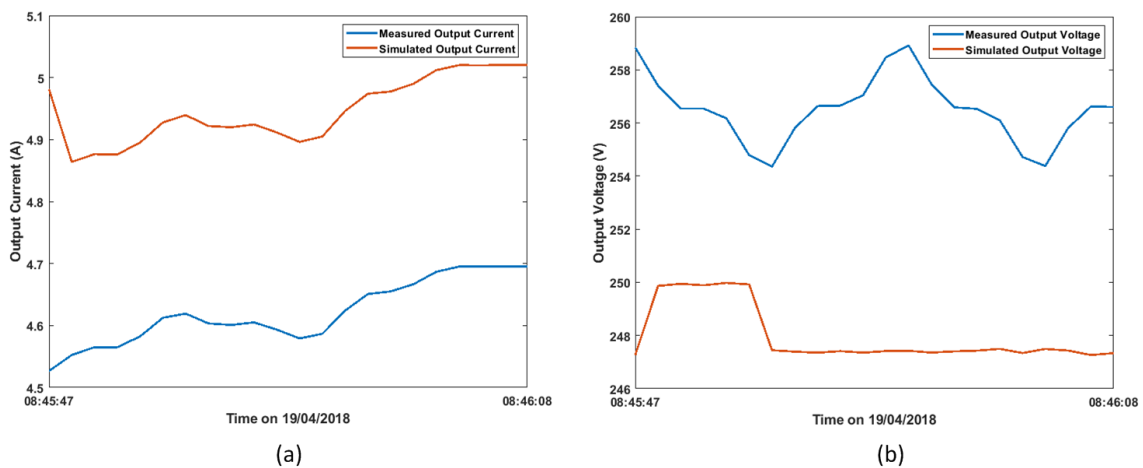
$$I_{out} \approx 0.004925 \cdot G \cdot [1 + 0.004666 \cdot (T_{cell} - 25)] \quad (6.3)$$





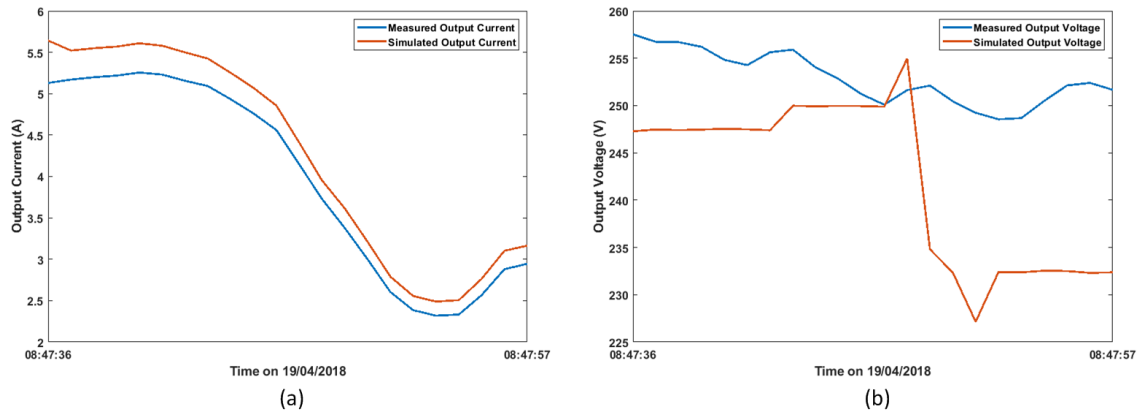
**Figure 6.4 – Linear regression model for the simplified irradiance-current relationship.**

Due to simulation constraints, the model is only verified over the 20 first seconds of the time period. However, results over that period, which spans between 8:45:47 am and 8:46:08 am show an nRMSE of 6.46% for the output current and 3.45% for the output voltage. The graphs comparing measured and simulated output currents and voltages for that period are respectively shown in Figure 6.5 (a) and (b).



**Figure 6.5 – First comparison between the simulated and measured output current (a) and voltage (b).**

Similarly, a shading episode occurred between 8:47:36 and 8:47:57 am, leading to an nRMSE of 6.57% for the output current and 5.42% for the output voltage. The graphs comparing measured and simulated output currents and voltages for that new period are respectively featured in Figure 6.6 (a) and (b).



**Figure 6.6 – Second comparison between the simulated and measured output current (a) and voltage (b).**

It can be seen in both Figure 6.5 and Figure 6.6 that the simulated current follows the exact same trend as the measured current even if an offset is present. Now that this model has been verified in a satisfactory manner, some parameters can be readjusted to analyse the effect of sunlight dimming.

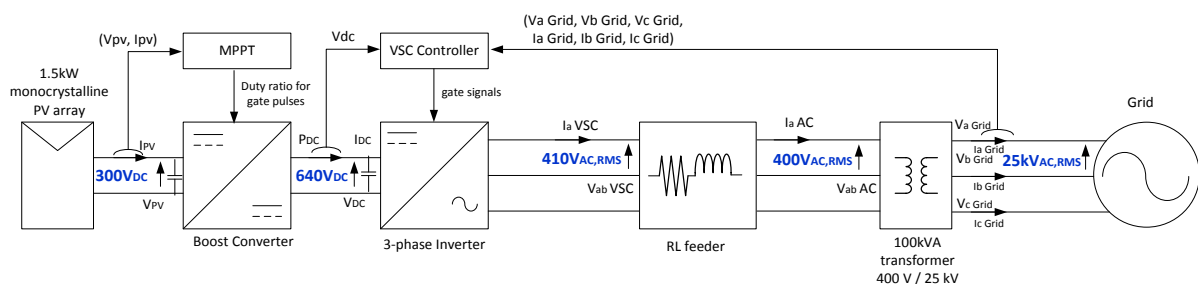
### 6.1.1.2 Adjusted Simulink Model

Some parameters have been adjusted in the SIMULINK model presented in Figure 6.2 as to obtain approximately the voltage levels given in Table 6.1 when referring to the voltages features in the diagram from Figure 6.1.

**Table 6.1 – Approximate target voltage levels for the readjusted Simulink model.**

$V_{PV}$	$V_{DC}$	$V_{VSC}$	$V_{AC}$	$V_{Grid}$
300 V <sub>DC</sub>	640 V <sub>DC</sub>	410 V <sub>AC, RMS</sub>	400 V <sub>AC, RMS</sub>	25,000 V <sub>AC, RMS</sub>

This configuration corresponds to the new diagram shown in Figure 6.7.



**Figure 6.7 – Diagram of the GEEP lab 1.5kW monocrystalline PV system with its target voltage levels.**

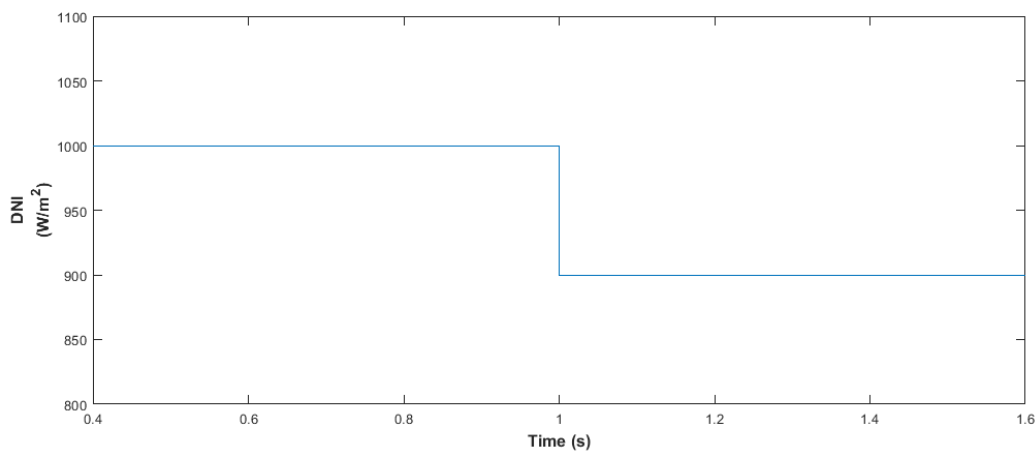
Also, for the sake of simplification, the DNI input fed to the readjusted PV system is initially set at 1000 W/m<sup>2</sup> and the cell temperature is set to a fixed 35°C throughout the whole simulation.

The system may require up to 0.4 seconds to reach a completely steady state after a cold start is performed. To apply an extreme stress to the system, the DNI is abruptly reduced by a given factor after

1 second, which is quite extreme compared with cloud-related shading. The result of the simulation is observed until 1.6 second and an analysis of the power flow through the system is carried out between 0.4 s and 1.6 s. In addition, the signals at every link of the system that has been annotated on the diagram in Figure 6.7 is analysed over the same observation period, for each sunlight reduction factor.

#### 6.1.1.2.1 Sunlight Reduction by 10%

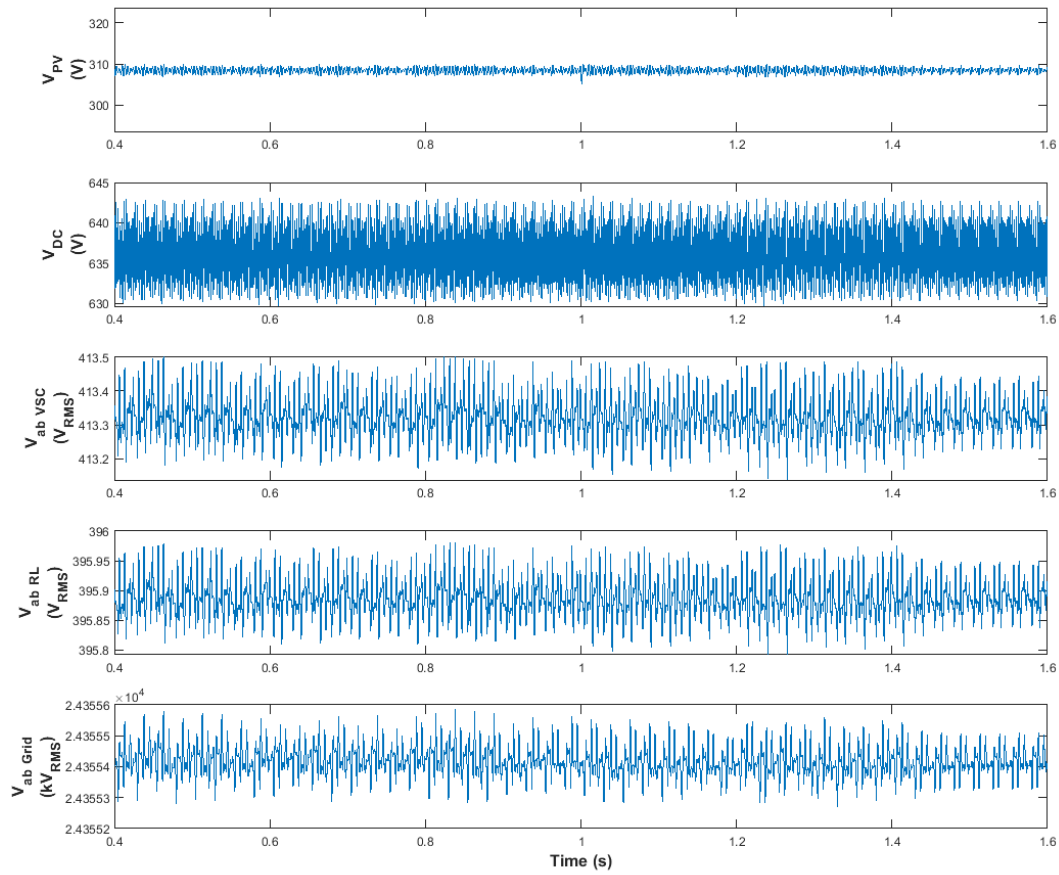
With a sunlight reduction factor of 10%, the evolution of DNI, the voltage at every link, the reactive power at the AC links, the current at all links, the power factor at the AC links and the active power at all links are respectively shown from Figure 6.8 to Figure 6.13.



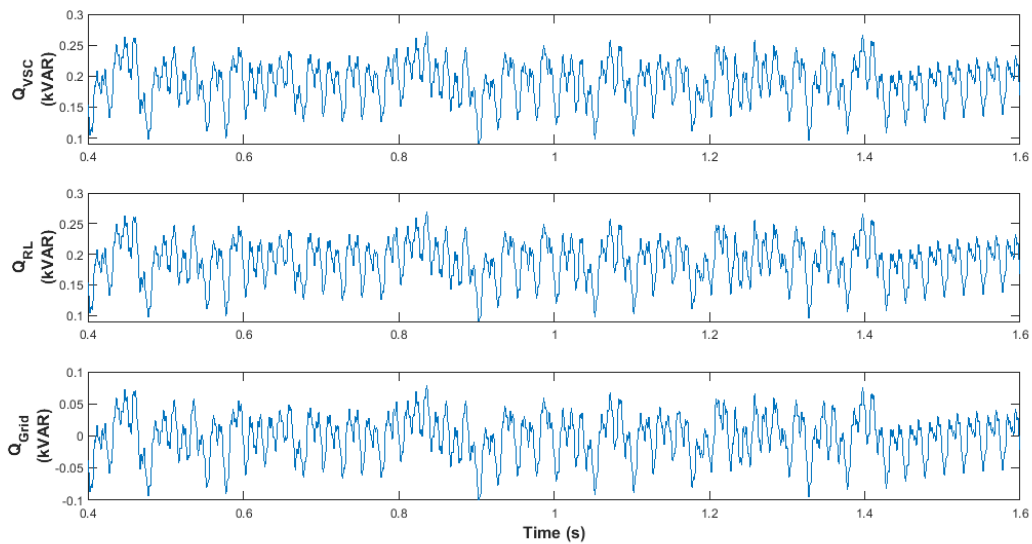
**Figure 6.8 – Evolution of the DNI over time when the sunlight dimming factor is 10%.**

It can be seen in Figure 6.9 that because of the MPPT, the output voltage of the PV array is barely affected by such a relatively small variation of DNI. It is actually the current output of the array that follows almost all the variation of power, as shown in Figure 6.11. This effect is propagated through the further links and the DAC link voltage as well as the line to line voltages at the VSC output, RL feeder and grid connection point are quite consistent, while the change of current at the same links reflect the change in power and corresponds to a 10% drop.

Further, as it can be seen in Figure 6.13, although the active power gets reduced by 10% under the effect of the dimmed sunlight, the power factor remains quite steady as shown in Figure 6.12.



**Figure 6.9 – Evolution of the voltage at every link over time when the sunlight dimming factor is 10%.**



**Figure 6.10 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 10%.**

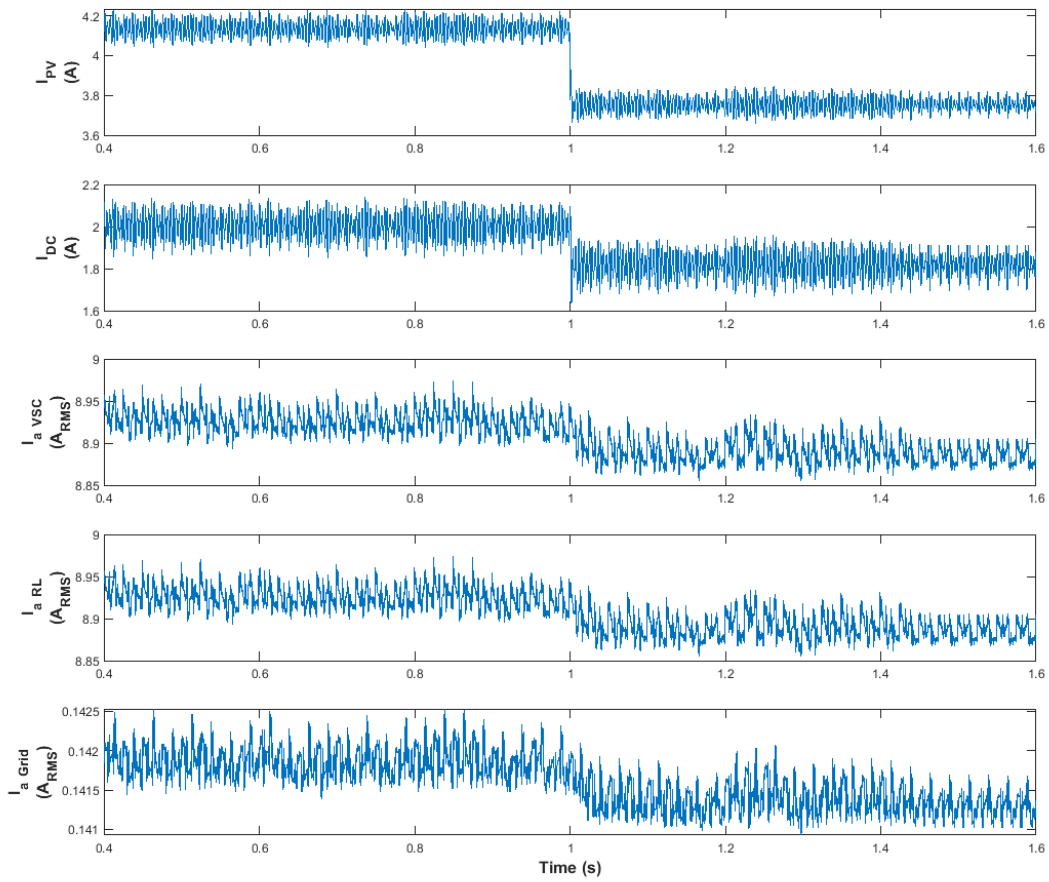


Figure 6.11 – Evolution of the current at every link over time when the sunlight dimming factor is 10%.

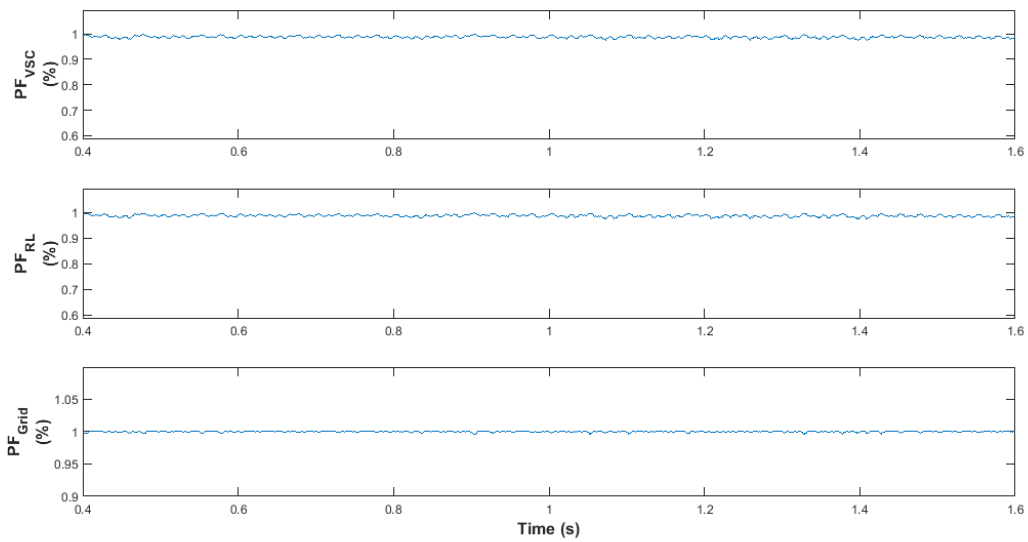
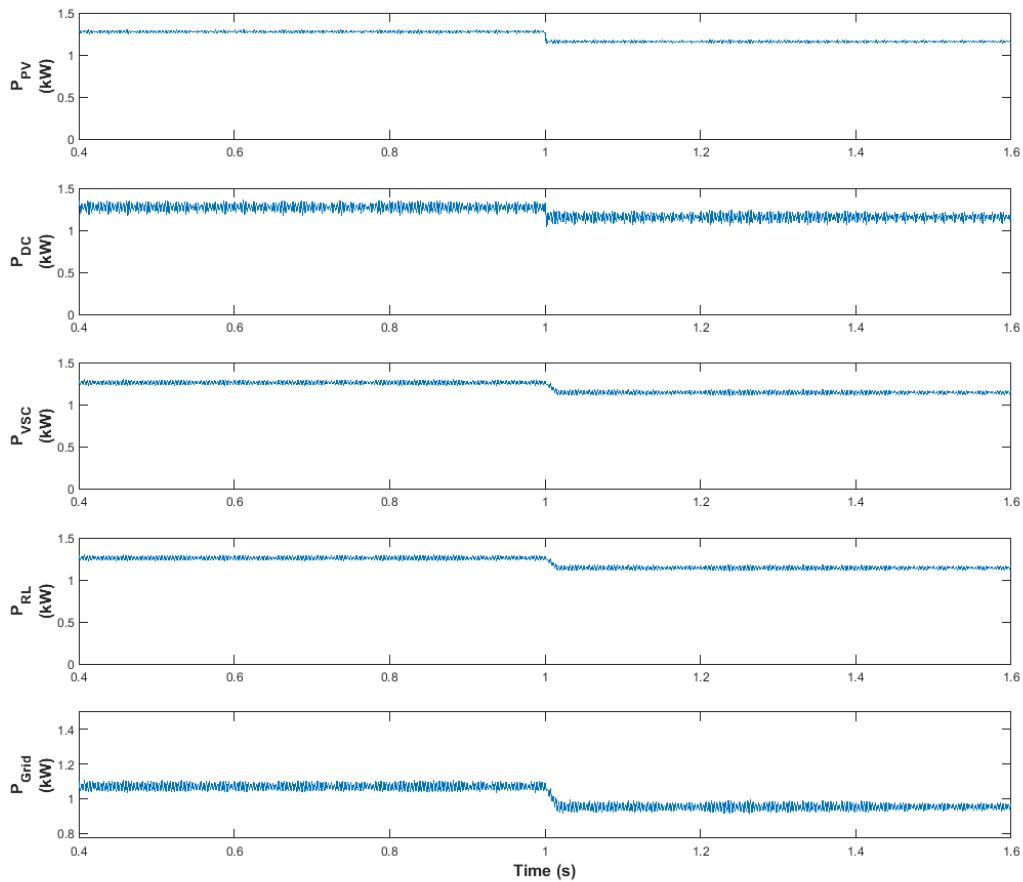


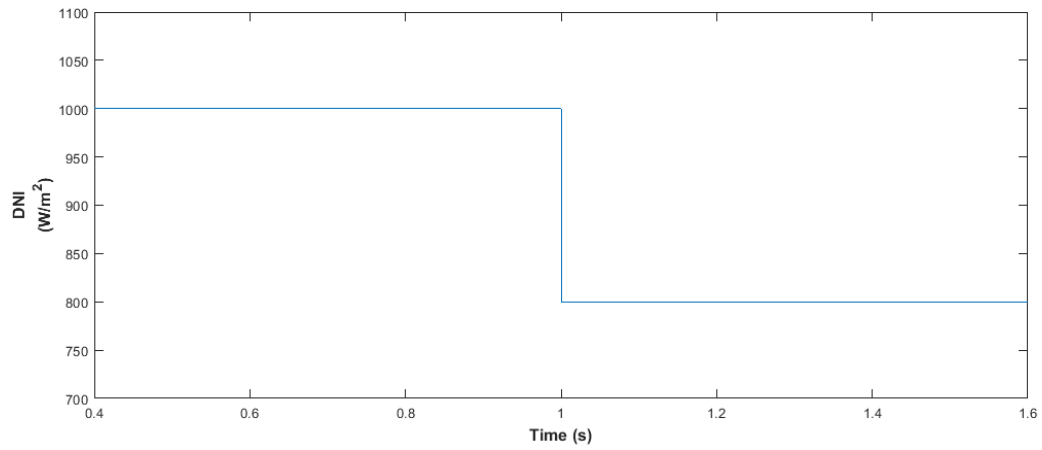
Figure 6.12 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 10%.



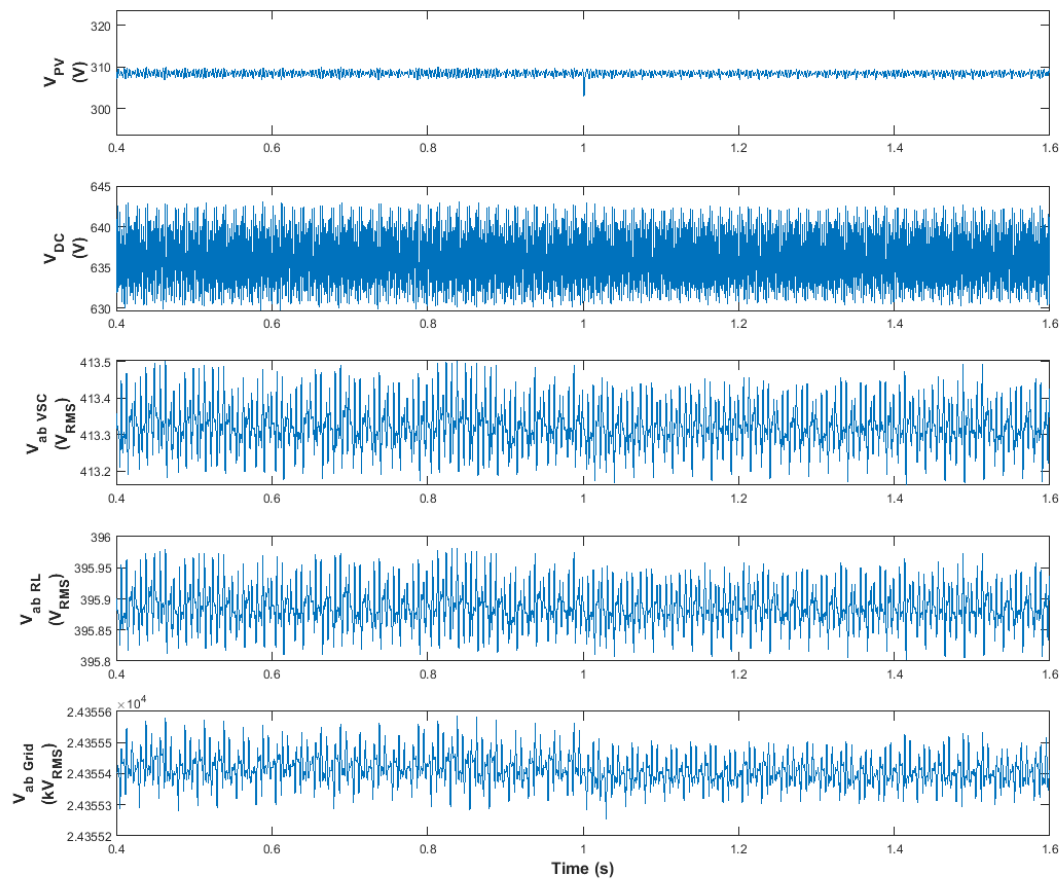
**Figure 6.13 – Evolution of the active power at every link over time when the sunlight dimming factor is 10%.**

#### 6.1.1.2.2 Sunlight Reduction by 20%

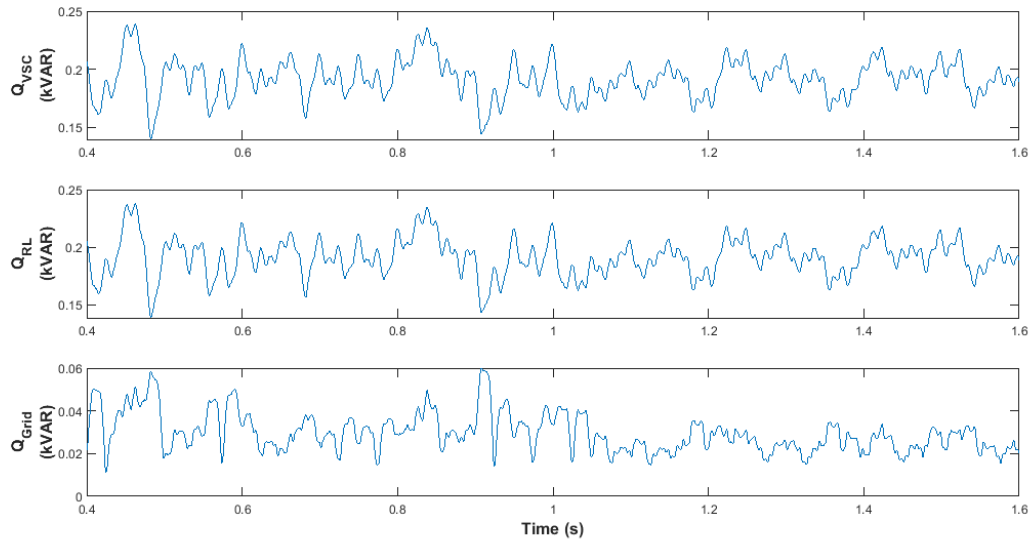
It could be seen that dimming the DNI by 10% seemingly did not have any visible effect on the link voltages of the system and only affected the currents, with virtually no effect on the power factor. Let us increase that dimming effect to a reduction of 20% of the original value of the DNI to see how doubling the reduction affects more the system than just 10% did. In those conditions, the change of DNI, the voltage at every link, the reactive power at the AC links, the current at all links, the power factor at the AC links and the active power at all links are respectively shown from Figure 6.14 to Figure 6.19. While the PV array output voltage remains unchanged, the PV output current and output power are cut by 19%, which makes in turn the power that is transmitted to the grid drop by 22%, as seen from Figure 6.17 to Figure 6.19.



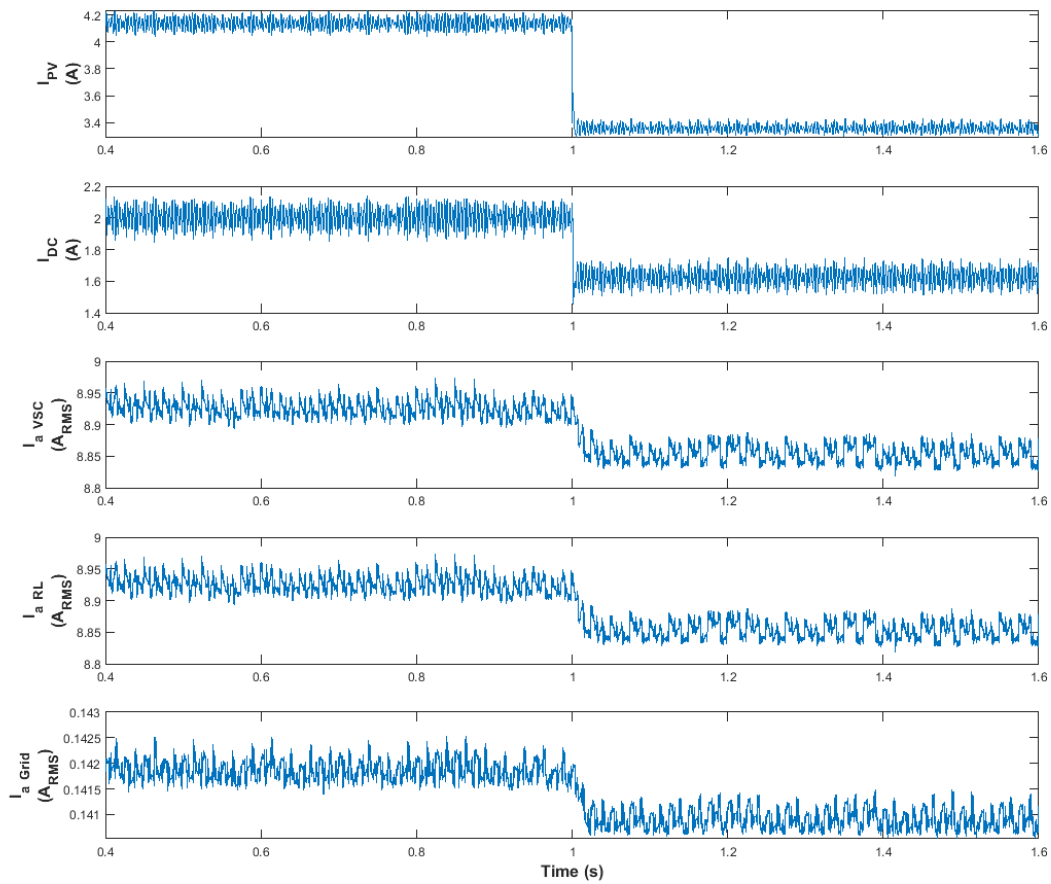
**Figure 6.14 – Evolution of the DNI over time when the sunlight dimming factor is 20%.**



**Figure 6.15 – Evolution of the voltage at every link over time when the sunlight dimming factor is 20%.**

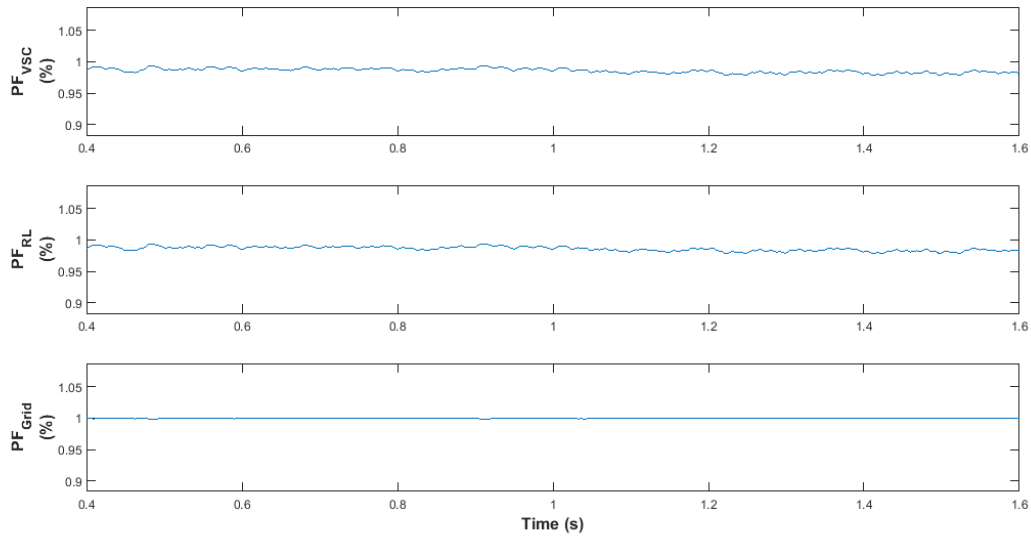


**Figure 6.16 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 20%.**

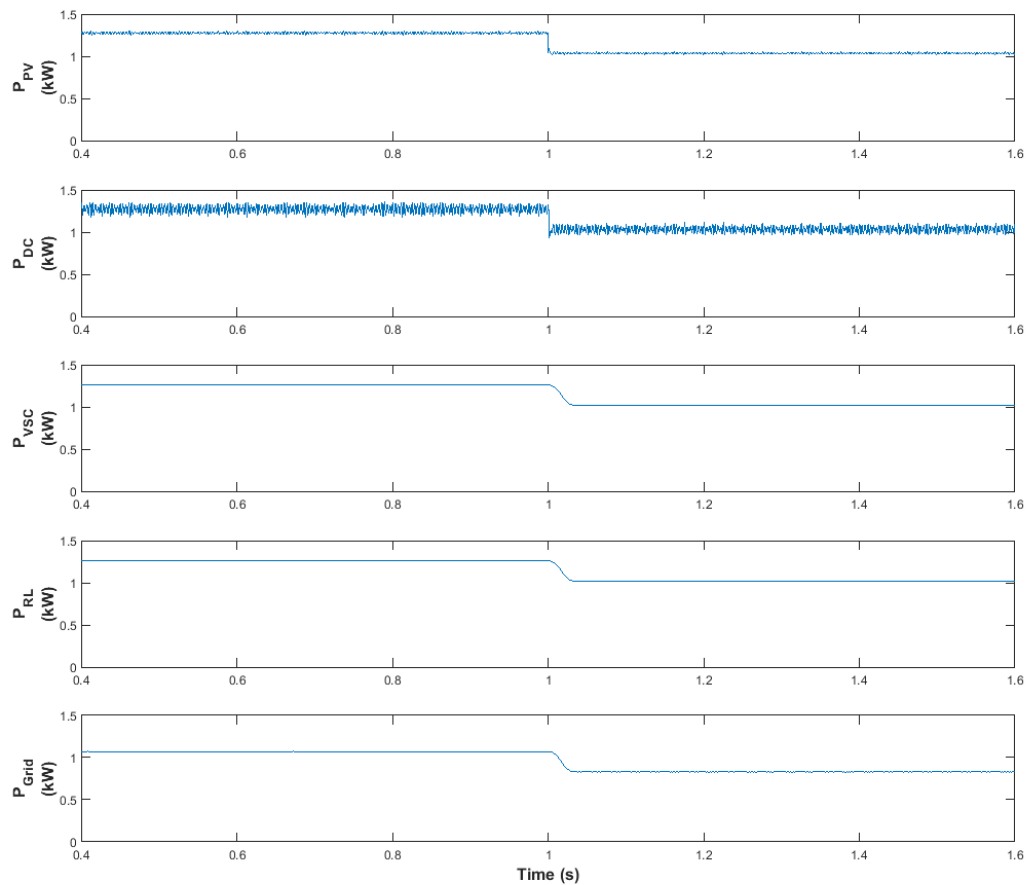


**Figure 6.17 – Evolution of the current at every link over time when the sunlight dimming factor is 20%.**





**Figure 6.18 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 20%.**

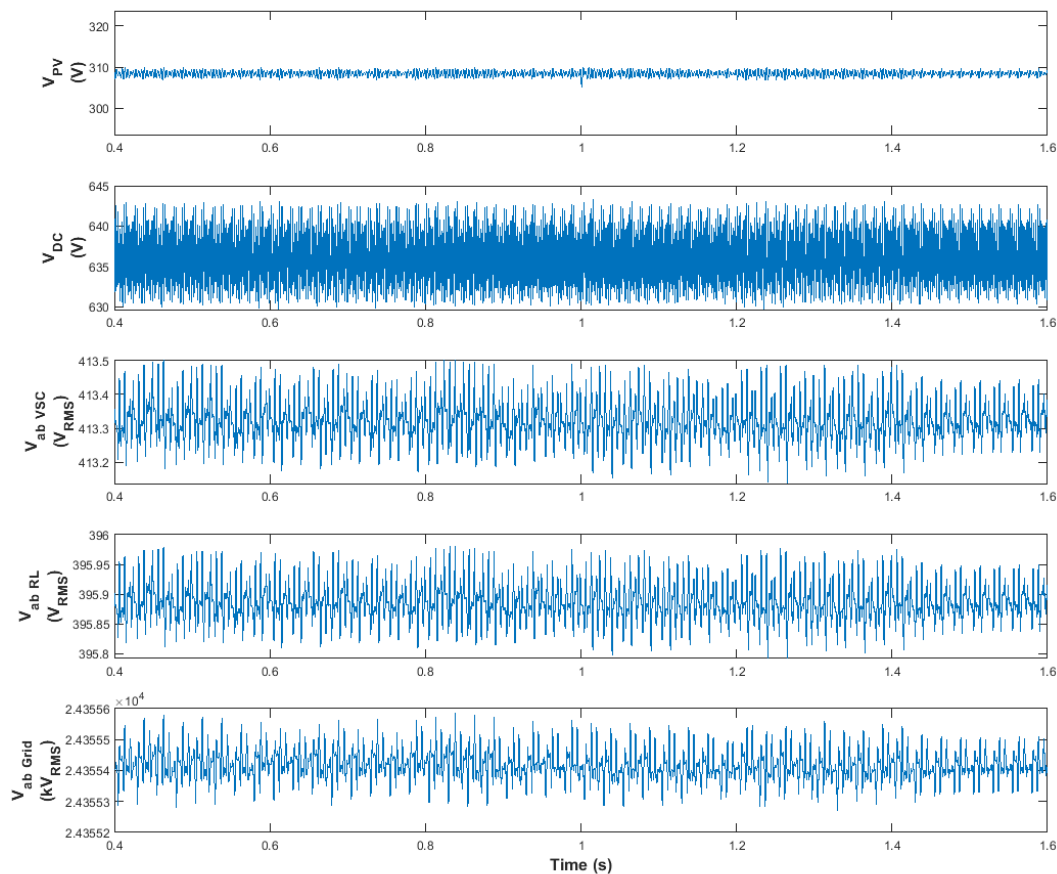


**Figure 6.19 – Evolution of the active power at every link over time when the sunlight dimming factor is 20%.**

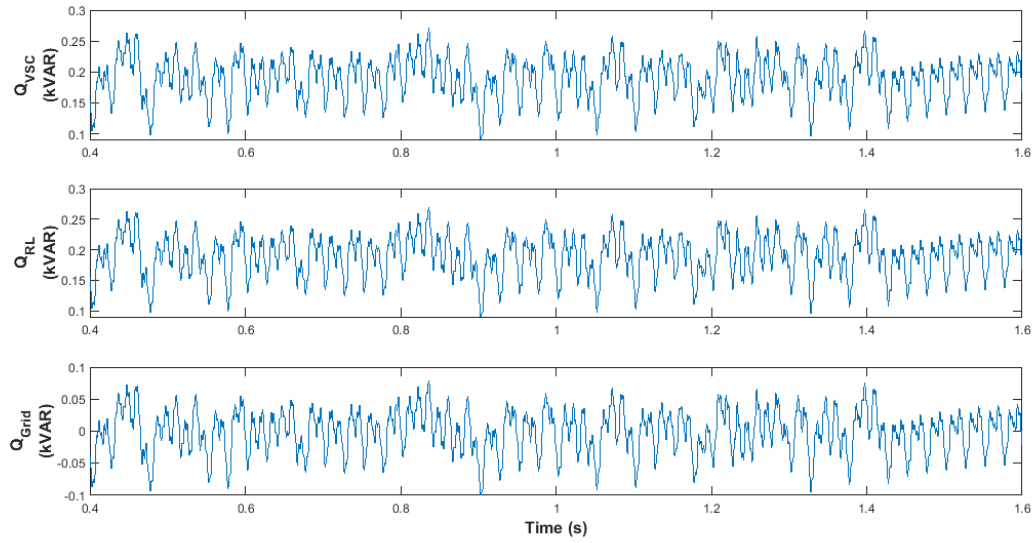
The exact same observations apply as for the previous dimming factor.

### 6.1.1.2.3 Sunlight Reduction by 30%

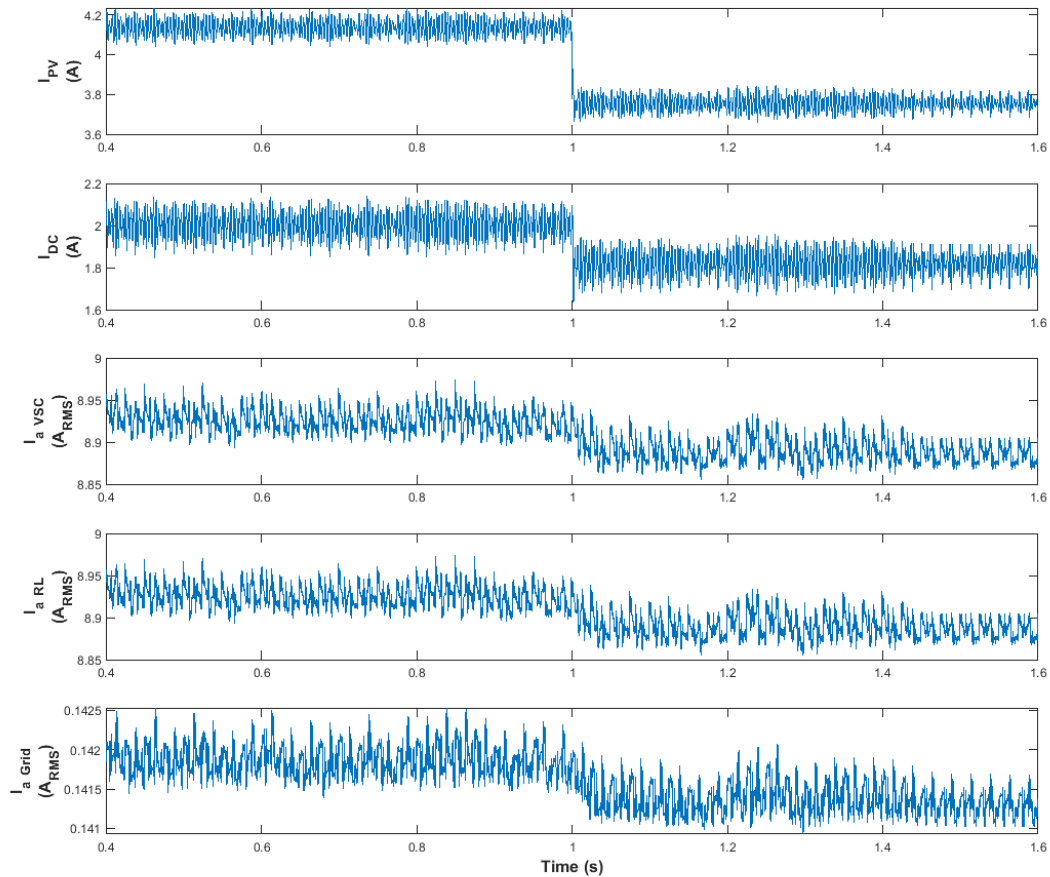
Since the system reacted very similarly for a DNI dimmed by 20% as it did when it was dimmed by 10%, the dimming factor is this time increased to 30%. In that situation, with the DNI dimmed after 1 second, the voltage at every link, the reactive power at the AC links, the current at all links, the power factor at the AC links and the active power at all links are respectively given from Figure 6.20 to Figure 6.24. This time, the decrease in DNI has a noticeable effect on the output voltage of the PV array, although it only drops by 1.3%. However, the VSC controller strongly mitigates this effect and manages to maintain the DC link voltage despite of the variations of the PV voltage. The currents and hence the active powers are more dramatically affected by the change of DNI. For instance, the PV output current and output power have been both cut by 25% as seen in Figure 6.22 and Figure 6.24. Consequently, the power transmitted to the grid has dropped by 31%. For its part, however, the power factor changes only very slightly, up to 0.2% on the output of VSC and through the feeder.



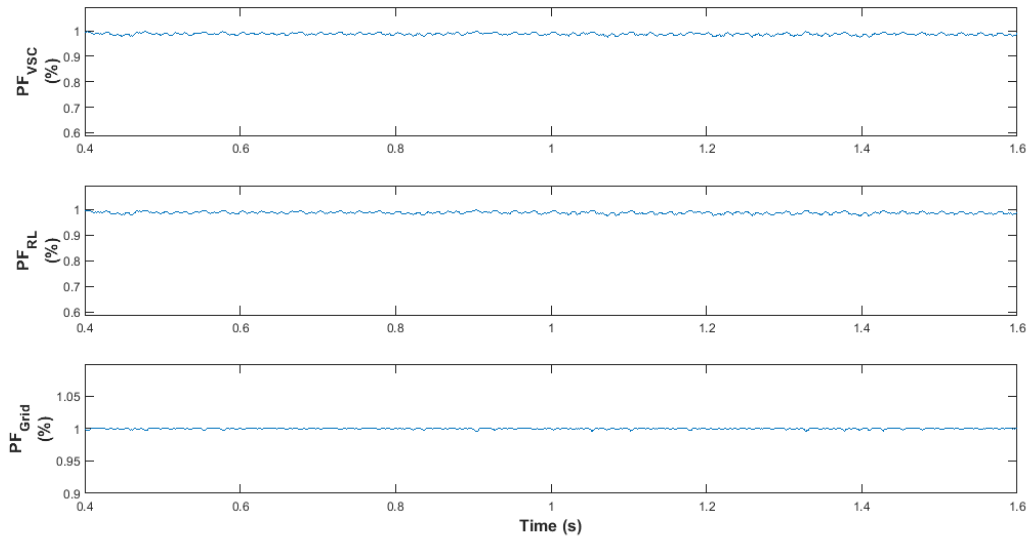
**Figure 6.20 – Evolution of the voltage at every link over time when the sunlight dimming factor is 30%.**



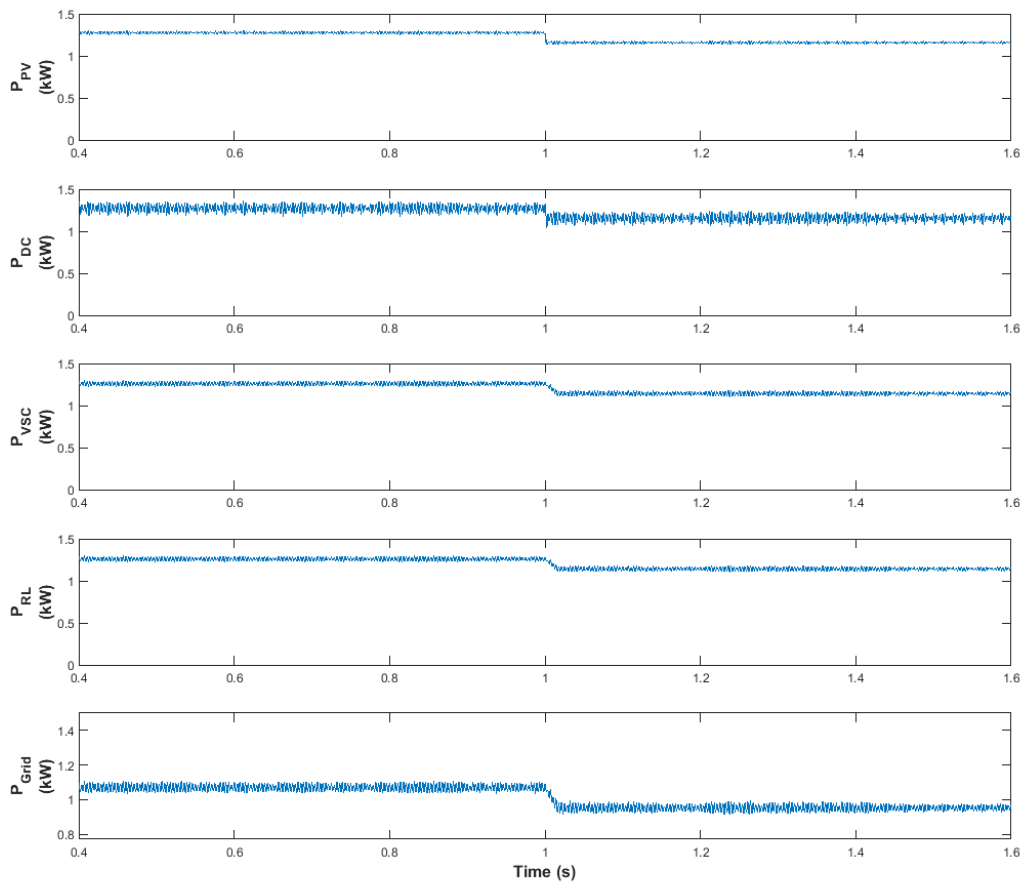
**Figure 6.21 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 30%.**



**Figure 6.22 – Evolution of the current at every link over time when the sunlight dimming factor is 30%.**



**Figure 6.23 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 30%.**



**Figure 6.24 – Evolution of the active power at every link over time when the sunlight dimming factor is 30%.**

Once more, the sunlight dimming factor will be increased to further study its effect on the system.

#### 6.1.1.2.4 Sunlight Reduction by 40%

The output voltage of PV array just started to change when the DNI dropped to a value located between 20% and 30% of the initial value. In order to observe a more dramatic effect on the system, the sunlight dimming factor has been further increased to 40%. With the DNI being decreased to 600 W/m<sup>2</sup> after 1 second of simulation, the voltage at every link, the reactive power at the AC links, the current at all links, the power factor at the AC links and the active power at all links are respectively given from Figure 6.25 to Figure 6.29. It can already be seen in Figure 6.25 that the drop of irradiance by 40% is accompanied by an immediate decrease of the PV array output voltage by about 5%. As it can be seen in Figure 6.27, the change in the output current of the PV array is still quite dramatic and corresponds to a 29% drop. It results from all of this that the PV array output active power is decreased by 32% while the amount of active power transmitted to the grid is reduced by 38.5%, as seen in Figure 6.29.

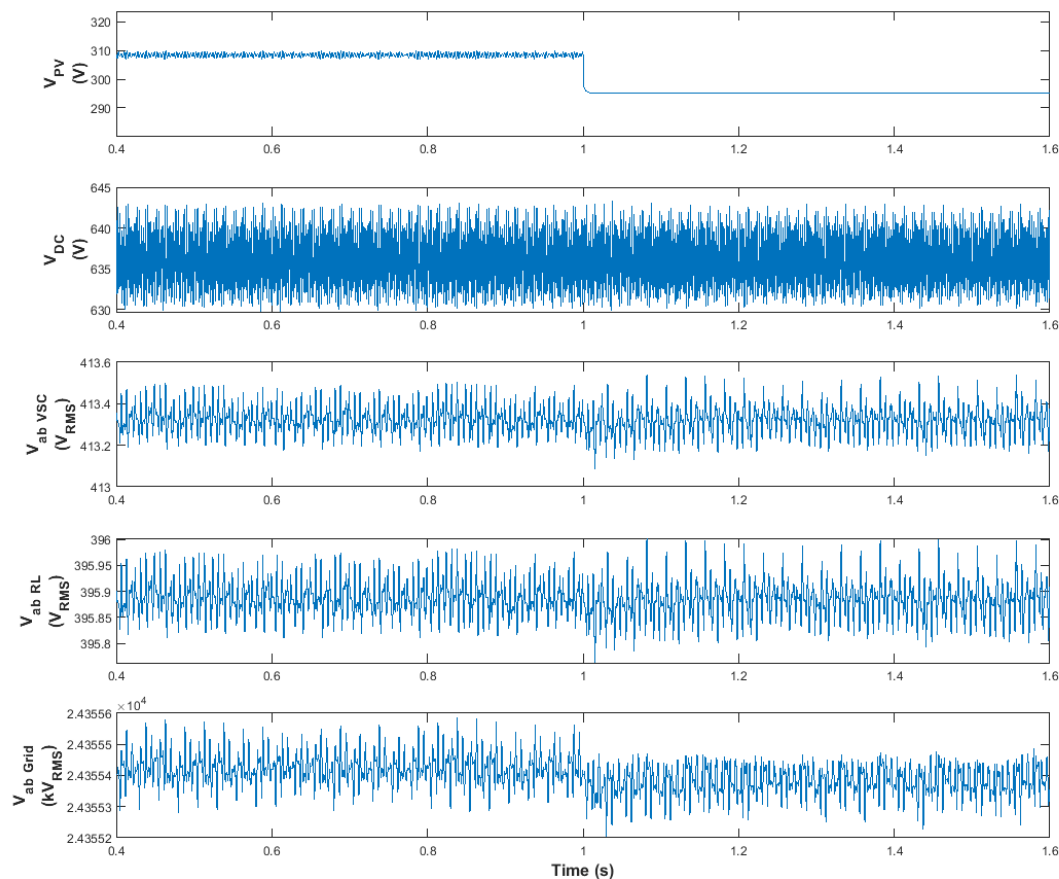


Figure 6.25 – Evolution of the voltage at every link over time when the sunlight dimming factor is 40%.

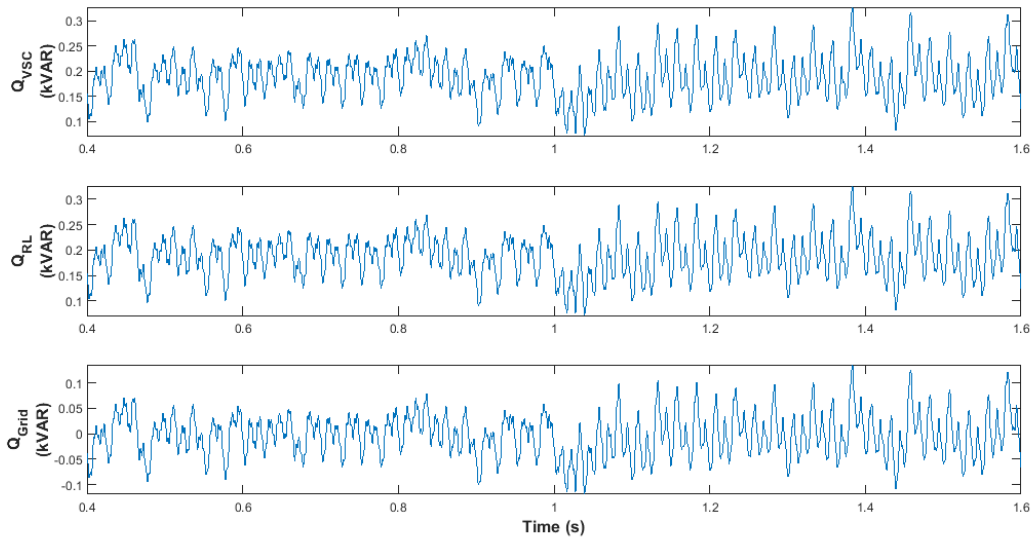


Figure 6.26 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 40%.

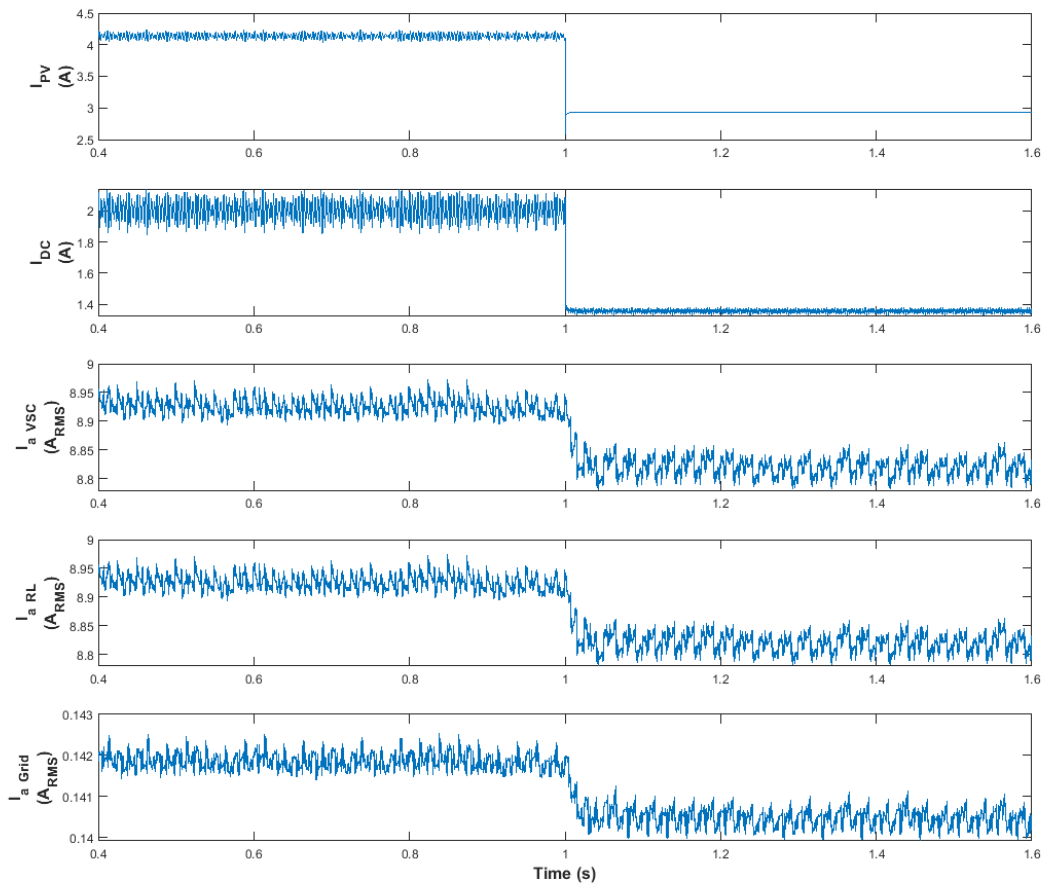
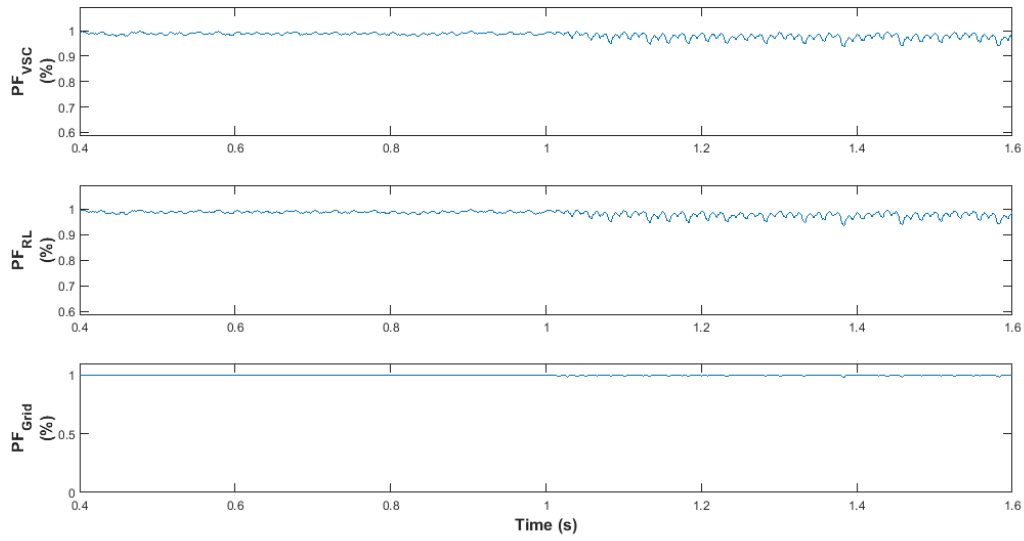
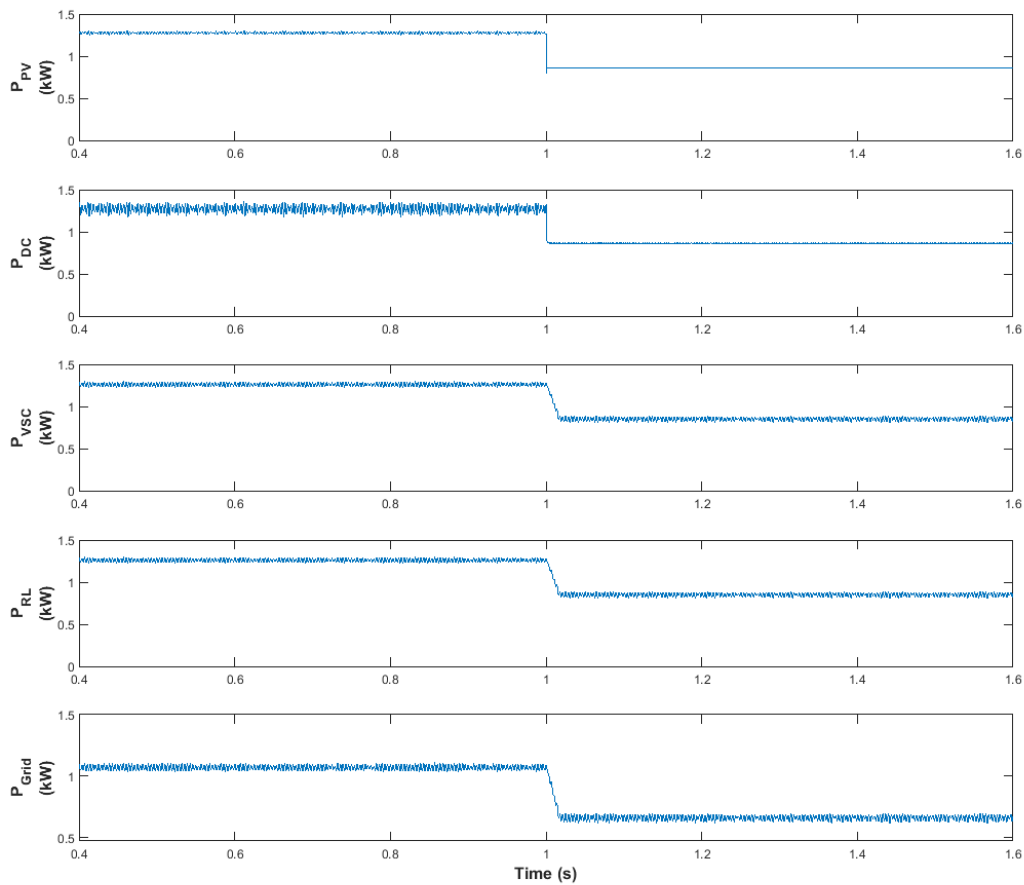


Figure 6.27 – Evolution of the current at every link over time when the sunlight dimming factor is 40%.



**Figure 6.28 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 40%.**

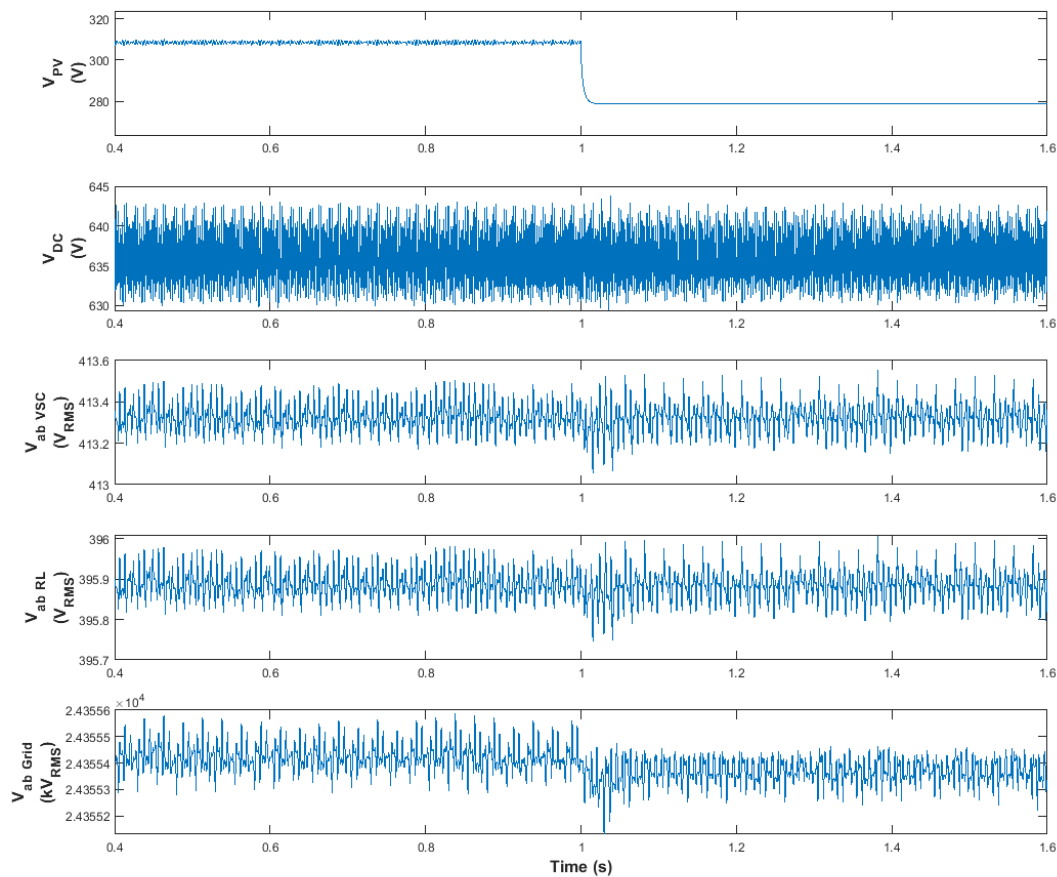


**Figure 6.29 – Evolution of the active power at every link over time when the sunlight dimming factor is 40%.**

The effect of dimming the sunlight even further is studied in the next cases.

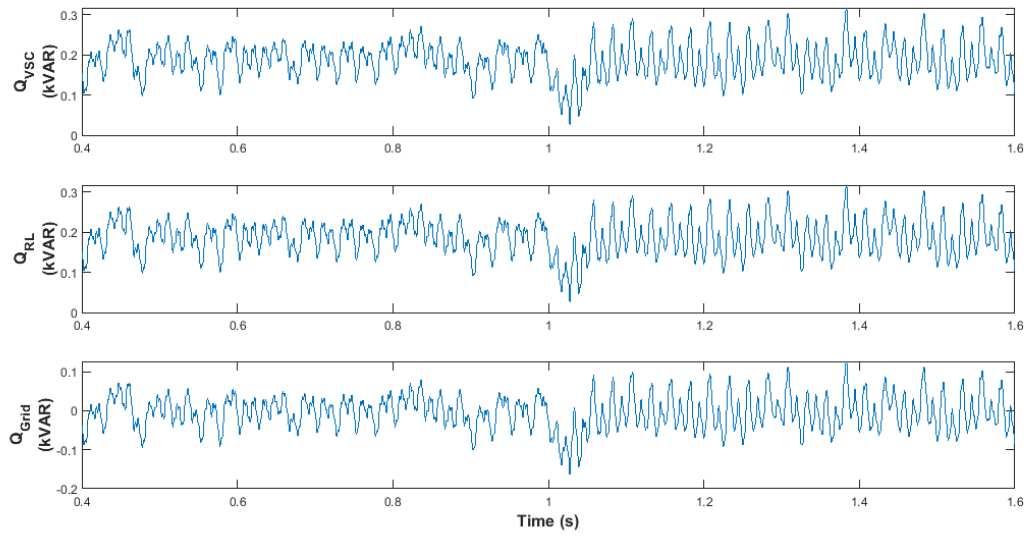
### 6.1.1.2.5 Sunlight Reduction by 50%

With the DNI decreased to  $600 \text{ W/m}^2$  or 40% of its original value, a drop was clearly visible concerning the output voltage of the PV array. However, this decrease was still relatively small and it can be anticipated that the output voltage of the PV array may further decrease as the DNI drops lower. For that purpose, this time, the DNI have been decreased by half after 1 second, which  $500 \text{ W/m}^2$ . This leads to the changes in terms of voltages, the reactive power, currents, power factor and active power as respectively shown from Figure 6.30 to Figure 6.34. As shown in Figure 6.30, the drop in terms of output voltage has doubled compared with the previous case and now represents about 10%. In comparison, the output current of the PV array varies by 36% in the same conditions, as seen in Figure 6.32. As shown in Figure 6.33, the power factor may fluctuate slightly but actually barely changes at all since it only losses up to 1.2% of its value at the point of interconnection between the system and the grid.

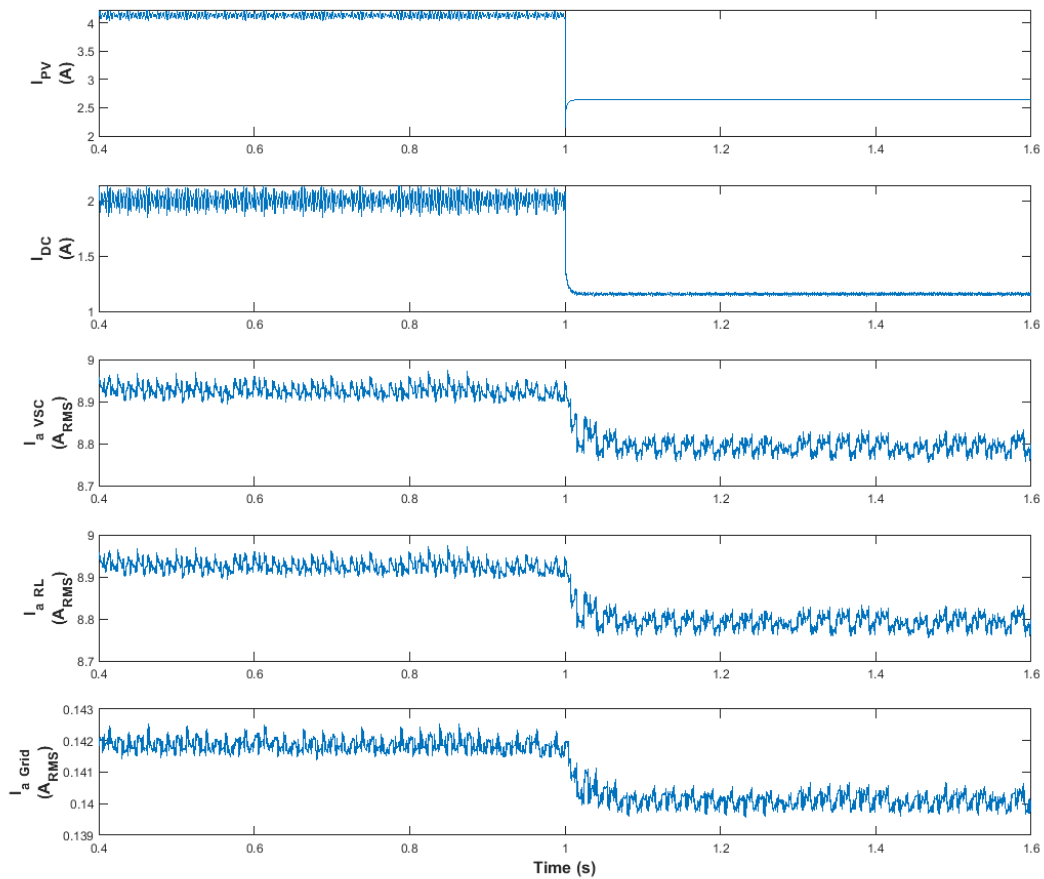


**Figure 6.30 – Evolution of the voltage at every link over time when the sunlight dimming factor is 50%.**

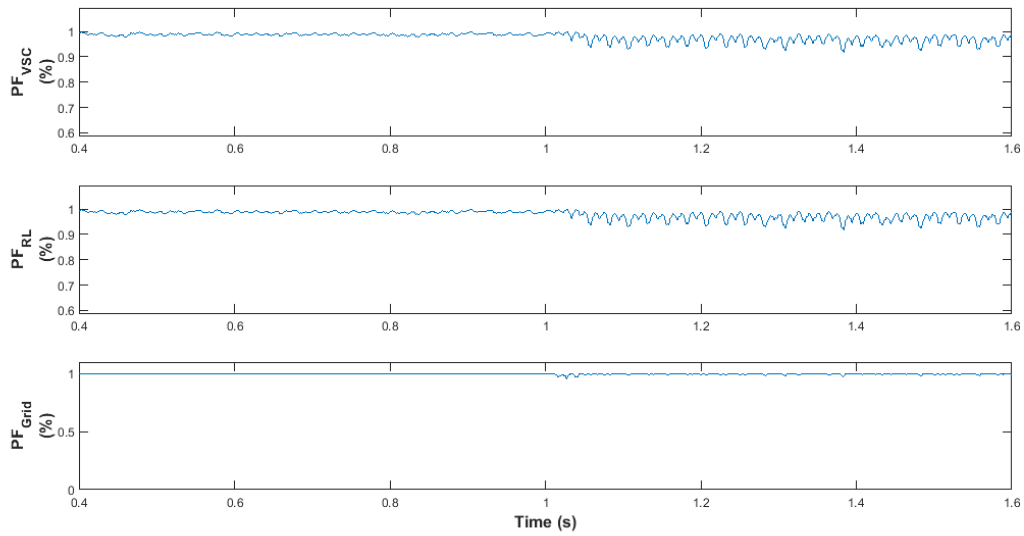




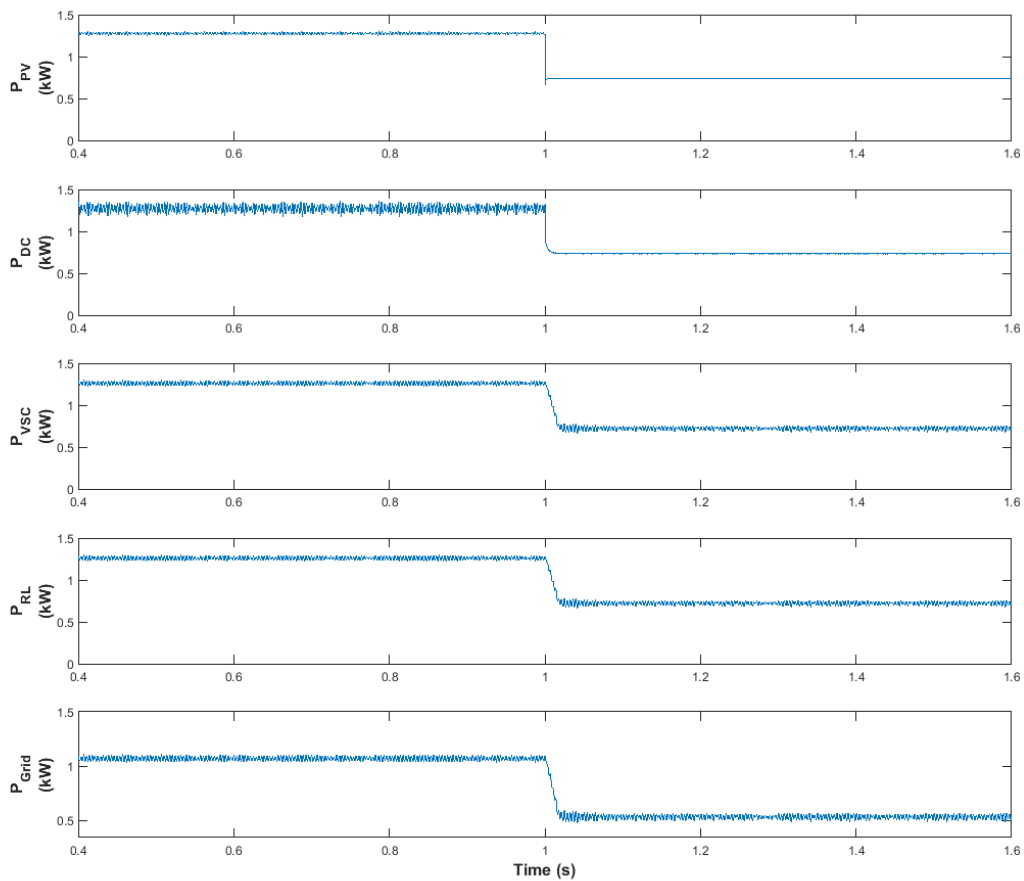
**Figure 6.31 – Evolution of the reactive power at the AC Links over time when the sunlight dimming factor is 50%.**



**Figure 6.32 – Evolution of the current at every link over time when the sunlight dimming factor is 50%.**



**Figure 6.33 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 50%.**



**Figure 6.34 – Evolution of the active power at every link over time when the sunlight dimming factor is 50%.**

Again, the effect of dimming the sunlight will be further studied in the following sections.

### 6.1.1.2.6 Sunlight Reduction by 60%

So the effect of decreased DNI can be further studied, the DNI is decreased to 60% of its original value, which is  $400 \text{ W/m}^2$  after 1 second. The resulting changes in terms of voltages, the reactive power, currents, power factor and active power as respectively shown from Figure 6.35 to Figure 6.39. As shown in Figure 6.35, the output voltage of the PV array drops by 19%. Conversely, the output current of the PV array is suddenly decreased by 50% in the same conditions, as seen in Figure 6.37. As a result, the output power of the array decreases by 57% as seen in Figure 6.39 while the total power injected to the grid drops by about 66%. As shown in Figure 6.38, the power factor decreases but remains high as it only loses up to 3% at the point of common connection between the PV system and the grid.

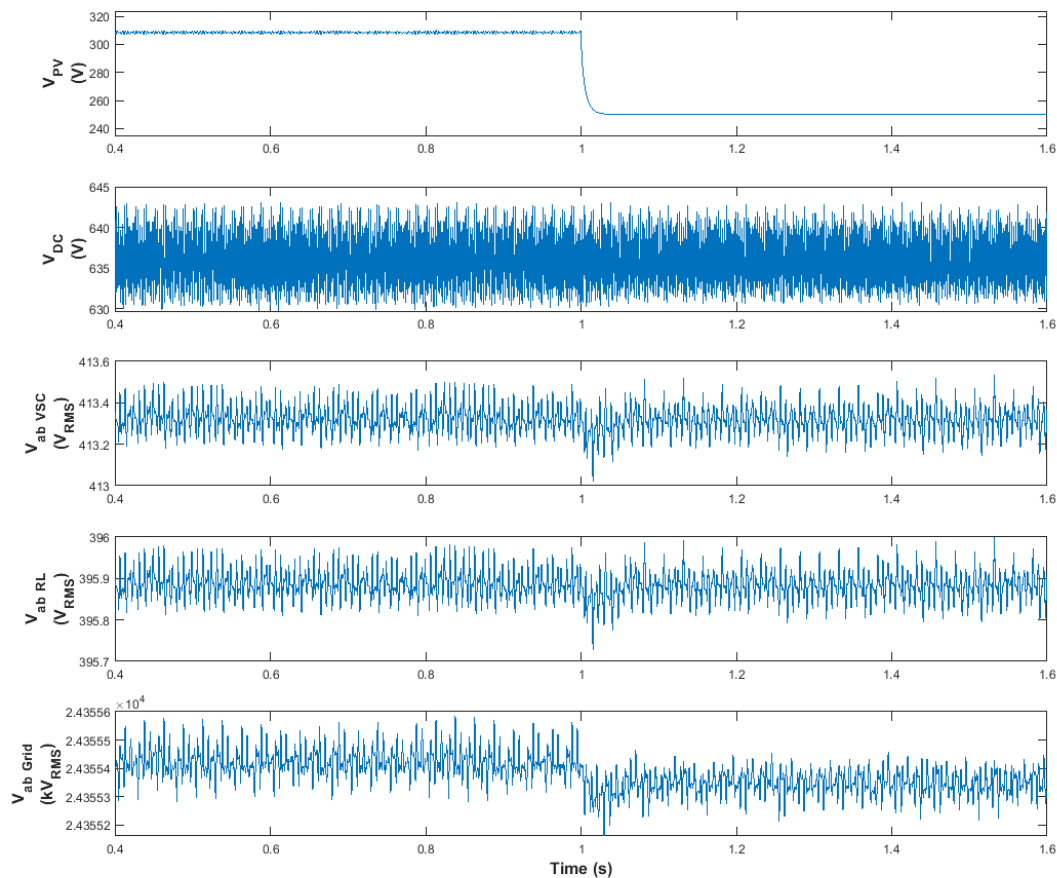
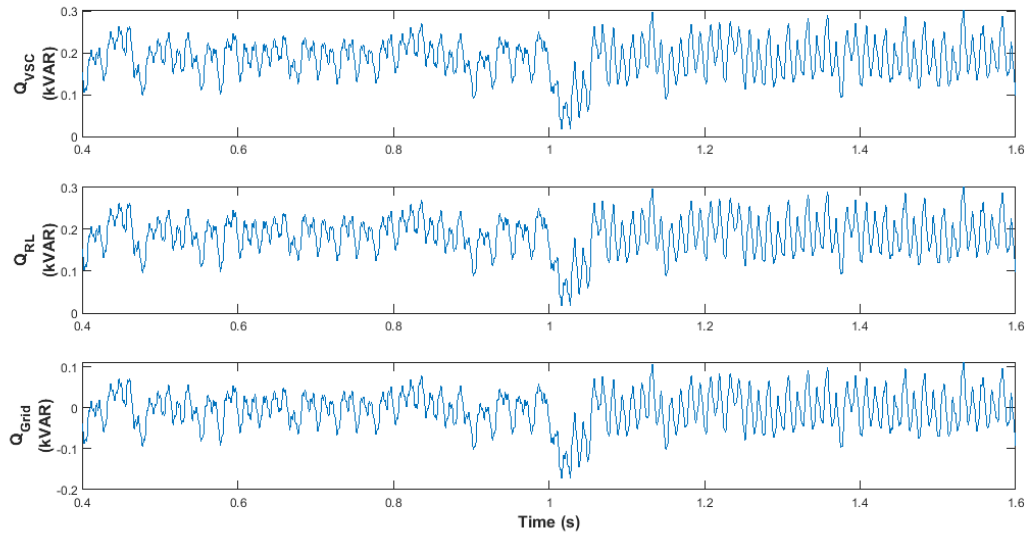
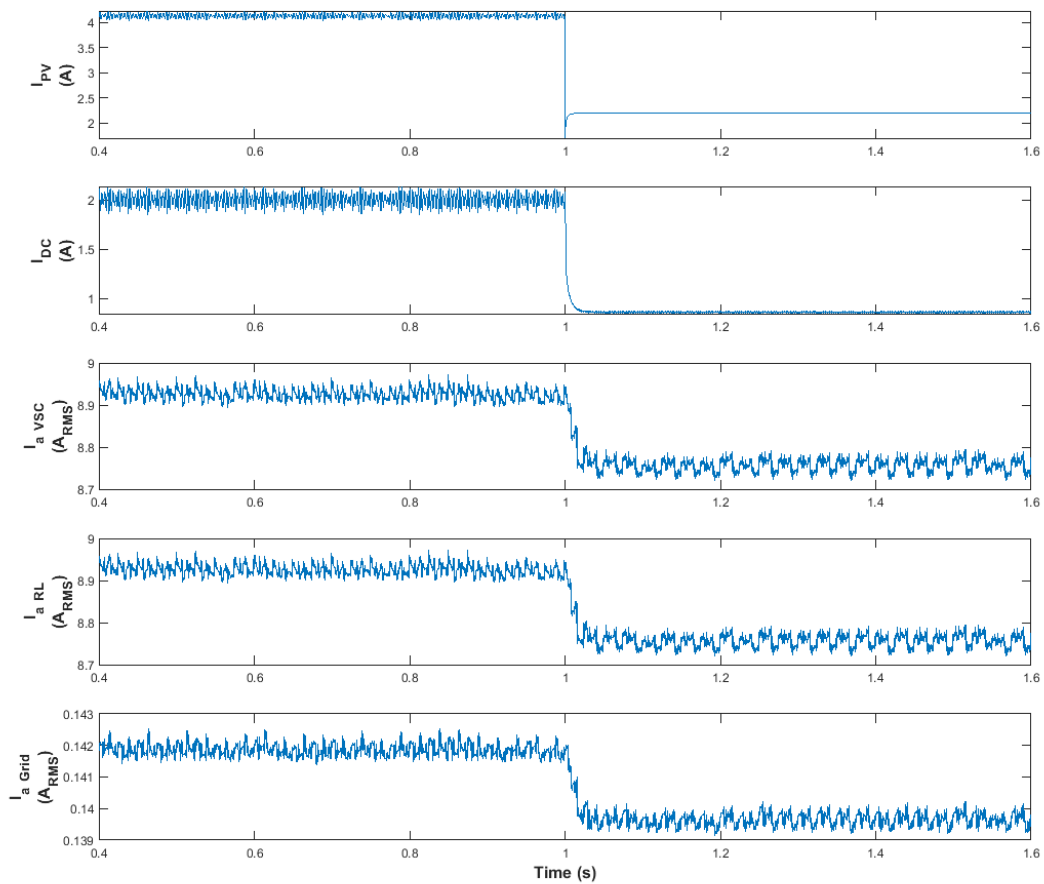


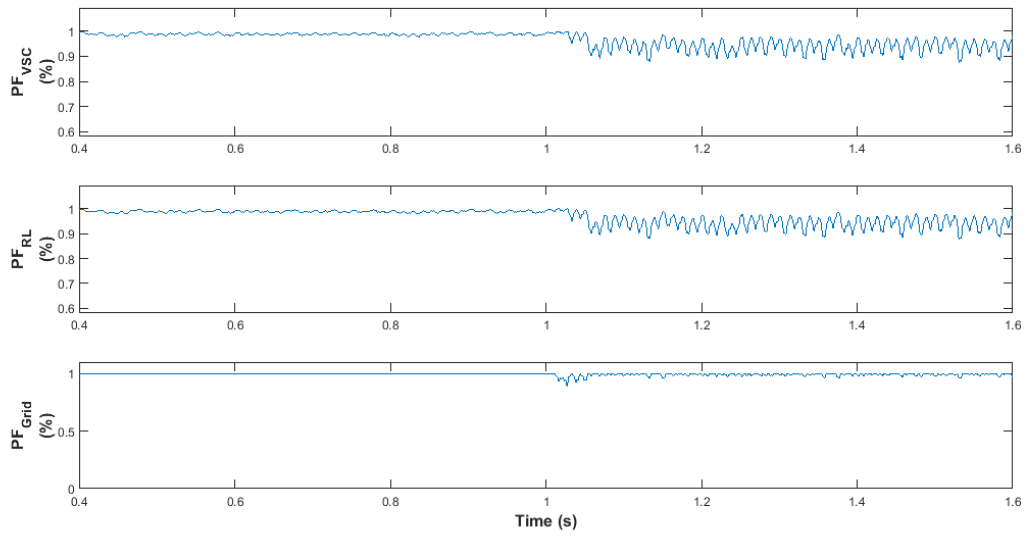
Figure 6.35 – Evolution of the voltage at every link over time when the sunlight dimming factor is 60%.



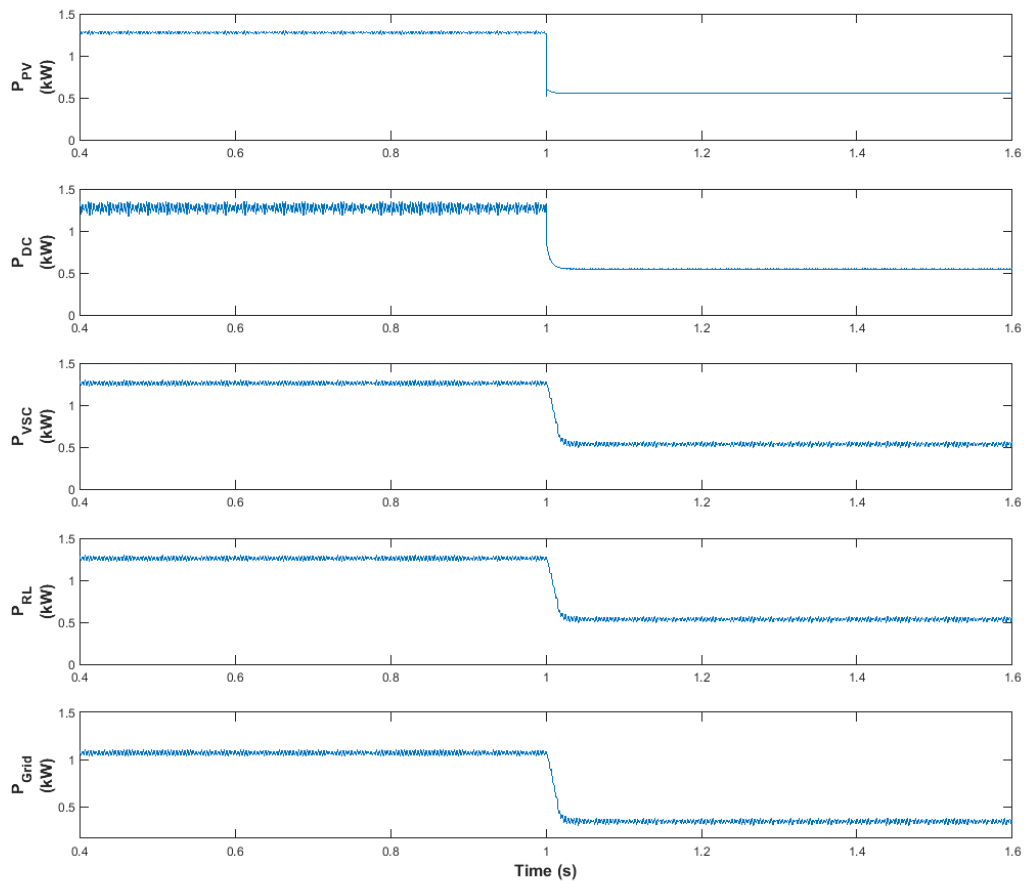
**Figure 6.36 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 60%.**



**Figure 6.37 – Evolution of the current at every link over time when the sunlight dimming factor is 60%.**



**Figure 6.38 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 60%.**



**Figure 6.39 – Evolution of the active power at every link over time when the sunlight dimming factor is 60%.**

The dimming factor could be even further increased to study its effects.

### 6.1.1.2.7 Sunlight Reduction by 70%

The DNI is now decreased to only 70% of its original value, which is  $300 \text{ W/m}^2$  after 1 second in order to see how the PV system reacts to this change. The signal of the system, such as the voltages, reactive powers, currents, power factors and active powers are respectively shown from Figure 6.40 to Figure 6.44. As shown in Figure 6.40, the output voltage of the PV array is decreased by 32%. At the same time, the output current of the PV array suddenly drops by 60% after 1 second, as it can be seen in Figure 6.42. Consequently, the output power of the array decreases by 72% while the total power injected to the grid drops by nearly 85% as shown in Figure 6.44. The power factor starts to decrease more significantly since it can drop by up to 7% as seen in Figure 6.43.

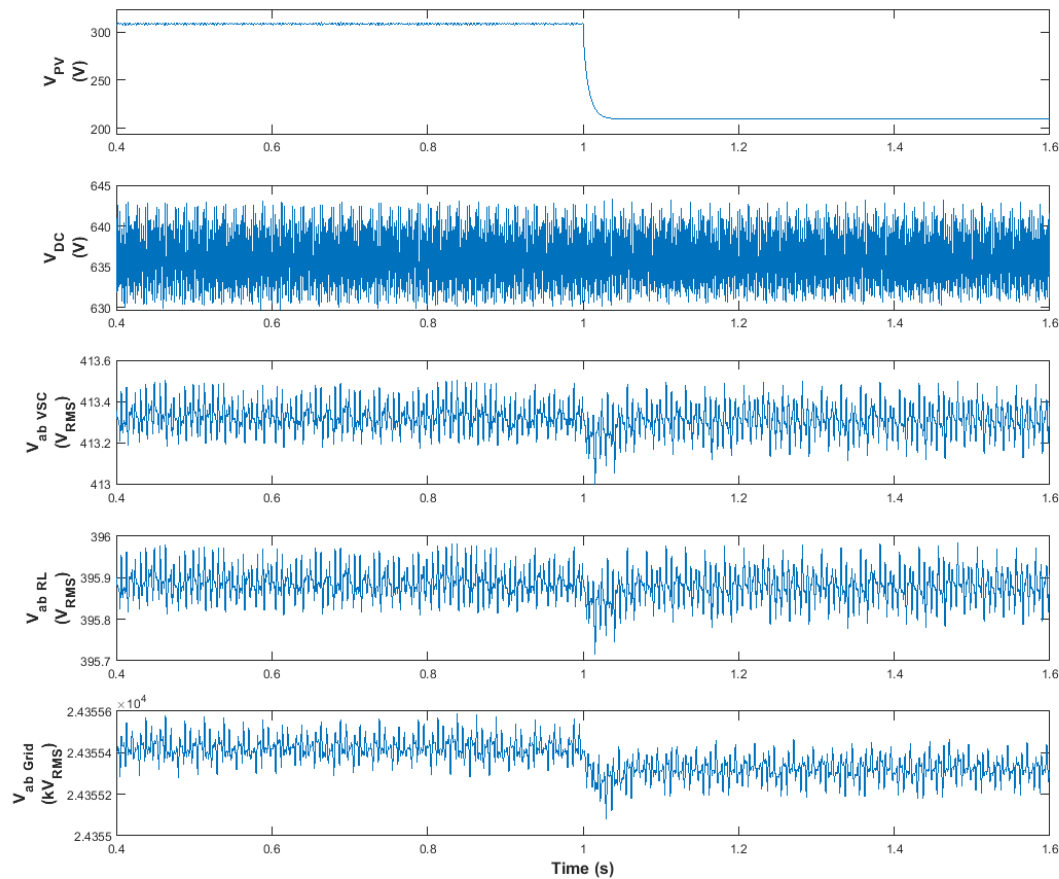
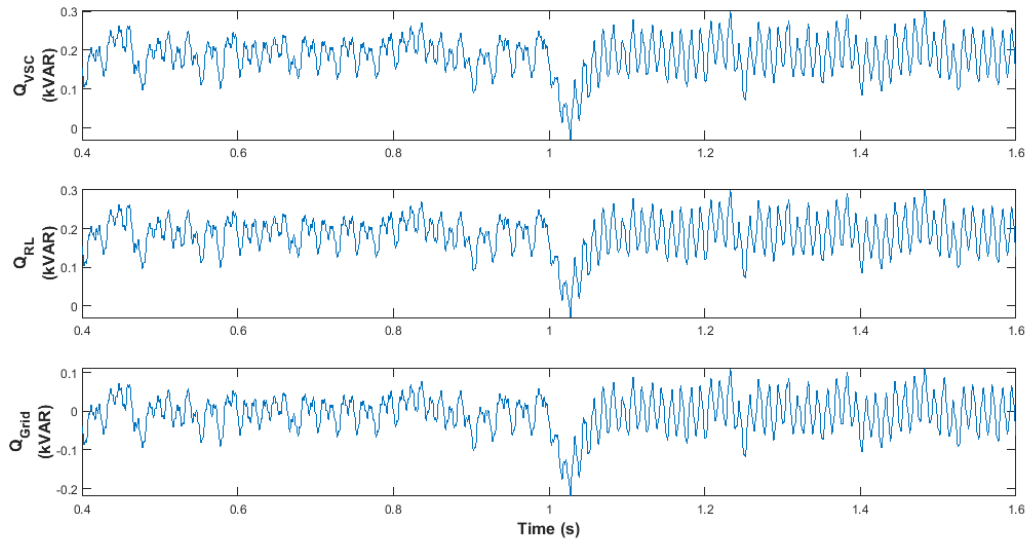
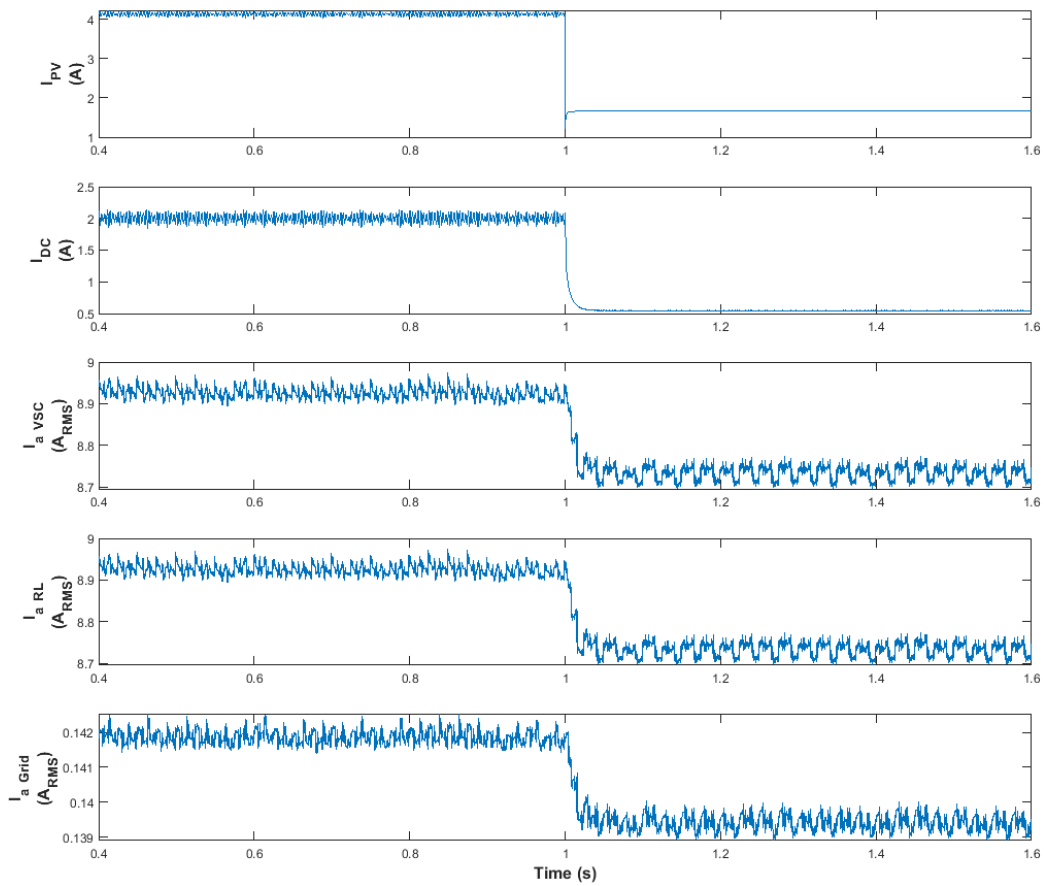


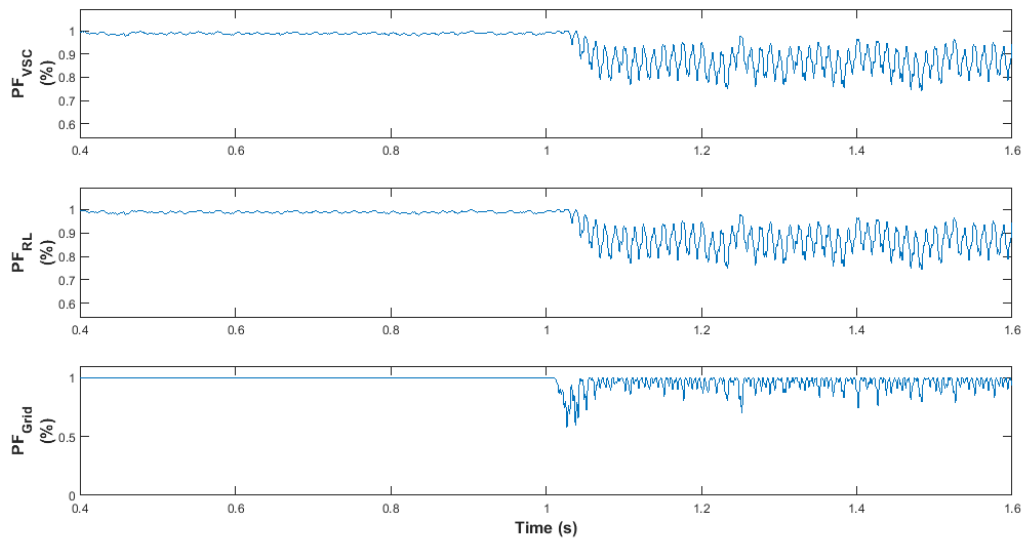
Figure 6.40 – Evolution of the voltage at every link over time when the sunlight dimming factor is 70%.



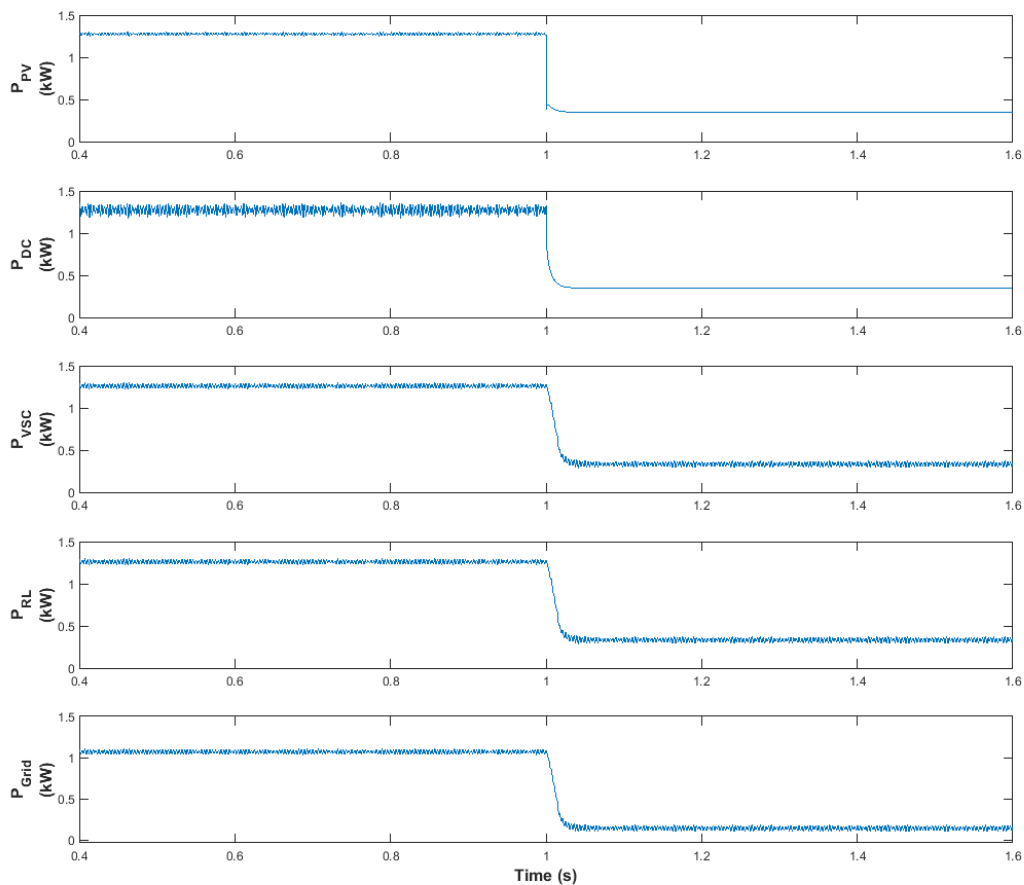
**Figure 6.41 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 70%.**



**Figure 6.42 – Evolution of the current at every link over time when the sunlight dimming factor is 70%.**



**Figure 6.43 – Evolution of the power factor at the AC Links over time when the sunlight dimming factor is 70%.**



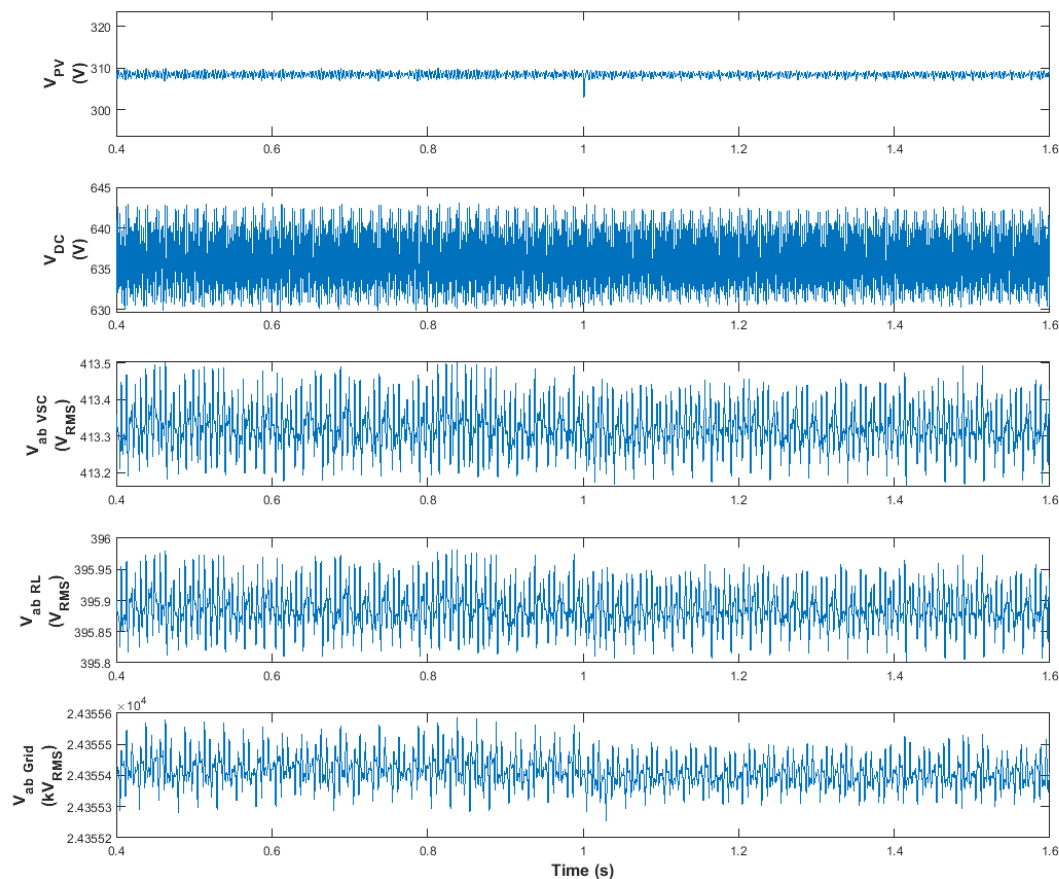
**Figure 6.44 – Evolution of the active power at every link over time when the sunlight dimming factor is 70%.**

Again, the dimming factor will be increased in the following section.

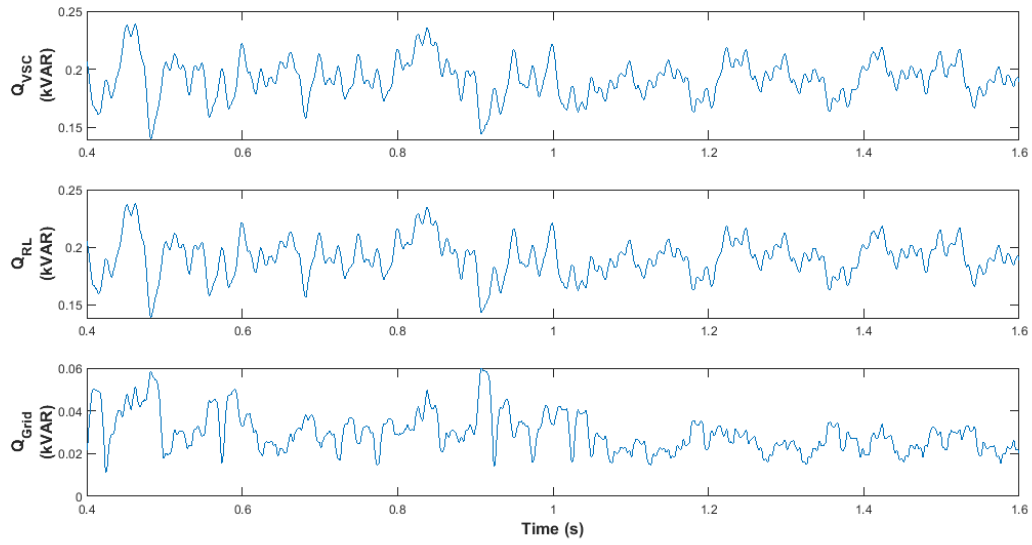


### 6.1.1.2.8 Sunlight Reduction by 80%

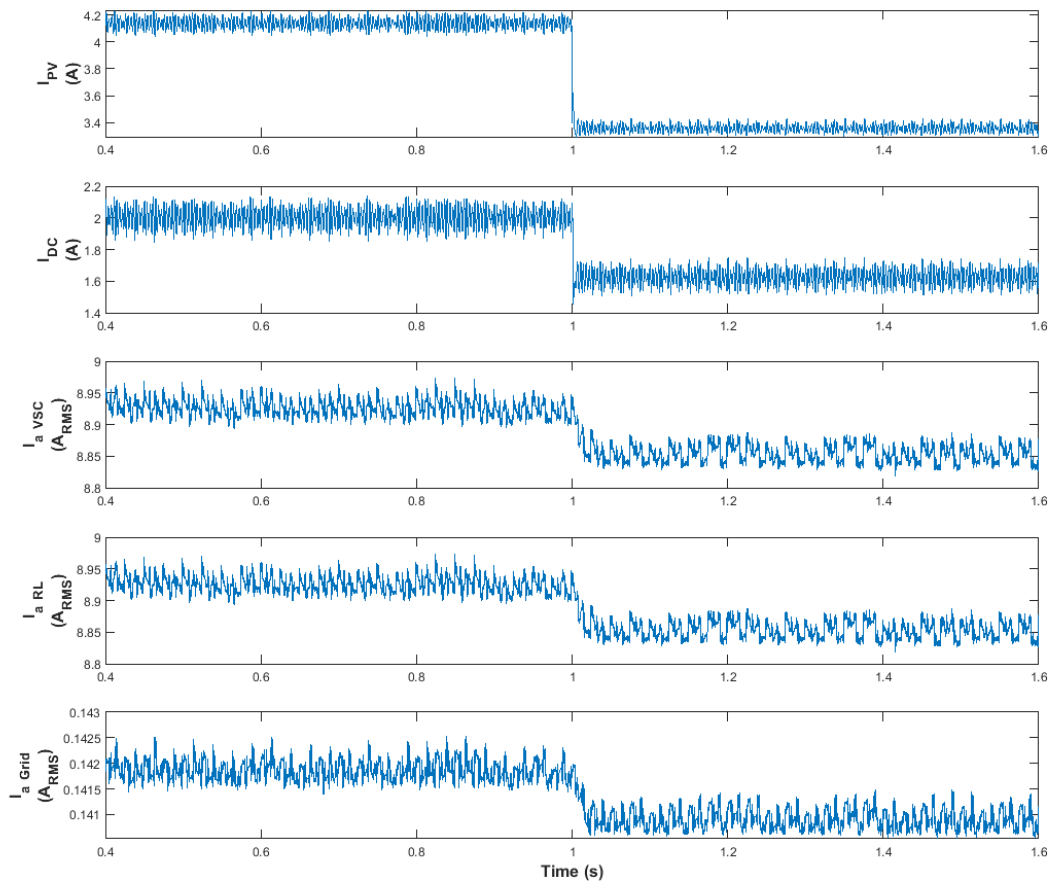
A more extreme situation is envisaged compared with the previous cases, with a decreased of the DNI by 80% after 1 second, or  $200 \text{ W/m}^2$ . The voltages, reactive power, currents, power factor and active power are respectively shown from Figure 6.45 to Figure 6.49. It is visible in Figure 6.45 that the output voltage of the PV array gets decreased by 49%. From Figure 6.47, it can also be seen that the output current of the PV array suddenly drops by a dramatic 73%. As a consequence, as shown in Figure 6.49, the output power of the array itself decreases by 86% and the total power injected to the grid drops becomes null. The power factor also plummets quite dramatically as it drops by 26% on the VSC and RL feeder links and even becomes close to 0 at the point of interconnection with the grid. Thus, in those conditions, the PV array does not supply any power whatsoever to the grid.



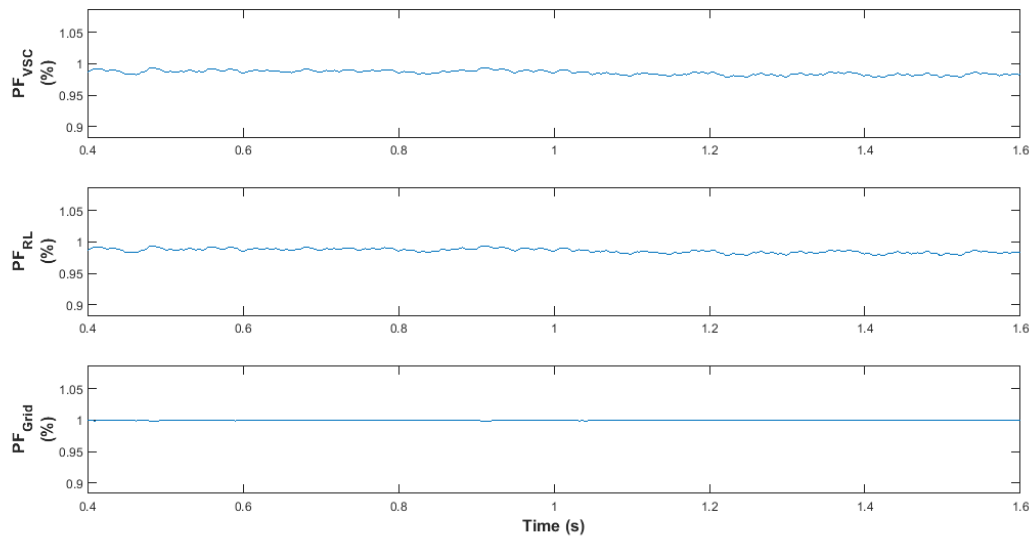
**Figure 6.45 – Evolution of the voltage at every link over time when the sunlight dimming factor is 80%.**



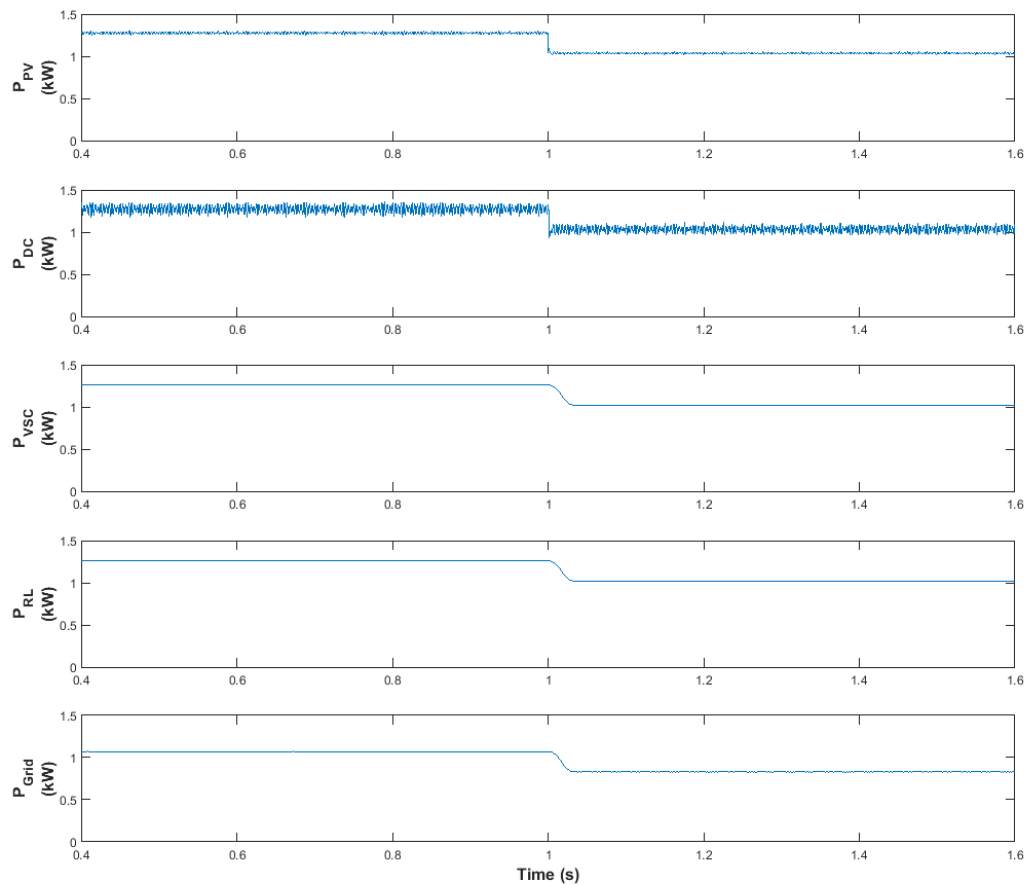
**Figure 6.46 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 80%.**



**Figure 6.47 – Evolution of the current at every link over time when the sunlight dimming factor is 80%.**



**Figure 6.48 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 80%.**

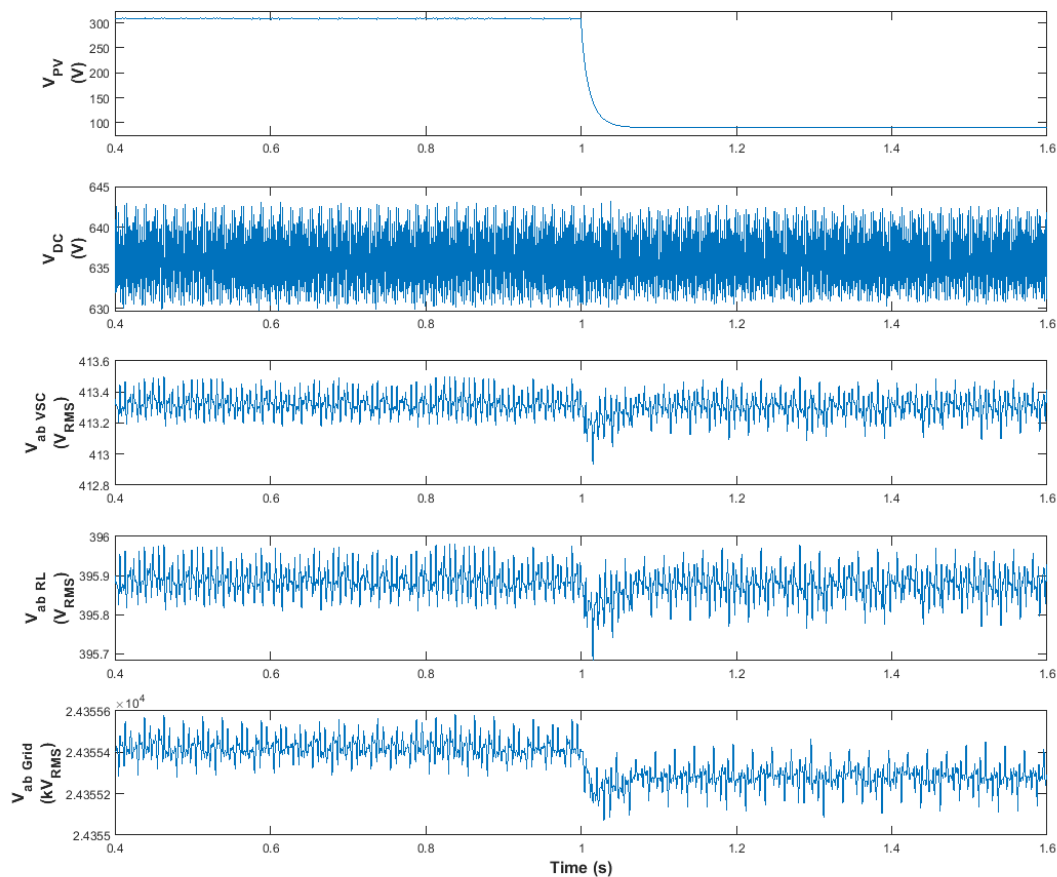


**Figure 6.49 – Evolution of the active power at every link over time when the sunlight dimming factor is 80%.**

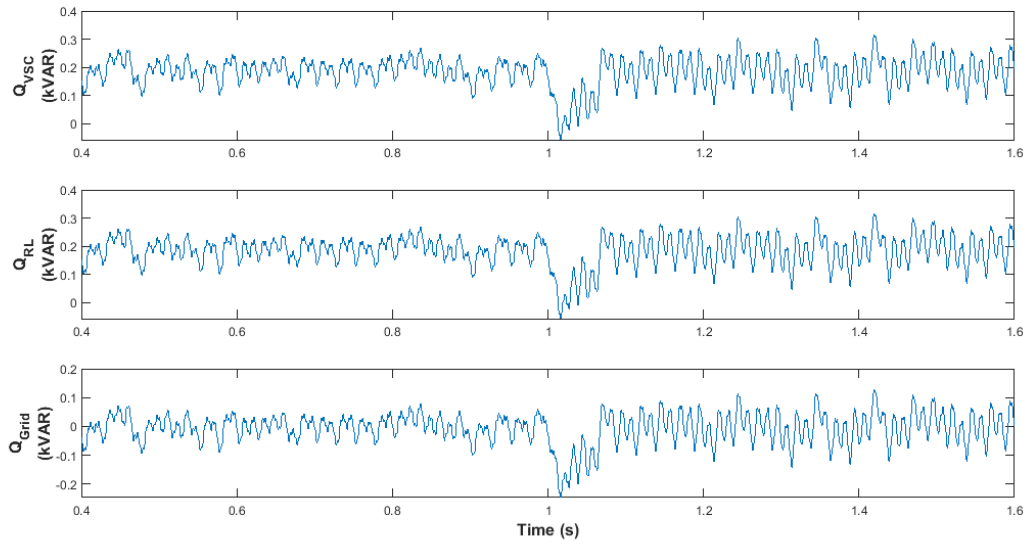
Although the PV array has stopped feeding the grid, there is still a little room to decrease the DNI.

### 6.1.1.2.9 Sunlight Reduction of 90%

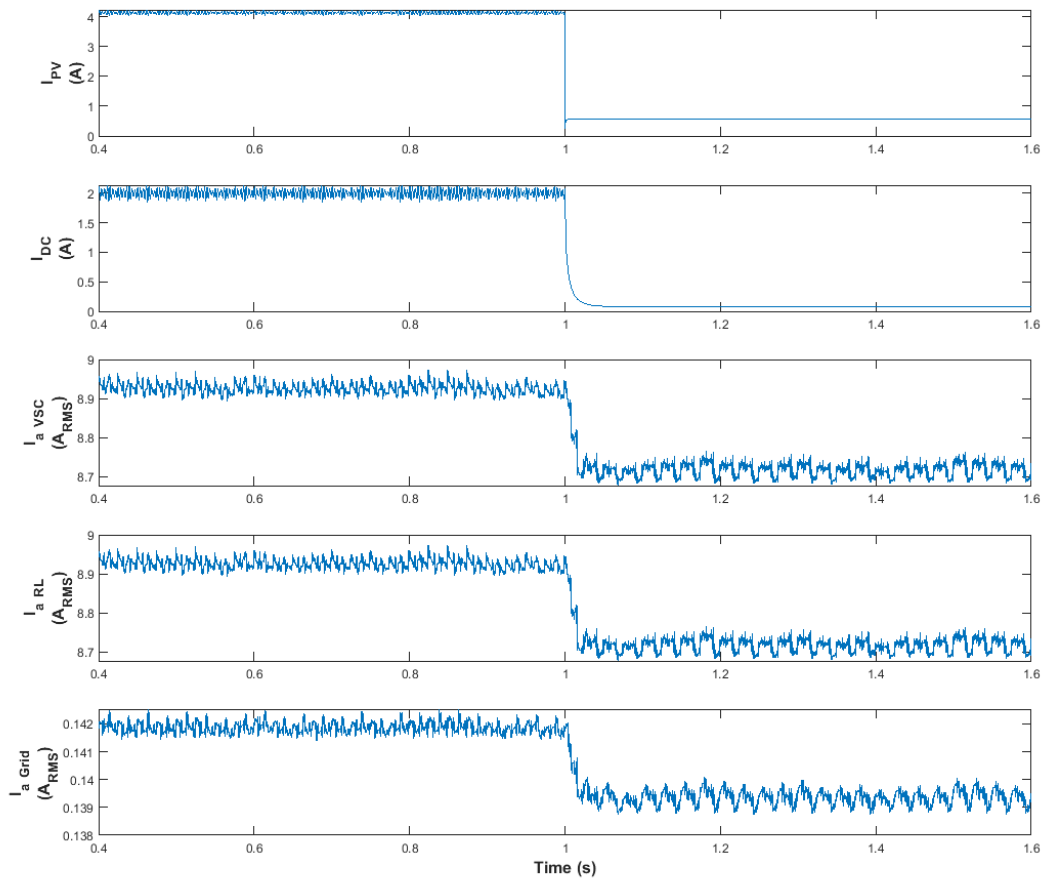
In the previous case, it could be seen that the PV array completely stopped feeding the grid with any active power when it was receiving a DNI of only  $200 \text{ W/m}^2$ . However, it could be interesting to see what happens when the received DNI is even lower and below this threshold value. If the DNI is suddenly cut by 90% after 1 second, and only represents  $100 \text{ W/m}^2$ , the resulting voltages, reactive power, currents, power factor and active power are respectively shown from Figure 6.50 to Figure 6.54. It can be seen in Figure 6.50 that the output voltage of the PV array gets decreased by 71%. In Figure 6.47, it can also be seen that the output current of the PV array drops by as much as 86%. The consequence is visible in Figure 6.54, where the output power of the array itself decreases by 96% and where some power is actually taken from the grid to maintain the DC link voltage. The power factor also plummets by 68 % on the VSC and RL feeder links and is equal to -1 at the point of interconnection with the grid.



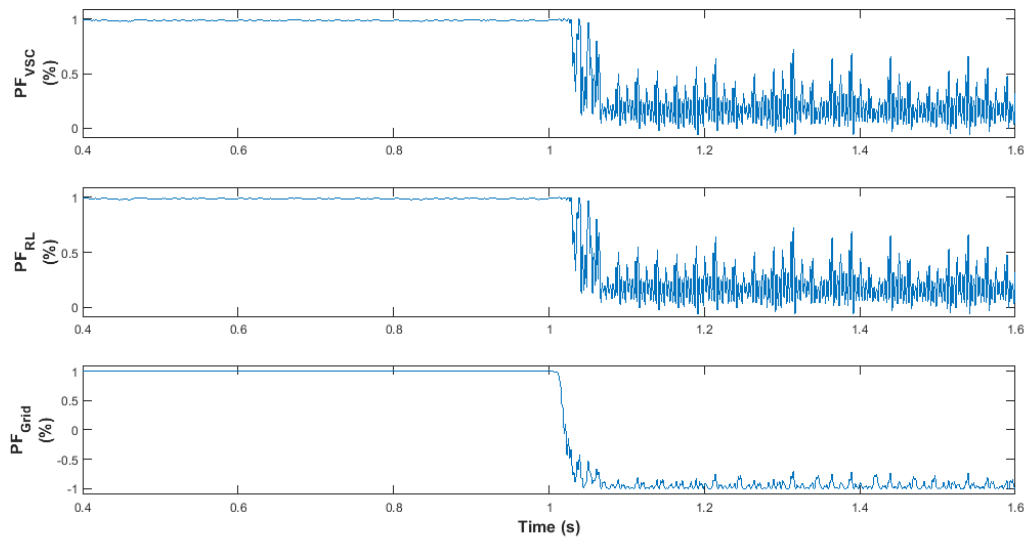
**Figure 6.50 – Evolution of the voltage at every link over time when the sunlight dimming factor is 90%.**



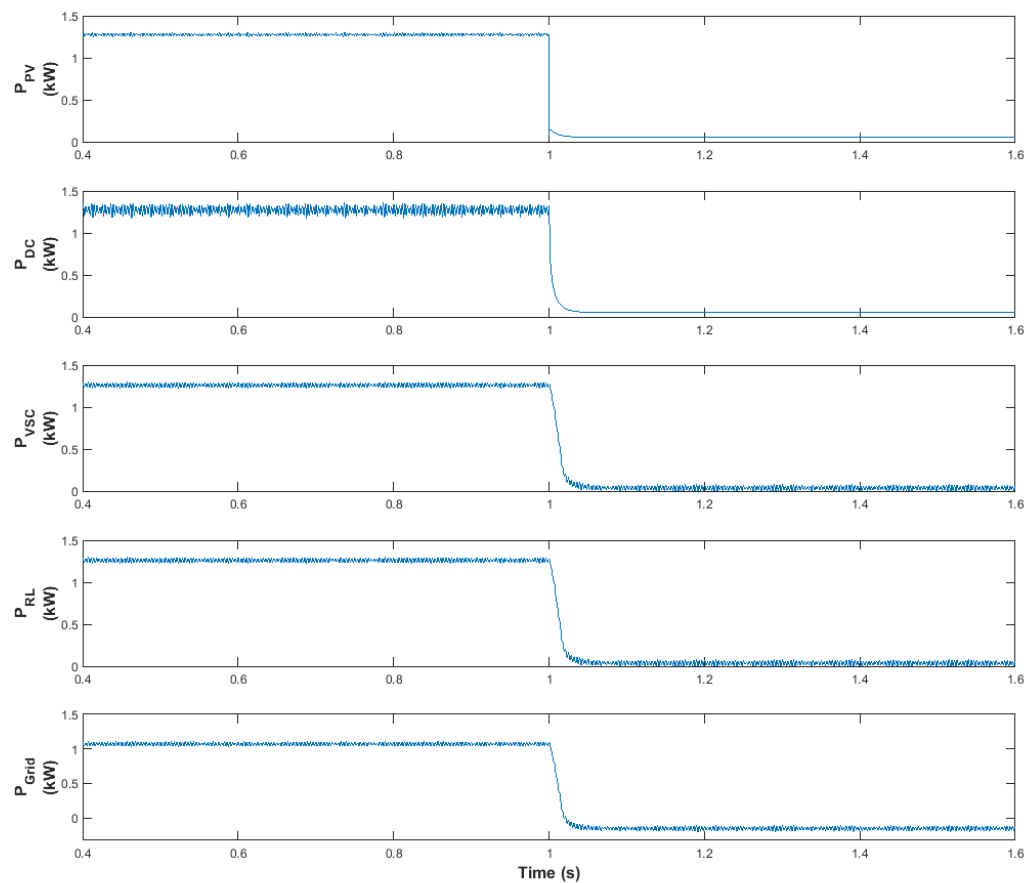
**Figure 6.51 – Evolution of the reactive power at the AC links over time when the sunlight dimming factor is 90%.**



**Figure 6.52 – Evolution of the current at every link over time when the sunlight dimming factor is 90%.**



**Figure 6.53 – Evolution of the power factor at the AC links over time when the sunlight dimming factor is 90%.**



**Figure 6.54 – Evolution of the active power at every link over time when the sunlight dimming factor is 90%.**

For better visibility, the results of all cases are compiled and discussed in the following section.

### 6.1.1.2.10 Summary and Discussion of Results

The settled values of the voltages, currents and powers for each sunlight dimming case are shown in Table 6.2. Observations show that for a constant cell temperature of 35°C, the output voltage of the array only starts dropping when the output power of the array drops below 1 kW, which happens when the value of the DNI gets below 770.5 W/m<sup>2</sup> for the tested cell temperature of 35°C.

Also, it can be observed that when the irradiance drops down to only 200 W/m<sup>2</sup>, the power generated by the array is insufficient to be transmitted to the grid. It is also just sufficient to maintain the DC link voltage to its level. Below the DNI threshold of 200 W/m<sup>2</sup>, or below the equivalent power of 175 W, the power becomes even insufficient to maintain the DC link voltage. In those conditions, not only cannot the PV array feed the grid with any active power but it actually also operates as a load and consumes some power from the grid in order to maintain the DC link voltage to its desired level.

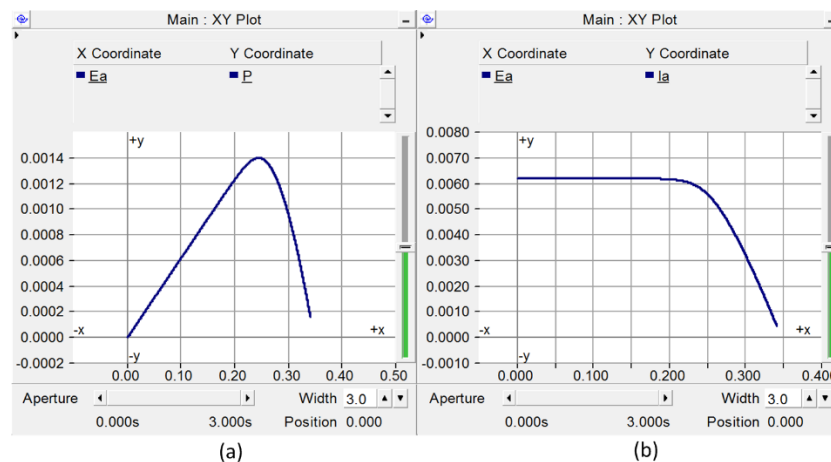
**Table 6.2 – Summary of the settled signal values of the Simulink PV model for each final value when Tcell = 35°C.**

Signal	Final DNI Value (W/m <sup>2</sup> )									
	1000	900	800	700	600	500	400	300	200	100
V <sub>PV</sub> (V <sub>DC</sub> )	308.63	308.62	308.67	<b>304.61</b>	294.24	278.88	250.15	209.41	157.46	90.36
V <sub>DC</sub> (V <sub>DC</sub> )	635.31	635.46	635.23	635.17	635.86	635.34	635.17	635.30	634.75	635.10
V <sub>ab_VSC</sub> (V <sub>AC,RMS</sub> )	413.36	413.34	413.35	413.36	413.37	413.36	413.36	413.36	413.37	413.38
V <sub>ab_AC</sub> (V <sub>AC,RMS</sub> )	395.90	395.90	395.90	395.90	395.91	395.90	395.90	395.90	395.91	395.91
V <sub>a_Grid</sub> (kV <sub>AC,RMS</sub> )	24.36	24.36	24.36	24.36	24.36	24.36	24.36	24.36	24.36	24.36
I <sub>PV</sub> (A <sub>DC</sub> )	4.13	3.74	3.34	3.11	2.93	2.64	2.19	1.66	1.11	0.56
I <sub>DC</sub> (A <sub>DC</sub> )	2.09	1.88	1.72	1.49	1.36	1.16	0.86	0.545	0.274	0.08
I <sub>a_VSC</sub> (A <sub>AC,RMS</sub> )	8.95	8.90	8.87	8.84	8.82	8.79	8.76	8.74	8.713	8.718
I <sub>a_AC</sub> (A <sub>AC,RMS</sub> )	8.95	8.90	8.87	8.84	8.82	8.79	8.76	8.74	8.713	8.718
I <sub>a_Grid</sub> (A <sub>AC,RMS</sub> )	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.139	0.139
P <sub>PV</sub> (kW)	1.28	1.16	1.04	<b>0.95</b>	0.87	0.737	0.555	0.355	<b>0.175</b>	<b>0.05</b>
P <sub>DC</sub> (kW)	1.28	1.16	1.04	0.95	0.87	0.737	0.555	0.355	0.175	0.05
P <sub>VSC</sub> (kW)	1.28	1.16	1.04	0.94	0.86	0.731	0.555	0.355	0.184	0.064
P <sub>AC</sub> (kW)	1.28	1.16	1.04	0.94	0.86	0.731	0.555	0.355	0.184	0.064
P <sub>Grid</sub> (kW)	1.09	0.97	0.85	0.75	0.67	0.542	0.365	0.165	<b>-0.006</b>	<b>-0.126</b>
Q <sub>VSC</sub> (kVAr)	0.194	0.168	0.170	0.155	0.125	0.0105	0.098	0.126	0.175	0.1916
Q <sub>AC</sub> (kVAr)	0.193	0.168	0.170	0.155	0.124	0.105	0.098	0.126	0.175	0.1915
Q <sub>Grid</sub> (kVAr)	0.0018	-0.022	-0.020	-0.035	-0.066	-0.085	-0.094	-0.064	-0.015	0.0017
PF <sub>VSC</sub>	0.9889	0.9896	0.9868	0.9866	0.9897	0.9898	0.9852	0.942	0.725	0.317
PF <sub>AC</sub>	0.9890	0.9897	0.9869	0.9867	0.9899	0.9899	0.9852	0.942	0.725	0.317
PF <sub>Grid</sub>	1.0000	0.9997	0.9997	0.9989	0.9952	0.9878	0.9687	0.933	<b>-0.373</b>	<b>-0.999</b>

### 6.1.1.3 Equivalent PSCAD Model

Although SIMULINK has some advantages like giving the possibility to import some measured data from the workspace, PSCAD is a preferred tool in our team. Also, PSCAD can run some faster simulations, which may lead to more results than SIMULINK in a timely manner. The 1.5 kW grid-connected model previously presented has thus been modelled in PSCAD version 4.5 with the same characteristics as for the adjusted SIMULINK model from the previous section.

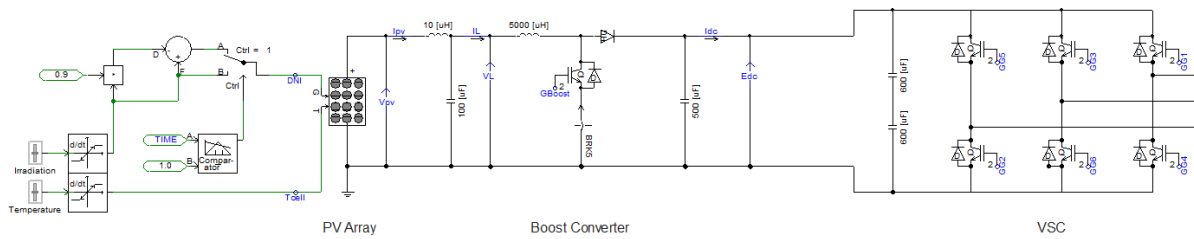
Although some NREL PV models are available both in MATLAB SIMULINK and in PSCAD, those models have been developed independently and may differ significantly. Thus, the PV array has been sized and its parameters have been set up in a way to match as closely as possible the array previously modelled in SIMULINK. This adjustments lead to the power-voltage (P-V) and current-voltage (I-V) characteristics featured in Figure 6.55, where the power is shown in kW, the voltage is expressed in kV and the current is in kA.



**Figure 6.55 – P-V (a) and I-V (b) characteristics of the PSCAD PV array model when receiving a DNI of 1000W/m<sup>2</sup> and operating at 25°C.**

The developed PSCAD model is composed of the same elements as the Simulink model was. Namely, it contains a PV array consisting of 8 monocrystalline modules in series, which is connected to a boost converter controlled by an MPPT using the incremental conductance method. It also contains a VSC along with its VSC controller. The VSC is connected to the grid via an RL feeder and a transformer. A partial view of the system can be seen in Figure 6.56.





**Figure 6.56 – Partial view of the grid-connected PV system.**

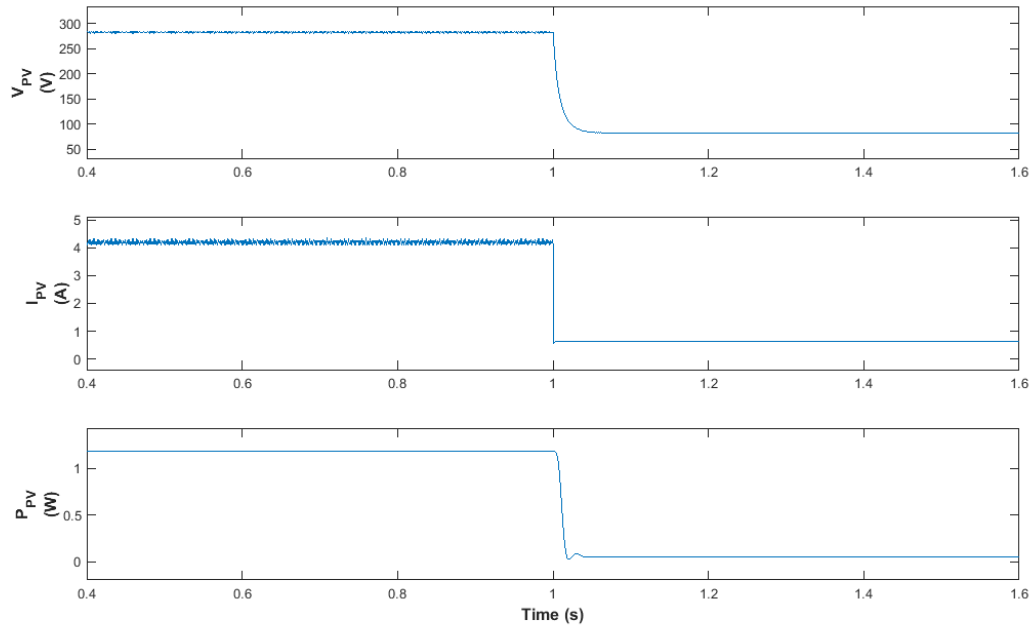
The system has been designed for the voltage and active power levels to be similar to the Simulink model. For practical reasons, though, the PV output voltage and the DC link voltage may be somewhat lower than in the Simulink model. Also, the AC currents and power factors may be different. However, the overall system has been designed so that the main variations of voltage, current and active power remain very similar to the Simulink model.

The values of the various signals of the system are compiled in Table 6.3 for various DNI values. The same comments apply concerning the trends of the voltages, currents and active powers as for the previously presented Simulink model. The output voltage of the PV array starts dropping by more than 1% and the PV array starts acting like a load for the same levels of DNIs as for the previous model.

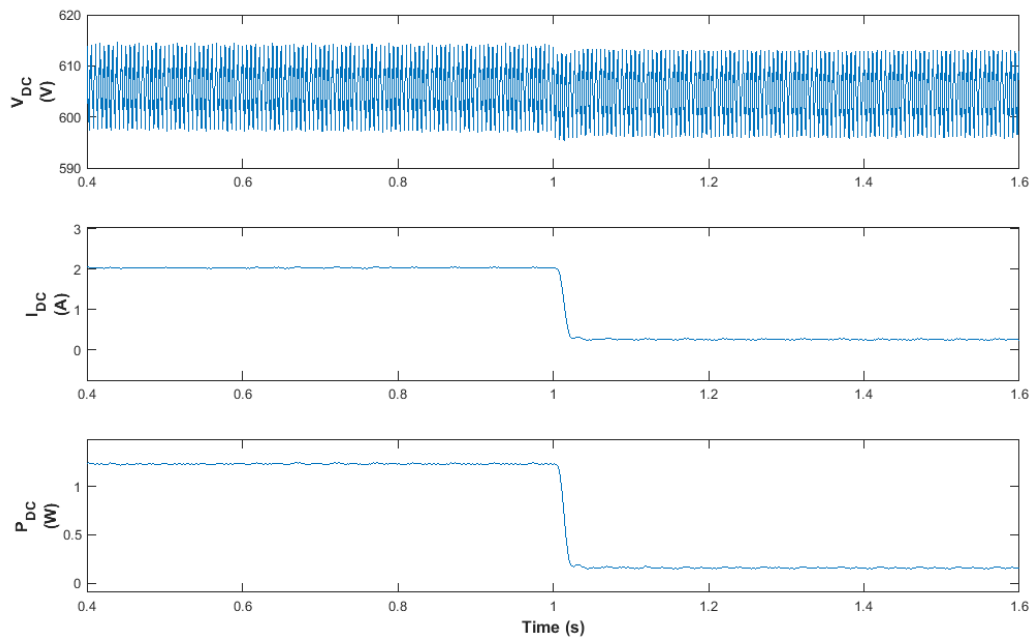
**Table 6.3 – Summary of the settled signal values of the PSCAD PV model for various DNI values when Tcell = 35°C.**

Signal	Final DNI Value (W/m <sup>2</sup> )									
	1000	900	800	700	600	500	400	300	200	100
V <sub>PV</sub> (V <sub>DC</sub> )	281.19	279.05	276.96	<b>274.03</b>	270.03	263.76	251.21	220.1	167.6	82.81
V <sub>PV</sub> change (%)	N/A	0.02	0.77	<b>1.82</b>	3.25	5.50	9.99	21.4	39.95	70.33
V <sub>DC</sub> (V <sub>DC</sub> )	609.64	602.8	608.11	608.8	609.22	609.44	607.5	608.4	608.4	608.6
V <sub>ab_VSC</sub> (V <sub>AC,RMS</sub> )	464.17	464.18	464.12	464.09	464.39	464.44	463.85	464.37	463.99	461.69
V <sub>ab_AC</sub> (V <sub>AC,RMS</sub> )	399.75	399.22	399.69	399.35	399.48	399.71	399.48	399.81	399.33	399.41
V <sub>a_Grid</sub> (kV <sub>AC,RMS</sub> )	24.99	24.99	24.99	24.99	24.99	24.99	24.97	24.99	24.99	24.98
I <sub>PV</sub> (A <sub>DC</sub> )	4.37	4.15	3.99	3.82	3.59	3.29	2.86	2.25	1.5	0.619
I <sub>PV</sub> change (%)	N/A	5.03	8.70	12.59	17.85	24.71	34.55	48.51	65.68	85.84
I <sub>DC</sub> (A <sub>DC</sub> )	2.08	1.98	1.9	1.82	1.68	1.5	1.26	0.9	0.5	0.08
I <sub>DC</sub> change (%)	N/A	4.81	8.65	12.50	19.23	27.88	39.42	56.73	75.96	96.15
I <sub>a_VSC</sub> (A <sub>AC,RMS</sub> )	28.59	28.55	28.56	28.52	28.61	28.71	28.71	28.59	28.51	28.46
I <sub>a_AC</sub> (A <sub>AC,RMS</sub> )	4.79	4.77	4.76	4.73	4.72	4.71	4.74	4.6	4.6	4.62
I <sub>a_Grid</sub> (A <sub>AC,RMS</sub> )	0.078	0.078	0.078	0.078	0.078	0.078	0.078	0.078	0.078	0.078
P <sub>PV</sub> (kW)	1.24	1.21	1.16	<b>1.10</b>	0.97	0.87	0.72	0.54	<b>0.3</b>	<b>0.16</b>
P <sub>PV</sub> change (%)	N/A	4.72	8.66	13.39	23.62	31.50	43.31	57.48	76.38	87.40
P <sub>DC</sub> (kW)	1.24	1.21	1.16	1.10	1.00	0.92	0.77	0.54	0.3	0.16
P <sub>DC</sub> change (%)	N/A	4.72	8.66	13.39	21.26	34.71	51.03	57.48	76.38	87.40
P <sub>VSC</sub> (kW)	1.13	1.11	1.05	0.99	0.92	0.82	0.66	0.44	0.21	0.09
P <sub>VSC</sub> change (%)	N/A	5.98	10.26	15.38	21.37	29.91	43.59	62.39	82.05	92.31
P <sub>AC</sub> (kW)	0.919	0.9	0.85	0.77	0.7	0.6	0.45	0.22	-0.07	-0.13
P <sub>AC</sub> change (%)	N/A	2.07	7.51	16.21	23.83	34.71	51.03	76.06	107.6	114.15
P <sub>Grid</sub> (kW)	0.919	0.9	0.85	0.79	0.68	0.58	0.45	0.21	<b>-0.07</b>	<b>-0.13</b>
P <sub>Grid</sub> change (%)	N/A	2.07	7.51	14.04	26.01	36.89	51.03	77.15	107.6	114.15
Q <sub>VSC</sub> (kVAr)	-17.87	-17.87	-17.87	-17.87	-17.87	-17.96	-17.87	-17.86	-17.86	-17.86
Q <sub>AC</sub> (kVAr)	0.018	0.018	0.016	0.012	0.013	0.013	0.015	0.01	0.016	0.014
Q <sub>Grid</sub> (kVAr)	-0.38	-0.38	-0.38	-0.38	-0.38	-0.38	-0.36	-0.38	-0.36	-0.37
PF <sub>VSC</sub>	0.063	0.061	0.059	0.055	0.051	0.046	0.037	0.025	1.175	0.005
PF <sub>AC</sub>	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.996	<b>-0.996</b>	<b>-0.994</b>
PF <sub>Grid</sub>	0.999	0.924	0.913	0.901	0.873	0.836	0.781	0.484	<b>-0.467</b>	<b>-0.331</b>

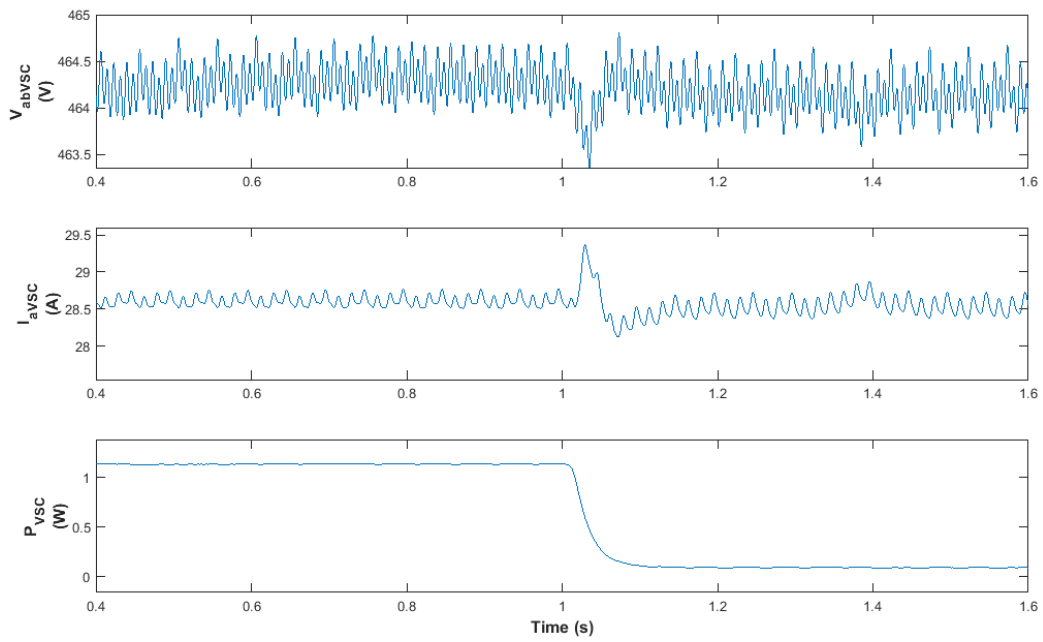
To illustrate the behaviour of the system when the DNI drops from  $1000 \text{ W/m}^2$  to  $100 \text{ W/m}^2$ , the values of the voltages, currents and active powers are shown for every link of the system respectively from Figure 6.57 to Figure 6.61.



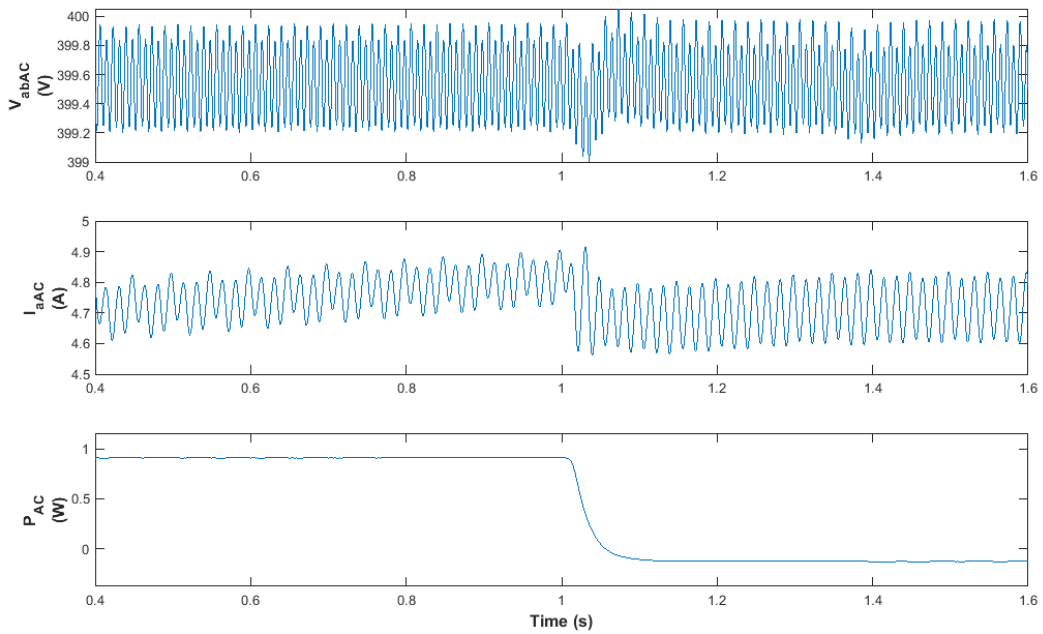
**Figure 6.57 – Evolution of the signals at the PV output over time when the sunlight dimming factor is 90%.**



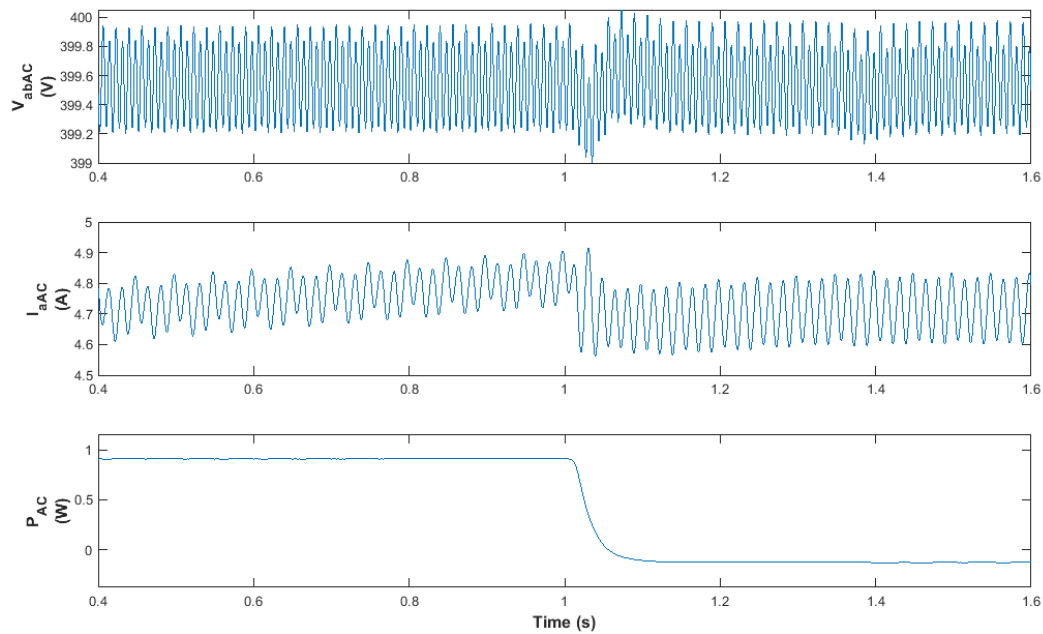
**Figure 6.58 – Evolution of the signals at the DC link over time when the sunlight dimming factor is 90%.**



**Figure 6.59 – Evolution of the signals at the VSC output over time when the sunlight dimming factor is 90%.**



**Figure 6.60 – Evolution of the signals at the RL feeder output over time when the sunlight dimming factor is 90%.**

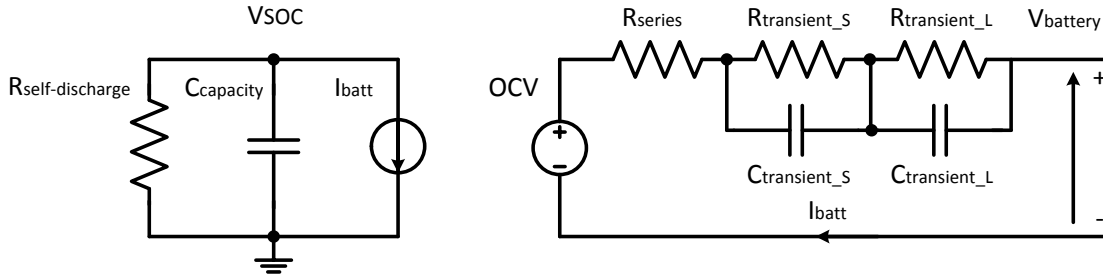


**Figure 6.61 – Evolution of the signals at the transformer output over time when the sunlight dimming factor is 90%.**

In order to mitigate the observed and discussed effects due to sudden drops of DNI, using a battery is a possible solution that would allow to smooth PV output power cuts by storing and shift historical excess power. This solution will be explored in the next sections. First, a realistic battery model is required. Then, a proper DC-DC converter is also needed with an adequate controller.

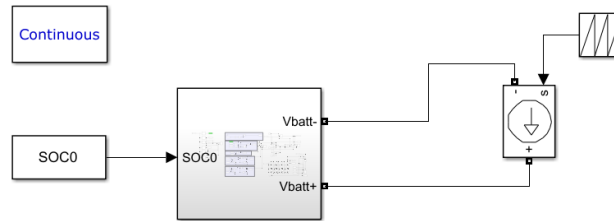
### 6.1.2 Battery Model

In the eventuality that the previously presented SIMULINK model was simulated with a battery bank, it would then matter to dispose of a proper battery model. Although MATLAB SIMULINK dispose of an existing battery model, some imperfections have been reported about this model in literature. In particular, from a short-term perspective, the built-in battery block is based on an equation that cannot render realistically the non-linear relationship that exists between the battery current and its output voltage [168]. Thus, for the sake of accuracy, a custom MATLAB/SIMULINK battery bank model has been designed by following an accurate model that has been created empirically based on some measurements of the internal resistance and capacitance values for a polymer lithium ion battery [169, 170]. The model in question is illustrated for a single cell in Figure 6.62.



**Figure 6.62 – Accurate electrical battery model as proposed in [168, 169].**

The top level view of the SIMULINK model is shown in Figure 6.63.



**Figure 6.63 – Top view of the MATLAB/SIMULINK battery bank model.**

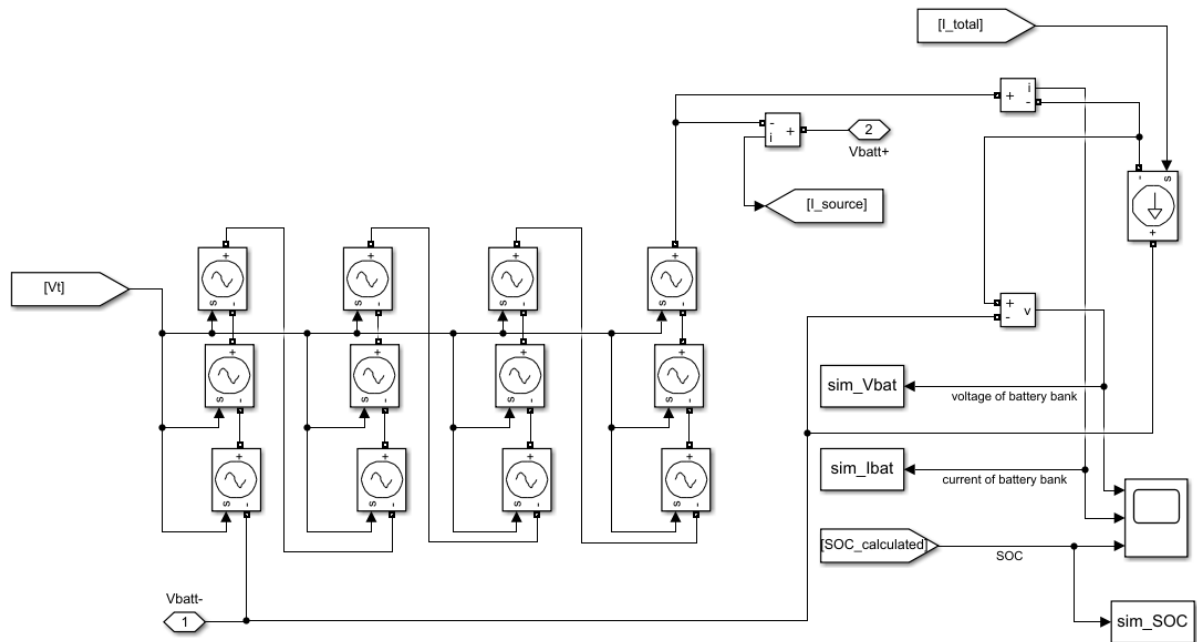
The state of charge (SOC) of the battery is calculated as per Equation 6.4 where  $C_0$  is the capacity of one cell of the battery, which is 2 Ah, and  $I_{total}$  is the controlled series charge/discharge current flowing through all battery cells. Effectively,  $I_{total}$  is the same series current as the current  $I_{batt}$  that is featured in Figure 6.62. The charge/discharge current  $I_{source}$  flowing to (or from) the battery is regulated through a custom made battery charge controller. This controller operates so that the battery stops charging as soon as its SOC reached 100% by making  $I_{total}$  equal to 0. Similarly, the battery is not allowed to discharge below 5% of SOC and the current  $I_{total}$  is also curtailed when that occurs.

In addition, the model also requires some information on the initial state of charge  $SOC_0$  that has been defined as 100%, which means that the battery bank is initially fully charged.

$$SOC = SOC_0 - \frac{1}{C_0} \int I_{total} dt \quad (6.4)$$

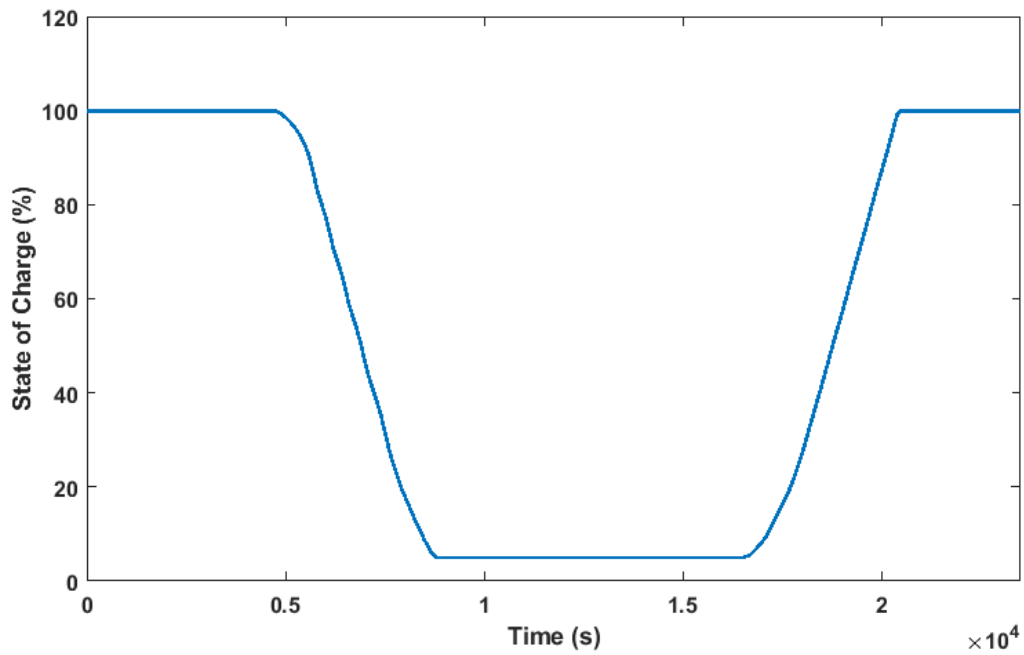
Further, it is to be noted that the model has been initially designed for a continuous simulation mode. A discrete adaptation is needed to integrate this model with the SIMULINK grid-connected PV system that has been mentioned in the previous section. However, running a continuous simulation rather than a discrete one enables to study a full cycle of discharge and charge, which is more appropriate to verify the good functionality of the model.

The target voltage of the battery bank is 48V. Given the fact that the open circuit voltage (OCV) of each cell is 4.1V, 12 cells are connected in series to constitute the target voltage as shown in Figure 6.64, sharing the same current  $I_{total}$ .



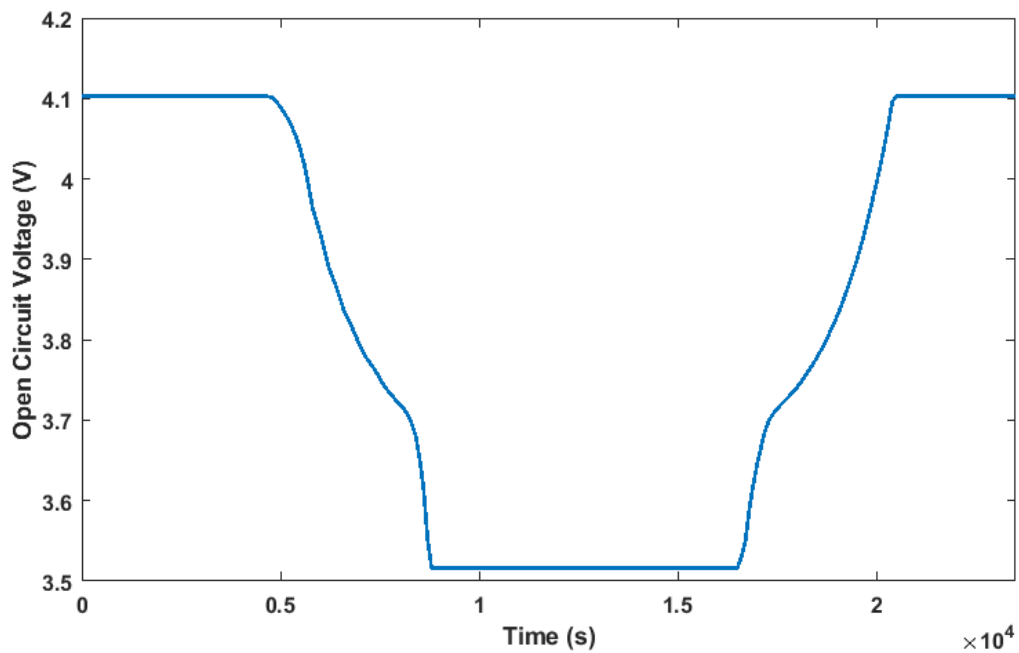
**Figure 6.64 – View of the power part of the MATLAB/SIMULINK battery bank model.**

Over a simulation time of 6 hours 40 minutes, the initially fully charged battery bank starts being discharged at  $t_1 = 1$  hour 40 minutes. The SOC is allowed to go down as low as 5% and is maintained at that level until the battery bank is charged up again at  $t_2 = 5$  hours and eventually fully recharges. When a continuous model is designed as per Figure 6.64 and is subsequently simulated over a full discharge and recharge cycle, assuming that the battery is initially full, then the evolution of the state of charge of the battery bank in those conditions is shown in Figure 6.65.

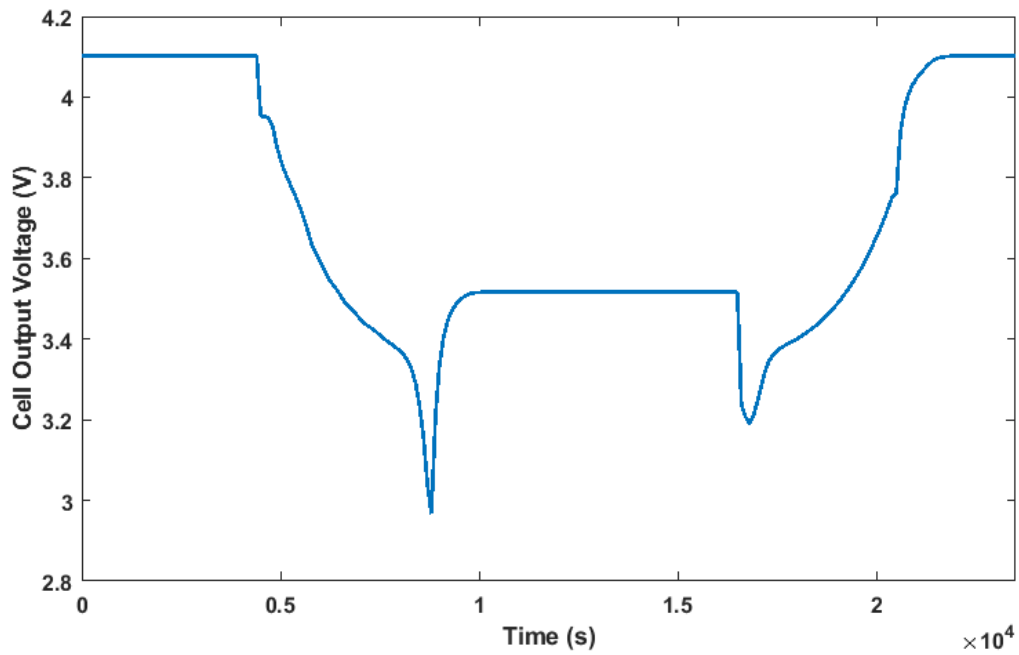


**Figure 6.65 – State of charge of the battery bank.**

The changes occurring within one single cell of the battery bank in the exact same conditions in terms of open circuit voltage and output voltage are shown in Figure 6.66 and Figure 6.67 respectively. As expected, the output voltage of the cell loosely follows the open circuit voltage although it plunges to a lower level for a relatively short period of time mostly due to the effect of  $C_{transient\_S}$ .

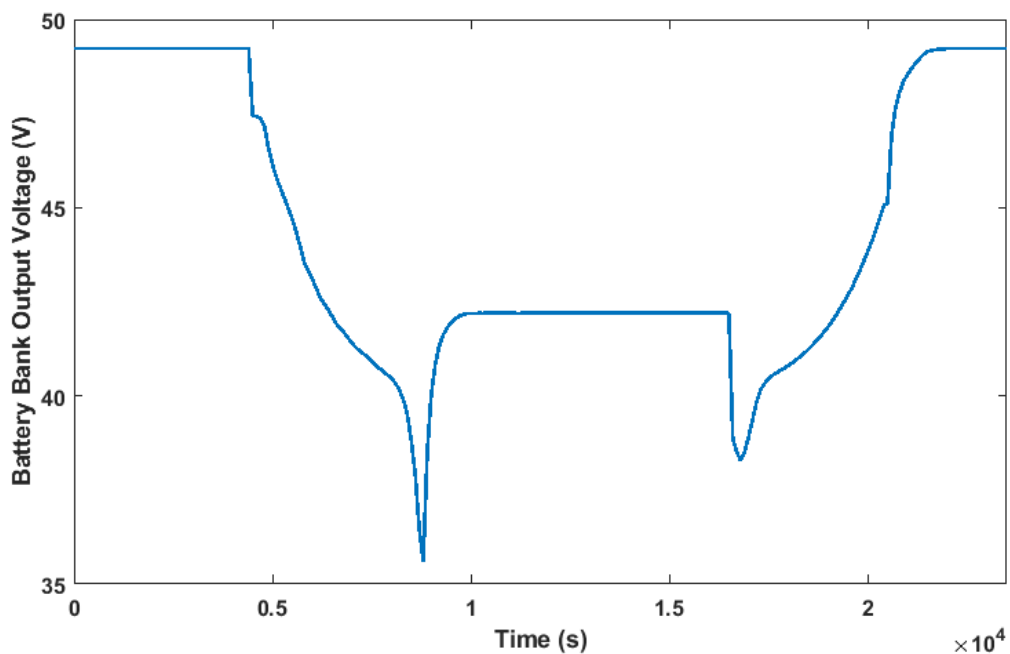


**Figure 6.66 – Open circuit voltage of one single cell in the battery bank.**



**Figure 6.67 – Output voltage of one single cell in the battery bank.**

As a result, the output voltage of the complete battery bank is simply proportional to the output voltage of a single cell, as seen in Figure 6.68, with a maximum voltage of 49.24V when fully charged and a minimum voltage of 35.60V when discharged at 5%.



**Figure 6.68 – Output voltage of the complete battery bank.**

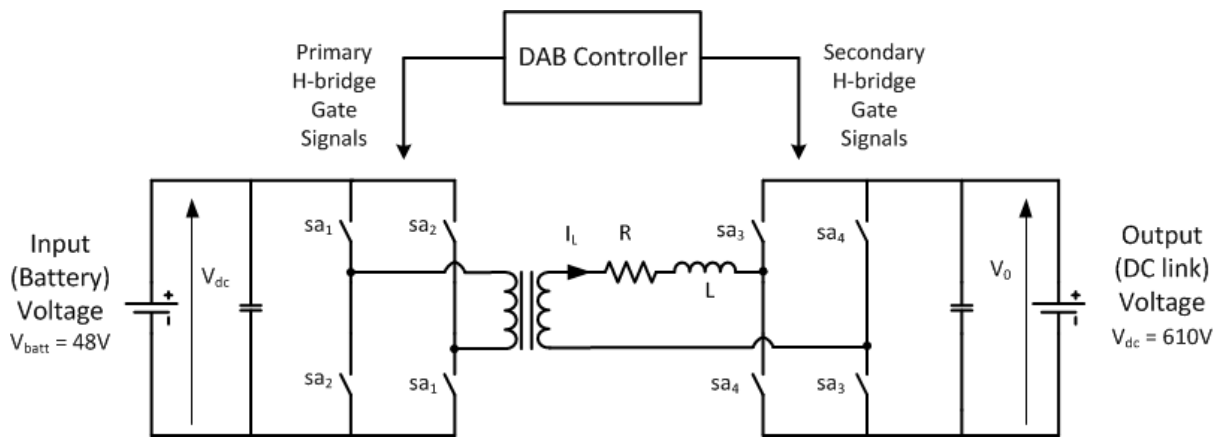


Simulation results prove the battery bank model to be fully functional over a full discharge/recharge cycle. The target voltage of 48% is achieved when the battery is fully charged. That voltage is even exceeded by 1.24V, which only corresponds to a relative error of 2.58%.

However, disposing of a battery model is a necessary step towards smoothing PV output power but is still insufficient on its own. A proper control is additionally needed to control the charge and discharge of the battery through the DC link of the grid-connected PV system.

### 6.1.3 Battery and Bidirectional DC/DC Converter with Controller

The functionality that is sought would ideally require to connect a bi-directional DC/DC converter between the DC link and the battery. A good example of a suitable bidirectional converter for that sort of application would be a dual active bridge (DAB) [171-174].



**Figure 6.69 – Diagram of the Dual Active Bridge (DAB) envisaged for this study.**

A dual active bridge consists of two H-bridge converters connected via a high-frequency transformer, as shown in Figure 6.69. This means that the converter has four switching combinations. Also, the gates in each H-bridge are to be switched with a phase shift angle  $\Phi$  between both bridges. It must be noted that the angle  $\Phi$  constitutes the control variable of the controller. Also, both bridges operate at a 50% duty cycle ratio. The full details of the architecture of a DAB can be found in [172].

If we consider the state vector  $x$  defined in Expression 6.5, where  $I_L$  is the current flowing through the inductor  $L$  and  $V_{dc}$  is the output voltage. An input perturbation would be noted  $\tilde{x}$  while an output perturbation would be noted  $\tilde{y}$ . In our specific application, the input voltage  $V_{batt}$  is the battery voltage, which is chosen as 48 V when the battery is completely charged, which means when the state of charge (SOC) of the battery is 100%. The output voltage  $V_{dc}$ , on the other hand is the DC link voltage of the PV system, which is maintained close to 610 V by the VSC controller.

$$x = \begin{bmatrix} V_{dc} \\ I_L \end{bmatrix} \quad (6.5)$$

A relationship between  $x(t_0)$  and  $x(t_4)$  can be established as per Equation 6.6:

$$x(t_4) = F \cdot x(t_0) + G \cdot V_{batt} \quad (6.6)$$

Over a switching cycle, the linearized model of a DAB is given in discrete-time domain by Equations 6.7 and 6.8, where  $A = \frac{\partial F}{\partial x}$  and  $B = \frac{\partial F}{\partial \Phi} + \frac{\partial G}{\partial \Phi}$ :

$$\tilde{x}(k+1) = A \cdot \tilde{x}(k) + B \cdot \tilde{\Phi}(k) \quad (6.7)$$

$$\tilde{y}(k) = [0 \quad 1] \cdot \tilde{x}(k) = C_1 \cdot \tilde{x}(k) \quad (6.8)$$

If a state feedback with an integral control action is applied to the DAB, then the discrete-time equivalent of the integral controller becomes as shown in Equation 6.9 where  $y_{ref}$  is the reference voltage.

$$\begin{cases} \tilde{e}(k) = \tilde{y}(k) - \tilde{y}_{ref}(k) \\ z(k) = z(k-1) - K_i \cdot \tilde{e}(k) \end{cases} \quad (6.9)$$

In Equation 6.9,  $K_i$  is an integral gain. From Equations 6.8 and 6.9, a new expression can be inferred as shown in Equation 6.10.

$$z(k+1) = z(k) + K_i \cdot C_1 \cdot \tilde{x}(k) - K_i \cdot \tilde{y}_{ref}(k) \quad (6.10)$$

Also, an extended state vector can be defined as per Equation 6.11.

$$x_e(k) = \begin{bmatrix} \tilde{x}(k) \\ z(k) \end{bmatrix} \quad (6.11)$$

It can be inferred from Equations 6.10 and 6.11 that the extended state space description of the system can be written as shown in Equation 6.12.

$$x_e(k+1) = \begin{bmatrix} A & 0 \\ K_i \cdot C_1 & 1 \end{bmatrix} \cdot x_e(k) + \begin{bmatrix} B \\ 0 \end{bmatrix} \cdot \tilde{\Phi}(k) - \begin{bmatrix} 0 \\ K_i \end{bmatrix} \cdot \tilde{y}_{ref}(k) \quad (6.12)$$

The DAB controller is then designed through a state feedback control as in Equation 6.13.

$$\tilde{\Phi}(k) = -K \cdot x_e(k) \quad (6.13)$$

The gain matrix can be computed using the linear quadratic regular (LQR) function in MATLAB. The above control law regulates the output voltage of the DAB. However for a power flow control, the output voltage needs to be modulated based on the power demand. To achieve this, a PI Controller is

employed, whose output is the voltage reference  $y_{ref}$  mentioned in Equation 6.9. The input to the PI controller is the difference between the desired power and the actual power output (or input) of (from) the DAB as given in 6.14, where  $P_{ref}$  is the reference power and  $P_{link}$  is the power flowing into the DAB or out of the DAB. It is to be noted that the positive direction for the power to flow is considered as into the DAB. In that case, when the power needs to flow out of the DAB, a negative  $P_{ref}$  is considered.

$$y_{ref} = K_p \cdot (P_{ref} - P_{link}) + K_i \cdot \int (P_{ref} - P_{link}) dt \quad (6.14)$$

Finally, in steady-state conditions, assuming that the transformer is ideal, the power transfer through the converter can be expressed as in Expression 6.15.

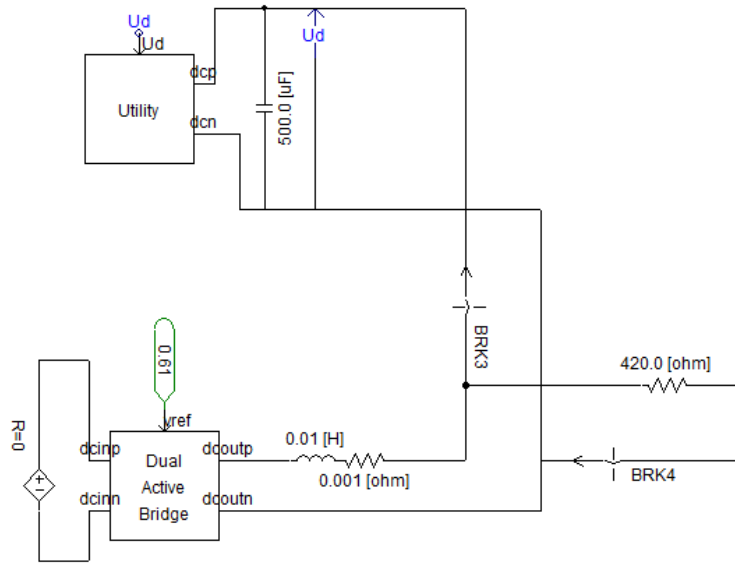
$$P = \frac{V_{batt} \cdot V_{dc} \cdot \Phi \cdot (\pi - \Phi)}{2 \cdot \pi^2 \cdot f \cdot L} \quad (6.15)$$

Hence, it can be said that the power transfer is maximum when  $\Phi = \pi/2$  and it is 0 when  $\Phi = 0$ .

### 6.1.3.1 DAB Modelling and Testing

The overall system for the grid connected bidirectional DC-DC converter has been modelled in PSCAD. The top view of the model is shown in Figure 6.70. It is composed of a DAB that is initially insulated and connected to a local load. After reaching a steady state, it is then disconnected from its load and connected to the grid instead. A simplified version of the grid previously shown has been used. The battery is assumed initially fully charge and thus its voltage  $V_{batt}$  is 48 V. Since the duration of the simulation is limited to a few second only, the battery has been replaced by an ideal voltage source.

After calculating the steady state conditions in MATLAB, the transformer of the DAB has been designed around a 20 kVA ideal transformer with a base operation frequency  $f = 10$  kHz, a leakage reactance  $X = 0.001$  pu, a 2 kV primary winding voltage and a 10 kV secondary voltage.



**Figure 6.70 – PSCAD DAB model connected to a utility grid block with a local load.**

Also, initially, the DAB is simply connected a load  $R_L = 420 \Omega$  to allow the converter to start up. After reaching a steady state, that resistive load is disconnected via the breaker BRK4 and instead, the DAB is connected to the utility grid via the breaker BRK3 with a  $500 \mu\text{F}$  filter capacitor.

The initial value of the phase angle, or nominal phase shift angle  $\Phi_0$  is chosen as  $1.9^\circ$ . Calculating the steady state vector  $x_0$  for a fixed duty ratio leads to the result shown in Expression 6.16.

$$x_0 = \begin{bmatrix} 659.44 \text{ V} \\ 128.05 \text{ A} \end{bmatrix} \quad (6.16)$$

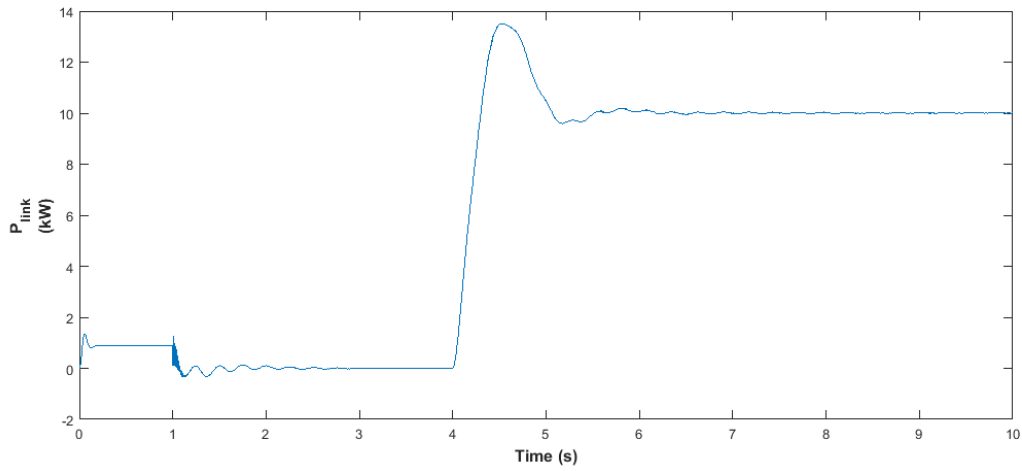
Based on this and the system parameters, the proportional gain matrix mentioned in Equation 6.13 and the integral gains are given in Expression 6.17.

$$\left\{ \begin{array}{l} K = \begin{bmatrix} 1.7635 \times 10^{-6} & -3.5468 \times 10^{-11} & 5.3019 \times 10^{-6} \\ & & \end{bmatrix} \\ K_i = 0.01 \end{array} \right. \quad (6.17)$$

Before being used to mitigate and compensate for power losses, the DAB is tested to inject some power into the grid and to draw some power from the grid. Effectively, this respectively corresponds to discharging and recharging the battery.

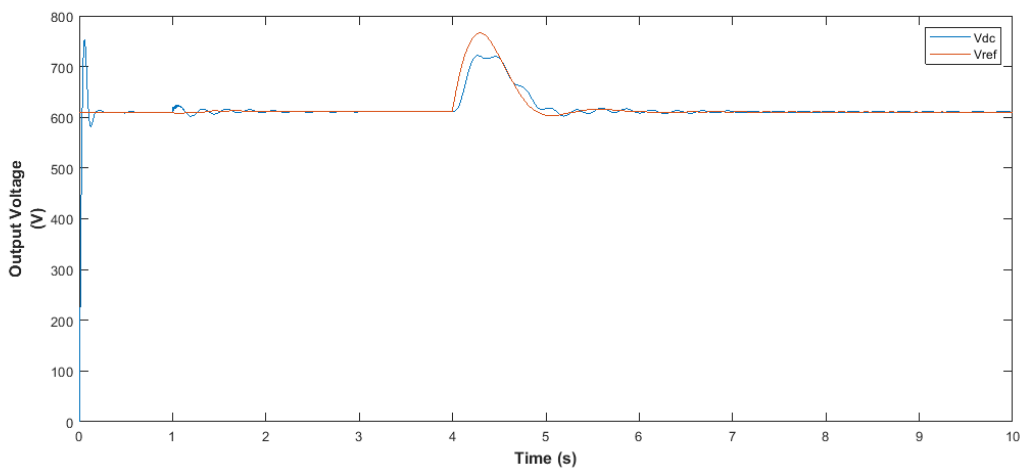
As seen in Figure 6.70, the system does a cold start at the initial time  $t = 0 \text{ s}$  with breaker BRK3 open and breaker BRK4 closed. Thus, effectively, the DAB is only connected to a local resistive load  $R_L = 420 \Omega$  to get started. The power flow controller is initially disabled and the power at the DC link junction is free to get the value  $P_{link} = 0.877 \text{ kW}$ . After one second, at  $t = 1 \text{ s}$ , the state of both breakers BRK3 and BRK4 is toggled to disconnect the local load and connect the DAB to the DC link. At the same

time, the power flow controller is activated with a power reference  $P_{ref}$  set to 0 kW. Finally, after 4 seconds, at  $t = 4$  s, the power reference  $P_{ref}$  is set to a different value, which is 10 kW.

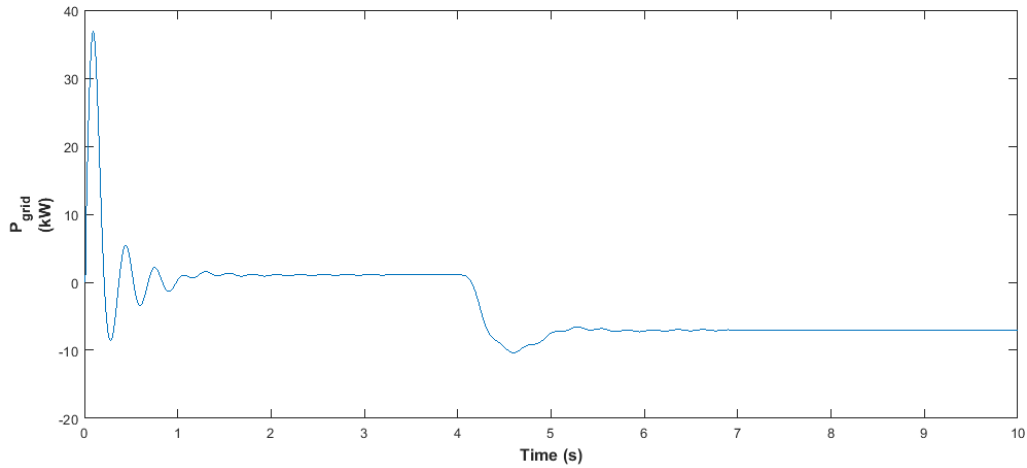


**Figure 6.71 – DC link power when the DAB supplies power to AC grid.**

The resulting output power of the DAB, which is also the power at the DC link can be seen in Figure 6.71. It is visible that the DAB starts supplying power to the grid via the VSC from  $t = 4$  s.



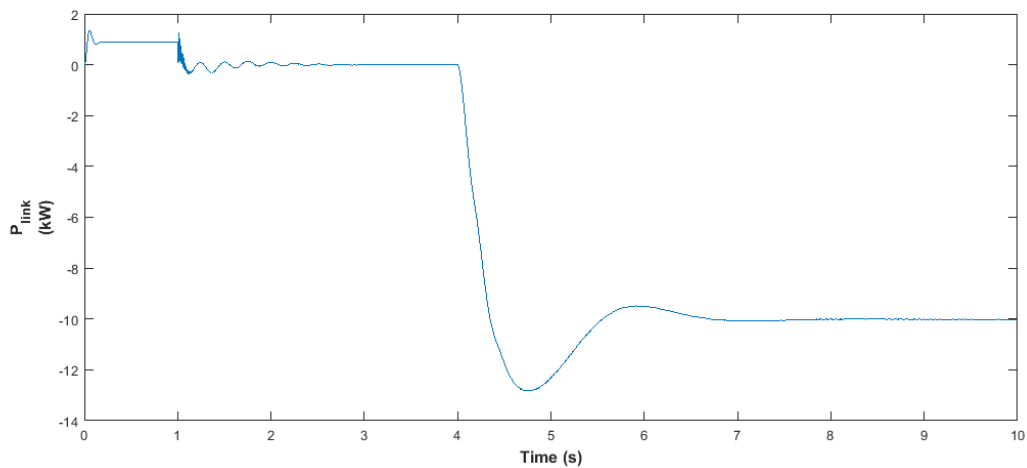
**Figure 6.72 – DAB voltage reference  $V_{ref}$  and actual output voltage  $V_{dc}$  when the DAB supplies power to AC grid.**



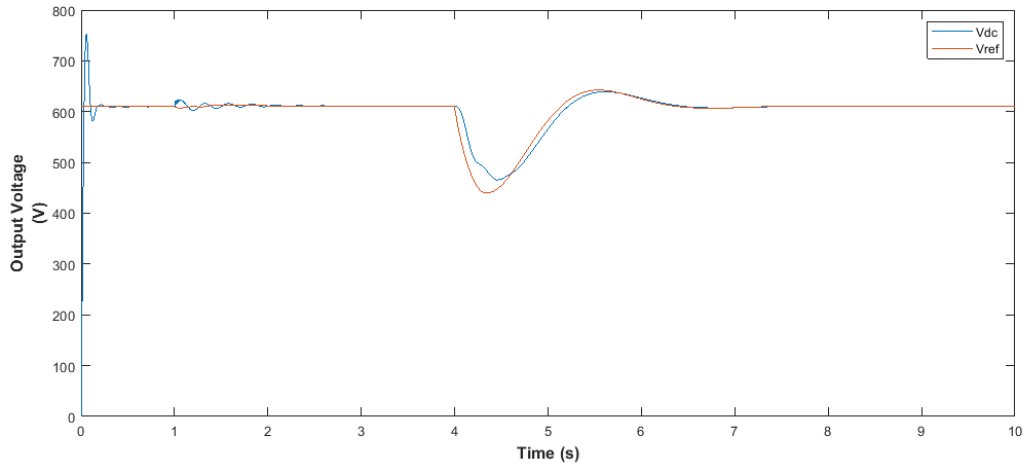
**Figure 6.73 – Grid Power when the DAB supplies power to AC grid.**

The actual output voltage of the DAB  $V_{dc}$  is compared the reference voltage  $V_{ref}$  in Figure 6.72. Also, it must be noted that the grid power in Figure 6.73 becomes negative since the grid is fed from the DAB and absorbs some power.

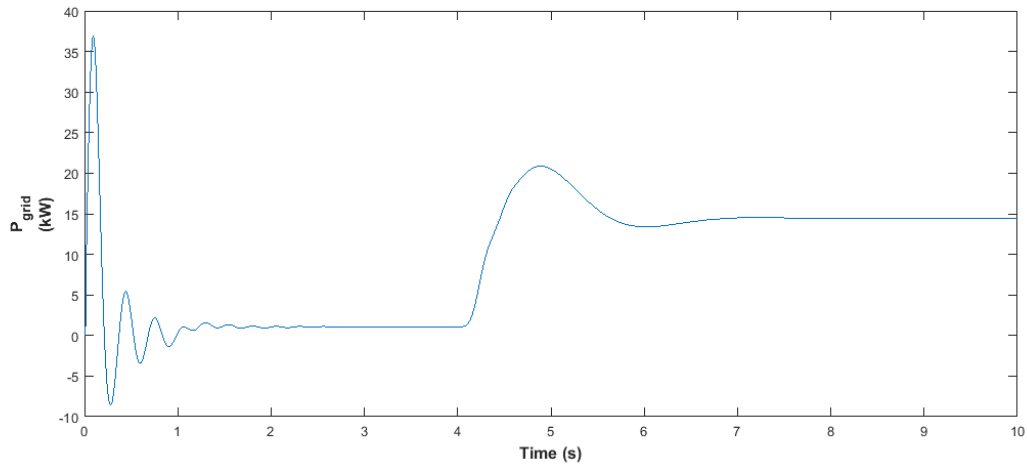
The power reference of the DAB can be changed to inverse the power and feature the case when the battery needs to charge. When  $P_{ref}$  is set so that 10 kW flows into the DAB after 4 s to charge up the battery, the corresponding DC link power is shown in Figure 6.74.



**Figure 6.74 – DC link power when the DAB absorbs some power from the grid.**



**Figure 6.75 – DAB voltage reference  $V_{ref}$  and actual output voltage  $V_{dc}$  when the DAB absorbs some power from the grid.**

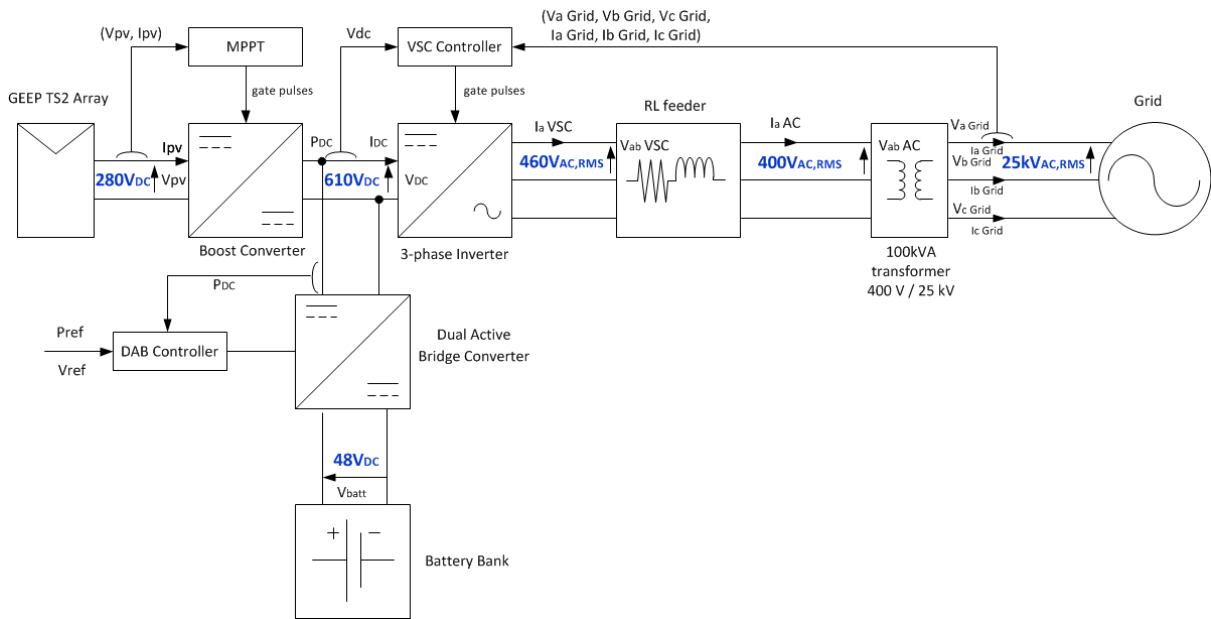


**Figure 6.76 – Grid power when the DAB absorbs some power from the grid.**

It can be seen in Figure 6.74 that the  $P_{link}$  is negative, which indicates that the battery is absorbing some power and thus charging up. As a result, the grid power is positive, which shows that the grid is supplying some power for the operation, as seen in Figure 6.76. The actual output voltage of the DAB and the reference voltage are also shown in Figure 6.75.

#### **6.1.4 PV Power Control**

In order to mitigate the effects of sudden drops of solar irradiance, the architecture shown in Figure 6.77 could be envisaged [79]. Having a bidirectional DC-DC converter with its controller and a battery bank connected to the DC link of the system offers at least two advantages. Firstly, connecting the battery with its converter to the PV link would disturb the MPPT. Secondly, the DC link enables the battery to charge from both the PV array and the grid since some power can flow from the grid to the DC link via the VSC. The complete system is shown in Figure 6.77.



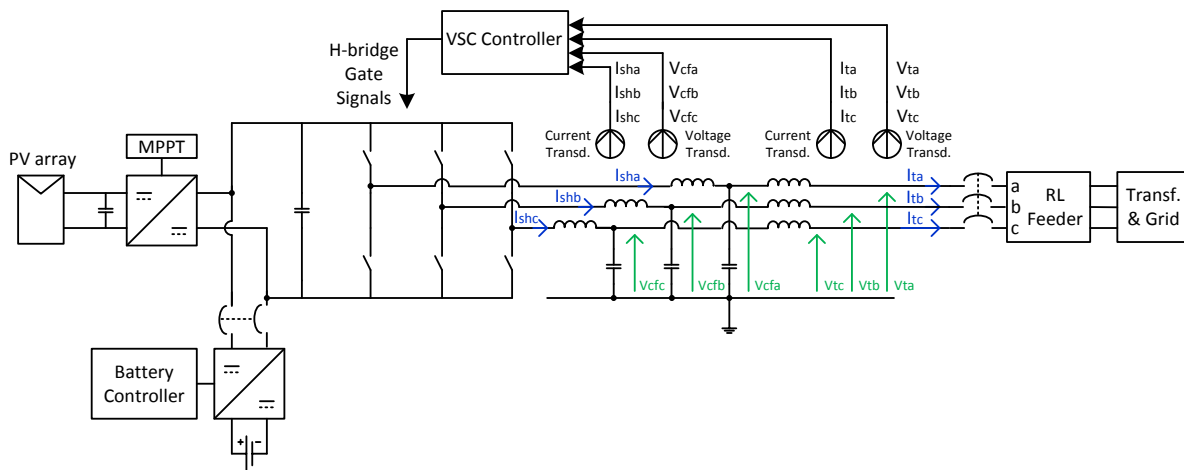
**Figure 6.77 – Complete grid-connected PV system with battery and bidirectional DC-DC converter.**

Forecasting the PV power loss with the method presented in Chapter 4 could enable to anticipate the power losses experienced at the level of the DC link. This could be done by updating the power reference of the DAB controller  $P_{ref}$  with the forecasted difference of power between the clear-sky condition DC link power and its equivalent when a cloud is forecasted to dim the sunlight.

## 6.2 VOLTAGE SOURCE CONVERTER CONTROL

The previously presented PSCAD model of the grid-connected PV system unfortunately features a poor VSC power factor. Hence, in order to improve and control the VSC power factor, a further improvement of the design could consist in optimising the control of the three-phase voltage source converter (VSC) with a topology as shown in Figure 6.78, where the signal on each phase is filtered by an LCL filter [175, 176]. The design could optionally function with a battery and its controller although this is not an absolute necessity. The whole PV system would need to have the option to be isolated from the grid to prevent the system from acting like a load if it does not generate enough power to maintain the VSC output voltage.





**Figure 6.78 – Diagram of the overall system with closed-loop controlled three-phase VSC.**

For an optimal performance, the gates of the switches of the VSC must be driven by a real-time closed-loop controller [177, 178]. For that purpose, some IGBTs with an optocoupler interface could be used to feature the switches of the H-bridge.

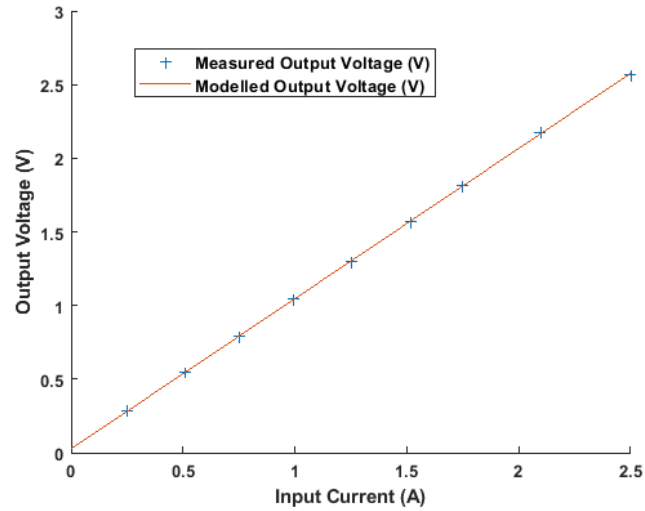
### 6.2.1 VSC Controller Proposed Equipment and Setup

The critical real-time nature of the control implies to make some design choices in terms of hardware equipment as well as in term of software architecture to that implement the proposed VSC controller. Some hardware choices will be discussed in the following part, followed by a presentation of a strategy to code the real-time functionality of the VSC controller, with some results that support that strategy.

#### 6.2.1.1 VSC Hardware

The VSC itself could consist in an H-bridge composed of IGBT switches that can operate at a maximum frequency of 20 kHz, a sampling period between 66  $\mu$ s and 100  $\mu$ s could be considered to refresh both input readings and output commands. More specifically, a sampling period of 80  $\mu$ s has been supposed as suitable and has subsequently verified to appropriate for such an application.

The DSP's embedded 12-bit analogue to digital converter (ADC) could be set up to operate at 12.5 MHz in sequential mode with a 0-3V input voltage range. Some high precision sensors could be used to measure the signals under control such as a LEM LV 25-P [179] to measure the output voltages of the VSC and a LEM HAIS 50-P hall-effect transducer to measure currents [180]. Since the current transducer need a given number of turns to measure the current, the transfer function of that transducer may vary accordingly and needs to be measured manually.

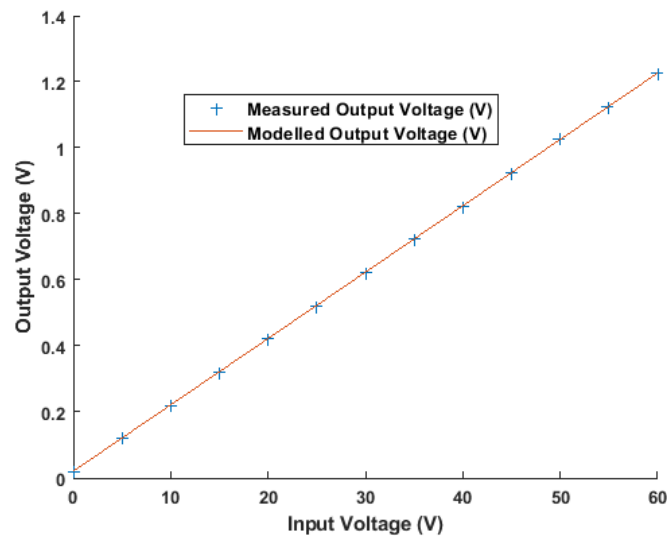


**Figure 6.79 – Characterisation of the current transducer LEM HAIS 50-P.**

For example, if the number of turns is chosen to be 9, then the output voltage of the LEM current transducer with respect to the input current is as shown in the scatter plot in Figure 6.79. Measurements show a non-linearity of 0.13% of the full scale when measuring currents from 0.25 to 2.5 A. Given the fact that the transfer function is indeed very linear, it is safe to approximate it with a linear function like the one given in Equation 6.18. This approximation is characterised by an RMSE of 0.26% a Pearson's correlation coefficient of 0.999992.

$$V_{out} = 1.02 \cdot I_{in} + 0.0264 \quad (6.18)$$

Similarly, the transfer function of the voltage transducer needs to be defined through measurements, as shown in Figure 6.80.

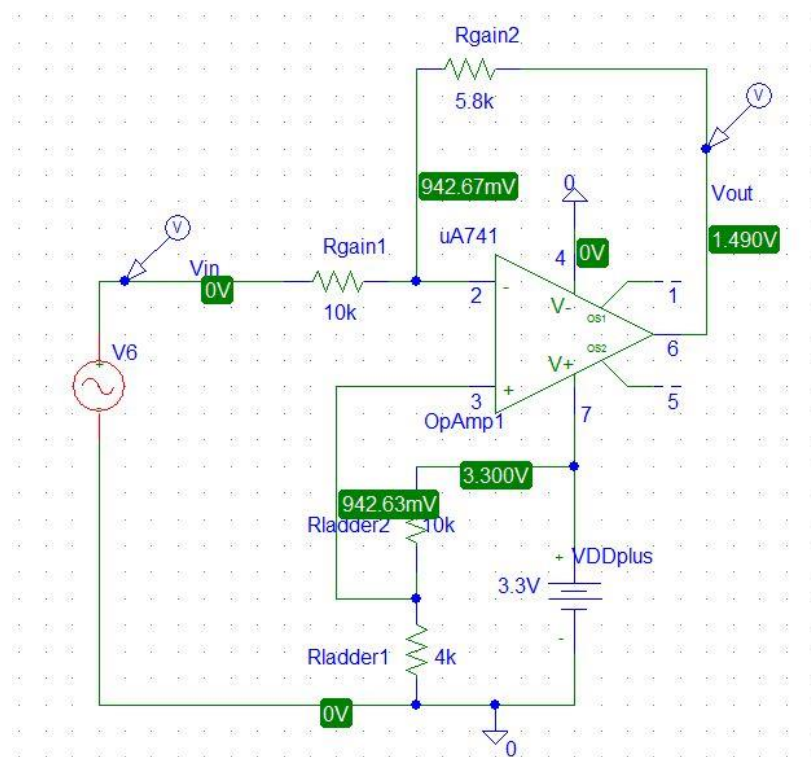


**Figure 6.80 – Characterisation of the voltage transducer LEM LV 25-P.**

The measurements feature a non-linearity is 0.48% of the full scale when measuring voltages ranging between 0 and 60V. Like for the current sensor, it is deemed safe to approximate the transfer function is a linear relationship. The linear approximation from Equation 6.19 is characterised by an RMSE of 4.76% and a Pearson correlation coefficient of 0.999995.

$$V_{out} = 0.0201 \cdot V_{in} + 0.0205 \quad (6.19)$$

It is necessary to condition the voltages corresponding to the measured voltage and the measured current so they can be read in full scale by the ADC of the digital signal processor (DSP). As to guarantee a safe margin, the full scale of the signal has been reconditioned to fit a range [0.25V - 2.75V]. For that purpose, a level shifter has been designed and built for each sensor. The design of both level shifter has been done using PSPICE and the PCB implementation has been made by using KiCad. The resistance have been chosen to match up some standard values for 0.5W resistors. Also, the level shifter have been designed around a general purpose operational amplifier Texas Instruments  $\mu$ A741 [181]. For comparison, Figure 6.81, Figure 6.82 and Figure 6.83 respectively show the PSPICE schematic, KiCad schematic and KiCad layout of the level shifting circuit that is meant to condition the output voltage of voltage sensors. Figure 6.84, Figure 6.85 and Figure 6.86 show respectively the same thing but this time for the level shifter that is intended for current sensors.



**Figure 6.81 – PSPICE schematic and DC analysis of the level shifter for voltage sensors.**

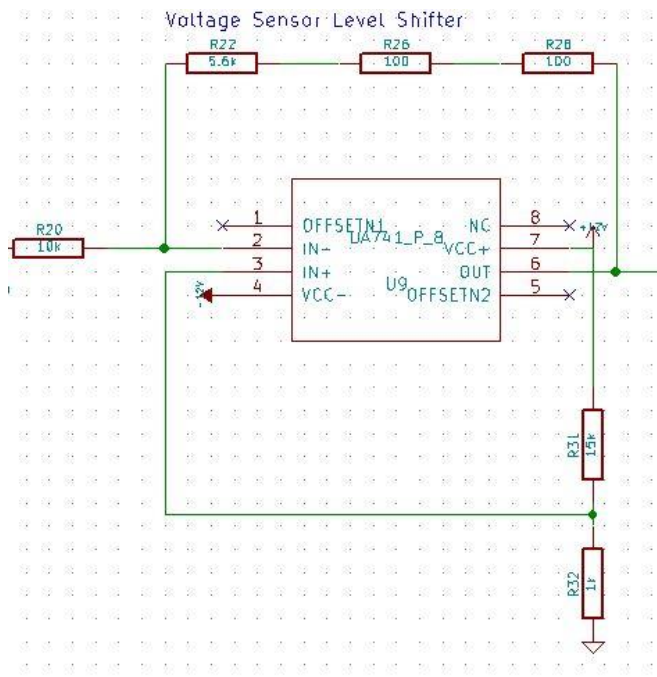


Figure 6.82 – KiCad schematic of the level shifter for voltage sensors.

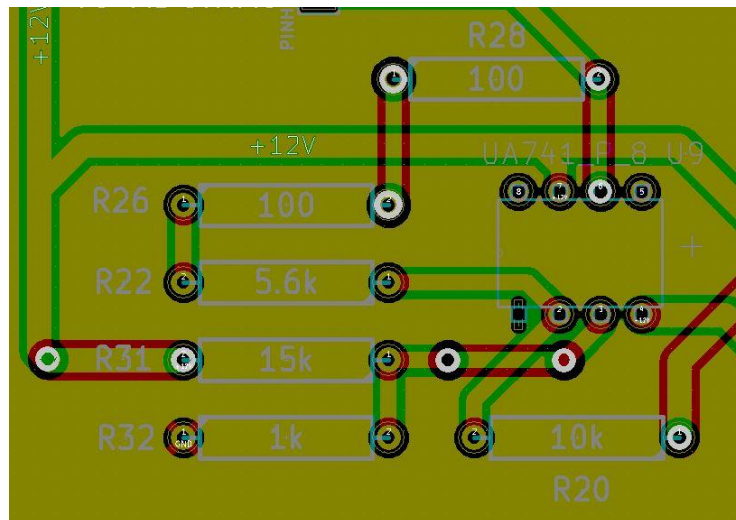


Figure 6.83 – KiCad layout of the level shifter for voltage sensors.

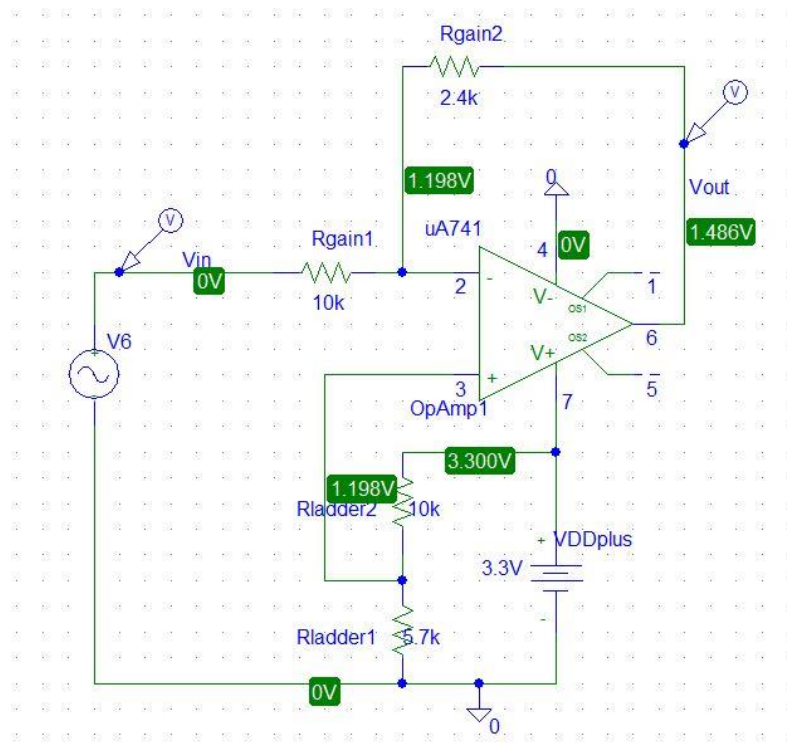


Figure 6.84 – PSPICE schematic and DC analysis of the level shifter for the current sensors.

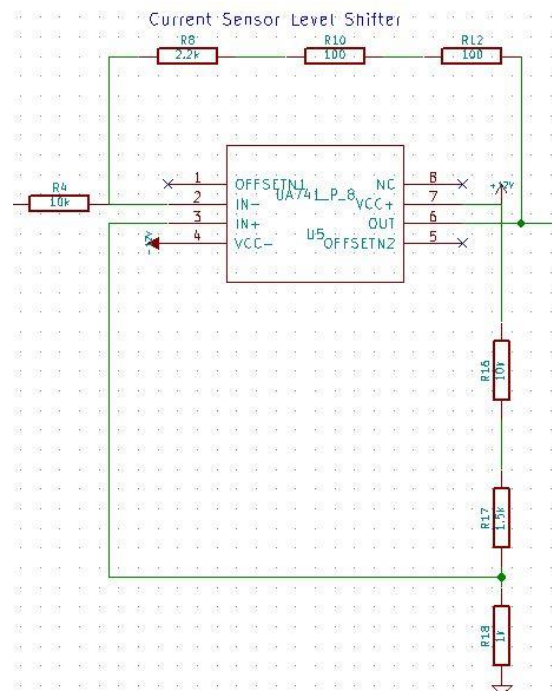


Figure 6.85 – KiCad schematic of the level shifter for current sensors.

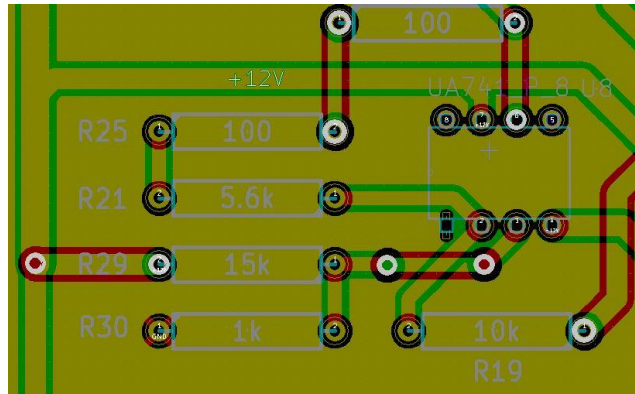


Figure 6.86 – KiCad layout of the level shifter for current sensors.

### 6.2.1.2 VSC Controller Implementation

According to Sepulveda et al [182], both field programmable grid arrays (FPGAs) and digital signal processors (DSPs) can be well suited for real-time control of VSCs and can achieve a nearly similar performance in terms of dc voltage dynamics. This conclusion was drawn by the same team after testing a 20% voltage sag on a 1kVA VSC prototype. After comparing a TMS320F2812 DSP on a Starter Kit eZdsp F2812 with a Spartan-3 XCS1000 FPGA on a Digilent Starter Board, Sepulveda et al have concluded that FPGAs could be recommended for critical applications that require a reduced computation time via a low-level control of hardware resources and through the use of parallel tasks. FPGAs can also be a better choice if a large number of general purpose inputs/outputs (GPIOs) need to be used. However, generally speaking, besides such specific applications, DSPs are a better choice because they allow a much faster development thanks a simplified programming that involves C rather than VHDL. Also, DSPs can give the access to some useful proprietary libraries.

Thus, a digital signal processor (DSP) such as a Texas Instrument TMS320F28335 with a fast modified Harvard architecture could be envisaged as an advantageous option, considering the fact it is a more advanced model than the TMS320F2812.

The real-time C program would heavily rely on the interruption service requests (ISRs) of the user accessible timer *Timer0* every 80μs. An extensive use of lookup tables, macro definitions and built-in functions could also be made to guarantee the real time operation of the program.

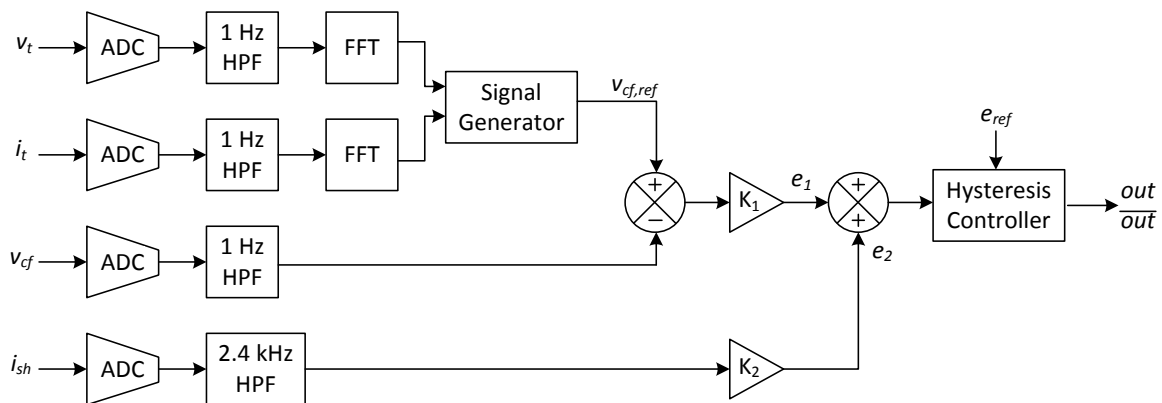
### 6.2.2 VSC Controller Program Testing

As illustrated in Figure 6.78, the input variables of the control system are the following for each phase:

- input current of the LC filter  $i_{sh}$
- voltage across the capacitor of the LC filter  $v_{cf}$
- output voltage of the LC filter and voltage at the point of common coupling with the grid  $v_t$

- output current of the LC filter and current flowing to the point of common coupling with the grid  $i_t$

After being digitised through an ADC, these input signals would need to be sensed and subsequently conditioned through rescaling, digitisation and filtering. For example, in particular, the specific signals  $v_t$ ,  $v_{cf}$  and  $i_t$  could be passed through a 1 Hz high-pass filter (HPF) prior to being processed by the controller to be stripped of their DC component. On the other hand, the high frequency components of the signal  $i_{sh}$ , which effectively corresponds to a noise, could be extracted through a 2.4 kHz HPF and amplified by a gain  $K_2$ . This noise effectively corresponds to a VSC output current error and is noted  $e_2$ . Meanwhile, an FFT analysis of  $v_t$  and  $i_t$  should be done via a wave analyser that would contain a PLL function. The resulting phasors could be used to generate a reference voltage  $v_{cf,ref}$ , to be compared with  $v_{cf}$ . The difference of the comparison is multiplied by a positive gain  $K_1$  and corresponds to the output current error of the LC filter, which is noted  $e_1$ . The sum of the two VSC errors  $e_1$  and  $e_2$  is compared with a reference value  $e_{ref}$  within a hysteresis control block. The output signals of the controller are a pair of complementary commands  $out$  and  $\overline{out}$  that are generated by a hysteresis function with a dead-band controller. Those outputs are fed to the driver circuit of the IGBT switches in the VSC via some intermediate optocouplers. The complete algorithm is shown in Figure 6.87.



**Figure 6.87 – Complete overall architecture of the VSC controller.**

To guarantee an effective real-time functionality of the overall controller, the overall operation including every block must be located within the Timer0 interrupt service request (ISR). The accumulated execution time must not exceed the sampling period of  $80 \mu\text{s}$ , which corresponds to 12,000 clock cycles. For that purpose, the execution time of every single block composing the overall architecture of the controller must be tested and optimised for performance.

### 6.2.2.1 Signal Generator

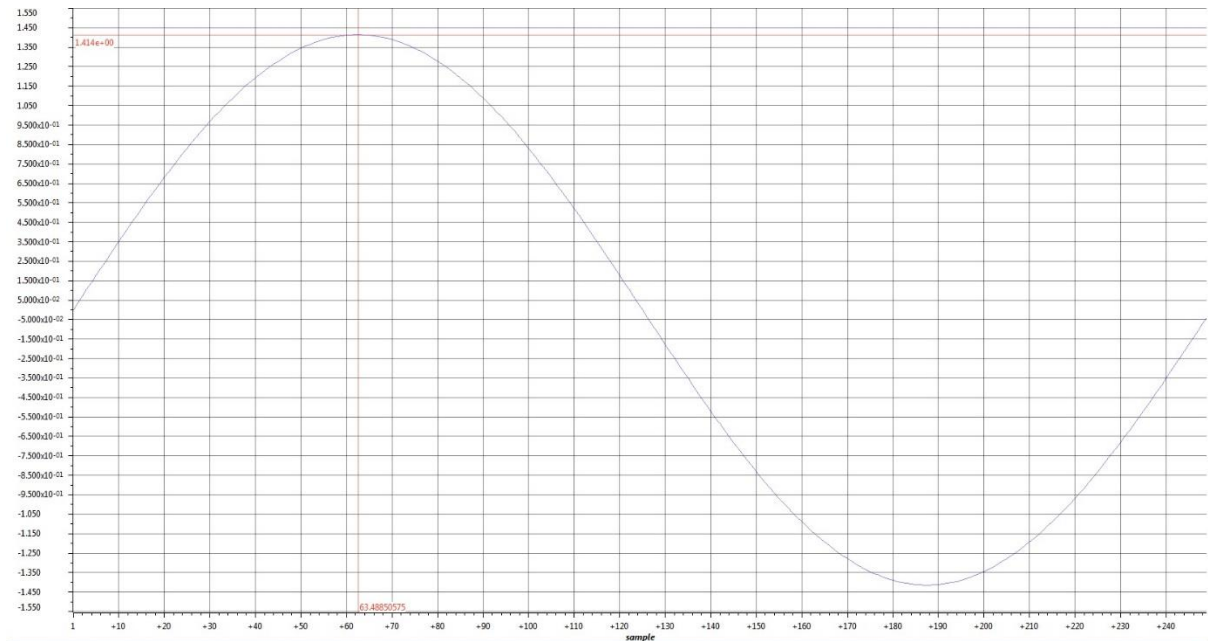
The purpose of the signal generator block is to generate a reference sine wave signal inside the DSP program with an amplitude and frequency that matches up with the desired setpoint.

An excessively large number of mathematical calculations may result in a long execution time. It is thus wise to try to avoid repetitive calculations and prepare the result of some easily foreseeable calculations during the setup stage of the program. Following this idea, a lookup table  $ref[nb]$ , where  $nb$  is the number of samples, is filled with the values corresponding to a full period in the main function of the program. The content of the table is simply read later on whenever needed in Timer0 ISR in order to optimize the execution time of the ISR.

$$ref[i] = V \cdot \sin(2 \cdot \pi \cdot f \cdot t + \varphi) \quad (6.20)$$

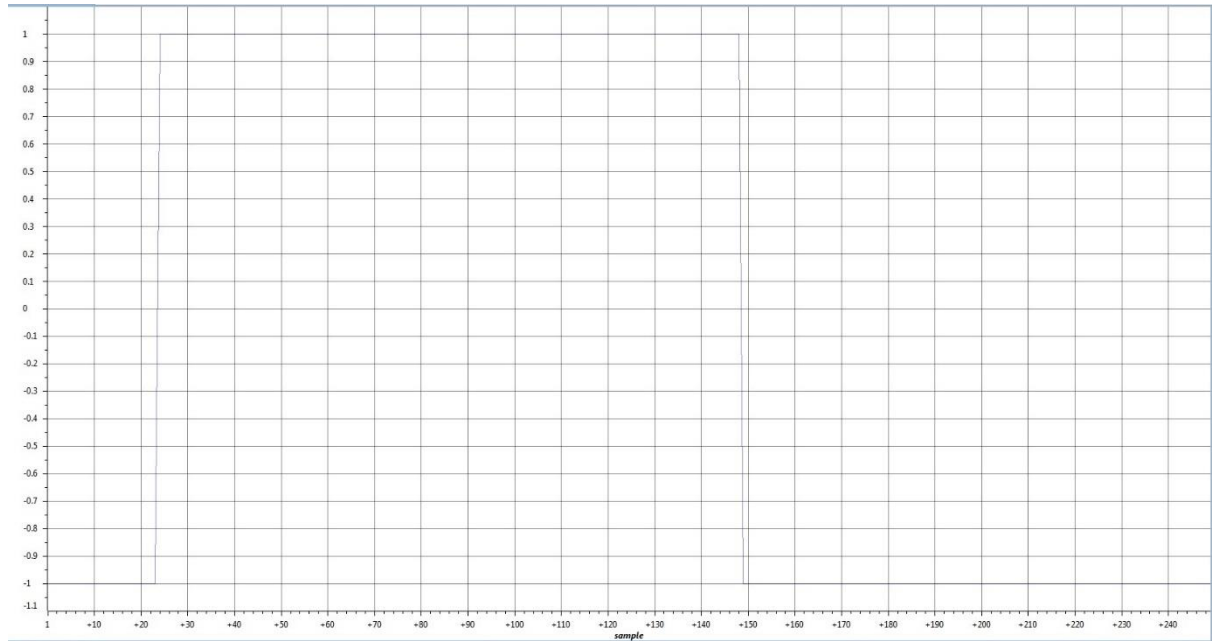
An output flag is then defined inside Timer0 ISR and is set to +1 when the values of the reference signal are greater than a threshold value  $h$  and set to -1 for values smaller than  $-h$ .

With such a strategy, it only takes 52 clock cycles, which is 346ns to run the signal generating function in the ISR. This execution time only corresponds to 3% of the total duration of 80μs.



**Figure 6.88 – Example of reference signal generated by the signal generator block.**





**Figure 6.89 – Output flag signal corresponding to the reference generated signal given in Figure 6.88.**

In Figure 6.88 and Figure 6.89 is shown an example where the gain has been defined as  $V = \sqrt{2}$ , the frequency is  $f = 50\text{Hz}$ , the phase angle is  $\varphi = 0$ , the threshold value is  $h = 0.8$  and the number of samples is  $nb = 250$ .

### 6.2.2.2 FFT Block

The FFT analysis block is actually composed of a few independent sub-blocks. The performance of each block needs to be assessed separately.

#### 6.2.2.2.1 Integration

The purpose of this block is to integrate an analogue input signal after conversion through the ADC of the DSP.

A counter is used to differentiate the initial conversion from the ADC from the further conversions. The integration of a currently sampled value and the preceding sampled value that was sampled during the last cycle is calculated thanks to a trapezoidal approximation. A circular buffer containing two elements is used to store and update the sampled converted values. An additional interrupt service request is used to acknowledge the interruption to sample and convert the signal through the ADC.

In the main function, the ADC and the GPIOs are initialised and set up and a PWM is used to trigger the start of conversion. In Timer0 ISR, when the ADC is ready for conversion, the first reading *Voltage\_VR1* from the ADC is stored. When the second reading *Voltage\_VR2* arrives, the area of the rectangle defined by the average of both values and spaced by the sampling period is calculated and

added to the end result of the integration. Also, after each cycle, the newest value of the previous cycle becomes the oldest value of the new cycle. A counter also keeps track of the number of cycle elapsed.

The resulting execution time needed to run the content of Timer0 ISR function is 21 clock cycles, which corresponds to 140 ns or 1.1% of the period  $T_s = 80\mu s$ .

Expression	Type	Value	Address
(*) sine_wave	unknown	Error: identifier not found: sine_wave	
(*) output	unknown	Error: identifier not found: output	
(*) ptr	unknown	Error: identifier not found: ptr	
(*) integration	float	1.74084	0x0000C020@Data
(*) Voltage_VR1	float	2175.0	0x0000C01C@Data
(*) Voltage_VR2	float	2175.0	0x0000C01E@Data
(*) counter	int	11	0x0000C01A@Data
+ Add new expression			

**Figure 6.90 – Example of result obtained with the integration block.**

The results shown in Figure 6.90 verify the integration formula from Equation 6.21, where counter has to be greater than 1, proving the block to function correctly after 11 iteration cycles.

$$integration = \frac{Voltage\_VR1 + Voltage\_VR2}{2} \times (counter - 1) \times T_s \quad (6.21)$$

#### 6.2.2.2.2 Low-Pass Filter / DC Value Extractor

The objective of this block is to extract the DC average value from an AC input signal after conversion through the ADC of the DSP. To achieve this goal, a circular buffer containing  $nb = 250$  elements is used to store and update the sampled converted values. Once again, in the main function, the ADC and the GPIOs are initialised and set up and a PWM is used to trigger the start of conversion. The implementation of the filter is done in two times. A counter is incremented at each cycle. For as long as the pointer remains smaller than  $nb$ , a value is sample via the ADC the instantaneous partial sum, noted  $sum$ , of all of the values is calculated. When the pointer becomes equal to  $nb$ , it is reset to 0 and the value of the oldest element is replaced by the newest element in the sum as well as in a table used to store all values. If the pointer has reached  $nb$  at least once, the average of the AC signal is calculated as in Equation 6.22.

$$average\_value = \frac{Full\ Scale\ ADC\ input\ voltage}{(2^n - 1) \times nb} \times sum \quad (6.22)$$

The program has been tested with an input sine wave characterised by a peak-to-peak amplitude of 1.02V, an offset DC voltage of 1.028V and a frequency of 50Hz. The time needed to execute this function Timer0 ISR function is 52 clock cycles, which is 346 ns or 2.8% of the period  $T_s = 80\mu s$ .

After 153 iterations, the error proves to be 0.9155%. Changing the frequency of the signal affects the value of this error. As a remedy, the number of samples need to be amended dynamically to match with the frequency of the signal.

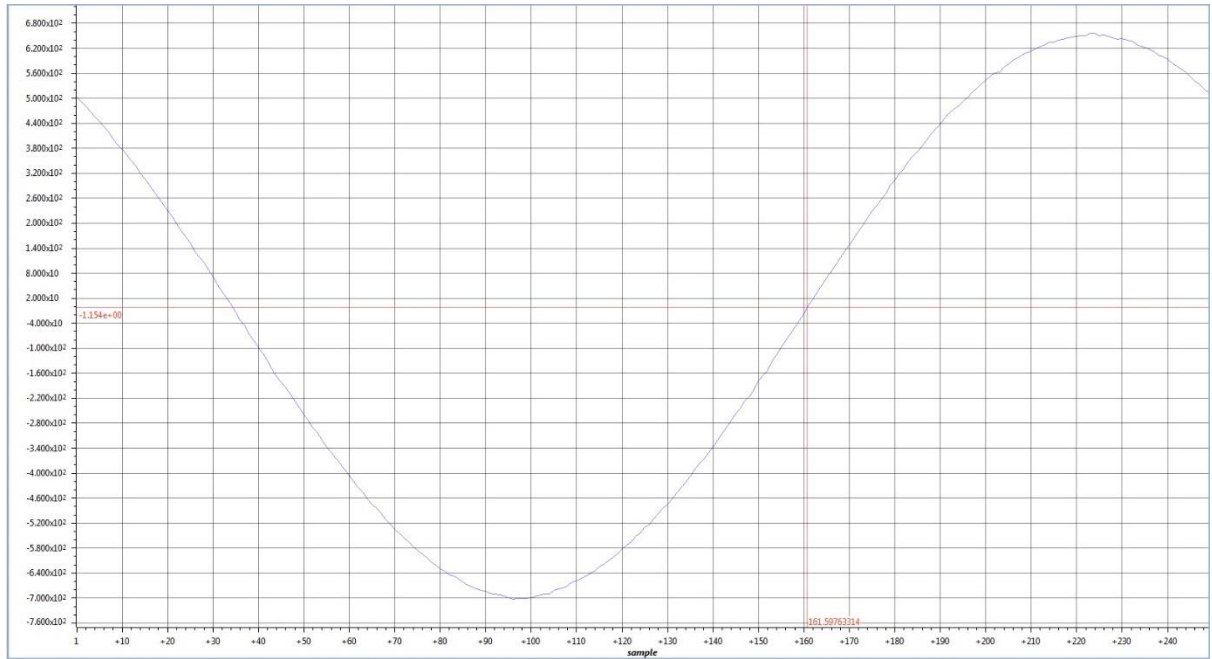
### 6.2.2.2.3 Discrete Fourier Transform (DFT)

This blocks is meant to obtain the magnitude and phase angle of the fundamental harmonic of an input signal after conversion through the ADC of the DSP. The amplitude of the input signal  $X$  is calculated according to the formula shown in Equation 6.23 where  $FS$  represents the full-scale of the ADC input, which is 3V,  $n$  represents the resolution of the ADC and  $N$  represents the number of samples. In order to effectively shrink the execution time in the ISR function of Timer0, the values of the cosine and sine function shown in Equation 6.23 are pre-calculated and stored in two lookup tables prior in the main function of the program to being used during the real-time execution of the Timer0 ISR. The amplitude of the signal is subsequently calculated the length of the diagonal formed by the real value and the imaginary value. The phase of the signal is simply used with a predefined arctangent function.

$$X = \frac{FS}{(2^n - 1) \times N} \times \sum_{n=0}^{N-1} x_n \cdot e^{-i2\pi n/N} = \frac{FS}{(2^n - 1) \times N} \times \sum_{n=0}^{N-1} x_n \cdot \left( \cos\left(\frac{2\pi n}{N}\right) - i \cdot \sin\left(\frac{2\pi n}{N}\right) \right) \quad (6.23)$$

The same input signal as for the previous block has been used for testing purposes. This signal is visible in Figure 6.91.

The maximum execution time needed to run the content of Timer0 ISR function has been measured as 872 clock cycles, which is 5.8  $\mu$ s, or 7.2% of  $T_s = 80\mu$ s. Like for the previous block, all times, the number of samples needs to be dynamically amended to match with the frequency of the signal.



**Figure 6.91 – Input signal used to test the DFT block.**

An error of 4.66% has been found regarding the measurement of the amplitude of the signal and an error of 1.01% regarding the measurement of the phase of the signal.

### 6.2.2.3 High Pass Filter

A high pass filter block is required to retrieve the high frequency component of an input signal after conversion through the ADC of the DSP. A digital filter is needed for that purpose. This means that a typical Laplace domain transfer function for high pass filter like the one shown in Equation 6.24 needs to be converted to an equivalent discrete time transfer function. This can be done by using the binomial approximation shown in Equation 6.25 where  $a = \omega_c = 15$  kHz is the cut-off frequency of the filter and  $T_s = 80\mu s$  is the sampling period. This results in z-domain transfer function from Equation 6.26.

$$H(s) = \frac{s}{s + a} \quad (6.24)$$

$$s = \frac{2}{T_s} \frac{z - 1}{z + 1} \quad (6.25)$$

$$H(z) = \frac{1}{a \cdot \frac{T_s}{2} + 1} \cdot \frac{z - 1}{1 + \frac{a \cdot \frac{T_s}{2} - 1}{a \cdot \frac{T_s}{2} + 1}} \quad (6.26)$$

Two circular buffers, one for the input of the digital filter and another for the output, are used to implement the filter iteratively by using the following pseudo-code within Timer0 ISR:

**step 1.** (initialisation)

**If** this is the first iteration **then:**

Write the current value of the input signal at the position 0 of a two-element input circular buffer  $x$

Write the current value of the input signal at the position 0 of a two-element output circular buffer  $y$

**Else:**

**step 2.** (running the filter)

**If** the input array is not full yet **then:**

Write the current value of the input signal at the position 1 of  $x$

Calculate the output value as  $y[1] = a0*x[1] + a1*x[0] + b1*y[0]$

**Else:**

**step 3.** (reset the filter)

**If** the input array is full **then:**

Reinitialise all pointers

Overwrite the value of the oldest element  $x[0]$  of the input buffer with the newest value  $x[1]$

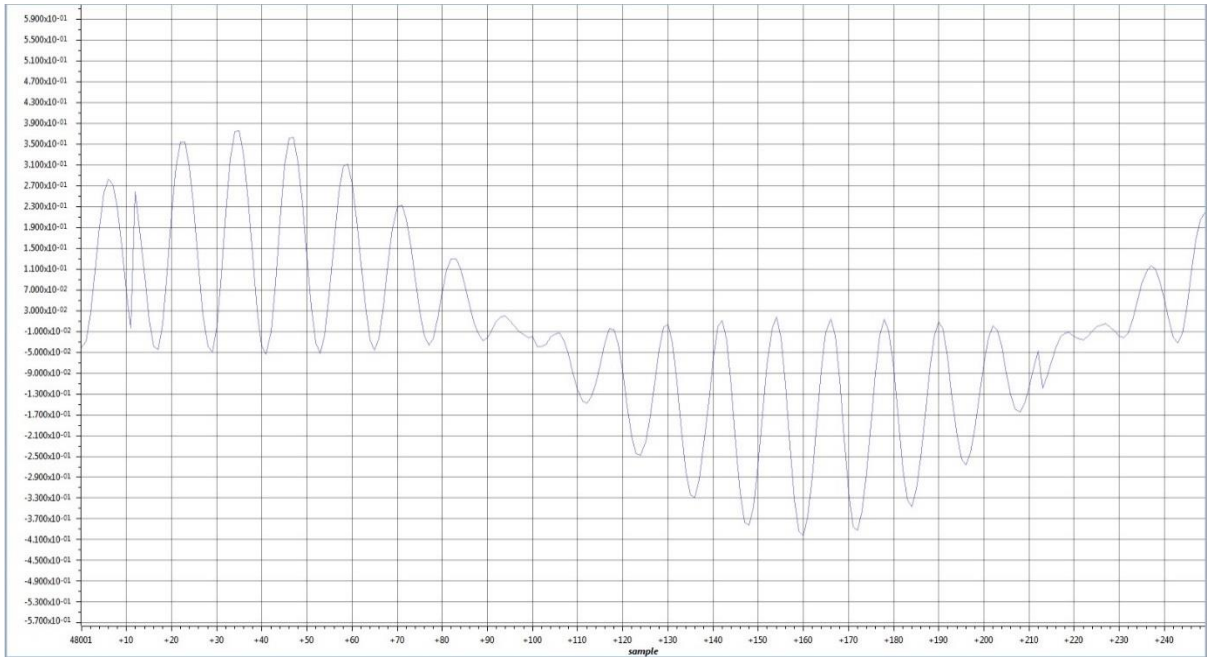
Overwrite the value of the oldest element  $y[0]$  of the input buffer with the newest value  $y[1]$

Write the current value of the input signal at the position 1 of  $x$

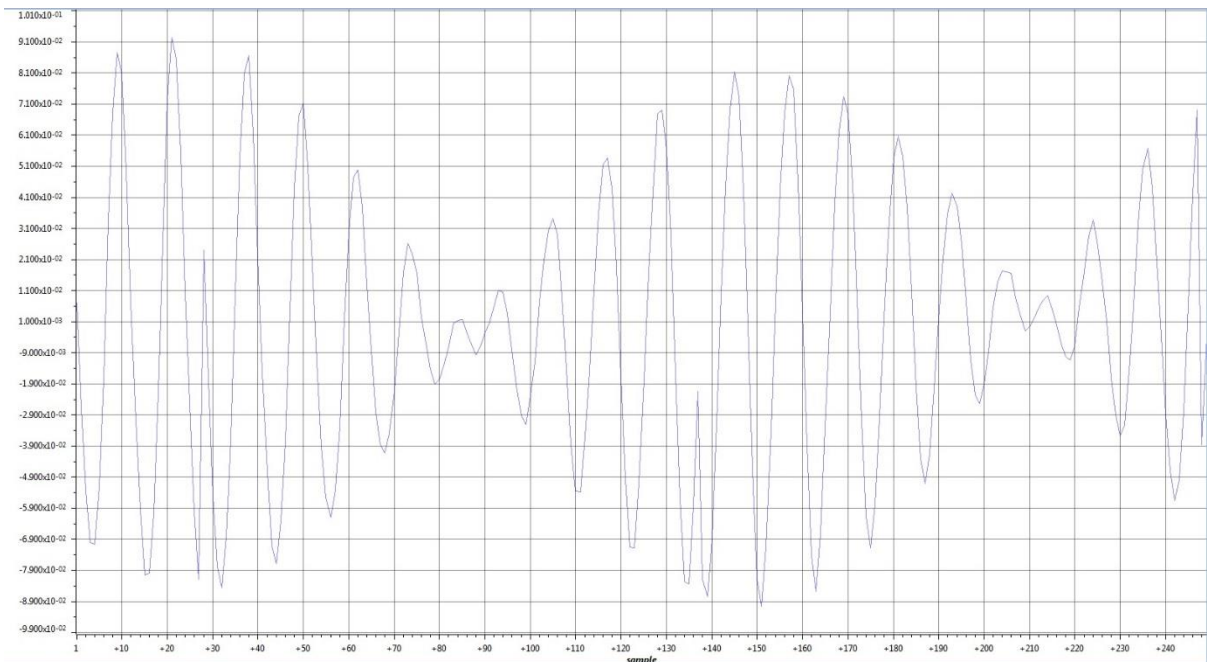
Update the output value as  $y[1] = a0*x[1] + a1*x[0] + b1*y[0]$

The execution time needed to run the content of Timer0 ISR function is 317 clock cycles, which is 2.1  $\mu\text{s}$  or 2.6% of  $T_s = 80\mu\text{s}$ .

When a modulated sine wave is applied at the input of the filter like in Figure 6.92, then the high frequency carrier can be retrieved, like in Figure 6.93.



**Figure 6.92 – Sine modulated input signal.**



**Figure 6.93 – Retrieved high frequency component of the sine modulated input signal.**

#### **6.2.2.4 Complete Real-time program with all blocks together**

After optimising the execution time of every block that is necessary for the controller, the real-time operation of the complete program is tested, excluding the droop and power calculation features. The content of all the blocks is included within the same Timer0 ISR while all necessary lookup tables are pre-calculated in the main function. The maximum execution time needed to run the content of Timer0

ISR function with all the blocks inside it is 1422 clock cycles, which is 9.5  $\mu$ s. Since this only corresponds to 11.8% of the 80 $\mu$ s sampling period, this leaves a very large safe margin for some further features.

### 6.3 CONCLUSIONS

An accurate grid-connected PV system model has been designed with MATLAB/SIMULINK. The accuracy of this model has been verified against some actual measurements taken on a real system of similar size. However, an equivalent model has been made with PSCAD with some equivalent voltage and active power levels as the verified SIMULINK model to study the effect of sunlight drops with various dimming factors. It has been found that the PV array output voltage does not significantly change for dimming factors between 10% and 30% of an initial DNI with a cell temperature of 35°C. The output PV array voltage only starts significantly decreasing when the irradiance is below 770.5 W/m<sup>2</sup>. However, the PV output current and output power drop along with the DNI. In the meantime, the DC voltage can still be maintained and the grid can still be fed by the PV array until the DNI drops by about 80% or more of its initial value. In that case, the output power of the PV array is less than 0.3 kW, which is insufficient to keep the DC link voltage constant. As a result, the grid has to supply the necessary power and the PV array acts like a load.

To remedy to this problem, compensating for the drop of PV output power has been thought by using a battery bank. Although a realistic bank model has been realised in SIMULINK, the overall strategy has been partially implemented by using PSCAD. A bidirectional DC-DC converter such as a dual active bridge is recommended on the DC link of the system. Such a DAC has been designed and tested both for charging and discharging the battery.

Also, a controller has been proposed to further control the VSC of the system and improve the overall dynamic performance. Several blocks of this controller have been tested individually to ensure the good real-time functionality of the overall embedded program. Eventually, the optimisation of each individual block has guaranteed a large safe margin in term of execution time of the whole program.

## 7 CONCLUSIONS AND FUTURE WORKS

---

### 7.1 CONCLUSIONS

#### 7.1.1 Conclusions about the Method

A Bayes filter based forecasting method has been proposed and tested for the purpose of very-short term cloud tracking. This method has been combined with a clear-sky regression model which has been obtained from empirical observations, after thorough comparison of a large number of existing models. In addition, a clear-sky library has been constituted and a method has been proposed to distinguish with high accuracy clear-sky pixels and cloudy pixels, accordingly with the chosen camera. Also, a full calibration methodology has been demonstrated to locate accurately the sun and measure angular distance with respect to the centre of the solar disk.

The forecasting method has been successfully verified with a simplified yet realistic dataset of images of the sky taken at regular intervals with a sky camera. The comparison of the proposed PHD/CPHD method with optical flow has highlighted the speed and accuracy of our approach. Since both speed and accuracy are critical with a very short-term timeframe, this gives a decisive advantage to the proposed method.

A simulation model has been designed to feature a small scale grid-connected PV array that has been verified against a real system. The effect of sudden sunlight disruptions have been thoroughly investigated on the PV system model. A battery-based power smoothing method is proposed with a bidirectional DC-DC converter. The method has been partially tested through simulation. It is envisaged to use the forecasting method to anticipate the DC link power drop and schedule an adequate amount of power to flow from the battery to the DC link at a scheduled moment as to maintain the DC link power constant. Also, in order to improve the performance of the VSC, a controller has been proposed. A hardware implementation of this controller has been done and a real-time embedded program has been written and tested.

### 7.2 FUTURE WORKS

#### 7.2.1 Full Validation of the Grid-Connected PV System with Battery Storage

Although a full system has been designed to mitigate the effects of sunlight disruption, this system could not be appropriately tested since it requires a lot of fine tuning. Also, many blocks of the system have been designed in two different environment, namely MATLAB/SIMULINK and PSCAD, due to the fact that each presents their own advantages that cannot be found in the other. Some further



modelling and simulations thus need to be done in order to assess the efficiency of the method. Additionally, a hardware implementation could be done to finalise the validation of the method and reassess its efficiency.

### **7.2.2 Further Testing of the Forecasting Method**

Some additional work would be required to improve the preliminary task of pre-processing sky images in order to discriminate efficiently clear-sky from cloudy pixels in the presence of strong scattering effect, in particular in the circumsolar region. A model could be developed on the basis of the clear sky library as to anticipate the chromatic values of the clear-sky pixels at any given point of the sky so that pixels that cannot possibly be some clouds could be correctly classified as clear-sky pixels. In addition, a deeper separate analysis of the circumsolar region of the sky could be carried out and a combination of several criteria could be used to improve the accuracy of the pixel discrimination. A particular stress could be given to distinguishing solar beams from overexposed nearby clouds. In addition to this, a fast and high resolution de-fisheying method would need to be used in order to extend the horizon of forecast while maintain the speed of the forecasting method.

### **7.2.3 More Accurate Clear-Sky Irradiance Modelling**

Since the accuracy of the very-short term PV power forecast depends heavily on the accuracy of the clear-sky irradiance, a comprehensive comparison of simple empirical clear-sky irradiance models readily available in literature has been carried out. A couple of regression models have been inferred from the two most accurate empirical models with respect to the global horizontal irradiance observed in Western Australia.

However, a significantly more accurate model could be investigated by using a more reliable source of data and a better method, as investigated by Engerer and Mills [183].

### **7.2.4 Automation of the Calibration**

A manual method has been proposed to calibrate the rotation matrix of the camera. Also, another manual method has been shown to choose a criterion and threshold values that are needed to apply to tell apart clear-sky and cloudy pixels. However, both of these manual methods require the selection of a relatively large number of images representing the sky in clear conditions. Instead of manually visualising and sorting sky images, some existing filters and methods could be applied to classify each sky image between the frames containing cloud and aerosols and the frames that are entirely free of them. Among these clear-sky detection methods, it is worth citing the Long-Ackerman filter [184, 185], Reno's time series-based clear sky period identification method [186], Inman's method [187], and Djafer's approach [188].

### **7.2.5 Application to a Large Solar Farm**

A large solar farm such as a photovoltaic farm with a 500 MW installed capacity is its size. According to NREL SAM [189], such farm may include around 64,500 monocrystalline modules, may have a total module area of about 105,200 m<sup>2</sup> and may stretch over 0.35 km<sup>2</sup> where the row spacing is considered to be 11 m. This means that passing broken cumulus clouds are very likely to create partial shading over some arrays or modules or the farm, simultaneously at various geographical locations. This may result either in significant reverse generation if the shaded strings are not short-circuited or in significant power generation losses and power oscillation if they are short-circuited by by-pass diodes. Thus, forecasting the upcoming position of the partial shades over solar modules could significantly optimise power generation and simultaneously avoid problems that may affect the quality of the power output. Our method could be used to achieve those goals in the case of a solar farm that does not contain any form of energy storage. If a battery storage or supercapacitors are locally available on the site of the farm as part of the farm, our method could be used to manage the energy storage and to plan when to charge it in order to smooth the power output and improve its quality.

### **7.2.6 Application to a Microgrid and Optimisation of Local Energy Resources**

To the upstream network, a microgrid (MG) looks like a controlled coordinated system of distributed energy resources (DERs), with three levels of control (local controllers, MG central controller (MGCC), and an inter-MG coordination controller). However, operating a MG in grid-connected mode leads to a problem of voltage control, which can be locally addressed through a power or voltage droop control. In islanded mode, the main concern becomes frequency control. A MG that gets disconnected from the main grid will experience a frequency deviation and will not reach its nominal frequency through the primary reaction of the droop control. This will require the secondary control to dispatch some power flow to correct the frequency deviation. In addition, an islanding MG will need an energy buffer in the form of batteries or supercapacitors to ensure an initial load-generation balance [190]. Thus, the relevance of very short-term solar forecasting and cloud tracking could be demonstrated as an improvement for the control of an MG, particularly at the moment when an MG gets disconnected from the grid.

Also, the approach used in this study could be applied to optimise the energy generation and storage, as well as minimise the cost and ecological impact of an MG that integrates PV generation. Our method could be used for forecasting horizon corresponding to the time needed for costly and polluting resources such a diesel generator to ramp up. This means that applying our method with the range of several minutes, in order to automate the decision of igniting or not a generator on the basis of the forecasted cloud cover.

Another possible application within an MG would consist in positioning cameras in regular intervals in residential areas in order to cover a significant part of the sky and forecast the cloud cover over a residential block. The advantage provided by this information would be to help the local DERs to coordinate one another and avoid load shedding in the case of prolonged dimmed sunlight, as well as excessive power oscillations in the case of a shading episode that would only affect one specific location in the block for a limited period of time. The application would be similar to the previously mentioned possible application to a large solar farm although on a smaller scale.

### **7.2.7 Hardware implementation**

Unfortunately, the camera that has been used has a limited lifetime, due to the fact that it necessarily utilises a physical shutter. This is motivated by the necessity to use proper exposure settings that can be compatible with shooting photos of the sun directly. However, such photos were a timelapse of single shots while it could be extremely useful to use the camera in video mode and extract the frames every second from the video feed. In order to capture useful information on the video feed, the driver of the camera would need to be overwritten though.

Another problem faced by the camera is the fact that the dome of the camera gets very quickly dirty and requires regular cleaning. Most of the dirt on the dome gets deposited on the outside of the dome and comes from the accumulation of dust particles sticking to rain drops after it rained. Thus, it would be a significant practical improvement if an automated cleaning could be programmed on the dome on a regular basis, with the option of remotely triggering the cleaning action.

### **7.2.8 Transactions**

The use of the forecasting method presented in this study could possibly be extended to helping the real-time control of energy transactions in a cluster of interconnected microgrids [191].

### **7.2.9 Scada Communications**

The forecast resulting from the proposed method could be used to improve the centralised management of grid-connected renewable energy resources through SCADA communications. The communication protocols in use by utility companies could be investigated and integrated into the overall method to exchange some information about excess generation and battery state of charge.

### **7.2.10 Composite Method**

The temporal and spatial resolution of our very-short solar power forecasting method could be broadened by combining it with a longer-term approach involving a larger area, such as satellite imagery or sensor networks.

## APPENDIX A – FORECASTING METHOD ASSESSMENT

The detailed results of the assessment of the accuracy of the four criteria that have been considered, across all 30 images of the dataset are shown in Table A.1.

**Table A.1 – Threshold values that have been found to minimize cloud detection error for every image and in average.**

<i>Image # k</i>	<i>Th<sub>1</sub></i>	<i>Th<sub>2</sub></i>	<i>Th<sub>3</sub></i>	<i>B/R Threshold</i>
<b>1</b>	21	10	56	0.77
<b>2</b>	21	9	55	0.77
<b>3</b>	21	9	55	0.77
<b>4</b>	22	12	57	0.77
<b>5</b>	22	12	56	0.77
<b>6</b>	22	12	56	0.77
<b>7</b>	22	12	57	0.77
<b>8</b>	22	12	57	0.77
<b>9</b>	22	0	56	0.77
<b>10</b>	22	0	56	0.77
<b>11</b>	22	0	57	0.77
<b>12</b>	22	0	57	0.77
<b>13</b>	22	0	57	0.77
<b>14</b>	21	0	55	0.77
<b>15</b>	22	0	57	0.77
<b>16</b>	21	0	56	0.77
<b>17</b>	21	0	56	0.77
<b>18</b>	21	1	55	0.77
<b>19</b>	21	1	55	0.77
<b>20</b>	21	1	55	0.77
<b>21</b>	21	1	55	0.77
<b>22</b>	21	9	55	0.77
<b>23</b>	21	9	55	0.77
<b>24</b>	20	10	55	0.77
<b>25</b>	21	10	55	0.77
<b>26</b>	21	10	55	0.77
<b>27</b>	21	10	55	0.77
<b>28</b>	20	10	55	0.77
<b>29</b>	20	10	55	0.77
<b>30</b>	20	10	55	0.77
<i>Average Values</i>	<b>21</b>	<b>6</b>	<b>56</b>	<b>0.77</b>

The training time for the simple Bayes filter has been measured as 79.2189 s. The time needed to assess each criterion is shown in Table A.2. Some additional comparison data is given in Table A.2 – A.6.

**Table A.2 – Time needed to evaluate each criterion based on the picture of every step.**

<i>Step k</i>	<i>R/B Criterion (s)</i>	<i>R-B Criterion (s)</i>	<i>Multicolour (s)</i>	<i>Bayes (s)</i>
<b>1</b>	0.0034227	0.0023884	0.0092603	10.3218
<b>2</b>	0.0033459	0.0028905	0.0096051	10.5962
<b>3</b>	0.0029099	0.0025214	0.0102	11.098
<b>4</b>	0.0038545	0.0022158	0.0097275	10.0274
<b>5</b>	0.0031326	0.0029977	0.012019	10.5148
<b>6</b>	0.0034341	0.0025043	0.0093417	10.018
<b>7</b>	0.0033322	0.0024545	0.010539	10.503
<b>8</b>	0.0031326	0.002464	0.010366	10.2577
<b>9</b>	0.0032782	0.0024317	0.012304	10.0857
<b>10</b>	0.0033695	0.0025594	0.0099936	10.1711
<b>11</b>	0.0033653	0.0026305	0.0094169	10.0826
<b>12</b>	0.0030787	0.0026126	0.0087707	10.5936
<b>13</b>	0.0037948	0.0027297	0.0092679	10.1019
<b>14</b>	0.0030661	0.0025237	0.0093675	10.1342
<b>15</b>	0.00293	0.0022853	0.0087529	10.5442
<b>16</b>	0.003347	0.0027871	0.0086225	10.3953
<b>17</b>	0.0045011	0.0030228	0.0095454	10.3543
<b>18</b>	0.003547	0.0029179	0.0090908	10.0867
<b>19</b>	0.0032638	0.002667	0.0098986	9.9484
<b>20</b>	0.00398	0.002261	0.011863	10.0392
<b>21</b>	0.0088528	0.0030669	0.0041362	10.4902
<b>22</b>	0.0034527	0.0022378	0.011749	10.3167
<b>23</b>	0.009233	0.0029779	0.008968	12.7337
<b>24</b>	0.0030798	0.0026354	0.011666	11.4509
<b>25</b>	0.0060479	0.0027955	0.0095728	11.6827
<b>26</b>	0.0035671	0.003753	0.026002	12.5649
<b>27</b>	0.0033208	0.0024309	0.0088304	12.5408
<b>28</b>	0.0032767	0.0026058	0.012285	11.033
<b>29</b>	0.0099655	0.0029247	0.0038279	11.0017
<b>30</b>	0.01018	0.0030152	0.0030426	10.6643
<b>Average</b>	<i>0.004302</i>	0.002677	0.009934	10.67843

Table A.3 – Relative classification error expressed as a percentage (%) for each criteria.

<i>Image #</i> <i>k</i>	<i>B-R</i>		<i>Multicolour</i>		<i>R/B</i>		<i>Simple Bayes Filter</i>	
	<i>Sky pixels</i>	<i>Cloud pixels</i>	<i>Sky pixels</i>	<i>Cloud pixels</i>	<i>Sky pixels</i>	<i>Cloud pixels</i>	<i>Sky pixels</i>	<i>Cloud pixels</i>
<b>1</b>	0.24	2.34	0.24	2.33	0.09	0.89	2.92	28.63
<b>2</b>	0.23	2.26	0.23	2.25	0.08	0.77	2.79	27.09
<b>3</b>	0.26	2.46	0.26	2.47	0.09	0.87	2.79	26.78
<b>4</b>	0.25	2.35	0.25	2.35	0.10	0.93	2.79	26.61
<b>5</b>	0.24	2.28	0.24	2.27	0.09	0.83	2.80	26.37
<b>6</b>	0.24	2.35	0.25	2.39	0.09	0.90	2.69	25.84
<b>7</b>	0.22	2.08	0.22	2.08	0.08	0.75	2.69	25.36
<b>8</b>	0.25	2.33	0.25	2.36	0.09	0.85	2.69	25.14
<b>9</b>	0.23	2.18	0.23	2.18	0.07	0.64	2.54	23.89
<b>10</b>	0.26	2.42	0.26	2.43	0.10	0.95	2.55	23.39
<b>11</b>	0.22	2.01	0.22	2.02	0.07	0.65	2.48	22.94
<b>12</b>	0.24	2.17	0.25	2.25	0.08	0.70	2.42	22.14
<b>13</b>	0.24	2.20	0.24	2.20	0.08	0.70	2.39	21.88
<b>14</b>	0.21	1.98	0.21	1.99	0.07	0.66	2.21	20.69
<b>15</b>	0.21	1.93	0.21	1.93	0.08	0.68	2.22	20.14
<b>16</b>	0.21	1.90	0.21	1.91	0.06	0.59	2.22	20.38
<b>17</b>	0.24	2.17	0.24	2.18	0.08	0.71	2.22	20.11
<b>18</b>	0.21	1.95	0.21	1.95	0.06	0.54	2.19	20.01
<b>19</b>	0.22	1.97	0.22	1.97	0.06	0.55	2.19	19.78
<b>20</b>	0.23	2.06	0.23	2.06	0.07	0.65	2.19	19.61
<b>21</b>	0.25	2.26	0.26	2.26	0.08	0.71	2.19	19.44
<b>22</b>	0.22	1.98	0.22	1.98	0.06	0.52	2.17	19.33
<b>23</b>	0.23	2.02	0.23	2.02	0.07	0.66	2.13	19.15
<b>24</b>	0.18	1.59	0.18	1.59	0.05	0.47	2.13	19.32
<b>25</b>	0.19	1.76	0.20	1.77	0.05	0.43	2.09	18.97
<b>26</b>	0.23	2.06	0.23	2.05	0.08	0.70	2.09	18.84
<b>27</b>	0.22	1.96	0.22	1.95	0.07	0.60	2.09	18.79
<b>28</b>	0.18	1.64	0.18	1.62	0.05	0.45	2.08	19.00
<b>29</b>	0.20	1.80	0.20	1.80	0.06	0.52	2.08	18.79
<b>30</b>	0.21	1.88	0.21	1.87	0.06	0.57	2.08	18.68
<b>Average</b>	0.23	2.08	0.23	2.08	0.07	<b>0.68</b>	2.37	21.90

**Table A.4 – Actual and estimated total cloud cover (TCC) expressed as a percentage (%) for each image of the dataset.**

<i>Image # k</i>	<i>Actual TCC (%)</i>	<i>(R/B) TCC (%)</i>	<i>(B-R) TCC (%)</i>	<i>Multicolour TCC (%)</i>	<i>Simple Bayes Filter (%)</i>
<b>1</b>	9.26	9.20	9.12	9.13	11.12
<b>2</b>	9.33	9.29	9.21	9.21	11.19
<b>3</b>	9.44	9.38	9.34	9.35	11.23
<b>4</b>	9.50	9.43	9.41	9.42	11.28
<b>5</b>	9.58	9.53	9.51	9.51	11.34
<b>6</b>	9.41	9.35	9.34	9.34	11.22
<b>7</b>	9.59	9.55	9.54	9.54	11.29
<b>8</b>	9.68	9.62	9.60	9.60	11.33
<b>9</b>	9.61	9.57	9.56	9.56	11.25
<b>10</b>	9.81	9.74	9.73	9.73	11.33
<b>11</b>	9.75	9.71	9.71	9.71	11.31
<b>12</b>	9.85	9.80	9.78	9.78	11.39
<b>13</b>	9.85	9.80	9.79	9.79	11.36
<b>14</b>	9.66	9.62	9.56	9.56	11.28
<b>15</b>	9.92	9.88	9.87	9.88	11.36
<b>16</b>	9.81	9.78	9.72	9.72	11.33
<b>17</b>	9.93	9.89	9.81	9.81	11.43
<b>18</b>	9.86	9.84	9.77	9.77	11.36
<b>19</b>	9.97	9.95	9.87	9.87	11.41
<b>20</b>	10.06	10.02	9.95	9.95	11.50
<b>21</b>	10.15	10.10	10.01	10.01	11.60
<b>22</b>	10.11	10.08	9.99	9.99	11.59
<b>23</b>	10.04	10.00	9.94	9.94	11.55
<b>24</b>	9.95	9.92	9.85	9.85	11.44
<b>25</b>	9.93	9.91	9.84	9.84	11.48
<b>26</b>	10.00	9.96	9.89	9.89	11.51
<b>27</b>	10.02	10.00	9.92	9.92	11.54
<b>28</b>	9.86	9.84	9.76	9.76	11.44
<b>29</b>	9.97	9.95	9.86	9.86	11.55
<b>30</b>	10.03	10.00	9.91	9.91	11.58

**Table A.5 – Relative error in the estimation of the total cloud cover by each criterion, expressed as a percentage.**

<i>Image #k</i>	<i>R/B (%)</i>	<i>B – R (%)</i>	<i>Multicolour (%)</i>	<i>Simple Bayes (%)</i>
<b>1</b>	-0.65	-1.45	-1.41	20.11
<b>2</b>	-0.48	-1.34	-1.30	19.94
<b>3</b>	-0.64	-0.99	-0.97	19.03
<b>4</b>	-0.68	-0.89	-0.84	18.73
<b>5</b>	-0.59	-0.80	-0.74	18.30
<b>6</b>	-0.66	-0.81	-0.76	19.16
<b>7</b>	-0.49	-0.59	-0.54	17.72
<b>8</b>	-0.61	-0.85	-0.81	17.05
<b>9</b>	-0.38	-0.53	-0.52	17.06
<b>10</b>	-0.71	-0.88	-0.88	15.51
<b>11</b>	-0.40	-0.48	-0.48	15.95
<b>12</b>	-0.45	-0.67	-0.66	15.67
<b>13</b>	-0.43	-0.58	-0.57	15.42
<b>14</b>	-0.35	-1.03	-1.03	16.82
<b>15</b>	-0.42	-0.48	-0.48	14.43
<b>16</b>	-0.31	-0.93	-0.92	15.56
<b>17</b>	-0.42	-1.15	-1.15	15.09
<b>18</b>	-0.25	-0.97	-0.96	15.21
<b>19</b>	-0.26	-1.00	-0.99	14.44
<b>20</b>	-0.33	-1.09	-1.08	14.34
<b>21</b>	-0.48	-1.38	-1.37	14.35
<b>22</b>	-0.27	-1.16	-1.14	14.62
<b>23</b>	-0.35	-0.98	-0.97	15.08
<b>24</b>	-0.22	-1.01	-0.98	15.06
<b>25</b>	-0.16	-0.89	-0.88	15.66
<b>26</b>	-0.36	-1.11	-1.10	15.08
<b>27</b>	-0.28	-1.03	-1.02	15.12
<b>28</b>	-0.17	-1.00	-0.97	16.05
<b>29</b>	-0.14	-1.10	-1.09	15.86
<b>30</b>	-0.25	-1.21	-1.20	15.50
<b>Average</b>	<b>-0.41</b>	<b>-0.95</b>	<b>-0.93</b>	<b>16.26</b>



**Table A.7.6 – Execution Time for each criterion and each image, expressed in seconds.**

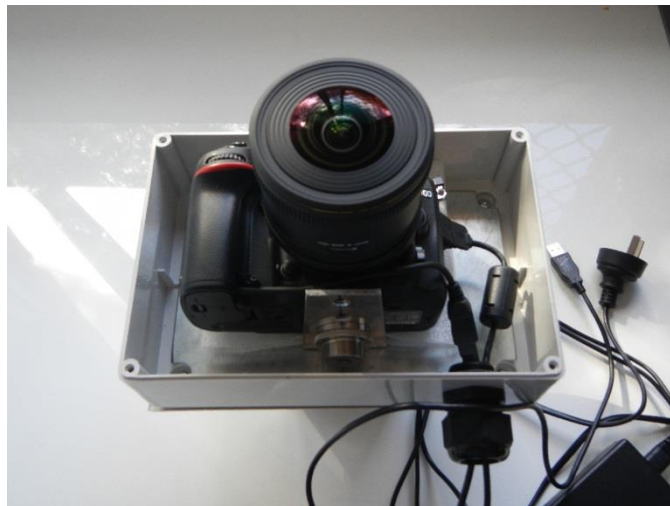
<b><i>Image #k</i></b>	<b><i>R/B</i></b>	<b><i>B – R</i></b>	<b><i>Multicolour</i></b>	<b><i>Simple Bayes</i></b>
<b>1</b>	0.013	0.012	0.009	3.856
<b>2</b>	0.010	0.009	0.012	4.267
<b>3</b>	0.016	0.014	0.012	3.816
<b>4</b>	0.010	0.011	0.015	3.986
<b>5</b>	0.010	0.011	0.011	5.174
<b>6</b>	0.010	0.009	0.011	3.876
<b>7</b>	0.010	0.012	0.010	4.590
<b>8</b>	0.011	0.010	0.011	3.781
<b>9</b>	0.010	0.009	0.011	3.789
<b>10</b>	0.011	0.009	0.010	3.909
<b>11</b>	0.010	0.011	0.009	3.917
<b>12</b>	0.010	0.011	0.010	3.831
<b>13</b>	0.011	0.011	0.009	3.856
<b>14</b>	0.010	0.009	0.021	3.851
<b>15</b>	0.012	0.010	0.010	3.854
<b>16</b>	0.011	0.010	0.010	3.807
<b>17</b>	0.009	0.009	0.009	3.844
<b>18</b>	0.012	0.011	0.010	4.041
<b>19</b>	0.010	0.010	0.010	3.816
<b>20</b>	0.011	0.011	0.009	3.737
<b>21</b>	0.009	0.010	0.010	3.812
<b>22</b>	0.011	0.011	0.010	3.823
<b>23</b>	0.009	0.009	0.011	4.300
<b>24</b>	0.012	0.009	0.010	3.970
<b>25</b>	0.011	0.009	0.009	3.894
<b>26</b>	0.010	0.009	0.011	3.800
<b>27</b>	0.012	0.011	0.009	4.112
<b>28</b>	0.012	0.010	0.010	3.941
<b>29</b>	0.011	0.009	0.009	3.965
<b>30</b>	0.011	0.009	0.009	4.006

## APPENDIX B – CAMERA INSTALLATION

The Nikon D7100 DSLR camera fitted with a Sigma 4.5 mm fisheye lens was installed on the 14th of March 2016 with the kind help of the technical staff. It has been placed next to the fence of the premises, only a few metres away from three different solar arrays, which use respectively monocrystalline, polycrystalline and amorphous technologies. A desktop computer tower running a version of DigiCamControl has been fitted in the cabinet at the bottom of the pole with a 3TB hard disk drive, onto which the sky images are saved. Also, that desktop computer is connected to the nearby building through an Ethernet cable so that it can be remotely accessed.



**Figure B.1 – View of the compartment enclosing the power supply of the camera.**



**Figure B.2 – View of the camera installed in its enclosure box with power and data cables.**



**Figure B.3 – Camera box closed and ready to be installed.**



**Figure B.4 – View of the cabinet containing the computer and the hard drive.**

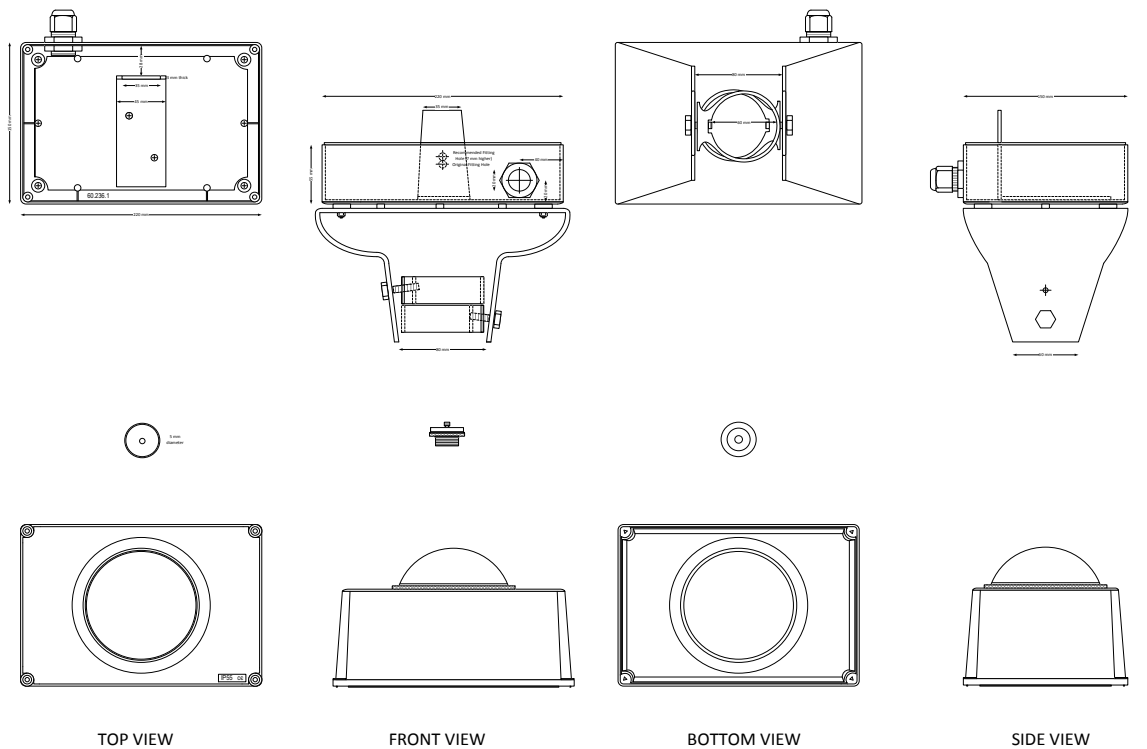


**Figure B.5 – Complete camera setup.**

## APPENDIX C – CAMERA BOX

Since we chose to use a DSLR camera, a waterproof enclosure needed to be prepared for the camera to be installed in the Green Energy Electric Park of the Wembley campus of Curtin University (32.01 S, 115.89 E, 18 m ASL). The blueprint of the box that was used to enclose the camera and fix it on a 4 metre high pole is shown in Figure C.1.

It consists of a bottom part inside of which a metal bracket has been set for the DSLR camera to be fixed. A cable gland has been fitted to let a power and a USB cable through and remotely control the camera. Also, the bottom of the box has been screwed on a robust metal bracket intended to keep the camera solidly fixed on the pole. In addition, the box consists of a lid or a top part on which a dome has been fitted. The dome does not only protect the camera but also enables the 4.5 mm fisheye lens of the camera to retain its full 180 degrees field of view.



SKY CAMERA ENCLOSURE BOX  
26/03/2016 Florian Benjamin Barbieri

**Figure C.1 – Blueprint of the top, front, bottom, and side view of the camera enclosure box.**

## APPENDIX D – AUTHORSHIP CONTRIBUTION STATEMENTS

### Statement of Contribution by Others

The purpose of these statements is to summarise and clearly identify the nature and extent of the intellectual input by the candidate and any co-authors.

[Student]



[Supervisor]



To whom it may concern,

I, Florian Benjamin Eric Barbieri, was the main contributor to the literature review work, methodology defined, dataset collected and processed, comparisons made, and recommendations given on the best choice of empirical clear-sky irradiance models for the state of Western Australia in the paper entitled “A Comparative Study of Clear-Sky Irradiance Models for Western Australia” that has been published at the occasion of the Power and Energy Society General Meeting conference in Boston in July 17 2016 (DOI. [10.1109/TSTE.2017.2733258](https://doi.org/10.1109/TSTE.2017.2733258))

(Signature of Candidate)



I, as a Co-Author, endorse that this level of contribution by the candidate indicated above is appropriate.

(Full Name of Co-Author 1)

(Signature of Co-Author 1)

Arindam Ghosh



(Full Name of Co-Author 2)

(Signature of Co-Author 2)

Sumedha Rajakaruna



(Full Name of Co-Author 4)

(Signature of Co-Author 4)

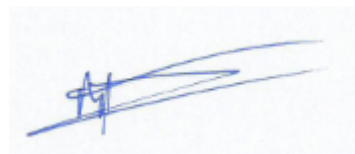
Ba-Tuong Vo



(Full Name of Co-Author 5)

(Signature of Co-Author 5)

Corentin Riffart



To whom it may concern,

I, Florian Benjamin Eric Barbieri, was the main contributor to the literature review work, methodology defined, dataset collected and processed, comparisons made between two different approaches with a state of the art method, assessment of performance of each method and recommendations given about the comparative use of a two different practical multi-target random finite set implementations carried out to track clouds with the intention of forecasting photovoltaic power discontinuities within a very-short timeframe, presented with a novel approach that helps select a cloud/clear-sky pixel discrimination criterion in the paper entitled “Intra-Hour Cloud Tracking Based on Probability Hypothesis Density Filtering” that has been published in the peer-review journal IEEE Transactions on Sustainable Energy (DOI. [10.1109/TSTE.2017.2733258](https://doi.org/10.1109/TSTE.2017.2733258))

(Signature of Candidate)



I, as a Co-Author, endorse that this level of contribution by the candidate indicated above is appropriate.

(Full Name of Co-Author 1)

(Signature of Co-Author 1)

Arindam Ghosh





(Full Name of Co-Author 2)

Sumedha Rajakaruna

(Signature of Co-Author 2)



(Full Name of Co-Author 4)

Ba-Tuong Vo


(Signature of Co-Author 4)



(Full Name of Co-Author 5)

Corentin Riffart

(Signature of Co-Author 5)



To whom it may concern,

I, Florian Benjamin Eric Barbieri, was the main contributor to the literature review work, the structure that has been defined, the comparisons that have been made, and the recommendations have been given in the sections concerning the methods used in the field of wind power forecasting, the statistical methods in use for photovoltaic power forecasting, the comparison of some popular solar power forecasting approaches, photovoltaic cell modelling, irradiance forecasting, cloud modelling and cloud classification, and the review of the state of the art methods used to forecast photovoltaic power within a very-short timeframe in the paper entitled “Very short-term photovoltaic power forecasting with cloud modeling: A review” published in the peer-to-peer journal Renewable and Sustainable Energy Reviews, Elsevier (DOI. [10.1016/j.rser.2016.10.068](https://doi.org/10.1016/j.rser.2016.10.068))

(Signature of Candidate)



I, as a Co-Author, endorse that this level of contribution by the candidate indicated above is appropriate.

(Full Name of Co-Author 1)

(Signature of Co-Author 1)

Arindam Ghosh



(Full Name of Co-Author 2)

Sumedha Rajakaruna

(Signature of Co-Author 2)

A handwritten signature in black ink, appearing to be 'S. Rajakaruna', written in a cursive style.

## APPENDIX E – COPYRIGHT OWNER PERMISSIONS

Permission from Elsevier to reuse the content of the article “Very short-term photovoltaic power forecasting with cloud modelling: A review” published in Renewable and Sustainable Energy Reviews:



The screenshot shows the Copyright Clearance Center RightsLink interface. At the top, there is a navigation bar with the Copyright Clearance Center logo, the RightsLink logo, and buttons for Home, Create Account, Help, and an email icon. The URL in the address bar is https://s100.copyright.com/AppDispatchServlet. Below the navigation bar, there is a section for article details. On the left is a thumbnail image of the journal cover for 'Renewable & Sustainable Energy Reviews'. To the right of the image, the following information is displayed: Title: Very short-term photovoltaic power forecasting with cloud modeling: A review; Author: Florian Barbieri, Sumedha Rajakaruna, Arindam Ghosh; Publication: Renewable and Sustainable Energy Reviews; Publisher: Elsevier; Date: August 2017. Below this information is a copyright notice: © 2016 Elsevier Ltd. All rights reserved. To the right of the article details is a LOGIN button and a text box that says: 'If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials. Already a RightsLink user or want to learn more?'. At the bottom of the interface, there are two buttons: BACK and CLOSE WINDOW. Below the buttons, there is a footer with the following text: Copyright © 2019 Copyright Clearance Center, Inc. All Rights Reserved. Privacy statement, Terms and Conditions. Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

BACK

CLOSE WINDOW

Copyright © 2019 Copyright Clearance Center, Inc. All Rights Reserved. [Privacy statement](#), [Terms and Conditions](#).

Comments? We would like to hear from you. E-mail us at [customercare@copyright.com](mailto:customercare@copyright.com)

Permission from IEEE to reuse the content from “A comparative study of clear-sky irradiance models for Western Australia” presented in 2016 IEEE PES General Meeting and from “Intra-hour cloud tracking based on probability hypothesis density filtering” published in IEEE Transactions on Sustainable Energy:

### IEEE COPYRIGHT AND CONSENT FORM

To ensure uniformity of treatment among all contributors, other forms may not be substituted for this form, nor may any wording of the form be changed. This form is intended for original material submitted to the IEEE and must accompany any such material in order to be published by the IEEE. Please read the form carefully and keep a copy for your files.

**A Comparative Study of Clear-Sky Irradiance Models for Western Australia**  
Florian Barbieri, Arindam Ghosh, Sumedha Rajakaruna, Ba-Tuong Vo, Coentim Riffart  
2016 IEEE PES General Meeting

**Intra-Hour Cloud Tracking Based on Probability Hypothesis Density Filtering**  
Barbieri, Florian; Riffart, Coentim; Vo, Ba-Tuong; Rajakaruna, Sumedha; Ghosh, Arindam  
IEEE Transactions on Sustainable Energy

## **COPYRIGHT TRANSFER**

The undersigned hereby assigns to The Institute of Electrical and Electronics Engineers, Incorporated (the "IEEE") all rights under copyright that may exist in and to: (a) the Work, including any revised or expanded derivative works submitted to the IEEE by the undersigned based on the Work; and (b) any associated written or multimedia components or other enhancements accompanying the Work.

## **GENERAL TERMS**

1. The undersigned represents that he/she has the power and authority to make and execute this form.
2. The undersigned agrees to indemnify and hold harmless the IEEE from any damage or expense that may arise in the event of a breach of any of the warranties set forth above.
3. The undersigned agrees that publication with IEEE is subject to the policies and procedures of the [IEEE PSPB Operations Manual](#).
4. In the event the above work is not accepted and published by the IEEE or is withdrawn by the author(s) before acceptance by the IEEE, the foregoing grant of rights shall become null and void and all materials embodying the Work submitted to the IEEE will be destroyed.
5. For jointly authored Works, all joint authors should sign, or one of the authors should sign as authorized agent for the others.
6. The author hereby warrants that the Work and Presentation (collectively, the "Materials") are original and that he/she is the author of the Materials. To the extent the Materials incorporate text passages, figures, data or other material from the works of others, the author has obtained any necessary permissions. Where necessary, the author has obtained all third party permissions and consents to grant the license above and has provided copies of such permissions and consents to IEEE

**You have indicated that you DO wish to have video/audio recordings made of your conference presentation under terms and conditions set forth in "Consent and Release."**

## **CONSENT AND RELEASE**

1. In the event the author makes a presentation based upon the Work at a conference hosted or sponsored in whole or in part by the IEEE, the author, in consideration for his/her participation in the conference, hereby grants the IEEE the unlimited, worldwide, irrevocable permission to use, distribute, publish, license, exhibit, record, digitize, broadcast, reproduce and archive, in any format or medium, whether now known or hereafter developed: (a) his/her presentation and comments at the conference; (b) any written materials or multimedia files used in connection with his/her presentation; and (c) any recorded interviews of him/her (collectively, the "Presentation"). The permission granted includes the transcription and reproduction of the Presentation for inclusion in products sold or distributed by IEEE and live or recorded broadcast of the Presentation during or after the conference.
2. In connection with the permission granted in Section 1, the author hereby grants IEEE the unlimited, worldwide, irrevocable right to use his/her name, picture, likeness, voice and biographical information as part of the advertisement, distribution and sale of products incorporating the Work or Presentation, and releases IEEE from any claim based on right of privacy or publicity.

BY TYPING IN YOUR FULL NAME BELOW AND CLICKING THE SUBMIT BUTTON, YOU CERTIFY THAT SUCH ACTION CONSTITUTES YOUR ELECTRONIC SIGNATURE TO THIS FORM IN ACCORDANCE WITH UNITED STATES LAW, WHICH AUTHORIZES ELECTRONIC SIGNATURE BY AUTHENTICATED REQUEST FROM A USER OVER THE INTERNET AS A VALID SUBSTITUTE FOR A WRITTEN SIGNATURE.

Florian Barbieri

11-11-2015

Signature

Date (dd-mm-yyyy)

Florian Benjamin Barbieri

26-07-2017

Signature

Date (dd-mm-yyyy)

### Information for Authors

#### AUTHOR RESPONSIBILITIES

The IEEE distributes its technical publications throughout the world and wants to ensure that the material submitted to its publications is properly available to the readership of those publications. Authors must ensure that their Work meets the requirements as stated in section 8.2.1 of the IEEE PSPB Operations Manual, including provisions covering originality,

authorship, author responsibilities and author misconduct. More information on IEEE's publishing policies may be found at [http://www.ieee.org/publications\\_standards/publications/rights/authrightsresponsibilities.html](http://www.ieee.org/publications_standards/publications/rights/authrightsresponsibilities.html) Authors are advised especially of IEEE PSPB Operations Manual section 8.2.1.B12: "It is the responsibility of the authors, not the IEEE, to determine whether disclosure of their material requires the prior consent of other parties and, if so, to obtain it." Authors are also advised of IEEE PSPB Operations Manual section 8.1.1B: "Statements and opinions given in work published by the IEEE are the expression of the authors."

#### RETAINED RIGHTS/TERMS AND CONDITIONS

- Authors/employers retain all proprietary rights in any process, procedure, or article of manufacture described in the Work.
- Authors/employers may reproduce or authorize others to reproduce the Work, material extracted verbatim from the Work, or derivative works for the author's personal use or for company use, provided that the source and the IEEE copyright notice are indicated, the copies are not used in any way that implies IEEE endorsement of a product or service of any employer, and the copies themselves are not offered for sale.
- Although authors are permitted to re-use all or portions of the Work in other works, this does not include granting third-party requests for reprinting, republishing, or other types of re-use. The IEEE Intellectual Property Rights office must handle all such third-party requests.
- Authors whose work was performed under a grant from a government funding agency are free to fulfill any deposit mandates from that funding agency.

## **AUTHOR ONLINE USE**

- **Personal Servers.** Authors and/or their employers shall have the right to post the accepted version of IEEE-copyrighted articles on their own personal servers or the servers of their institutions or employers without permission from IEEE, provided that the posted version includes a prominently displayed IEEE copyright notice and, when published, a full citation to the original IEEE publication, including a link to the article abstract in IEEE Xplore. Authors shall not post the final, published versions of their papers.
- **Classroom or Internal Training Use.** An author is expressly permitted to post any portion of the accepted version of his/her own IEEE-copyrighted articles on the author's personal web site or the servers of the author's institution or company in connection with the author's teaching, training, or work responsibilities, provided that the appropriate copyright, credit, and reuse notices appear prominently with the posted material. Examples of permitted uses are lecture materials, course packs, e-reserves, conference presentations, or in-house training courses.
- **Electronic Preprints.** Before submitting an article to an IEEE publication, authors frequently post their manuscripts to their own web site, their employer's site, or to another server that invites constructive comment from colleagues. Upon submission of an article to IEEE, an author is required to transfer copyright in the article to IEEE, and the author must update any previously posted version of the article with a prominently displayed IEEE copyright notice. Upon publication of an article by the IEEE, the author must replace any previously posted electronic versions of the article with either (1) the full citation to the IEEE work with a Digital Object Identifier (DOI) or link to the article abstract in IEEE Xplore, or (2) the accepted version only (not the IEEE-published version), including the IEEE copyright notice and full citation, with a link to the final, published article in IEEE Xplore.

**Questions about the submission of the form or manuscript must be sent to the publication's editor.**

**Please direct all questions about IEEE copyright policy to:**

**IEEE Intellectual Property Rights Office, [copyrights@ieee.org](mailto:copyrights@ieee.org), +1-732-562-3966**

## BIBLIOGRAPHICAL REFERENCES

1. Barbieri, F., S. Rajakaruna, and A. Ghosh, *Very short-term photovoltaic power forecasting with cloud modeling: A review*. Renewable and Sustainable Energy Reviews.
2. Barbieri, F., C. Riffart, B-T Vo, S. Rajakaruna, A. Ghosh, *A comparative study of clear-sky irradiance models for Western Australia*. in *2016 IEEE Power and Energy Society General Meeting (PESGM)*. 2016.
3. Barbieri, F., C. Riffart, B-T Vo, S. Rajakaruna, A. Ghosh, *Intrahour Cloud Tracking Based on Probability Hypothesis Density Filtering*. IEEE Transactions on Sustainable Energy, 2018. **9**(1): p. 340-349.
4. Keller, B. and A. Costa, *A Matlab GUI for calculating the solar radiation and shading of surfaces on the earth*. Computer Applications in Engineering Education, 2011. **19**(1): p. 161-170.
5. Mellit, A. and S.A. Kalogirou, *Artificial intelligence techniques for photovoltaic applications: A review*. Progress in Energy and Combustion Science, 2008. **34**(5): p. 574-632.
6. <http://arena.gov.au/files/2013/08/Chapter-10-Solar-Energy.pdf>
7. Masters, G.M., *Renewable and efficient electric power systems / Gilbert M. Masters*. Vol. Chapter 7. 2004, Hoboken, NJ: Hoboken, NJ : John Wiley & Sons.
8. Chakraborty, S., *Power Electronics for Renewable and Distributed Energy Systems : A Sourcebook of Topologies, Control and Integration / edited by Sudipta Chakraborty, Marcelo G. Simões, William E. Kramer*, ed. M.G. Simões, W.E. Kramer, and SpringerLink. Vol. Chapter 3. 2013, London: London : Springer London : Imprint: Springer.
9. Kopp, G. and J.L. Lean, *A new, lower value of total solar irradiance: Evidence and climate significance*. Geophysical Research Letters, 2011. **38**(1).
10. Morjaria, M., et al., *A Grid-Friendly Plant: The Role of Utility-Scale Photovoltaic Plants in Grid Stability and Reliability*. Power and Energy Magazine, IEEE, 2014. **12**(3): p. 87-95.
11. Barnes, A.K., J.C. Balda, and J.K. Hayes. *Modelling PV Clouding Effects Using a Semi-Markov Process with Application to Energy Storage*. in *World Congress*. 2014.
12. Manz, D., et al., *The Grid of the Future: Ten Trends That Will Shape the Grid Over the Next Decade*. Power and Energy Magazine, IEEE, 2014. **12**(3): p. 26-36.
13. Western Power Website [Online]. Available: <http://wiki-solar.org/analysis/index.html>
14. Australian PV Institute Website [Online]. Available: <http://pv-map.apvi.org.au/analyses>
15. Cai, T., S. Duan, and C. Chen. *Forecasting power output for grid-connected photovoltaic power system without using solar radiation measurement*. in *Power Electronics for Distributed Generation Systems (PEDG), 2010 2nd IEEE International Symposium on*. 2010.



16. Xiangjun, L., H. Dong, and L. Xiaokang, *Battery Energy Storage Station (BESS)-Based Smoothing Control of Photovoltaic (PV) and Wind Power Generation Fluctuations*. Sustainable Energy, IEEE Transactions on, 2013. **4**(2): p. 464-473.
17. Hill, C.A., et al., *Battery Energy Storage for Enabling Integration of Distributed Solar Power Generation*. Smart Grid, IEEE Transactions on, 2012. **3**(2): p. 850-857.
18. Teymour, H.R., et al., *Solar PV and Battery Storage Integration using a New Configuration of a Three-Level NPC Inverter With Advanced Control Strategy*. IEEE Transactions on Energy Conversion, 2014. **29**(2): p. 354-365.
19. Barton, J.P. and D.G. Infield, *Energy storage and its use with intermittent renewable energy*. IEEE Transactions on Energy Conversion, 2004. **19**(2): p. 441-448.
20. Horizon Power Website [Online]. Available: [www.horizonpower.com.au](http://www.horizonpower.com.au)
21. Synergy Website [Online]. Available: <https://www.synergy.net.au/About-us/Who-we-are>
22. Western Power Website [Online]. Available: [www.westernpower.com.au/corporate-information-about-us.html](http://www.westernpower.com.au/corporate-information-about-us.html)
23. Marra, F., et al., *A Decentralized Storage Strategy for Residential Feeders With Photovoltaics*. Smart Grid, IEEE Transactions on, 2014. **5**(2): p. 974-981.
24. Tuohy, A., et al., *Solar Forecasting: Methods, Challenges, and Performance*. Power and Energy Magazine, IEEE, 2015. **13**(6): p. 50-59.
25. Barnes, A.K., et al., *Placement of energy storage coordinated with smart PV inverters*. 2012. p. 1-7.
26. Cormode, D., et al. *Comparing ramp rates from large and small PV systems, and selection of batteries for ramp rate control*. in *Photovoltaic Specialists Conference (PVSC), 2013 IEEE 39th*. 2013.
27. Lave, M., J. Kleissl, and J.S. Stein, *A Wavelet-Based Variability Model (WVM) for Solar PV Power Plants*. Sustainable Energy, IEEE Transactions on, 2013. **4**(2): p. 501-509.
28. Jayadevan, V.T., et al. *Forecasting solar power intermittency using ground-based cloud imaging*. in *World Renewable Energy Forum 2012*. 2012.
29. Diagne, M., et al., *Review of solar irradiance forecasting methods and a proposition for small-scale insular grids*. Renewable and Sustainable Energy Reviews, 2013. **27**(0): p. 65-76.
30. Chow, C.W., et al., *Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed*. Solar Energy, 2011. **85**(11): p. 2881-2893.
31. Jewell, W. and R. Ramakumar, *The Effects of Moving Clouds on Electric Utilities with Dispersed Photovoltaic Generation*. Energy Conversion, IEEE Transactions on, 1987. **EC-2**(4): p. 570-576.
32. Jewell, W.T., R. Ramakumar, and S.R. Hill, *A study of dispersed photovoltaic generation on the PSO system*. Energy Conversion, IEEE Transactions on, 1988. **3**(3): p. 473-478.

33. Chengrui, C. and D.C. Aliprantis, *Cumulus Cloud Shadow Model for Analysis of Power Systems With Photovoltaics*. Power Systems, IEEE Transactions on, 2013. **28**(4): p. 4496-4506.
34. Jakhрани, A., et al., *Comparison of solar photovoltaic module temperature models*. World Applied Sciences Journal, 2011. **14**: p. 1-8.
35. Hashimoto, T., *Prediction of output power variation of solar power plant by image measurement of cloud movement*. Journal of Advanced Research in Physics, 2012. **2**(2).
36. <http://www.bom.gov.au/climate/austmaps/solar-radiation-glossary.shtml>
37. Chenni, R., et al., *A detailed modeling method for photovoltaic cells*. Energy, 2007. **32**(9): p. 1724-1730.
38. Lonij, V.P.A., et al. *Forecasts of PV power output using power measurements of 80 residential PV installs*. in *Photovoltaic Specialists Conference (PVSC), 2012 38th IEEE*. 2012.
39. Soman, S.S., et al. *A review of wind power and wind speed forecasting methods with different time horizons*. in *North American Power Symposium (NAPS), 2010*. 2010.
40. [http://www.imowa.com.au/docs/default-source/Reserve-Capacity/soo/imo\\_2014-swiss-electricity-demand-outlook5EA7F337C272.pdf?sfvrsn=0](http://www.imowa.com.au/docs/default-source/Reserve-Capacity/soo/imo_2014-swiss-electricity-demand-outlook5EA7F337C272.pdf?sfvrsn=0)
41. <https://nwis.com.au/the-nwis/>
42. [http://ds.data.jma.go.jp/mscweb/data/sat\\_dat/img/reg/sat\\_img\\_au.html](http://ds.data.jma.go.jp/mscweb/data/sat_dat/img/reg/sat_img_au.html)
43. <https://nwis.com.au/the-nwis/>
44. [www.horizonpower.com.au](http://www.horizonpower.com.au)
45. <https://www.synergy.net.au/About-us/Who-we-are>
46. [www.westernpower.com.au/corporate-information-about-us.html](http://www.westernpower.com.au/corporate-information-about-us.html)
47. [http://wa.aemo.com.au/docs/default-source/Reserve-Capacity/soo/imo\\_2014-swiss-electricity-demand-outlook.pdf?sfvrsn=0](http://wa.aemo.com.au/docs/default-source/Reserve-Capacity/soo/imo_2014-swiss-electricity-demand-outlook.pdf?sfvrsn=0)
48. Jahangiri, P. and D.C. Aliprantis, *Distributed Volt/VAr Control by PV Inverters*. IEEE Transactions on Power Systems, 2013. **28**(3): p. 3429-3439.
49. Baran, M.E., et al., *Accommodating High PV Penetration on Distribution Feeders*. IEEE Transactions on Smart Grid, 2012. **3**(2): p. 1039-1046.
50. Adhikari, S., L. Fangxing, and L. Huijuan, *P-Q and P-V Control of Photovoltaic Generators in Distribution Systems*. IEEE Transactions on Smart Grid, 2015. **6**(6): p. 2929-2941.
51. Niknam, T., M. Zare, and J. Aghaei, *Scenario-Based Multiobjective Volt/Var Control in Distribution Networks Including Renewable Energy Sources*. IEEE Transactions on Power Delivery, 2012. **27**(4): p. 2004-2019.
52. Lawan, S.M., et al., *Different Models of Wind Speed Prediction; A Comprehensive Review*. International Journal of Scientific & Engineering Research, 2014. **5**(1): p. 1760-1768.

53. Espino, I. and M. Hernandez, *Nowcasting of Wind Speed using Vector Regression, Experiments with Time Series from Gran Canaria*. Renewable Energy and Power Quality Journal, 2011(N9).
54. <http://w1.weather.gov/glossary>
55. Hassanzadeh, M., M. Etezadi-Amoli, and M.S. Fadali. *Practical approach for sub-hourly and hourly prediction of PV power output*. in *North American Power Symposium (NAPS), 2010*. 2010.
56. Ciulla, G., V. Lo Brano, and E. Moreci, *Forecasting the Cell Temperature of PV Modules with an Adaptive System*. International Journal of Photoenergy, 2013. **2013**: p. 1-10.
57. Mellit, A. and A.M. Pavan, *A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy*. Solar Energy, 2010. **84**(5): p. 807-821.
58. Heinle, A., A. Macke, and A. Srivastav, *Automatic cloud classification of whole sky images*. Atmospheric Measurement Techniques, 2010. **3**(3): p. 557-567.
59. Kazantzidis, A., et al., *Cloud detection and classification with the use of whole-sky ground-based images*. Atmospheric Research, 2012. **113**: p. 80-88.
60. Jie, S., et al., *Forecasting Power Output of Photovoltaic Systems Based on Weather Classification and Support Vector Machines*. Industry Applications, IEEE Transactions on, 2012. **48**(3): p. 1064-1069.
61. Yuehui, H., et al. *Comparative study of power forecasting methods for PV stations*. in *Power System Technology (POWERCON), International Conference on Power System Technology*. 2010.
62. Martins, F., E. Pereira, and R. Guarnieri, *Solar Radiation Forecast Using Artificial Neural Networks*. International Journal of Energy Science, 2012. **2**(6): p. 217-217.
63. Evans, D., *Simplified method for predicting photovoltaic array output*. Solar energy, 1981. **27**(6): p. 555-560.
64. Ross Jr, R. *Interface design considerations for terrestrial solar cell modules*. in *12th Photovoltaic Specialists Conference*. 1976.
65. Rauschenbach, H.S., *Solar cell array design handbook-The principles and technology of photovoltaic energy conversion*. 1980.
66. Risser, V. and M. Fuentes. *Linear regression analysis of flat-plate photovoltaic system performance data*. in *5th Photovoltaic Solar Energy Conference*. 1984.
67. Schott, T. *Operation temperatures of PV modules: a theoretical and experimental approach*. in *EC Photovoltaic solar energy conference*. 6. 1985.
68. Servant, J. *Calculation of the cell temperature for photovoltaic modules from climatic data*. in *Proceedings of the Ninth Biennial Congress of the ISES—Intersol*. 1985.

69. Ross Jr, R. and M. Smokler, *Flat-Plate Solar Array Project: Final Report: Volume 6, Engineering Sciences and Reliability*. 1986, Jet Propulsion Lab., Pasadena, CA (USA).
70. Lasnier, F., *Photovoltaic engineering handbook*. 1990: CRC Press.
71. Hove, T., *A method for predicting long-term average performance of photovoltaic systems*. *Renewable Energy*, 2000. **21**(2): p. 207-229.
72. Krauter, S.C., *Development of an integrated solar home system*. *Solar energy materials and solar cells*, 2004. **82**(1): p. 119-130.
73. Mondol, J.D., et al., *Long-term validated simulation of a building integrated photovoltaic system*. *Solar Energy*, 2005. **78**(2): p. 163-176.
74. Mondol, J.D., Y.G. Yohanis, and B. Norton, *Comparison of measured and predicted long term performance of grid a connected photovoltaic system*. *Energy Conversion and Management*, 2007. **48**(4): p. 1065-1080.
75. Mondol, J.D., Y.G. Yohanis, and B. Norton, *The effect of low insolation conditions and inverter oversizing on the long-term performance of a grid-connected photovoltaic system*. *Progress in Photovoltaics: Research and Applications*, 2007. **15**(4): p. 353-368.
76. Duffie, J.A., *William A. Beckman Solar Engineering of Thermal Processes*. 2006, JOHN WILEY & SONS, INC. NY.
77. *Solar electricity / edited by Tomas Markvart, with Klaus Bogus ... [et al.]*. 2nd ed.. ed, ed. K. Bogus and T. Markvart. 2000, Chichester: Chichester : John Wiley & Sons.
78. TamizhMani, G., et al., *PHOTOVOLTAIC MODULE THERMAL/WIND PERFORMANCE: Long -Term Monitoring and Model Development For Energy Rating*. NCPV and Solar Program Review Meeting, 2003, 2003: p. 936 - 939.
79. Riffonneau, Y., et al., *Optimal Power Flow Management for Grid Connected PV Systems With Batteries*. *Sustainable Energy, IEEE Transactions on*, 2011. **2**(3): p. 309-320.
80. Zieland, L., S. Rajakaruna, and D. Liyanage. *Modelling and performance assessment of large scale solar photovoltaic plants in rural Australia*. in *Power Engineering Conference (AUPEC), 2013 Australasian Universities*. 2013.
81. Niimura, T., et al. *Profiling residential PV output based on weekly weather forecast for home energy management system*. in *Power and Energy Society General Meeting, 2012 IEEE*. 2012.
82. Masters, G.M., *Renewable and efficient electric power systems / Gilbert M. Masters*. Vol. Chapter 8. 2004, Hoboken, NJ: Hoboken, NJ : John Wiley & Sons.
83. Lonij, V.P., et al. *Analysis of 80 rooftop PV systems in the Tucson, AZ area*. in *Photovoltaic Specialists Conference (PVSC), 2012 38th IEEE*. 2012. IEEE.
84. Kasten, F., *A simple parameterization of the pyr heliometric formula for determining the Linke turbidity factor*. *Meteorologische Rundschau*, 1980. **33**: p. 124-127.

85. Ineichen, P. and R. Perez, *A new air mass independent formulation for the Linke turbidity coefficient*. Solar Energy, 2002. **73**(3): p. 151-157.
86. Remund, J., et al. *Worldwide Linke turbidity information*. in *ISES Solar World Congress 2003*. 2003. International Solar Energy Society (ISES).
87. Linke, F., *Transmissions-koeffizient und Trübungsfaktor*. Beitr. Phys. Fr. Atmos, 1922. **10**: p. 91-103.
88. Bird, R.E., *A simple, solar spectral model for direct-normal and diffuse horizontal irradiance*. Solar energy, 1984. **32**(4): p. 461-471.
89. Bird, R.E. and C. Riordan, *Simple solar spectral model for direct and diffuse irradiance on horizontal and tilted planes at the earth's surface for cloudless atmospheres*. Journal of Climate and Applied Meteorology, 1986. **25**(1): p. 87-97.
90. Davies, J.A. and D.C. McKay, *Estimating solar irradiance and components*. Solar Energy, 1982. **29**(1): p. 55-64.
91. Davies, J. and D. McKay, *Evaluation of selected models for estimating solar radiation on horizontal surfaces*. Solar Energy, 1989. **43**(3): p. 153-168.
92. Iqbal, M., *An introduction to solar radiation*. 2012: Elsevier.
93. American Society of Heating, R., et al., *Procedure for Determining Heating and Cooling Loads for Computerized Energy Calculations: Algorithms for Building Heat Transfer Sub Routines*. 1971: ASHRAE.
94. Gueymard, C.A., *REST2: High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation – Validation with a benchmark dataset*. Solar Energy, 2008. **82**(3): p. 272-285.
95. Gueymard, C.A., *Clear-sky irradiance predictions for solar resource mapping and large-scale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models*. Solar Energy, 2011. **86**(8): p. 2145-2169.
96. <http://aeronet.gsfc.nasa.gov/>
97. Reno, M.J., C.W. Hansen, and J.S. Stein, *Global horizontal irradiance clear sky models: Implementation and analysis*. SANDIA report SAND2012-2389, 2012.
98. Masters, G.M., *Renewable and efficient electric power systems*, in *Second Edition*. 2013, Hoboken, New Jersey : John Wiley & Sons Inc. p. 186-252.
99. Annear, R.L. and S.A. Wells, *A comparison of five models for estimating clear-sky solar radiation*. Water resources research, 2007. **43**(10).
100. Kasten, F. and G. Czeplak, *Solar and terrestrial radiation dependent on the amount and type of cloud*. Solar Energy, 1980. **24**(2): p. 177-189.
101. Haurwitz, B., *Isolation in Relation to Cloud Type*. Journal of Meteorology, 1948. **5**(3): p. 110-113.

102. Haurwitz, B., *INSOLATION IN RELATION TO CLOUD TYPE*. Journal of Meteorology, 1946. **3**(4): p. 123-124.
103. Threlkeld, J. and R. Jordan, *Direct solar radiation available on clear days*. Heat., Piping Air Cond., 1957. **29**(12).
104. Berger, X., *Etude du Climat en Region Nicoise en vue d'Applications a l'Habitat Solaire*. CNRS, Paris, 1979.
105. Adnot, J., et al., *Utilisation de courbes de fréquences cumulées d'irradiation solaire globale pour le calcul des installations solaires*, in *Analyse statistique des processus météorologiques appliquée à l'énergie solaire*, R. Lestienne, Editor. 1979, Editions du CNRS. p. 9-40.
106. Robledo, L. and A. Soler, *Luminous efficacy of global solar radiation for clear skies*. Energy Conversion and Management, 2000. **41**(16): p. 1769-1779.
107. Laue, E.G., *The measurement of solar spectral irradiance at different terrestrial elevations*. Solar Energy, 1970. **13**(1): p. 43,IN1,51-50,IN4,57.
108. Meinel, A.B., *Applied solar energy : an introduction / by Aden B. Meinel and Marjorie P. Meinel*, ed. M.P. Meinel. 1976, Reading, Mass.: Reading, Mass. : Addison-Wesley.
109. Yang, D., W.M. Walsh, and P. Jirutitijaroen, *Estimation and Applications of Clear Sky Global Horizontal Irradiance at the Equator*. Journal of Solar Energy Engineering, 2014.
110. Dazhi, Y., P. Jirutitijaroen, and W.M. Walsh, *The Estimation of Clear Sky Global Horizontal Irradiance at the Equator*. Energy Procedia, 2012. **25**(0): p. 141-148.
111. Lorenz, E., et al., *Irradiance Forecasting for the Power Prediction of Grid-Connected Photovoltaic Systems*. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, 2009. **2**(1): p. 2-10.
112. Bacher, P., H. Madsen, and H.A. Nielsen, *Online short-term solar power forecasting*. Solar Energy, 2009. **83**(10): p. 1772-1783.
113. Reda, I. and A. Andreas, *Solar position algorithm for solar radiation applications*. Solar energy, 2004. **76**(5): p. 577-589.
114. Meeus, J.H., *Astronomical algorithms*. 1991: Willmann-Bell, Incorporated.
115. Bird, R.E. and R.L. Hulstrom, *Simplified clear sky model for direct and diffuse insolation on horizontal surfaces*. 1981, Solar Energy Research Inst., Golden, CO (USA).
116. [http://oceanservice.noaa.gov/education/yos/resource/JetStream/synoptic/clouds\\_max.htm](http://oceanservice.noaa.gov/education/yos/resource/JetStream/synoptic/clouds_max.htm)
117. Raschke, R.A. and S.K. Cox, *The Determination of Cloud Optical Depth from Multiple Fields of View Pyrheliometric Measurements*. 1982: Department of Atmospheric Science, Colorado State University.
118. <http://isccp.giss.nasa.gov/products/fieldvars.html>
119. [http://www.crh.noaa.gov/lmk/?n=cloud\\_classification](http://www.crh.noaa.gov/lmk/?n=cloud_classification)
120. <http://www.srh.noaa.gov/srh/jetstream/clouds/cloudwise/types.html>
121. Great Britain. Meteorological, O., *Meteorological glossary / Meteorological Office*

- compiled by D.H. McIntosh. 5th ed., ed, ed. D.H. McIntosh. 1972, London: London : H.M.S.O.
122. Kazantzidis, A., et al., *Cloud detection and classification with the use of whole-sky ground-based images*. Atmospheric Research, 2012. **113**(0): p. 80-88.
  123. Chen, C., et al., *Online 24-h solar power forecasting based on weather type classification using artificial neural network*. Solar Energy, 2011. **85**(11): p. 2856-2870.
  124. Adesso, P., et al., *MAP-MRF Cloud Detection Based on PHD Filtering*. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens., 2012. **5**(3): p. 919-929.
  125. Wood-Bradley, P., J. Zapata, and J. Pye, , *Cloud tracking with optical flow for short-term solar forecasting* “. Solar Thermal Group, Australian National University, Canberra, Australija, 2012.
  126. Linfoot, A. and R.J. Alliss, *J4. 2 A CLOUD DETECTION ALGORITHM APPLIED TO A WHOLE SKY IMAGER INSTRUMENT USING NEURAL NETWORKS*. 2008.
  127. Gonzalez, Y., C. López, and E. Cuevas. *Automatic observation of cloudiness: Analysis of all-sky images*. in *WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation. Session*. 2012.
  128. [http://www.icem2011.org/presentations2011/4\\_Thursday/3C/1100\\_Heinemann.pdf](http://www.icem2011.org/presentations2011/4_Thursday/3C/1100_Heinemann.pdf)
  129. Dambreville, R., et al., *Very short term forecasting of the Global Horizontal Irradiance using a spatio-temporal autoregressive model*. Renewable Energy, 2014. **72**(0): p. 291-300.
  130. <http://www.eumetsat.int/website/home/Data/DataDelivery/EUMETCast/index.html#>
  131. Lonij, V.P.A., et al., *Intra-hour forecasts of solar power production using measurements from a network of irradiance sensors*. Solar Energy, 2013. **97**(0): p. 58-66.
  132. <http://www.nrel.gov/midc/solpos/spa.html>
  133. Spencer, J., *Fourier series representation of the position of the sun*. Search, 1971. **2**(5): p. 172-172.
  134. Paulescu, M., et al., *Weather modeling and forecasting of PV systems operation*. Vol. Chapter 2. 2012: Springer Science & Business Media.
  135. Paltridge, G.W. and D. Proctor, *Monthly mean solar radiation statistics for Australia*. Solar Energy, 1976. **18**(3): p. 235-243.
  136. Daneshyar, M., *Solar radiation statistics for Iran*. Solar Energy, 1978. **21**(4): p. 345-349.
  137. <http://aeronet.gsfc.nasa.gov/>
  138. "A Quick Derivation relating altitude to air pressure", Portland State Aerospace Society, Version 1.03, 12/22/2004.
  139. Dev, S., et al., *WAHRISIS: A low-cost high-resolution whole sky imager with near-infrared capabilities*. 2014.
  140. <https://sites.google.com/site/scarabotix/ocamcalib-toolbox>
  141. [http://www.vision.caltech.edu/bouguetj/calib\\_doc/](http://www.vision.caltech.edu/bouguetj/calib_doc/)

142. <https://au.mathworks.com/matlabcentral/fileexchange/4698-simulating-photoshop-s-magic-wand-tool>
143. Gauchet, C., et al. *Surface solar irradiance estimation with low-cost fish-eye camera*. in *Workshop on "Remote Sensing Measurements for Renewable Energy"*. 2012.
144. Savoy, F.M., et al. *Geo-referencing and stereo calibration of ground-based whole sky imagers using the sun trajectory*. in *Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International*. 2016. IEEE.
145. <https://www.mathworks.com/matlabcentral/fileexchange/31524-colorgradient-generate-custom-linear-colormaps>
146. Youtube Website [Online]. Available: [https://www.youtube.com/watch?v=899DV\\_VCIDM](https://www.youtube.com/watch?v=899DV_VCIDM)
147. Long, C.N., et al., *Retrieving cloud characteristics from ground-based daytime color all-sky images*. *Journal of Atmospheric and Oceanic Technology*, 2006. **23**(5): p. 633-652.
148. Russell, S.J., *Artificial intelligence : a modern approach / Stuart J. Russell and Peter Norvig contributing writers, Ernest Davis ... [et al.]*. 3rd ed.. ed, ed. P. Norvig. 2010, Upper Saddle River, N.J.: Upper Saddle River, N.J. : Prentice Hall.
149. Haralick, R.M., K. Shanmugam, and I.H. Dinstein, *Textural Features for Image Classification*. *Systems, Man and Cybernetics, IEEE Transactions on*, 1973. **3**(6): p. 610-621.
150. Jayadevan, V.T., et al. *Forecasting solar power intermittency using ground-based cloud imaging*.
151. Vo, B.-N. and W.-K. Ma, *The Gaussian mixture probability hypothesis density filter*. *Signal Processing, IEEE Transactions on*, 2006. **54**(11): p. 4091-4104.
152. Mahler, R. , *Advances in Statistical Multisource-Multitarget Information Fusion*. Artech House Publishers, first edition, 2014. Chapt. 8. p. 181-216.
153. Vo, B.-N., S. Singh, and A. Doucet. *Sequential Monte Carlo implementation of the PHD filter for multi-target tracking*. in *Proc. Int'l Conf. on Information Fusion*. 2003.
154. Punithakumar, K., T. Kirubarajan, and A. Sinha, *Multiple-model probability hypothesis density filter for tracking maneuvering targets*. *Aerospace and Electronic Systems, IEEE Transactions on*, 2008. **44**(1): p. 87-98.
155. Mahler, R.P.S., *Statistical multisource-multitarget information fusion / Ronald P. S. Mahler*. Vol. Chapter 16. 2007, Boston: Boston : Artech House.
156. Mahler, R. *A theory of PHD filters of higher order in target number*. in *Defense and Security Symposium*. 2006. International Society for Optics and Photonics.
157. Vo, B.-T., B.-N. Vo, and A. Cantoni, *Analytic implementations of the cardinalized probability hypothesis density filter*. *Signal Processing, IEEE Transactions on*, 2007. **55**(7): p. 3553-3567.
158. <https://www.davisinstruments.com/solution/vantage-pro2/>
159. <http://www.suntech-power.com/productInfo.html?type=1726>



160. <https://www.sma-australia.com.au/products/solarinverters.html>
161. <http://digicamcontrol.com/>
162. <https://github.com/dukus/digiCamControl/>
163. [https://github.com/FloBar/digicamcontrol\\_amended](https://github.com/FloBar/digicamcontrol_amended)
164. [https://github.com/FloBar/digicam\\_amended\\_new2/blob/master/CameraControl/ViewModel/TimelapseViewModel.cs](https://github.com/FloBar/digicam_amended_new2/blob/master/CameraControl/ViewModel/TimelapseViewModel.cs)
165. Chow, C.W., *Cloud Characterization for Solar Forecasting Using Sky Imagery*, J. Kleissl, et al., Editors. 2015, ProQuest Dissertations Publishing.
166. Ina, N., et al., *Smoothing of PV system output by tuning MPPT control*. Electrical Engineering in Japan, 2005. **152**(2): p. 10-17.
167. Omran, W.A., M. Kazerani, and M.M.A. Salama, *Investigation of Methods for Reduction of Power Fluctuations Generated From Large Grid-Connected Photovoltaic Systems*. IEEE Transactions on Energy Conversion, 2011. **26**(1): p. 318-327.
168. Yao, L.W., et al. *Modeling of lithium-ion battery using MATLAB/simulink*. in *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society*. 2013.
169. Min, C. and G.A. Rincon-Mora, *Accurate electrical battery model capable of predicting runtime and I-V performance*. IEEE Transactions on Energy Conversion, 2006. **21**(2): p. 504-511.
170. Notten, P.H.L., H.J. Bergveld, and W.S. Kruijt, *Battery management systems : design by modeling / edited by Henk Jan Bergveld, Wanda S. Kruijt, Peter H.L. Notten*. 2002, Dordrecht: Dordrecht : Kluwer Academic.
171. Krismer, F. and J.W. Kolar, *Accurate Small-Signal Model for the Digital Control of an Automotive Bidirectional Dual Active Bridge*. IEEE Transactions on Power Electronics, 2009. **24**(12): p. 2756-2768.
172. Datta, A.J., et al., *Bidirectional power sharing in an ac/dc system with a dual active bridge converter (Book review)*. IET Generation, Transmission & Distribution, 2019. **13**(4): p. 495-501.
173. Kumar, D., F. Zare, and A. Ghosh, *DC Microgrid Technology: System Architectures, AC Grid Interfaces, Grounding Schemes, Power Quality, Communication Networks, Applications, and Standardizations Aspects*. IEEE Access, 2017. **5**: p. 12230-12256.
174. De Doncker, R.W.A.A., D.M. Divan, and M.H. Kheraluwala, *A three-phase soft-switched high-power-density DC/DC converter for high-power applications*. IEEE Transactions on Industry Applications, 1991. **27**(1): p. 63-73.
175. Barbieri, F., et al. *Application notes and recommendations on using TMS320F28335 digital Signal Processor to control voltage source converters*. in *2014 Australasian Universities Power Engineering Conference (AUPEC)*. 2014.

176. Chandrasena, R.P.S., et al. *Developing the guidelines for fabrication of laboratory prototype voltage source converters*. in *2014 Australasian Universities Power Engineering Conference (AUPEC)*. 2014.
177. Feldman, R., et al., *A Hybrid Modular Multilevel Voltage Source Converter for HVDC Power Transmission*. Industry Applications, IEEE Transactions on, 2013. **49**(4): p. 1577-1588.
178. Milasi, R.M., A.F. Lynch, and Y.W. Li, *Adaptive vector control for voltage source converters*. IET Control Theory and Applications, 2013. **7**(8): p. 1110-1119.
179. [https://www.lem.com/sites/default/files/products\\_datasheets/lv\\_25-p.pdf](https://www.lem.com/sites/default/files/products_datasheets/lv_25-p.pdf)
180. [https://www.lem.com/sites/default/files/products\\_datasheets/hais\\_50\\_400-p\\_and\\_50\\_150-tp.pdf](https://www.lem.com/sites/default/files/products_datasheets/hais_50_400-p_and_50_150-tp.pdf)
181. [www.ti.com/lit/ds/symlink/ua741.pdf](http://www.ti.com/lit/ds/symlink/ua741.pdf)
182. Sepulveda, C.A., et al., *FPGA v/s DSP Performance Comparison for a VSC-Based STATCOM Control Application*. Industrial Informatics, IEEE Transactions on, 2013. **9**(3): p. 1351-1360.
183. Engerer, N.A. and F.P. Mills, *Validating nine clear sky radiation models in Australia*. Solar Energy, 2015. **120**: p. 9-24.
184. Gueymard, C.A. and J.A. Ruiz-Arias, *Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance*. Solar Energy, 2016. **128**: p. 1-30.
185. Long, C.N. and T.P. Ackerman, *Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects*. Journal of Geophysical Research: Atmospheres (1984–2012), 2000. **105**(D12): p. 15609-15626.
186. Reno, M.J. and C.W. Hansen, *Identification of periods of clear sky irradiance in time series of GHI measurements*. Renewable Energy, 2016. **90**: p. 520-531.
187. Inman, R.H., J.G. Edson, and C.F.M. Coimbra, *Impact of local broadband turbidity estimation on forecasting of clear sky direct normal irradiance*. Solar Energy, 2015. **117**: p. 125-138.
188. Djafer, D., A. Irbah, and M. Zaiani, *Identification of clear days from solar irradiance observations using a new method based on the wavelet transform*. Renewable Energy, 2017. **101**: p. 347-355.
189. <https://sam.nrel.gov/>
190. Strbac, G., et al., *Microgrids: Enhancing the Resilience of the European Megagrid*. Power and Energy Magazine, IEEE, 2015. **13**(3): p. 35-43.
191. Vaahedi, E., et al., *The Emerging Transactive Microgrid Controller: Illustrating Its Concept, Functionality, and Business Case*. Power and Energy Magazine, IEEE, 2017. **15**(4): p. 80-87.

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.