



# Fault detection and isolation for low voltage distribution grids with distributed generation

Nikolaos Sapountzoglou

## ► To cite this version:

Nikolaos Sapountzoglou. Fault detection and isolation for low voltage distribution grids with distributed generation. Electric power. Université Grenoble Alpes, 2019. English. NNT : 2019GREAT083 . tel-02632242

**HAL Id: tel-02632242**

**<https://theses.hal.science/tel-02632242>**

Submitted on 27 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

### **DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES**

Spécialité : **Génie Electrique**

Arrêté ministériel : 25 mai 2016

Présentée par

**Nikolaos SAPOUNTZOGLU**

Thèse dirigée par **Bertrand RAISON**

préparée au sein du **Laboratoire de Génie Electrique**  
dans l'**École Doctorale Electronique, Electrotechnique, Automatique,**  
**Traitement du Signal (EEATS)**

### **Détection et localisation des défauts dans les réseaux de distribution basse tension en présence de production décentralisée**

### **Fault detection and isolation for low voltage distribution grids with distributed generation**

Thèse soutenue publiquement le **29 Novembre 2019**,  
devant le jury composé de :

**Monsieur Nicolas RETIERE**

Professeur, Communauté Université Grenoble Alpes, Président

**Monsieur Mohamed BENBOUZID**

Professeur, Université de Bretagne Occidentale, Rapporteur

**Monsieur Marc PETIT**

Professeur, CentraleSupélec, Rapporteur

**Monsieur Patrick SCHWEITZER**

Maître de conférences HDR, Université de Lorraine, Examineur

**Monsieur Ramon COSTA-CASTELLO**

Profesor Titular de Universidad, Universitat Politècnica de Catalunya, Invité

**Monsieur Bertrand RAISON**

Professeur, Communauté Université Grenoble Alpes, Directeur de thèse





# Fault detection and isolation for low voltage distribution grids with distributed generation

**Nikolaos Sapountzoglou**

Supervisor: Prof. Bertrand Raison

G2ELab

Université Grenoble Alpes

This dissertation is submitted for the degree of

*Doctor of Philosophy*

November 2019







This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 675318 (INCITE).





## Acknowledgements

First of all, I would like to thank the esteemed members of the jury before whom I will defend my thesis and my supervisor, Professor Bertrand Raison, whose guidance and fruitful comments were paramount to completing this thesis. I would also like to thank him for helping me settle in Grenoble.

I am also grateful to my cousin Dr. Dimitrios Christaras whose advice and support proved crucial to securing a PhD position.

Moreover, I would also like to thank my colleagues and friends Mr. Jesus Lago and Mr. Konstantinos Kotsalos, both Early Stage Researchers of the INCITE–ITN. Jesus, for proposing an artificial intelligence approach to tackle the fault diagnosis issue in low voltage distribution grids and for helping in their design and implementation. Konstantinos on the other hand, for his constant support all these three years of the project and especially for his hospitality during my secondment in Efacec. I am very glad that through the INCITE–ITN I was able to make some new friends like Kostas, Tomas, Jesus and more.

Additionally, I would also like to thank all the guys from the SYREL cluster for the joyful lab environment we created especially during the first year of my PhD. Special thanks to the Geoguessers: Dr. Bhargav Swaminathan, Dr. Stéphane Allard, Dr. Kevin Marojahan and Dr. Felix Koeth.

Most importantly though, I would like to thank my parents for sacrificing so much for many years to provide me with all the necessary tools to advance in life. Finally, following the chain of events that brought me here today, I would like to thank Stavroula who has facilitated my arrival in France two times, one back in 2011 and a second one in 2016; I am happy to consider her as family.



# Abstract

In this thesis, three different methods of fault diagnosis for low voltage (LV) distribution grids and two methods of fault isolation for grid-connected photovoltaic systems (GCPVs) are proposed. The proposed tools for fault diagnosis in LV grids are: a) a conventional method based on overcurrent monitoring and sparse voltage measurements across the faulty branch, b) gradient boosting trees and c) deep neural networks which are the most reliable solution demonstrating a 100 % accuracy in fault detection and an average of 12 % of error in distance estimation. Moreover, under limited available measurements their accuracy is decreased by only 4.5 %. Furthermore, two algorithms based on a signal approach are proposed for fault isolation in GCPVs. They use current and voltage measurements at the output of the inverter, examining faults occurring both in the dc and the ac side. Finally, the proposed algorithms achieve an isolation of 15 out of the 19 studied fault cases in less than 100 *ms*.

**keywords:** fault detection, fault location, fault diagnosis, machine learning, low-voltage distribution grids, distributed generation



## Résumé

Dans cette thèse, trois méthodes différentes de diagnostic des défauts pour les réseaux de distribution basse tension (BT) et deux méthodes de localisation des défauts pour les systèmes photovoltaïques raccordés au réseau (GCPV) sont proposées. Les outils proposés pour le diagnostic des défauts dans les réseaux BT sont: a) une méthode conventionnelle basée sur la surveillance de courant et de tension, b) des gradient boosting trees et c) des réseaux de neurones profonds qui sont la solution la plus fiable démontrant une précision de 100 % dans la détection des défauts et 12 % d'erreur moyenne dans l'estimation de la distance. De plus, sous des mesures disponibles limitées, leur précision n'est réduite que de 4.5 %. En outre, deux algorithmes basés sur une approche signal sont proposés pour la localisation de défauts dans les GCPV. Ils utilisent des mesures de courant et de tension à la sortie de l'onduleur, en examinant les défauts apparaissant aussi bien du côté dc que du côté ac. Enfin, les algorithmes proposés permettent la localisation de 15 sur 19 de cas de défauts étudiés en moins de 100 *ms*.

**mots-clés:** détection des défauts, localisation des défauts, diagnostic des défauts, apprentissage automatique, réseaux de distribution basse tension, production décentralisé





# Contents

List of Figures	xvii
-----------------	------

List of Tables	xxiii
----------------	-------

<b>1 General introduction</b>	<b>1</b>
-------------------------------	----------

<b>2 Introduction</b>	<b>5</b>
-----------------------	----------

2.1 Faults in power systems . . . . .	6
---------------------------------------	---

2.1.1 Classification of threats . . . . .	6
---	---

2.1.2 Costs and notable events . . . . .	9
--	---

2.1.3 System average interruption duration index . . . . .	10
--	----

2.1.4 Affected grid elements . . . . .	11
--	----

2.2 Faults in distribution grids . . . . .	13
--	----

2.2.1 Low voltage grids . . . . .	13
-----------------------------------	----

2.2.2 Smart grids . . . . .	15
-----------------------------	----

2.2.3 Fault diagnosis in MV distribution grids . . . . .	17
--	----

2.3 Faults in photovoltaics . . . . .	21
---------------------------------------	----

2.3.1 Solar energy . . . . .	21
------------------------------	----

2.3.2 Photovoltaics . . . . .	21
-------------------------------	----

2.3.3 Grid-connected photovoltaic system topologies . . . . .	22
---	----

2.3.4 Modeling of photovoltaic systems . . . . .	24
--	----

2.3.4.1 PV cell: Bishop's model . . . . .	24
---	----

2.3.4.2 PV array . . . . .	25
----------------------------	----

2.3.4.3 PV system . . . . .	26
-----------------------------	----

2.3.5 Fault types . . . . .	26
-----------------------------	----

2.3.6 Fault diagnosis in photovoltaics . . . . .	27
--	----

2.4 Thesis outline . . . . .	30
------------------------------	----

<b>I</b>	<b>Fault diagnosis in low voltage smart distribution grids</b>	<b>31</b>
<b>3</b>	<b>Conventional method</b>	<b>33</b>
3.1	Method description . . . . .	34
3.1.1	Fault detection . . . . .	34
3.1.2	Fault location . . . . .	37
3.1.2.1	Faulty branch identification . . . . .	37
3.1.2.2	Faulty sector localization . . . . .	38
3.1.2.3	Fault distance estimation . . . . .	42
3.2	Application . . . . .	44
3.2.1	Case study . . . . .	44
3.2.1.1	Grid characteristics . . . . .	44
3.2.1.2	Influencing parameters . . . . .	45
3.2.2	Fault detection . . . . .	48
3.2.3	Fault location . . . . .	49
3.2.3.1	Branch identification . . . . .	49
3.2.3.2	Sector localization . . . . .	53
3.2.3.3	Distance estimation . . . . .	56
3.2.4	Measurement uncertainty . . . . .	60
3.2.5	Less available measurements . . . . .	62
3.3	Conclusion . . . . .	65
<b>4</b>	<b>Gradient boosting trees</b>	<b>67</b>
4.1	Method description . . . . .	68
4.1.1	Model definition . . . . .	68
4.1.2	Algorithm functionality . . . . .	68
4.1.3	Working principle . . . . .	68
4.1.4	Training, validation & test . . . . .	69
4.1.5	Input features . . . . .	70
4.2	Application . . . . .	72
4.2.1	Case study . . . . .	72
4.2.1.1	Grid characteristics . . . . .	72
4.2.1.2	Influencing parameters . . . . .	72
4.2.2	Implementation . . . . .	74

4.2.3	Model training and evaluation . . . . .	74
4.2.4	Fault detection . . . . .	74
4.2.5	Fault location . . . . .	75
4.2.6	Fault identification . . . . .	77
4.3	Conclusion . . . . .	80
<b>5</b>	<b>Deep Neural Networks</b>	<b>81</b>
5.1	Method description . . . . .	82
5.1.1	Deep neural networks . . . . .	82
5.1.2	Network inputs . . . . .	83
5.1.3	Feature and hyperparameters selection . . . . .	84
5.1.4	DNN estimation . . . . .	84
5.1.5	Network outputs and cost function . . . . .	85
5.2	Application . . . . .	87
5.2.1	Case study . . . . .	87
5.2.1.1	Grid characteristics . . . . .	87
5.2.1.2	Influencing parameters . . . . .	87
5.2.2	Fault detection . . . . .	89
5.2.3	Fault location . . . . .	90
5.2.3.1	Branch identification . . . . .	90
5.2.3.2	Distance estimation . . . . .	93
5.2.4	Measurement uncertainty . . . . .	96
5.2.5	Less available measurements . . . . .	97
5.3	Conclusion . . . . .	99
<b>6</b>	<b>Comparative analysis</b>	<b>101</b>
6.1	Common case study . . . . .	102
6.2	Fault detection . . . . .	102
6.3	Fault location . . . . .	103
6.3.1	Branch identification . . . . .	103
6.3.2	Sector localization . . . . .	105
6.3.3	Distance estimation . . . . .	106
6.4	Conclusion . . . . .	110

<b>II</b>	<b>Fault detection and isolation in distributed generators</b>	<b>111</b>
<b>7</b>	<b>Fault detection and isolation in grid-connected photovoltaic systems</b>	<b>113</b>
7.1	Method description . . . . .	114
7.1.1	Fault detection . . . . .	114
7.1.2	Fault isolation . . . . .	115
7.2	Application . . . . .	118
7.2.1	Case study . . . . .	118
7.2.2	Studied fault types . . . . .	119
7.2.3	Fault detection . . . . .	120
7.2.4	Fault isolation algorithm A: dc side faults . . . . .	121
7.2.5	Fault isolation algorithm B: dc and ac side faults . . . . .	127
7.2.6	Algorithm performance . . . . .	136
7.3	Measurement uncertainty . . . . .	138
7.4	Conclusion . . . . .	141
<b>8</b>	<b>General conclusion</b>	<b>143</b>
8.1	Results & Contributions . . . . .	144
8.1.1	Part I: Fault diagnosis in low voltage smart distribution grids . . . . .	144
8.1.2	Part II: Fault detection and isolation in distributed generators	146
8.2	Future work & Perspectives . . . . .	148
8.2.1	Short term . . . . .	148
8.2.2	Mid term . . . . .	149
8.2.3	Long term . . . . .	149
	<b>Bibliography</b>	<b>151</b>
	<b>Appendix A LV grid characteristics</b>	<b>173</b>
	<b>Appendix B Grid-connected PV fault isolation figures</b>	<b>183</b>
	<b>Appendix C Dependency of the neural network on the dataset size</b>	<b>191</b>

---

<b>Appendix D Fault detection and isolation in grid-connected PEM</b>	
<b>fuel cells</b>	<b>193</b>
D.1 Faults in fuel cells . . . . .	194
D.2 Case study . . . . .	194
D.3 Results . . . . .	195
D.4 Conculsion . . . . .	198
<b>Résumé de la thèse</b>	<b>199</b>
<b>Publications</b>	<b>207</b>



# List of Figures

2.1	Causes of historical blackouts from 1974 to 2019. . . . .	8
2.2	Percentages of causes of outages. . . . .	8
2.3	Number of major historical blackouts per country and cause. . .	11
2.4	Unplanned SAIDI including exceptional events for 28 countries of Europe. The shaded area shows the zone of recorded SAIDI values. The blue line represents the average value per year. . . .	12
2.5	Fault occurrence in different grid parts. . . . .	12
2.6	Downed electric towers resulting in power outages. . . . .	14
2.7	Fault detection, isolation and restoration steps time frame [27]. Icons desgined by [28]. . . . .	16
2.8	Impedance-based method. . . . .	18
2.9	Traveling wave method. . . . .	18
2.10	Fault location research trend. . . . .	19
2.11	Cumulative evolution of PV installations [3]. . . . .	23
2.12	Cumulative evolution of GCPVs installations [3]. . . . .	23
2.13	Bishop's model [141]. . . . .	25
2.14	Fault detection and location for GCPV systems research trend. .	29
3.1	Flowchart of the horizontal branch identification algorithm. . . .	39
3.2	Theoretical voltage profile across a faulty line for a single-phase to ground SC fault for two cases of fault resistance values: zero ("0") and non-zero ("x"). The sector where the fault occurred is marked with orange color. . . . .	40



3.3	Voltage profile and example of the usage of the sector localization criteria in a faulty branch. After the faulty sector (colored in orange) the voltage will either: slightly increase (6a) or stabilize to the value of the last nodal measurement (6b) or slightly decrease (6c). Criterion (1) will be triggered in the case of (6a) since the voltage difference ( $\Delta V_{56} = V_6 - V_5 > 0$ ) will be positive and criterion (2) in all three cases of this example as $\Delta V_{56}$ is the smallest voltage difference. . . . .	41
3.4	Theoretical estimation of the fault location. . . . .	42
3.5	Fault detection and isolation method flowchart. . . . .	43
3.6	Single line diagram of the semi-rural Portuguese LV distribution grid. . . . .	44
3.7	Generation and load profiles for one day. . . . .	46
3.8	Simultaneity factor per number of consumers in distribution grids [235]. . . . .	48
3.9	Sensitivity of fault detection criteria per fault resistance for single phase to ground faults. . . . .	50
3.10	Sensitivity of fault detection criteria per fault resistance for three phase faults . . . . .	50
3.11	Superimposed phase rms current increase per fault resistance in each grid feeder. . . . .	51
3.12	Branch identification accuracy. . . . .	53
3.13	Sector localization accuracy. . . . .	55
3.14	Effect of fault resistance under 0 % load and 0 % PV penetration on the distance estimation method for single phase to ground fault located in branch 5 at 270 m from the beginning of feeder 2. . .	57
3.15	Distance estimation error. . . . .	59
3.16	Measurement error effect on fault detection and fault location methods. . . . .	61
3.17	Grid schematic with the reduced number of available nodal voltage measurements. . . . .	63
3.18	Less available measurements effect on the fault location method.	64

4.1	Generation and load profiles for one day. The shaded areas show the pool from which values were drawn to create multiple healthy operation scenarios. . . . .	73
4.2	Fault detection accuracy. . . . .	76
4.3	Branch identification accuracy: all branches as part of the training set vs. out-of-sample branches. . . . .	76
4.4	Phase identification accuracy. . . . .	78
5.1	Example of a DNN. . . . .	82
5.2	Generation and load profiles for one day. The shaded areas show the pool from which values were drawn to create multiple healthy and faulty operation scenarios. Figure (c) presents the scenarios that were considered in this study. . . . .	88
5.3	Branch identification accuracy. . . . .	92
5.4	Distance estimation error. . . . .	94
5.5	Mean absolute percentage error for distance estimation for different $SF$ throughout the day. The main generation and load profiles are provided in the secondary y axis. Three distinct regions (I, II and III) of different performance are marked. The dominant effect of the load demand over the PV generation is shown as the $SF$ curves follow the trend of the load curve and not so much that of generation. . . . .	96
5.6	Class 1 accuracy vs. 2% underestimation error in measurements for branch identification (a) and the distance estimation method (b). . . . .	97
5.7	Less available measurements effect on the faulty branch identification (a) and the distance estimation method (b). . . . .	98
6.1	Branch identification accuracy per fault resistance values for the three developed methods compared to two from the literature. . .	105
6.2	Faulty sector localization accuracy per fault resistance and method. .	106
6.3	Fault distance estimation error in meters per fault resistance values for the two developed methods compared to two from the literature. .	109
6.4	Fault distance estimation error per fault resistance values for the conventional method and the DNN method (in general and within a correctly localized sector). . . . .	109

7.1	Evolution over time of the phase current $I_a$ under a SC fault between phases A, B and the ground. . . . .	116
7.2	GCPV system schematic. . . . .	118
7.3	Flowchart of the fault isolation algorithm A. . . . .	125
7.4	Spectrograms of the lowpass filters for the SC inverter leg fault case (f12) around a) 10 $kHz$ and b) 50 $Hz$ show the disappearance of the pre-fault existing frequency components depicted in red. .	127
7.5	Monitored variables under various faults that were used for the symptom generation. Fault inception is at 100 $ms$ . . . . .	129
7.6	Fault isolation algorithm B flowchart. . . . .	133
7.7	Moving average of the phase angle of the positive sequence component of the voltage (s7) for a three phase SC fault (f16) at different irradiance levels. Fault inception at 100 $ms$ . . . . .	135
7.8	Moving average of the phase angle of the negative sequence component of the voltage (s8) for a single phase to ground SC fault (f13) at different irradiance levels. For NO curve shape refer to Fig. 7.5g. Fault inception at 100 $ms$ . . . . .	135
7.9	Signal distortion modeling of the measured $I$ or $V$ signal . . . .	138
7.10	Erroneous (dashed line) vs. ideal measurements (solid line) for symptom s7 under a f16 fault at 1000 and 200 $W/m^2$ . The chosen threshold at $-25^\circ$ appears in a red dashed line. Fault inception at 100 $ms$ . . . . .	139
7.11	Erroneous (dashed line) vs. ideal measurements (solid line) for symptom s8 under a f13 fault at 1000 and 200 $W/m^2$ . The chosen threshold at $-105^\circ$ appears in a red dashed line. Fault inception at 100 $ms$ . . . . .	140
7.12	Erroneous (dashed line) vs. ideal measurements (solid line) for symptom s9 under a f18 fault at 1000 and 200 $W/m^2$ . The chosen threshold at $110^\circ$ appears in a red dashed line. Fault inception at 100 $ms$ . . . . .	140
B.1	Decrease of the phase a rms current (s1) for an inverse bypass diode fault (f02) at different irradiance levels. Fault inception at 100 $ms$ . Under this fault all three phases decrease but only phase a is shown here. . . . .	184

B.2	Decrease of the phase a rms current (s1) for a bypass diode breakdown fault (f04) at different irradiance levels. Fault inception at 100 <i>ms</i> . Under this fault all three phases decrease but only phase a is shown here. . . . .	184
B.3	Decrease of the phase a rms current to zero for an open-circuited boost IGBT fault (f11) at different irradiance levels. Fault inception at 100 <i>ms</i> . Under this fault all three phases drop to zero but only phase a is shown here. . . . .	185
B.4	Decrease of the phase a rms current to zero (s5) for an open-circuited inverter leg fault (f10) at different irradiance levels. Fault inception at 100 <i>ms</i> . The current on the other phases decreases during FO but doesn't drop to zero. . . . .	185
B.5	Decrease of the phase a rms voltage to zero for a short-circuited inverter leg fault (f11) at different irradiance levels. Fault inception at 100 <i>ms</i> . Under this fault all three phases drop to zero but only phase a is shown here. . . . .	186
B.6	Decrease of the phase rms voltage in all three phases (s5) for a short-circuited inverter IGBT fault (f11) at different irradiance levels (for phase a only). Fault inception at 100 <i>ms</i> . . . . .	186
B.7	Sum of phase currents (s5) for a positive pole to ground SC fault (f05) at different irradiance levels. Fault inception at 100 <i>ms</i> . . .	187
B.8	Sum of phase currents (s5) for a negative pole to ground SC fault (f06) at different irradiance levels. Fault inception at 100 <i>ms</i> . . .	187
B.9	Moving average of the zero sequence component of the voltage (s6) for a single phase voltage sag fault (f17) at different irradiance levels. Fault inception at 100 <i>ms</i> . . . . .	188
B.10	Moving average of the zero sequence component of the voltage (s6) for a two phase voltage sag fault (f18) at different irradiance levels. Fault inception at 100 <i>ms</i> . . . . .	188
B.11	Moving average of the phase angle of the positive sequence component of the voltage (s7) for a double phase SC fault (f15) at different irradiance levels. Fault inception at 100 <i>ms</i> . . . . .	189

B.12	Moving average of the phase angle of the negative sequence component of the voltage (s8) for a double phase SC fault (f15) at different irradiance levels. For NO curve shape refer to Fig. 7.5g. Fault inception at 100 <i>ms</i> . . . . .	189
C.1	Branch identification accuracy per dataset size and number of layers.	192
C.2	Distance estimation error per dataset size and number of layers.	192
D.1	Schematic of the grid connected PEMFC system. . . . .	194
D.2	Fuel cell output voltage profile compared to the increase of temperature inside the fuel cell. Three distinct regions (I, II and III) of output voltage increase or decrease can be identified. . . . .	196
D.3	Boost converter output voltage profile as the drying effect progresses. DC bus voltage was set at 800 <i>V</i> . . . . .	196
D.4	Inverter output voltage difference of phase A, $\Delta V$ , between faulty and normal operation as the drying effect progresses. The same behavior is noticed in all three phases. . . . .	197

# List of Tables

3.1	Confusion matrix. . . . .	36
3.2	Total per feeder and phase ( $a, b, c$ ) contracted (loads) and installed (PVs) power. . . . .	45
3.3	Simulation scenarios. . . . .	47
3.4	Distance estimation error ( %) in homogeneous and heterogeneous lines. . . . .	60
4.1	Dataset sizes. . . . .	75
6.1	Comparison of studied parameters between branch identification methods. . . . .	104
6.2	Comparison of studied parameters between fault location methods.	108
7.1	PV system characteristics. . . . .	119
7.2	Fault detectability table of the monitored electrical variables. .	121
7.3	Fault signature table. . . . .	123
7.4	Fault signature table . . . . .	131
7.5	Algorithms isolation speed in $ms$ . . . . .	137
A.1	Characteristics of the eleven different types of conductors connecting the nodes of the grid. . . . .	174
A.2	Cont. characteristics of the eleven different types of conductors connecting the nodes of the grid. . . . .	175
A.3	Characteristics of the different types of conductors per branch. .	176
A.4	Cont. characteristics of the different types of conductors per branch.	177
A.5	Total per branch and phase ( $a, b, c$ ) contracted (loads) and installed (PVs) power in the first feeder. . . . .	178

A.6	Total per branch and phase ( $a, b, c$ ) contracted (loads) and installed (PVs) power in the second feeder. . . . .	179
A.7	Total per branch and phase ( $a, b, c$ ) contracted (loads) and installed (PVs) power in the thirid feeder. . . . .	180
A.8	Total per feeder and phase ( $a, b, c$ ) contracted (loads) and installed (PVs) power. . . . .	181
A.9	Cont. Total per feeder and phase ( $a, b, c$ ) contracted (loads) and installed (PVs) power. . . . .	182

# Chapter 1

## General introduction

New smart meters, distributed generation, renewable energy sources and the concern about the environment are redefining the way to conceive and operate electrical grids. To take full advantage of the new electrical smart grids we need to monitor and protect them. The capability of self-healing is thus important in smart grids in order to ensure a proper behavior under faults and reduce the duration of power outages. Moreover, in electrical power distribution systems, faults are responsible for 80 % of customer interruptions [1]. For this purpose, this thesis proposes three different methods of fault diagnosis for *low voltage (LV)* distribution grids and two methods of fault isolation for *grid-connected photovoltaic systems (GCPVs)*.

Initially, *low voltage (LV)* distribution grids were designed to follow the traditional “fit and forget” doctrine which allowed only a unidirectional flow of power from the distribution transformer to the end consumers. However, the necessary for the fight against climate change installation of photovoltaics and other green types of microgeneration units in the LV grid, obliges operators to shift their attention to the monitoring and control of the LV grid.

While several fault location methods for distribution grids exist in the literature, the majority of them focuses on *medium voltage (MV)* grids. Only a very limited number of studies have considered the LV grid, probably because of its increased complexity compared to the MV ones. A LV grid presents five basic characteristics that hinder the fault location methods: a) an increased number of laterals, b) multi-phase and unbalanced operation, c) unbalanced distribution of loads, d) various types of conductors connecting the nodes and e) up to now, limited number of available measurements (a fact that is now changing with



the transition to smart grids). Moreover, fault location methods for MV grids usually consider only low fault resistance values that rarely surpass the  $100\ \Omega$ . Taking into account that: a) faults of a higher fault resistance can occur, e.g. when a downed conductor touches the earth, fault resistances vary from  $90\ \Omega$  (concrete as ground) to  $1500\ \Omega$  (wet sand as ground), and b) that distribution system operators usually rely on customer phone calls to detect and locate faults in LV grids, the need for fault detection and isolation techniques that cover these cases, i.e. large fault resistances and LV distribution grids, is evident.

The three fault detection and isolation methods proposed in this thesis are:

Chapter 3: A conventional fault detection method based on overcurrent monitoring in combination with a method that uses sparse voltage measurements to build the voltage profile across the faulty branch for fault location.

Chapter 4: *Gradient boosting trees (GBT)*, a method that has been proven to excel in many applications the last few years.

Chapter 5: *Deep neural networks (DNN)*, a method that improve the traditional neural network architecture by taking advantage of an increased number of hidden layers.

Simulations on a real semi-rural LV distribution grid of Portugal are performed to validate the results. A common case study is used to compare the three methods. The considered influencing parameters of this case study are: a) a big variety of fault resistance values (63,772 values between 1 and  $1000\ \Omega$ ), b) nine different fault locations within each sector, c) two fault types (single phase to ground and three phase faults), d) a simultaneity factor of 0.5, e) a big spectrum of PV generation and load demand scenarios with 70,334 studied combinations and f) a 2% underestimation error in measurements.

The contributions of this first part of the thesis are summarized below:

1. Three different methods of fault detection and fault isolation are proposed for the quite unexplored case of LV grids: one that was initially conceived for the MV case and is now extended to the LV grid [2] and two artificial intelligence methods. It is, to the author's knowledge, the first application of artificial intelligence for fault detection and fault isolation in the LV grid case.

2. Fault resistance values of up to 1000  $\Omega$  were considered in this study, something unprecedented for the LV grid and quite rare for the MV one.
3. With the use of GBT and DNN the fault detection problem is tackled with an accuracy of 99.15 % and 100 % respectively.
4. The use of DNN for fault location is the most reliable solution of the three with an average error of only 12 % considering erroneous measurements.
5. All three of the fault location methods are immune to the per phase distribution of loads and microgeneration units in the LV grid.
6. Under the case of extremely limited measurements, one in the beginning of the feeder and one at each terminal node, the fault location accuracy of the DNN is decreased by only 4.5 % while the fault detection accuracy remains intact.

The second part of the thesis, Chapter 7, studies faults in GCPVs, a field that presents an increased interest the last decade due to the rapid growth of installation of renewable energy sources. In 2017 alone, the global PV installed capacity increased by 99 *GW* reaching a total of 403 *GW* which represents the 2.5 % of the global energy demand and 55 % of the installed capacity of new renewable energy sources in 2017 with wind farms following in the second place with 29 % [3].

A big variety of faults can disrupt the operation of a GCPV power plant. Based on their location faults can appear: a) in the PV array, b) in the power converters, c) on the dc bus and d) in the grid side. The development of fast, efficient and reliable fault detection and isolation methods for GCPVs, capable of dealing with the different types of faults, is a recognized necessity from the scientific community and a prerequisite for their integration in the smart grids.

So far, to the author's knowledge, and although different approaches have been tested, no research has been found to monitor the GCPV as a complete system, i.e. isolating faults in every part of the plant with a single method. Moreover, several methods tend to fail in very low irradiance levels. For this reason, two algorithms based on a signal approach, one considering faults only on the dc side of the plant and a second one considering faults on both the dc

and the ac side, that consider various irradiance levels, are proposed as a fault isolation strategy.

The algorithms use current and voltage measurements at the output of the inverter, examining faults occurring on all four of the aforementioned possible locations. The choice of the output of the inverter, i.e. the point of common coupling, as the monitoring source of the status of the GCPV system is in accordance with the location of voltage sensors used in the previous case of fault location methods in the LV distribution grid. The robustness of the algorithms is tested against various irradiance levels (1000, 800, 600, 400 and 200  $W/m^2$ ). Finally, a sensitivity analysis against erroneous measurements is performed.

As part of this second part of the thesis, a preliminary study of fault isolation in grid-connected *proton exchange membrane fuel cells (PEMFC)* was also conducted. In detail, the same fault types with the case study of photovoltaics were studied; in the place of faults at the source of the renewable energy system, in this case, only the drying fault is considered. The isolation of the drying fault is achieved without interfering with the isolation of the faults from the other parts of the system. The complete analysis is presented in Appendix D.

Finally, the contributions of this second part of the thesis are summarized below:

1. For the first time, to the author's knowledge, a study considers faults on both the dc and the ac side with measurements taken only from the ac side. This limitation minimizes the amount of necessary sensors and hence decreases the monitoring cost.
2. The second algorithm, the one that considers faults in both the dc and the ac side, achieves an isolation of 15 out of the 19 studied fault cases.
3. The algorithms are unaffected by different levels of solar irradiance.
4. Robust algorithms, without the need to readjust threshold settings, against distorted measurement signals.
5. The isolation of the faults is achieved in the first 100  $ms$  after their occurrence, time inferior to the reaction margin of the inverter which is expected to trip at approximately 200  $ms$  [4].

# Chapter 2

## Introduction

New smart meters, distributed generation, renewable energy sources and the concern about the environment are redefining the way to conceive and to operate electrical grids. To take full advantage of the new electrical smart grids we need to monitor and protect them. The capability of self-healing is thus important in smart grids in order to ensure a proper behavior under faults and reduce power outage times. For this purpose, this thesis proposes three different methods of fault diagnosis for *low voltage (LV)* distribution grids and two methods of fault isolation for *grid-connected photovoltaic systems (GCPVs)*.

In this chapter, the necessary context and motivation regarding faults and the means to tackle them is provided. Beginning with an outline of the potential threats to the electrical grid, the first section introduces the notion of various threats in an electrical grid, the impact that power outages can have on economies and societies and their effects on the grid itself. Further on, faults in LV distribution grids and GCPVs are analyzed and recent developments in the field of fault detection and isolation are discussed for both cases. Moreover, the research gaps are identified and justified. Finally, in the last section, a description of the structure of this thesis is also provided.

## 2.1 Faults in power systems

In fault diagnosis theory, a fault is defined as the deviation of one of the features of the system from the usual, normal operating condition [5]. In power systems, the occurrence of a fault can, under certain conditions, result to the interruption of electricity supply, i.e. power outage. Power outages can lead to serious consequences of both economic and societal nature, ranging from production losses to risk to health and safety. More specifically, production losses along with restart costs, equipment damage and raw materials spoilage can be very costly. At the same time, uncomfortable temperatures at work or home, loss of leisure time and risk to health and safety, e.g. interrupting hospital service, transportation or industrial operations, are some of the aspects of societal impact of power outages [6].

### 2.1.1 Classification of threats

In order to avoid blackouts, i.e large-scale power outages, and to guarantee the secure and reliable operation of a power system, a thorough analysis of all the influencing parameters and possible causes of such events, both endogenous and exogenous, has been carried out by electrical engineers [7]. Threats causing blackouts have been classified in two big categories: a) conventional and b) unconventional. These two categories can be further partitioned into two sub-categories each: a1) natural threats, a2) accidental threats, b1) malicious threats and b2) emerging threats.

The geographic location of the power systems exposes them to a big variety of natural threats around the globe. Those threats, in order of frequency of appearance as a cause for a blackout, are: windstorm, rainstorm, thunderstorm, blizzard, cyclonic storm, ice storm, cold storm, heat wave, lighting strike, earthquake, tornado, drought, flood, hailstorm, landslide, geomagnetic storm, tsunami, wildfire and others. The other type of conventional threats, accidental threats, in order of frequency of occurrence, include: technical failure, fire/explosion, human or animal interference, operation mistake, equipment defect or aging and maintenance error.

To the already big number of possible causes of blackouts listed above, malicious and emerging threats come to add up to the list. As the dependence

of our societies on electricity increases, malicious threats including hacking, sabotage, terrorist attacks and acts of war are becoming more and more popular. With the evolution of the traditional electrical grid to a *Smart Grid*, through the technological advancements in telecommunications supporting electrical power systems and the increased number of hardware being installed on them, the electrical grid is becoming vulnerable to new types of threats such as cyber-attacks and more specifically to: a) network availability, b) data integrity and c) information privacy [8]. Furthermore, the increasing interconnectivity of our societies' infrastructures, i.e. electric power, water, oil, natural gas, transportation and telecommunications, makes it possible for faults to transpose from one system to another. Finally, an example of interconnectivity as an emerging threat, is the increased vulnerability of the electrical grids to geomagnetic storms as their interconnection between different countries increases the length of the transmission lines and therefore renders them more exposed to larger geomagnetically induced currents.

As power outages are a frequent phenomenon, it is impossible to record all of them. For this purpose, three criteria were proposed by [7], to distinguish the most important events from the less significant ones:

- a) *affected population*  $> 1,000$  *residents*
- b) *duration*  $> 1$  *h*
- c) *affected population*  $\cdot$  *duration*  $> 1,000,000$  *resident-hour*

Extending the analysis done in [7], which included blackouts until 2011, a total of 188 blackouts that satisfy the above criteria have been recorded to date and their causes are presented in Fig. 2.1. Two important conclusions can be drawn from Fig. 2.1. Firstly, an increase of faults due to natural causes is noticed over the last decade; this is yet another evidence of how climate change can affect power systems and our lives. Secondly, as described above, the appearance of malicious and emerging threats is observed after the year 2000. In addition, Fig. 2.2 presents the number of blackouts for which each category was considered responsible over the total studied cases.

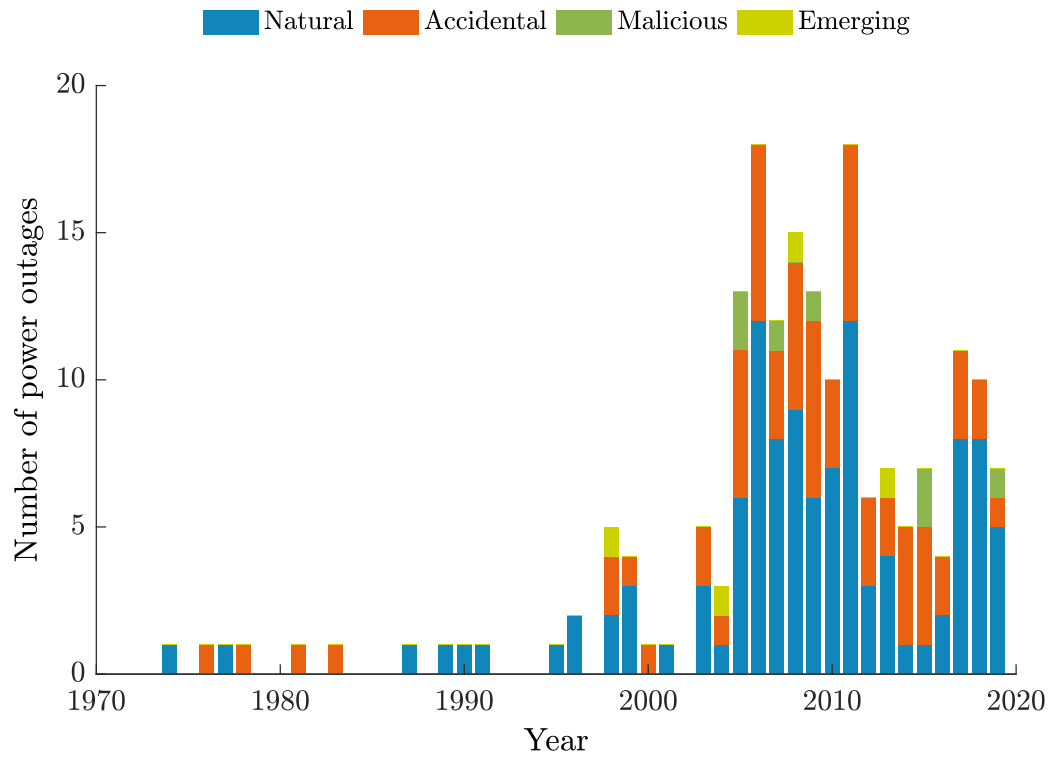


Figure 2.1 Causes of historical blackouts from 1974 to 2019.

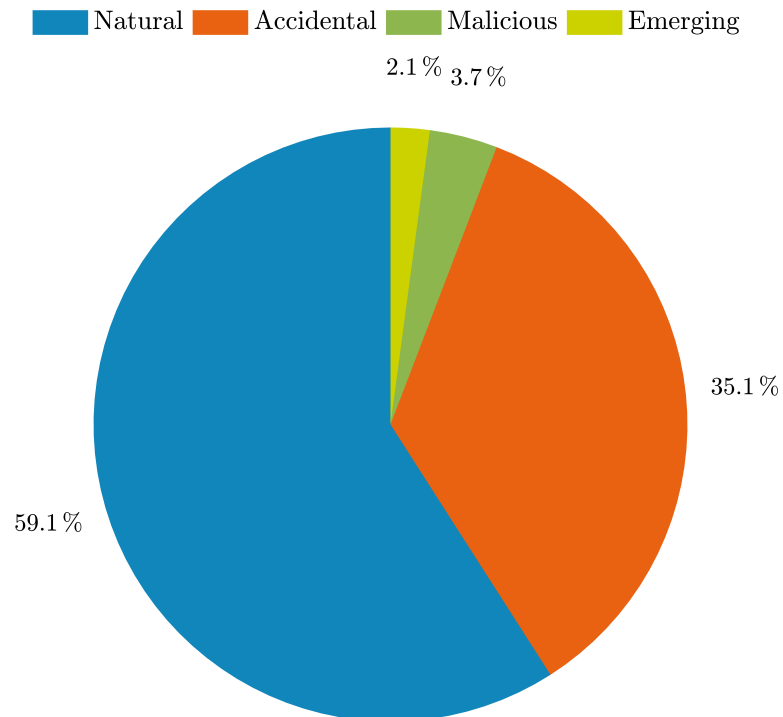


Figure 2.2 Percentages of causes of outages.

### 2.1.2 Costs and notable events

While the economic and societal consequences of blackouts are huge, it is very difficult to calculate the cost of a power outage as it is a multivariant equation with many factors that are difficult to estimate, e.g. customer behavior or company reliability loss [9]. In an attempt to measure the impact of the faults on the customers, two metrics are proposed: a) the willingness to pay and b) the *value of lost load (VoLL)*. The willingness of a residential customer to pay in order to avoid an one hour outage event, in USA, was found to be \$ 3 on average [10]. The VoLL (€/kWh) on the other hand, is defined as the ratio of the economic value of leisure in households over the total household consumption. An annual average of 8.37 €/kWh was measured in Europe in 2013 [11].

In order to get a better grasp of the magnitude of the cost that blackouts can have, five examples are provided below: the first three caused by natural phenomena and the last two by accidents.

Caused by natural phenomena:

- A threat to which we are exposed to and still quite unprotected is the geomagnetic storms. Solar flares from the Sun release charged particles that upon reaching the Earth's magnetosphere cause disturbances in the magnetic field and create geomagnetically induced currents that can harm the electrical grid. From the power systems point of view, the most notable event took place on 13 March 1989, when the whole province of Quebec, experienced a major nine hours blackout. The problems faced by electric utilities during this storm served as the basis for future studies. The estimated cost of an extreme geomagnetic storm that would hit the USA could potentially reach \$ 2.6 trillion [12]. On 23 July 2012, a powerful coronal mass ejection took place. The unleashed solar wind missed the Earth by nine days. If it had reached the Earth it would have been one of the biggest and most catastrophic geomagnetic storms in history. As we leave 2019, a year that marks the beginning of the 25<sup>th</sup> solar cycle, utilities should remain vigilant as even numbered solar cycles tend to be cycles with increased solar activity [13].
- The tsunami that was caused after an earthquake in Japan and hit the Fukushima Daiichi power plant on 11 March 2011 caused a nuclear accident.



The financial cost of this accident was huge; \$ 118 billion cost for the Japanese taxpayers [14] and another \$ 47.3 million [15] that the government and the Tokyo Electric Power Company were forced to compensate with 137 families for their negligence. Moreover, the ecological disaster was also enormous with the radioactive decontamination of the area projected to require a few decades. Although unexpected, the Fukushima event could have been avoided.

- The largest blackout in the European continent was experienced on 28 September 2003 when a power line that connected Italy to Switzerland was damaged by windstorms leaving millions of people in both countries without power for twelve hours. People got trapped in elevators, trains and underground metros on a cold rainy night. The event was reported to have caused a damage of € 120 million to the local economy [9].

Caused by accidents:

- Some technical faults in combination with a contact of some power lines with trees led to the famous blackout of 2003 in the USA and Canada, affecting 145 million people with a total estimated cost of \$ 6 billion.
- The largest blackout so far is attributed to a series of technical faults. It happened in India on 30-31 July 2012 and left 700 million people without power for about fifteen hours and businesses with \$ 107 million losses [16].

### 2.1.3 System average interruption duration index

Another tool to measure the impact of power outages is the *system average interruption duration index (SAIDI)*. The SAIDI is measured in minutes per customer and is defined as the average duration of all interruptions (sustained outages) per utility customer in a yearly basis.

$$SAIDI = \frac{\sum N_{int} \cdot D_{int}}{N_{total}} \quad (2.1)$$

where  $N_{int}$  is the number of customers that experienced an electricity interruption,  $D_i$  is the duration of each interruption and  $N_{total}$  the total number of customers.

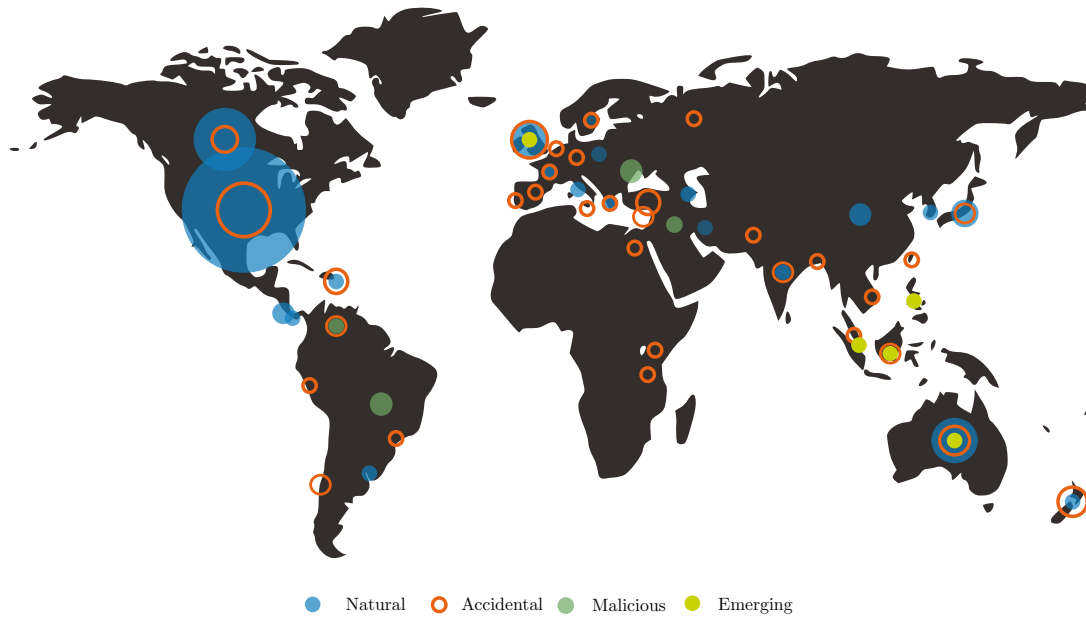


Figure 2.3 Number of major historical blackouts per country and cause.

The maximum, minimum and average unplanned SAIDI values (including exceptional events) for the countries in Europe, between 2002 and 2016 are presented in Fig. 2.4. A decreasing tendency of the SAIDI can be easily observed. From a value of 177.1 *min/customer* in 2002, a value of 92.45 has been attained in 2016, where the SAIDI presented a minimum value of 9 *min* in Switzerland and a maximum of 371 *min* in Romania [17]. In 2018, according to ENEDIS [18], France achieved a total SAIDI of 63.87 *min/customer* (50.78 of unplanned and 13.09 of planned outages) for its LV customers.

#### 2.1.4 Affected grid elements

The most popular root causes of fault occurrences are severe storms damaging overhead lines and strong winds bending trees thus bringing them in contact with one or more of the overhead lines. The percentage occurrence of faults in different parts of the electrical power distribution system is presented in Fig. 2.5 [19]. It is shown that faults in overhead lines cover 50 % of the fault cases followed by faults in switches, various electrical equipment, transformers, underground cables and finally generators.

In this thesis faults in overhead lines or underground cables and generators will be examined thus covering 66 % of the cases.

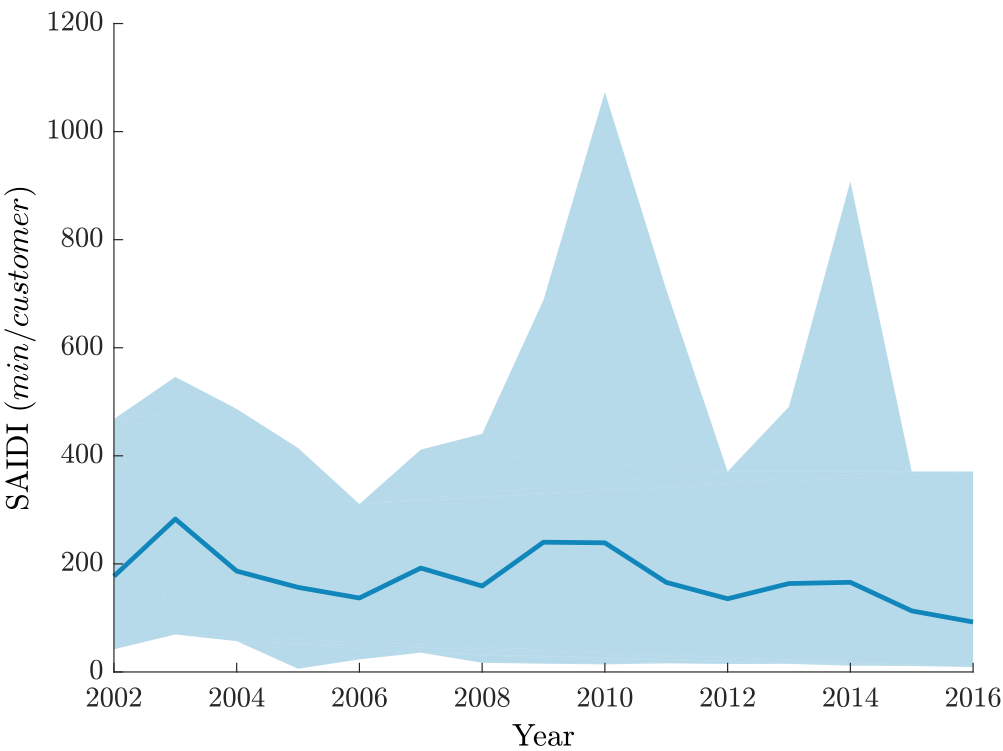


Figure 2.4 Unplanned SAIDI including exceptional events for 28 countries of Europe. The shaded area shows the zone of recorded SAIDI values. The blue line represents the average value per year.

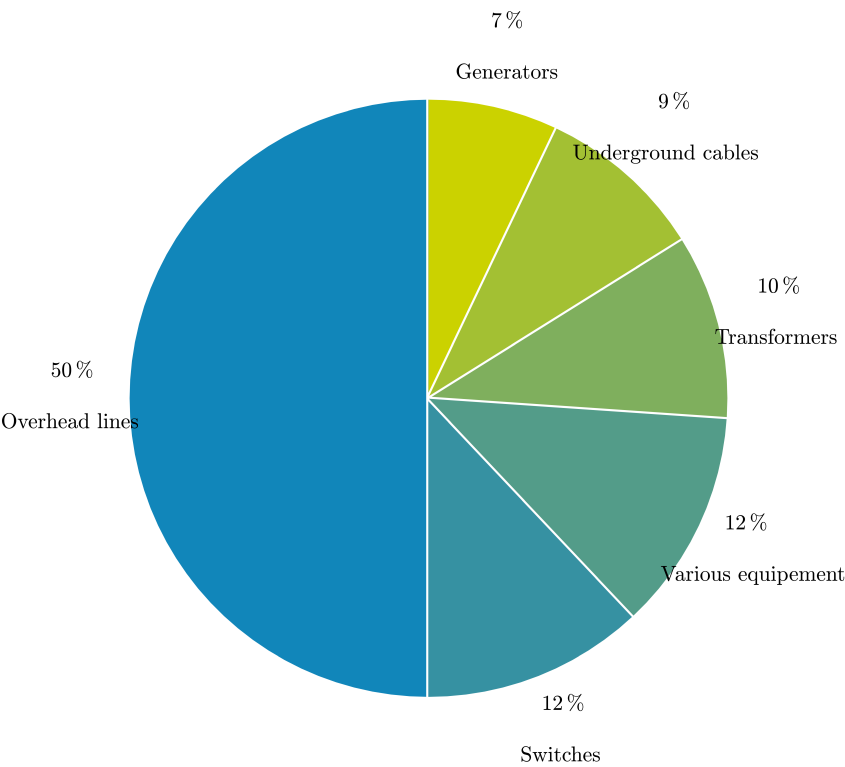


Figure 2.5 Fault occurrence in different grid parts.

## 2.2 Faults in distribution grids

In electrical power distribution systems, faults are responsible for 80 % of customer interruptions [1]. As mentioned in the previous section, faults in distribution grids are basically located in overhead lines. Three examples of damaged electrical distribution poles after exposure to extreme weather conditions are provided in Fig. 2.6.

The most common faults in distribution systems are the single phase to ground *short circuit (SC)* faults which account for 70 % of the total fault cases. Such faults occur for example when one phase comes in contact with a tree or the ground if the tower is damaged. Other types of faults can be phase to phase faults, e.g. when strong wind brings one phase in physical contact with another, double phase to ground, and three phase faults. The most severe, the three phase faults, account for 5 % of the total fault cases. The rest 25 % is divided among the phase to phase faults (15 %) and double phase to ground faults (10 %) [1]. At this point, the notion of fault resistance,  $R_f$ , should be also introduced, as the fault impedance path between the phase and the ground [20].

### 2.2.1 Low voltage grids

The LV distribution grid is the final link of the electricity supply chain that connects customers with the distribution substation. Initially, LV distribution grids were designed to follow the traditional “fit and forget” doctrine which allowed only a unidirectional flow of power from the distribution transformer to the end consumers. With the integration of renewable energy sources into the grid, bidirectional power flow is becoming a reality and *distribution system operators (DSOs)* face several problems such as congestion, voltage rises and decrease of power quality [21]. Until recently, the focus of operators was to improve the transmission and *medium voltage (MV)* distribution parts of the grid, leaving the LV part unmonitored and uncontrolled. However, the necessary for the fight against climate change installation of *photovoltaics (PV)* and other green types of microgeneration units in the LV grid, obliges operators to shift their attention to the monitoring and control of the LV grid increasing its functionalities through installation of smart meters, e.g the Linky smart meter in France, and implementing self-healing strategies [22].



(a) More than 30,000 people were left without power after an ice storm in the Altus area, Oklahoma on January 28, 2010 [U.S. Air Force photo by Senior Airman Leandra D. Hernandez/Released].



(b) A power pole leans over a burned property during the Carr fire near Redding, California on July 27, 2018 [JOSH EDELSON / AFP #000\_17Z6A6].



(c) Storm downed power lines and trees in the aftermath of Storm Sandy that hit Arlington, Virginia on October 29, 2012 [<https://www.flickr.com/photos/arlingtonva/8138919297>].

Figure 2.6 Downed electric towers resulting in power outages.

In contrast to the MV grid, the LV one is more complex and difficult to manage due to some of its unique characteristics [23]. The most important of those are cited below:

- radial structure with a high number of laterals (branches) and nodes
- multi-phase and unbalanced operation
- unbalanced distribution of loads
- different type of conductors connecting the nodes with different characteristics and lengths hence wide range of resistance,  $R$ , and reactance,  $X$ , values.

Distribution lines present a high  $R/X$  ratio ( $R/X > 1$ ) which can vary a lot due to the different type of conductors connecting the nodes of the LV grid. All the above characteristics attribute an unbalanced and heterogeneous nature to the LV grid and complexify its analysis. As it will be analyzed in the next chapters, all four of these features are important obstacles in the fault detection and location processes.

The path to an adequately inventoried LV grid and its smartification is paved with the advancements in telecommunication technologies and the installation of advanced metering infrastructure in the LV grid, especially smart meters, establishing thus bidirectional communication channels between *supervisory control and data acquisition (SCADA)* control centers and prosumers (i.e. consumers who also generate energy). Through the necessity of more information and control over the LV grid and the availability of new tools and metering infrastructure the *smart grid* concept emerged.

### 2.2.2 Smart grids

According to the different definitions of a smart grid [24], one of its key features is the ability to self-heal with the aid of advanced metering and communication tools, and intelligent monitoring, aiming to a more secure, cost-effective and reliable operation. In Europe, a smart grid is defined as “*an electricity network that can intelligently integrate the actions of all users connected to it – generators, consumers and those that do both – in order to efficiently deliver sustainable, economic and secure electricity supplies*” [25].

The backbone of self-healing strategies are the *fault detection, fault isolation and fault restoration processes (FDIR)*. When a fault occurs in a distribution grid, a certain strategy is needed to deal with it and minimize its negative effects. Before going into more details, let us give the definition of some basic terms:

- *Fault detection* is the process of recognizing that a fault occurred and that the grid is not operating as intended.
- *Fault isolation* is the process of identifying the fault, i.e. what kind of fault has occurred, and localizing it, i.e. finding where exactly it happened.
- *Fault diagnosis* is the combined effort of both fault detection and isolation.
- *Fault restoration* is the process of reconfiguration of the grid that is set in action after fault diagnosis in order for the grid to return back to operating condition.

At this point, it should be underlined that despite the serious effects of power outages described in the previous sections, many utilities still rely on customer phone calls to detect a fault [26]. Once a fault is reported, operators send a crew to locate it and fix it. All the actions from fault occurrence to fault restoration and the average time frame for each step are provided in Fig. 2.7. With a minimum time of 45 *min* without power, this very time consuming process might leave customers without power for several minutes or even hours.

However, in the context of smart grids, new solutions arise proposing the automation of the FDIR processes and minimize human interference [29]. These functionalities are reducing the outage time by up to 45 times [27] and the operational cost of the grid thus increasing its reliability. In a recent study [30], ENEDIS, the main French DSO, in collaboration with ADEeF, the French DSO







events	fault inception	customer phone call	send patrol crew	fault located	restoration of healthy feeder sections	normal operation
						
time frame	5-10 min	15-30 min	15-20 min	10-15 min	1-4 hours	

Figure 2.7 Fault detection, isolation and restoration steps time frame [27]. Icons designed by [28].

association, reported profits of € 3.3 million per year from the implementation of self-healing tools in a MV smart distribution grid. At the same time, the benefits of just a situation awareness tool incorporating a distribution management system and automatic meter reading devices in a Finnish DSO substantially reduced the cost of power outages [31].

### 2.2.3 Fault diagnosis in MV distribution grids

In the literature, several attempts have been made to automatize the FDIR process. The available fault location methods can be divided in four main categories: a) the conventional methods, including impedance-based [32–58] and traveling wave methods [59–74], b) methods based on sparse measurements [75–81, 2, 82], c) the knowledge-based methods that use artificial intelligence [83–112] and d) hybrid methods [113–128].

From the first category (conventional methods), impedance-based methods are the most widely used for their simplicity and low implementation cost. They are a whole family of methods with many variations. The most popular one is the one-end technique. From current and voltage measurements only at the substation level and sufficient knowledge of the distribution line model, the estimation of voltages and currents in different parts of the line becomes possible. A simplified example is provided in Fig. 2.8. By applying Kirchoff's law, the fault location,  $f_d$ , is obtained from the following equation:

$$f_d = \frac{V_s - R_f \cdot I_f}{Z_l} \quad (2.2)$$

where  $R_f$  is the fault resistance,  $I_f$  the fault current,  $V_s$  the source voltage and  $Z_l$  the line impedance in per unit length.

Traveling wave methods, although initially applied to transmission systems, have been used for distribution systems as well, as they can be very fast and quite accurate. After a fault occurrence, transient voltages and currents (traveling waves) are generated at the fault location and propagate through the distribution line in both directions. Once the wave reaches the end of the line, part of it will get reflected back towards the fault location. Upon reaching the fault location, a part of the wave will get refracted and continue towards the other end, and another part will get reflected. The time difference,  $\Delta t$ , between the time of



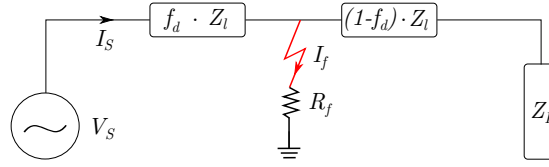


Figure 2.8 Impedance-based method.

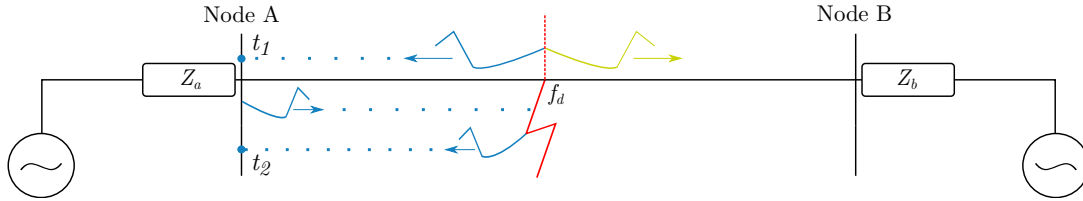


Figure 2.9 Traveling wave method.

arrival of the first wave in one end of the line,  $t_1$ , and the arrival of the second (its reflection),  $t_2$ , can be used to calculate the fault location since the speed of the traveling wave is equal to the speed of light,  $c$ .

$$f_d = c \cdot \frac{\Delta t}{2} \quad (2.3)$$

With the increased availability of measurements that the smart grid concept brings, new methods have been tested against the fault location problem. The first category of such methods is those that are based on sparse measurements. Although quite different from one another, most of them are based on the common principle of a voltage sag creation with different characteristics at different nodes after a fault occurrence. In detail, as a first step, considering every node to be under fault, a set of calculated voltage sags is obtained. Then, by measuring voltage sags in various nodes and comparing those measurements with the pre-calculated values, the location of the faulty node is becoming possible: the one presenting the minimum difference between the calculated and the measured values is the one under fault.

Another approach that emerged with the smartification of the grid, is the use of artificial intelligence. Different aspects of artificial intelligence have been employed to tackle the fault location problem including: a) artificial neural networks [83–86, 89, 90, 95, 96, 99, 101, 103, 104, 107, 111, 112], b) support vector machine neural networks [92–94], c) fuzzy logic [87, 91, 106, 108, 110], d) genetic algorithms [102, 109] and e) decision trees [97, 105]. All these methods, also called learning or knowledge-based, are trained offline from a big dataset

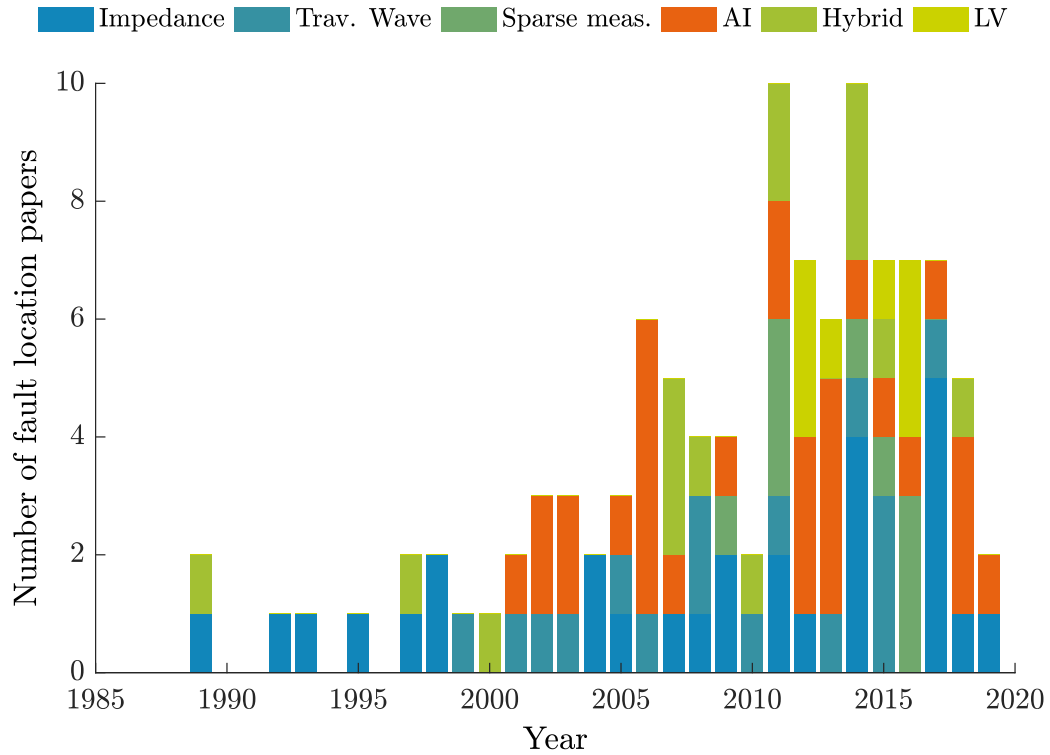


Figure 2.10 Fault location research trend.

including different fault scenarios and variables depending on the availability of measurements, the grid topology and the method itself. They are capable of identifying links and patterns between the input features that conventional methods cannot not. When exposed to the real fault cases, based on their training, they try to estimate the fault location.

Finally, neural networks and fuzzy logic have been used in combination with conventional methods to create hybrid tools [108, 116, 126].

While several methods exist in the literature, the majority of them focuses on MV distribution grids with only a very limited number of studies being performed on LV grids [129–135]. The research trend of all the 107 papers cited above is presented in Fig. 2.10. There, the changes the smart grid brought to distribution grids, can be observed. First of all, the appearance of unconventional methods can be noticed after the year 2000 and especially those of sparse measurements (after 2009). Additionally, as stated before, the shift of the research interest to the LV grid is happening now. With the oldest method dating back to 2012, the LV grids are likely to shape the research trend of fault location in the near future.

It should be also noted, that the existing literature of fault location, besides focusing mainly on MV grids, it has been also limited to low resistance faults, with fault resistance values rarely surpassing  $100\ \Omega$ . Taking into account that faults in LV grids are usually located manually [130], and that faults of a higher fault resistance can occur, e.g. when a downed conductor touches the earth fault resistances vary from  $90\ \Omega$  (concrete as ground) to  $1500\ \Omega$  (wet sand as ground), the need for fault detection and isolation techniques that cover these cases, i.e. large fault resistances and LV distribution grids, is evident.

A second problem of the existing literature methods is the disadvantages they present and the fact that some of them could be amplified if applied to LV grids. More specifically, impedance-based methods present limited accuracy, require the knowledge of the fault resistance and may identify multiple possible locations for a fault [26, 29]. In the case of the LV grid, which as mentioned in a previous section, consists of a high number of branches, the estimation of multiple possible fault locations is a big problem. Additionally, traveling wave methods can be as well seriously affected by the presence of multiple branches which hinder the distinction between the different traveling waves [26, 29]. Moreover, traveling wave methods require high speed data acquisition, increasing significantly their cost. Furthermore, both impedance-based and traveling wave methods, depend on line parameters which in distribution grids vary a lot thus affecting their accuracy.

Most of the aforementioned problems could be overcome with the use of unconventional methods, either based on sparse measurements or artificial intelligence. It should be noted that to the author's knowledge there has not yet been such an implementation in the LV grid. For this reason, three methods of fault detection and fault isolation were developed in this thesis:

1. A conventional fault detection method based on overcurrent monitoring in combination with a fault location method based on sparse voltage measurements.
2. Gradient boosting trees, a method that has been proven to excel in many applications the last few years.
3. Deep neural networks, a method that improves the traditional neural network architecture by taking advantage of an increased number of hidden layers.

## 2.3 Faults in photovoltaics

### 2.3.1 Solar energy

Solar energy is inexhaustible. The amount of solar irradiation that reaches the Earth's surface is tremendous: approximately 10,000 times more than the current humanity's energy needs for a whole year [136]. Besides solar power, many other forms of renewable energy such as wind energy, hydropower and biomass depend on the Sun. Wind, used by wind farms to produce electricity, is created by the differential heating of different regions of the Earth; water evaporates because of Sun's heat, forming clouds and then through rain water masses are transported at higher altitudes thus creating a potential energy difference of water that hydropower plants take advantage of; organic matter stores energy through the process of photosynthesis which would not be possible without the Sun.

The Ancient Greeks were the first to demystify the Sun and tried to understand its nature, developing different theories over the centuries. The 3<sup>rd</sup> century B.C., Aristarchos the Samian (“Ἀρίσταρχος ο Σάμιος”), followed the teachings of Pythagoreans who were the first - back in the 6<sup>th</sup> century B.C.- to question the geocentric theory and formed the hypothesis of the heliocentric system. He also proposed the simultaneous rotation of the Earth around its axis on a daily basis and its movement in a circular orbit around the Sun on an annual basis [137].

Many centuries passed and several solar applications emerged until the year 1839, when the biggest milestone in solar energy was reached. The French physicist Alexandre-Edmond Becquerel discovered the *photovoltaic (PV)* effect, a way to produce electric current from sunlight. In 1883, Fritts constructed the first PV device made by Selenium (Se) [136]. Seven decades later, in 1954, Bell Laboratories announced the creation of a 6% efficient silicon (Si) solar cell [138]. Another milestone was reached in 1958 when Vanguard I, the first solar-powered satellite, was launched at Cape Canaveral, Florida. Since then and especially during the last two decades, the solar sector has experienced a remarkable growth.

### 2.3.2 Photovoltaics

A list of the major applications of PV in power systems is provided by the International Energy Agency (IEA) [3]:

- Pico PV systems (lighting, phone charging, powering small computers)
- Off-grid domestic (household loads: lighting, refrigerator, washing machines etc.)
- Off-grid non-domestic (telecommunications, water pumping, navigation etc.)
- Hybrid systems (PV and diesel generators)
- Grid-connected distributed (power either directly to the customer or to the electrical network)
- Grid-connected centralized (supply bulk power)

Grid-connected photovoltaic systems is a remarkably growing sector representing 55 % of the installed capacity of new renewable energy sources in 2017 with wind farms following in the second place with 29 % [3]. In 2017 alone, the global PV installed capacity increased by 99 *GW* reaching a total of 403 *GW* which represents the 2.5% of the global energy demand. Figure 2.11 shows the cumulative increase of PV installed capacity. It is remarkable that within just three years (2014-2017) the installed capacity was doubled! In Fig. 2.12 the cumulative evolution of the distributed and centralized GCPVs installations is presented. It is shown that, although distributed GCPVs were dominating the PV field until 2014, centralized GCPVs are rapidly becoming more popular, with China leading this trend [3].

### 2.3.3 Grid-connected photovoltaic system topologies

In GCPVs a converter is necessary to convert the *direct current (dc)* electricity produced by the PV array to *alternating current (ac)* that is supplied to the grid. Different grid-connected arrangements of PV modules and converters have been proposed over the years by various researchers [139]:

- *cell inverter*: directly connect the PV cell to the grid via an inverter
- *module inverter*: one inverter per module

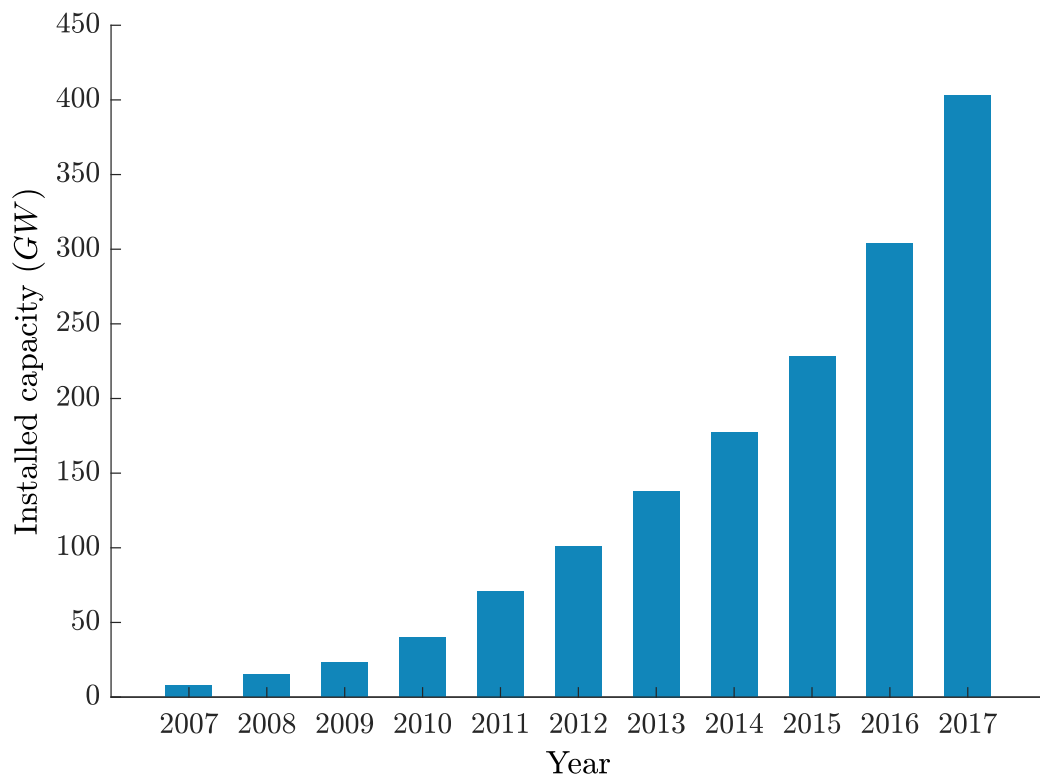


Figure 2.11 Cumulative evolution of PV installations [3].



Figure 2.12 Cumulative evolution of GCPVs installations [3].

- *series connected dc-dc converter*: one dc-dc converter per module but modules are connected in series and all of them are connected to a common inverter
- *parallel connected dc-dc converter*: one dc-dc converter per module connected to a common inverter
- *multi-string inverter*: one dc-dc converter per string connected to a common inverter
- *string inverter*: using one inverter per PV string of the array
- *centralized inverter*: using a single inverter to connect the PV array to the grid.

Among the different topologies the one that presents the best trade-off between cost and efficiency is the one of the centralized inverter as reported in [139].

### 2.3.4 Modeling of photovoltaic systems

The modeling of a GCPV system can be decomposed into three simple steps. First of all, the use of the mathematical equations of the PV cell's equivalent circuit is necessary for the extraction of the characteristic parameters; the calculations at this stage are subject to the limited available data provided by the manufacturer's data sheets [140]. After the extraction of the parameters and based on their values, follows the solution of the transcendental current-voltage characteristic equation [139]. The last step is the design of the PV array by connecting in series and in parallel all the necessary elements and choosing the mode of its connection to the grid.

#### 2.3.4.1 PV cell: Bishop's model

An alternative to the widely used one-diode model [140] and a more appropriate one for fault detection simulations, is the Bishop's model [142], an upgraded version of the one-diode model. Its equivalent electrical circuit is given in Fig. 2.13a. The additional element connected in series with the shunt resistance is a non-linear multiplication factor. In that way, the Bishop's model takes into account the avalanche effect of the diode thus permitting the monitoring of the

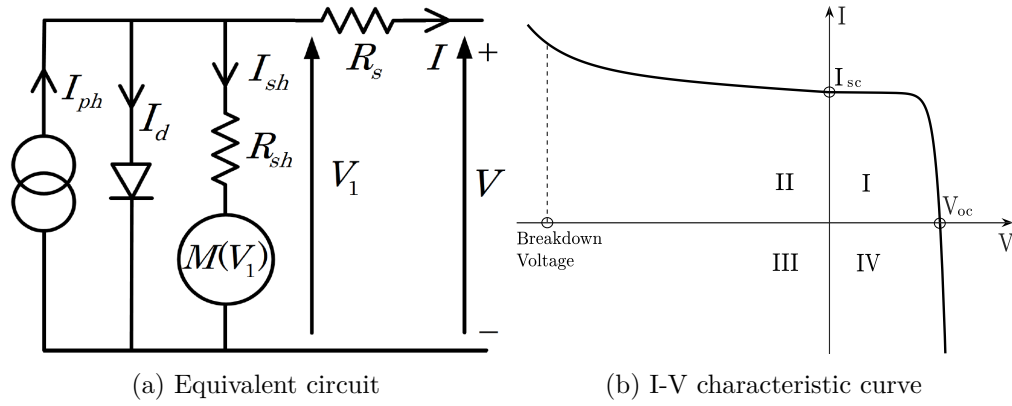


Figure 2.13 Bishop's model [141].

cell's response in the three quadrants - the first (I), the second (II) and the fourth (IV) - of the  $I$ - $V$  characteristic curve as presented in Fig. 2.13b.

The transcendental equation of the equivalent circuit is provided by 2.4.

$$I = I_{ph} - I_0 \cdot \left[ \exp\left(\frac{V + R_s \cdot I}{V_t}\right) - 1 \right] - \frac{V + R_s \cdot I}{R_{sh}} \cdot \left[ 1 + k \cdot \left( 1 - \frac{V + R_s \cdot I}{V_b} \right)^{-n} \right] \quad (2.4)$$

Eight parameters can be distinguished from (2.4). These parameters are: the photo current  $I_{ph}$ , the diode reverse saturation current  $I_0$ , the diode's thermal voltage  $V_t$ , the series resistance  $R_s$ , the shunt resistance  $R_{sh}$ , the Bishop's coefficients  $k$  ( $\sim 0.1$ ) and  $n$  (3.4-4) and the breakdown voltage  $V_b$ . Since (2.4) is a transcendental equation, a numerical solution is required. The Newton-Raphson algorithm [139] is usually used to solve the equation, although the Lambert-W function could be an alternative solution [143].

#### 2.3.4.2 PV array

Having solved the characteristic  $I$ - $V$  equation, the last remaining step in order to complete the modeling of the PV array, is the connection of several elements together. Multiple cells are connected in series to form a PV module. PV modules are also connected in series forming a PV string which when connected in parallel create a PV array. The design of the PV array is completed with the use of blocking and bypass diodes as protective elements against the flow of inverse currents and inverse polarization respectively.



### 2.3.4.3 PV system

In this work, a slight variation of the centralized mode of connection of the PV array to the grid, described in the previous section, is used: a two-stage conversion centralized mode. The first stage of conversion is achieved with a dc-dc boost converter which is used to control the dc output of the PV array and extract the maximum power of the PV array in coordination with a maximum power point tracker (in this case the perturb-and-observe algorithm is chosen). The second stage of conversion involves an inverter that converts the dc current to ac and can ensure ancillary services if required.

### 2.3.5 Fault types

Faults in PVs can be divided in three main categories [144]: a) physical (damaged bypass diode, cracks in PV modules, degradation, etc.), b) environmental (shading) and c) electrical (*open circuit (OC)*, SC and arc faults). Since in a GCPV system faults can be encountered in every component, a classification based on their location is more fitting. A few representative examples that present special interest due to their frequency of appearance and their severity are provided below.

First of all, different kinds of faults can appear inside the PV array itself. An extensive analysis of such faults is provided in [141] and includes four subcategories: a) faults of the PV generator (cell deterioration, module SC etc.), b) faults inside the junction box, c) faults of the connecting cables (OC, SC, inverse module etc.) and d) faults of bypass and blocking diodes (destruction of the diodes, inverse polarization etc.).

The second location where faults could appear, is inside the power converters, in this case the boost converter and the inverter. Again an extensive list of possible faults is provided in [141] ranging from faulty power electronic elements to sizing and compatibility problems.

The third location where faults appear, is on the dc bus connecting the two power converters. There, a SC can appear between either the positive or the negative pole and the ground, posing a threat to the system itself.

The final location where we can encounter faults, is the grid side of the system. Faults occurring anywhere on the ac side can also affect its performance. Such

faults are mainly SCs between the lines and the ground or SCs between the lines themselves. The possible causes of such faults have been explained in the previous sections.

### 2.3.6 Fault diagnosis in photovoltaics

The development of fast, efficient and reliable fault detection and isolation methods for GCPVs, capable of dealing with the different types of possible errors, is a recognized necessity from the scientific community and a prerequisite for their integration in the smart grids.

Among the numerous fault detection and isolation approaches, six big categories can be distinguished: a) comparative (usually model-based) [145–170], b) artificial intelligence [171–190], c) signal-based [191–203], d) time domain reflectometry [204–208], e) statistical [209–215] and f) others [216–226].

The first category, comparative methods, includes mostly model based fault detection and isolation techniques. The measured monitored variables (usually power, current and voltage) or other quantities, are compared with their pre-calculated/expected values for the given operating conditions. The residuals are then used to establish the necessary thresholds. A common fault detection technique of this category is the tracing of module  $I$ - $V$  curves in order to detect any anomalies and locate accurately the faulty modules. An alternative approach proposes the use of the  $(dI/dV)$ - $V$  curve instead. These methods, are often preferred due to their simplicity.

The second most popular category in the literature concerning GCPV is the lately growing field of artificial intelligence. Machine learning algorithms are trained offline based on various fault and healthy operation scenarios to detect and isolate faults. Then when exposed to real operating conditions they are able to provide a diagnosis. As stated before, the merit of this approach is that these algorithms are able to create links and patterns between the available data that the other methods cannot identify.

Signal-based methods are yet another proposed approach that aims at identifying changes in the output current or voltage waveforms. The most common tools used in this category are the time domain discrete wavelet transform and short-time Fourier transform. These tools are used to extract certain features which are used as a basis for the creation of the fault signature.

A less popular method, time domain reflectometry, is also used for fault detection and isolation in GCPVs. The method is based on the injection of a test signal in the system. In the presence of a fault, the signal will get reflected back (similarly to the traveling wave method described in the previous section) and delays and changes of its waveform are used for fault diagnosis. The merit of this method is that it does not require the measurement of any variable.

Furthermore, statistical methods use various tools to analyze available data such as the standard deviation, mean and rms values, as well as an exponentially weighted moving average control chart.

Finally, various alternative methods have been also proposed by researchers using cutting edge technology such as satellite image processing or flying drones over the PV array to monitor the GCPVs operation and detect certain types of anomalies.

The research trend of the 82 papers cited above is presented in Fig. 2.14. It is shown that the research interest of a reliable and secure operation of the GCPV system has increased substantially after 2011. A sudden shift of the researchers' interest in artificial intelligence techniques is also noticeable the last three years. Moreover, interestingly enough, signal-based approaches are continuing, steadily for the last decade, to being proposed as a solution to the fault detection and isolation problem of GCPVs.

Although many times efficient, the proposed methods in the literature present several disadvantages [144]. First of all, the comparative methods are strongly depended on the accuracy of the designed models and they are quite vulnerable to mismatch losses, partial shading and low irradiance, and under such conditions they can lead to the activation of false alarms. Moreover, artificial intelligence and statistical methods strongly depend on the availability of measured data and they can be quite costly since they require a lot of measurements. Furthermore, signal-based approaches can be quite limited on the variety of faults they consider. Finally, in a complex system such as the GCPV, the accuracy of reflectometry methods, especially when trying to isolate a fault, can be easily affected as they are prone to errors caused by noise in the signals.

So far, to the author's knowledge, no research has been found that monitored the GCPV as a complete system; each one of them is concentrated on a specific part of the GCPV in order to detect and isolate faults. Additionally, up to

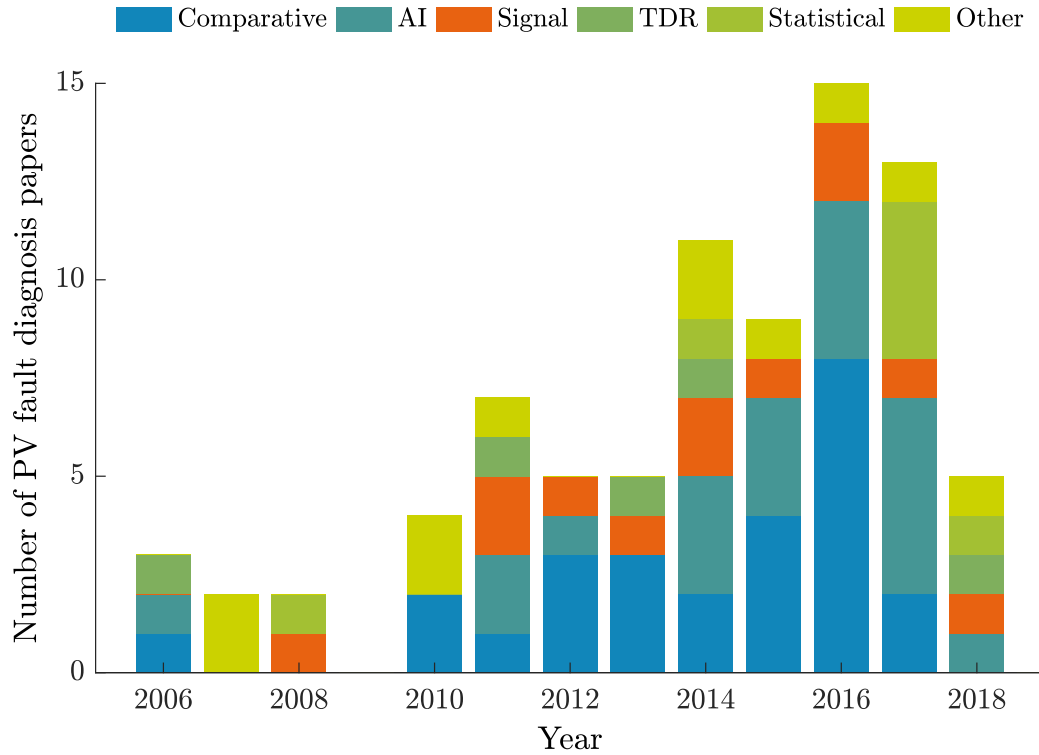


Figure 2.14 Fault detection and location for GCPV systems research trend.

now, no study has focused on the possibility of detecting and isolating faults on both the dc and the ac side with measurements taken only from the ac side. Such a possibility would significantly reduce the number of sensors required for the supervision of the system and thus the total system cost. Finally, several methods tend to fail at very low irradiance levels.

For this reason, in this work, the signal approach was selected to develop a fault isolation method for GCPV systems based on current and voltage measurements at the output of the inverter, examining faults occurring on both the dc and the ac side of the plant at different irradiance levels. The choice of the output of the inverter, i.e. the point of common coupling, as the monitoring source of the status of the GCPV system is in accordance with the location of voltage sensors used in the previous studied case of fault location methods in the LV distribution grid.

## 2.4 Thesis outline

This thesis is divided in two parts: part I) fault diagnosis in *low voltage (LV)* smart distribution grids and part II) fault isolation in *grid-connected photovoltaic systems (GCPVs)*.

The first part is divided in four chapters, one for each of the three different proposed methods and a fourth one for a comparative analysis. Chapter 3 is dedicated to a more conventional and simple fault detection and isolation method. This method uses *root-mean-square (rms)* current measurements at the beginning of each feeder and voltage measurements spread across the grid to detect and localize a fault respectively. Chapter 4, introduces the first of the two artificial intelligence methods that were developed: the gradient boosting trees. Gradient boosting trees were designed to perform three different tasks: a) fault detection and faulty feeder identification, b) fault type identification and c) faulty branch identification. Furthermore, in Chapter 5, the use of deep neural networks is proposed as the second artificial intelligence method of this study. Deep neural networks with two, three and four layers are designed for: a) fault detection, b) faulty branch identification and c) fault distance estimation respectively. Finally, Chapter 6 compares all three of the developed methods with others in the literature and among themselves underlining the merits of deep neural networks.

In the second part of the thesis, in Chapter 7 two algorithms based on a signal approach are proposed as a fault isolation strategy for GCPVs. A big variety of faults can occur in a PV power plant. Based on their location faults can appear: a) in the PV array, b) in the power converters, c) on the dc bus and d) in the grid side. The algorithms use current and voltage measurements at the output of the inverter, examining faults occurring on all four of the aforementioned possible locations and at different irradiance levels. The choice of the output of the inverter, i.e. the point of common coupling, as the monitoring source of the status of the GCPV system is in accordance with the location of voltage sensors used in the previous case of fault location methods in the LV distribution grid.

The thesis concludes with a general conclusion chapter and the appendices where the fuel cell case is also presented.

# Part I

## Fault diagnosis in low voltage smart distribution grids



# Chapter 3

## Conventional method

In an attempt to shed light on a quite unexplored field, as underlined in the Introduction, a fault detection method, based on current measurements, and a fault location method, based on voltage measurements, for *low voltage (LV)* distribution grids are presented in this chapter. Although, the basic idea of the fault location method was initially conceived for *medium voltage (MV)* distribution grids [2], in this case it is tested in the more complex case of LV grids, as explained in Chapter 2. The performances of the two methods were tested against: a) different types of measurements: phase *root mean square (rms)* measurements and transformed rms quantities via a Fortescue analysis, b) various fault resistance values ranging from 0.1 to 1000  $\Omega$ , c) various fault locations within the grid, d) two fault types and e) various hours within the day. Regarding the fault types, both single phase to ground and three phase *short circuit (SC)*<sup>1</sup> faults were studied. In a distribution system, 70 % of the faults are single phase to ground faults and only 5 % are three phase faults [1]. However, three phase faults are the most severe. By selecting those two types, the most frequent and most severe faults were studied. Moreover, a sensitivity analysis is also presented in this chapter including measurement uncertainty (three different types of measurement errors) and a limited measurement availability scenario. Finally, a real semi-rural LV distribution grid of Portugal provided by the Efacec was used as a basis for the case study.

---

<sup>1</sup>from hereafter the term short circuit will be omitted and considered as a given for grid faults.



## 3.1 Method description

### 3.1.1 Fault detection

In LV distribution grids there are two indicators of a fault occurrence: a) a significant **voltage drop** and b) a significant **current increase**. But both of them can be affected by various factors either related or unrelated to the existence of a fault. Some common factors are:

- a) *a sudden load connection*: the sudden connection of a load will cause a current increase.
- b) *topological criteria*: the location of the monitored point inside the grid, i.e. in the beginning or in the end of the feeder, plays a significant role since voltage is also affected by the line losses resulting in lower voltage levels for consumers located at the end of the feeder.
- c) *fault resistance*: high fault resistance values will decrease the fault current flowing through the faulty feeder thus making it more difficult to differentiate normal from faulty operating condition.
- d) *penetration level of distributed generation (DG)*: the presence of interconnected DG in the residential level of smart grids that allow bidirectional flow of energy, can increase the voltage levels of nearby consumers, thus mitigating the voltage drop caused by either faults or line losses. Moreover, in smart grids with inverter-interfaced renewable sources, fault currents can be very limited because of the inverter control actions. Traditionally, utilities use overcurrent as a basis for fault detection, which with the presence of DG can lead quite often to false alarms or even worse to amplification of their contribution to the fault current if they remain connected during the fault [227].

For the reasons listed above, six fault detection criteria, three based on voltage and three on current, were proposed by two international standards [228, 229] of power quality and four methods based on superimposed quantities. Superimposed quantities are nothing more but a comparison of the pre-fault with the post-fault values of voltages or currents that can serve in the detection of any incremental changes. This last approach is immune to the impact of loads as any observed

change will be attributed solely to the existence of a fault. More specifically, those criteria are:

### Voltage

1. *Voltage drop*: The European standard EN50160-2010 [228] defines as voltage of normal operating conditions in the LV grid, any voltage that does not drop below 90 % of the nominal value.
2. *Negative sequence voltage,  $V_{neg}$* : The negative sequence voltage component is expected to increase significantly during unbalanced faults thus providing a relatively safe fault detection tool [230]. On the other hand, this criterion can not be used for three phase faults. Taking into consideration that only 5 % [1] of the faults in distribution grids are three phase faults, this criterion could theoretically cover 95 % of fault cases. It should be noted that negative sequence voltage component might also exist because of various imbalances between the phases.
3. *Positive sequence voltage,  $V_{pos}$* : A complementary method to the previous one would be the monitoring of the positive sequence voltage pre- and post-fault values as a decrease of the post-fault value would indicate the presence of a three phase fault. This method can be also used for the detection of single phase faults but in that case it would be less accurate as information from two healthy phases and one faulty would be combined for the voltage transformation thus mitigating the voltage decrease of the positive sequence.

### Current

1. *Negative over positive sequence current ratio*: The IEEE 1159-2009 standard [229] proposes the use of another index based on the symmetrical components of the current for the monitoring of a normal operation:

$$NO : \frac{|I_{neg}|}{|I_{pos}|} < 0.3 \quad (3.1)$$

As described by the equation above, an abrupt change in the negative sequence of the current would be considered as an imbalance. Hence, if this ratio exceeds the threshold of 0.3 this would be an indication of a

fault occurrence. However, in LV grids, which as described in Chapter 2, are usually quite unbalanced, it is possible to encounter negative sequence component of the current even under normal operating conditions.

2. *Positive sequence current,  $I_{pos}$* : Especially for grids where the DG are connected to it through inverters, the positive sequence current can be used as a monitoring tool since changes in this quantity would be attributed to the behavior of the grid and not to inverter control actions [230].
3. *Increase of phase current,  $I_{ph}$* : An increase of the phase current could also indicate the presence of a fault. However, changes in the current can occur under various operational conditions. Especially for high fault resistance values where the amplitude of the fault current is very small, as it will be later explained, it is extremely difficult to detect with certainty a fault occurrence. Finally, this criterion has the advantage of being useful for a second task. Through the use of phase current, fault type identification is also possible, i.e. identifying which phase(s) is under fault, since the phase(s) under fault would be the one(s) presenting an increase in their current.

The output of a fault detection algorithm is the trigger of an alarm signal when a fault is detected. However, such an alarm signal can be falsely triggered under normal operating conditions. In order to define all possible conditions and to properly measure the performance of each method, the confusion matrix, presented in Table 3.1 is used :

Table 3.1 Confusion matrix.

Predicted Condition	Actual Condition	
	Positive	Negative
Positive	true positive (tp)	false positive (fp)
Negative	false negative (fn)	true negative (tn)

Any fault detection method is considered successful when it correctly activates an alarm signal when there is a fault, *true positive*, and when it does not activate

the alarm under normal operating conditions, *true negative*. On the other hand, the method fails when it activates the alarm when there is no fault, *false positive*, and when it does not activate the alarm when there is a fault, *false negative*.

Though there are different ways of measuring the performance of a method (sensitivity, specificity, false positive rate, false negative rate, precision, accuracy, F1 score etc.) [231], only two tools are chosen for this study:

$$Sensitivity(Recall) = \frac{Detection}{Fault} = \frac{tp}{tp + fn} \quad (3.2)$$

$$Accuracy = \frac{Correct\ detection}{Total\ data\ sample} = \frac{tp + tn}{tp + fp + tn + fn} \quad (3.3)$$

Although the *accuracy* term is introduced here, it will only be used in Chapters 4,5. For the conventional method, only the *sensitivity*, tool is used.

### 3.1.2 Fault location

Once an alarm signal indicates the occurrence of a fault in one of the grid feeders, the fault location process is initiated. This process is divided in three distinct steps: a) faulty branch identification, b) faulty sector localization and c) fault distance estimation. All these steps are solely based on nodal rms voltage measurements.

#### 3.1.2.1 Faulty branch identification

Identifying the faulty branch within a feeder with multiple branches is the first and most important step of the fault location process. It is expected that the branch under fault would present the highest voltage drop within the faulty feeder. For this reason two methods are proposed to identify a faulty branch, both of them presented in Algorithm 1:

1. *Vertical*: The minimum voltages of each branch are compared and the branch with the lowest voltage is considered to be the one under fault.
2. *Horizontal*: A step by step comparison of the available measurements within each branch of the same feeder is implemented. For example, the

**Algorithm 1** Fault Location Algorithm - Step I

---

```

1: procedure BRANCH IDENTIFICATION - VERTICAL
2:    $min\_v \leftarrow min\_v(faulty\_feeder)$ 
3:   for  $i = first\_br : last\_br$  do  $\triangleright br: branch$ 
4:     if  $v(i) = min\_v$  then
5:        $faulty\_br \leftarrow i$ 
6: procedure BRANCH IDENTIFICATION - HORIZONTAL
7:    $\triangleright msp: measurement point$ 
8:   for  $j = first\_msp + 1 : last\_msp$  do
9:      $min\_count \leftarrow 0$ 
10:     $min\_v2 \leftarrow min(v(first\_br : last\_br, j))$ 
11:    for  $i = first\_br : last\_br$  do
12:      if  $v(i, j) = min\_v2$  then
13:         $min\_count \leftarrow min\_count + 1$ 
14:        if  $min\_count = 1$  then
15:           $faulty\_br \leftarrow i$ 
16:    if  $min\_count = 1$  then
17:      break

```

---

first available measurements in each branch are compared and the branch where the minimum voltage belongs is the one under fault. In case of voltage equality or multiple equal minimums, the second available voltage measurements within each branch are compared and then if a minimum is found the branch to which it belongs is identified as faulty. The process continues until a unique voltage minimum is found. The branch to which this minimum belongs is considered as the one under fault. The horizontal method is illustrated in Fig. 3.1.

### 3.1.2.2 Faulty sector localization

In transmission networks, according to the equal transfer process of transmission lines, the voltage should be approximately linearly distributed between two line segments. The same principle applies to LV grids. However, since in the LV grid the lines connecting the different nodes within a branch can present quite different characteristics (resistance, reactance and length), attributing thus an heterogeneous nature to the grid, the above principle, though still applicable, will present some deviations. In any case, when a ground fault occurs, the voltage at the fault point will decrease to zero [2, 232].

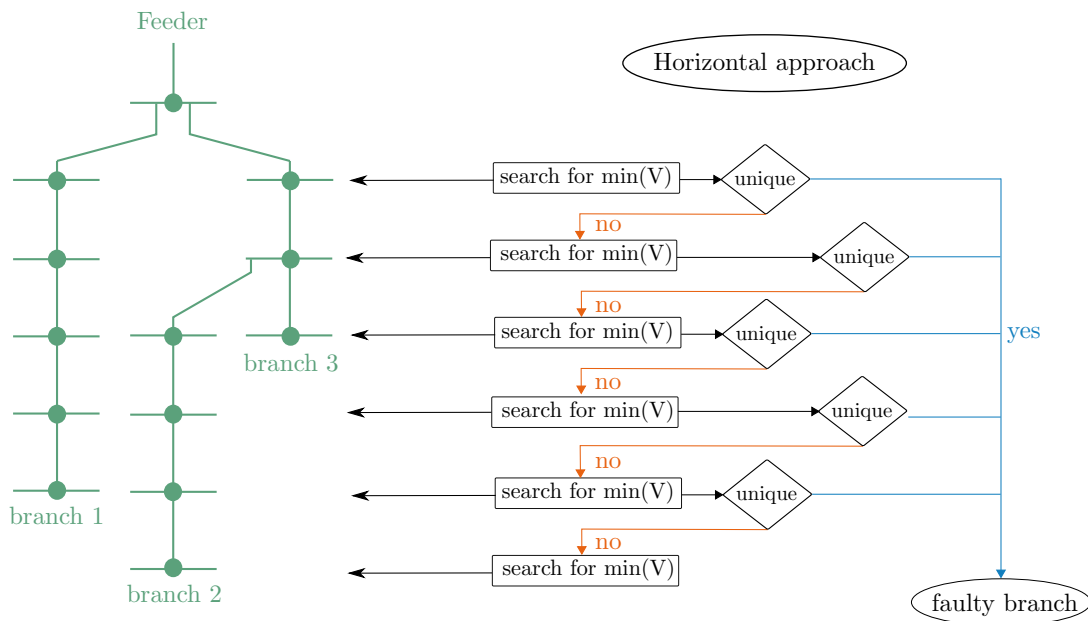


Figure 3.1 Flowchart of the horizontal branch identification algorithm.

The basic principle is based on the fact that across the line of a faulty branch, voltage is expected to drop linearly up until the faulty sector and stabilize to a certain value after it. In Fig. 3.2, the theoretical voltage profile and the form of its curve is presented. For  $0 \Omega$  of fault resistance, the voltage will drop to zero after the faulty sector (marked with orange color) since there will be no current flowing through this segment of the grid; all the current will flow from the line to the ground through their established connection during the fault leaving no circulating current in the segment of the line after the fault location. However, for any value “x” of fault resistance, the voltage is expected to stabilize to a higher value.

This idea was conceived in the context of the MV grid which is less complex than the LV grid [2]. Although in reality the unbalanced and heterogeneous nature of the LV grid can affect the form of the voltage profile, the basic principle of the voltage stabilizing to a value after the faulty sector, should theoretically still apply.

Linear interpolation can be used to create the lines connecting the voltage measurement points across the branch. The change in the slope of the voltage curve in Fig. 3.2 is obvious as it decreases to almost zero after the faulty sector.

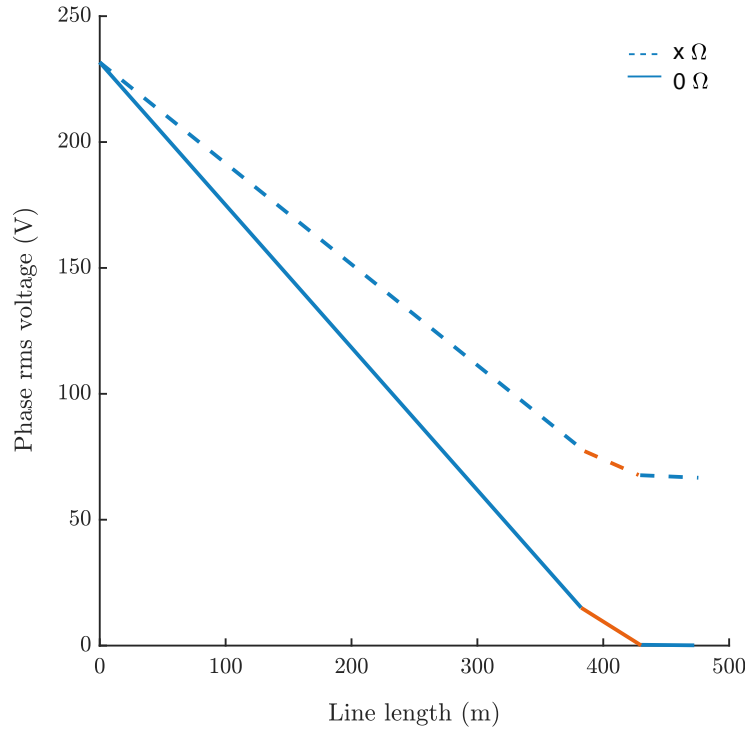


Figure 3.2 Theoretical voltage profile across a faulty line for a single-phase to ground SC fault for two cases of fault resistance values: zero (“0”) and non-zero (“x”). The sector where the fault occurred is marked with orange color.

In order to transform the above critical observation regarding the slope into an algorithm (Algorithm 2) and identify the faulty sector, the two following criteria were developed:

- 1) if the difference between two consecutive voltage measurements was positive, signifying a change in the sign of the slope, then the previous sector was the one under fault and
- 2) if the absolute value of the difference between two adjacent voltage measurements was the lowest within the branch, signifying a stabilization of the curve, then the previous sector was the one under fault.

An example is provided in Fig. 3.3. The slope of the curve is almost always negative since voltage is decreasing. However, a positive value of the slope after the faulty sector is possible under the presence of microgeneration units in this part of the grid, thus explaining the choice of the first criterion; the effect of microgeneration units and loads will be explained later.

**Algorithm 2** Fault Location Algorithm - Step II

---

```

1: procedure SECTOR LOCALIZATION
2:   accross the faulty branch
3:   for  $i = first\_msp + 1 : last\_msp$  do
4:      $\delta v(i) \leftarrow v(i) - v(i - 1)$ 
5:      $fs \leftarrow 0$   $\triangleright fs$ : faulty sector
6:      $min\_v3 \leftarrow min(|\delta v|)$ 
7:      $\triangleright$  first voltage criterion
8:     for  $i = first\_msp + 1 : last\_msp$  do
9:       if  $fs = 0$  and  $\delta v(i - 1) < 0$  and  $\delta v(i) > 0$  then
10:         $fs \leftarrow i - 1$ 
11:        break
12:      $\triangleright$  second voltage criterion
13:   for  $i = first\_msp : last\_msp - 1$  do
14:     if  $fs = 0$  and  $|\delta v(i)| = min\_v3$  then
15:        $fs \leftarrow i - 1$ 
16:       break

```

---

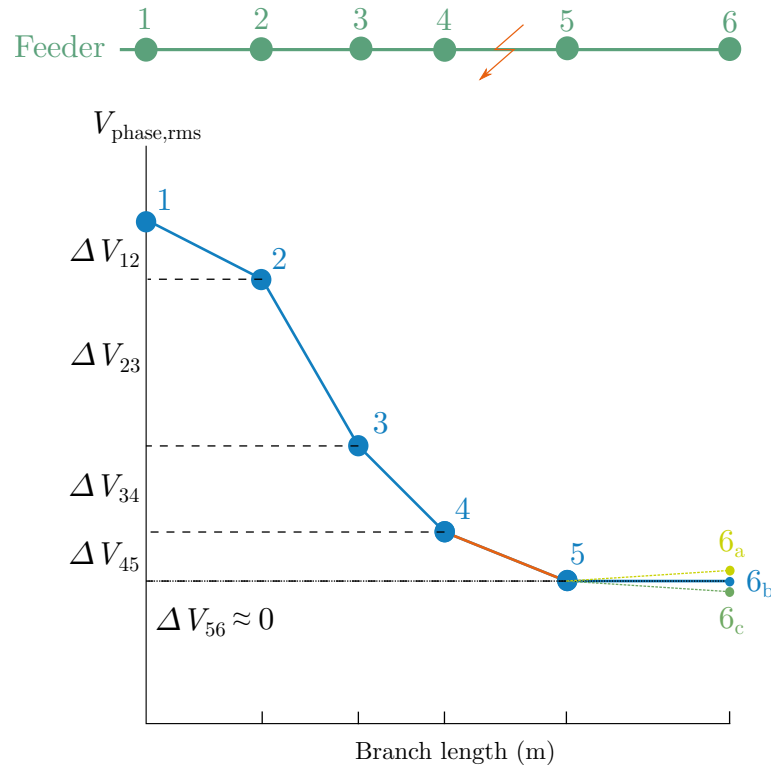


Figure 3.3 Voltage profile and example of the usage of the sector localization criteria in a faulty branch. After the faulty sector (colored in orange) the voltage will either: slightly increase (6a) or stabilize to the value of the last nodal measurement (6b) or slightly decrease (6c). Criterion (1) will be triggered in the case of (6a) since the voltage difference ( $\Delta V_{56} = V_6 - V_5 > 0$ ) will be positive and criterion (2) in all three cases of this example as  $\Delta V_{56}$  is the smallest voltage difference.



### 3.1.2.3 Fault distance estimation

The last step in localizing the fault, after having identified the faulty branch and sector, is to estimate its location within the faulty sector. To achieve that, a graphic method is implemented. From the linearly interpolated curve of Fig. 3.4, the lines of the sectors adjacent to the one under fault are linearly extrapolated (green dashed lines) and their intersection point is used to estimate the location of the fault inside the sector. Another possibility would be to use all the remaining sectors before and after the faulty one for the linear extrapolation. A discussion on which of the two methods is better will follow in the next section.

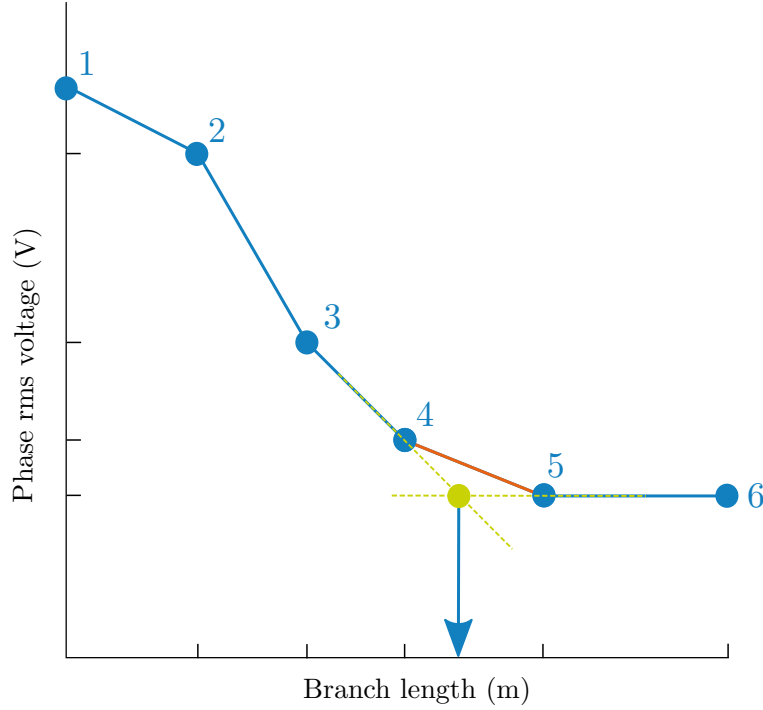


Figure 3.4 Theoretical estimation of the fault location.

An obvious observation is that this method is incapable of locating a fault that occurred in the first or the last sector of the grid since it requires a minimum of four measurement points (three sectors). To measure the method accuracy, the most popular formula in the literature, described in [233] is used:

$$error (\%) = \frac{|d_{estimated} - d_{actual}|}{l_{total}} \cdot 100 \quad (3.4)$$

where the estimated distance is compared to the actual one and the result is normalized over the total length of the line.

As a last step, a threshold crossing check is applied to ensure that the estimated location falls within the limits of the identified faulty sector. This corrective process minimizes false estimations. The fault distance estimation process is summarized in Algorithm 3 and the results of the entire fault location method are presented in the following section.

---

**Algorithm 3** Fault Location Algorithm - Step III
 

---

- 1: **procedure** DISTANCE ESTIMATION
  - 2:   from extrapolation of the lines of the adjacent sectors
  - 3:    $d_{est} \leftarrow \text{distance of the intersection point}$
  - 4:   **if**  $d_{est} > fs\_upper\_limit$  **then**
  - 5:      $d_{est} \leftarrow fs\_upper\_limit$
  - 6:   **else if**  $d_{est} < fs\_lower\_limit$  **then**
  - 7:      $d_{est} \leftarrow fs\_lower\_limit$
  - 8:   the fault is located at  $d_{est}(m)$
- 

The complete fault detection and location process based on the steps described in the previous section, is summarized in the following flowchart in Fig. 3.5.

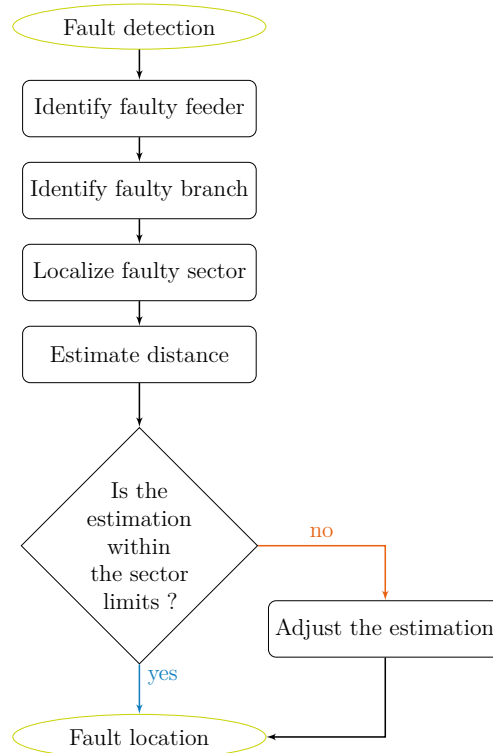


Figure 3.5 Fault detection and isolation method flowchart.

## 3.2 Application

### 3.2.1 Case study

#### 3.2.1.1 Grid characteristics

A real semi-rural radial LV distribution grid of Portugal provided by Efacec [234], was used as the basis of the case study. The connection to the MV is achieved through a distribution transformer. The grid is a three-phase-four-wire one with a solidly grounded neutral. A total of forty eight consumers and eighteen *photovoltaic systems (PV)* are connected to the grid via single phase connections. The single line diagram of the LV distribution grid is presented in Fig. 3.6.

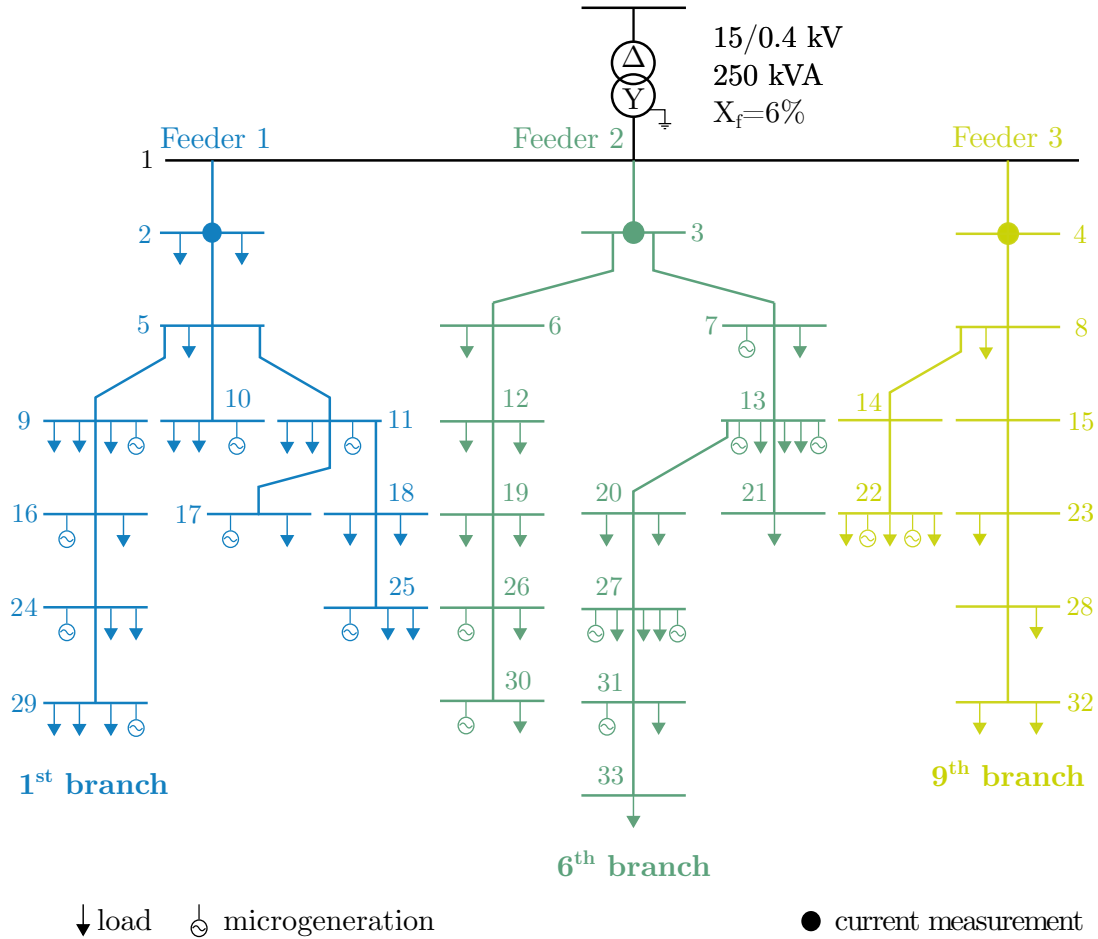


Figure 3.6 Single line diagram of the semi-rural Portuguese LV distribution grid.

There are two characteristics that define this grid and complexify its analysis:

- its *heterogeneity*: eleven different types of conductors in terms of resistance and reactance, with lengths ranging from 35 to a maximum of 210 *m* are used

to connect the nodes, and b) its *imbalance*: asymmetrical, both topologically and per phase, distribution of the loads and PVs. For more details on the heterogeneity of the grid, the specifications of the conductors are provided in Tables A.1 to A.3 in the Appendix. Concerning its unbalanced nature, Table 3.2 summarizes the total per feeder and phase contracted and installed power in the grid of loads and PVs respectively; again, Tables A.4 to A.9 in the Appendix provide the full details.

Table 3.2 Total per feeder and phase ( $a, b, c$ ) contracted (loads) and installed (PVs) power.

Feeder	Contracted Power ( $kVA$ )			Installed Power ( $kW$ )		
	$S_a$	$S_b$	$S_c$	$P_a$	$P_b$	$P_c$
1	31.05	41.40	48.30	3.68	8.83	14.49
2	34.50	41.40	24.15	10.58	7.13	7.08
3	10.35	13.80	17.25	3.68	0.00	1.70
<b>Total</b>	<b>75.90</b>	<b>96.60</b>	<b>89.70</b>	<b>17.94</b>	<b>15.96</b>	<b>23.27</b>

As explained before, the fault detection and location methods are based only on rms measurements. Hence, the phasor mode was selected to perform the simulations in **Matlab/Simulink**, in order to reduce the computational time. Furthermore, the faults were studied 150  $ms$  after their occurrence. This time frame ensured: a) that the fault would appear on a steady-state or at least very close to it and b) that no protective device has isolated the installed PV units; an action from the protective elements to isolate PVs is expected around 200  $ms$  from the fault occurrence [4]. This enabled the analysis of the contribution of the PV units to the faults.

Finally, it was assumed that the measurements were synchronized and as a first step, accurate and available in every node; aspects encountered only in smart grids.

### 3.2.1.2 Influencing parameters

The following influencing parameters were considered for this case study:

1. *Fault resistance:* As stated in Chapter 2, fault resistances up to  $1000 \Omega$  can be encountered in distribution grids. Hence, the following fault resistance values were considered: 0.1, 1, 5, 10, 50, 100, 500 and  $1000 \Omega$ , covering all the ranges of values, from very low (0.1) to extra high (1000).
2. *Voltage measurement:* The choice between using phase rms voltages or the positive, negative and zero components after a symmetrical (Fortescue) analysis was analyzed.
3. *Fault location:* Every sector of the grid, was divided in ten subsections and nine locations were chosen at a distance of 10, 20, 30, 40, 50, 60, 70, 80 and 90 % from the beginning of each sector. Due to the graphic method limitation described earlier, the first and last sectors were excluded during the sector localization and distance estimation processes.
4. *Fault types:* Single phase to ground faults (AG, BG and CG) were chosen because they are the most frequent and three phase faults (ABC) because they are the most severe.
5. *Time of the day:* The generation and load profiles that were considered in this study are provided in Fig. 3.7 . From these profiles four hours of the day were selected as presented in Table 3.3. Scenarios (1,2) and (3,4) were used to monitor how the developed method is affected by an increase of load or PV generation respectively.

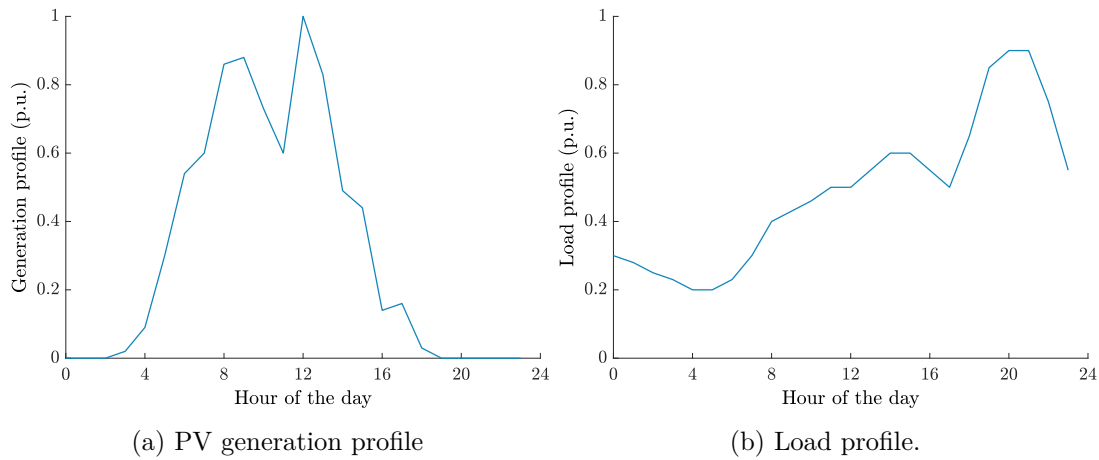


Figure 3.7 Generation and load profiles for one day.

Table 3.3 Simulation scenarios.

	Time	PV generation	Load
1	01:00	0 %	28 %
2	20:00	0 %	90 %
3	14:00	49 %	60 %
4	12:00	100 %	50 %

6. *Simultaneity factor*: Since not all the loads of every consumer are going to be activated simultaneously, a coincident-simultaneity factor was considered [235]:

$$SF = \frac{\max(D_{\text{system}})}{\sum_{i=1}^N D_i^{\max}} \quad (3.5)$$

where  $D$  is the load demand and  $N$  the number of loads.

The sum of the individual peak loads will always be bigger than the peak load of the system. Thus, the simultaneity factor is always smaller than unity. It can only take the unity value if all the loads reach their peak at the same time.  $SF$  is statistically stabilizing around 0.5 in residential areas of developed countries [235]. In general, the  $SF$  will depend: a) on the type of loads connected to the system, b) the number of consumers and c) as stated in [236] on weather diversity for larger systems; since the grid of this case study covers a small area, weather diversity was not considered as an influencing factor. However, the smaller the amount of consumers the more likely it is to notice deviations as the  $SF$  will not necessarily converge to 0.5. In the grid of Fig. 3.6, forty eight consumers are served; from Fig. 3.8, this would correspond to a  $SF$  a bit over 0.3. For that purpose, a  $SF = 0.3$  was used as a basis for this analysis. A  $SF = 0.5$  however, was used for the comparison analysis in Chapter 6 to cover a broader range of possible scenarios.

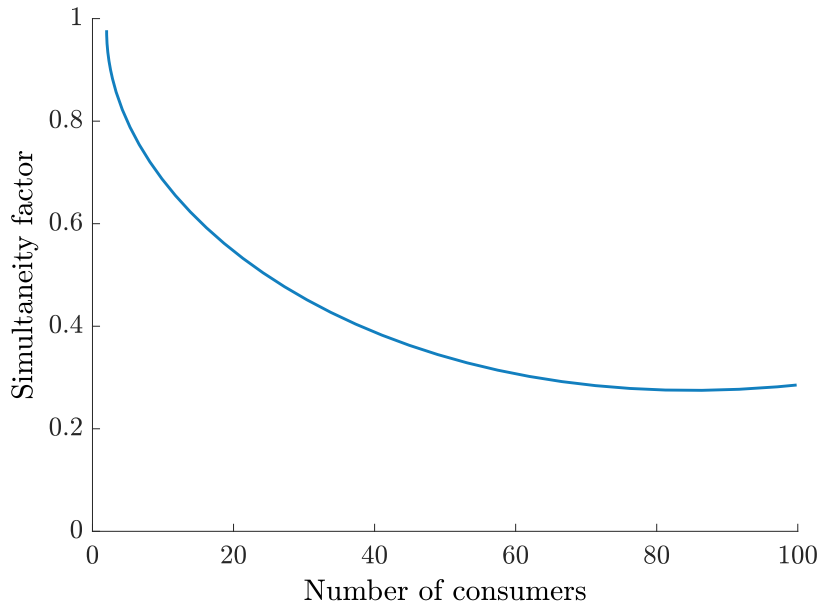


Figure 3.8 Simultaneity factor per number of consumers in distribution grids [235].

### 3.2.2 Fault detection

The six criteria that were described in the method description section were tested for their ability to accurately detect the occurrence of a fault. For the creation of the superimposed values (comparison between post- and pre-fault values), the values of faulty operation 150 *ms* after the occurrence of the fault were compared with those of normal operation 5 *min* before the occurrence of the fault<sup>2</sup>. The performance of the chosen criteria, measured by sensitivity, is presented in Fig. 3.9 for single phase to ground faults and in Fig. 3.10 for three phase faults.

An obvious conclusion from these two figures is that the increase of fault resistance decreases the chances of a correct fault detection. An increase of the fault resistance decreases the current and increases the voltage across a faulty branch. This effect seems to affect more the voltage criteria than the ones based on current in the case of single phase to ground faults (Fig. 3.9) where the two most reliable criteria are five (superimposed positive sequence current increase) and six (superimposed phase current increase). At the same time, some criteria seem more suitable than others for three phase faults with three (superimposed positive sequence voltage drop) and six outperforming the rest. Overall, the

<sup>2</sup>The time frame of 5 *min* was selected for consistency purposes with the methods of the next chapters. A complete justification of this choice is provided in Chapter 4.

monitoring of the phase current was deemed the best choice among the rest with an average sensitivity of 79.87 % for single phase to ground faults and 79.61 % for three phase faults, offering in parallel, the possibility to identify the fault type (phase under fault) and the faulty feeder.

However, as mentioned in the previous section, the phase current is prone to changes from different factors and although the use of the superimposed value decreases this chance, it does not eliminate it. Moreover, for fault resistance values higher than  $10 \Omega$ , as shown in Fig. 3.11, the superimposed current approaches the unity value, meaning that the current under faulty operation is almost the same under normal operation. This last fact, makes it even more vulnerable to other exogenous factors that can affect it.

In addition, the effect of the grid imbalance is illustrated in the same figure. In detail, especially for low resistance values, the differences of the current increase in the three feeders are quite substantial; the less elements connected to a feeder the easier it is to detect a fault (higher superimposed value). The latter is verified by Table 3.2 according to which Feeder 1 concentrates a total of 121.2  $kVA$  of contracted power, Feeder 2, 100.05  $kVA$  and Feeder 3, 41.4  $kVA$ .

In general, phase rms current measurements in the beginning of each feeder are considered as a reliable fault detection criterion for very low and low fault resistance values (lower than  $10 \Omega$ ). The installation of more current sensors throughout the grid would improve the performance of this method but the need for a more reliable method especially for high fault resistance faults is evident.

### 3.2.3 Fault location

Following the fault detection and faulty feeder identification, the fault location process, composed of three distinct steps: a) branch identification, b) sector localization and c) distance estimation, is initiated as shown in Fig. 3.5.

#### 3.2.3.1 Branch identification

The results of the branch identification analysis are gathered in Fig. 3.12. An analysis of the influencing parameters follows:

1. *Fault resistance*: With the increase of fault resistance, the total voltage drop across the faulty branch becomes more and more insignificant thus making



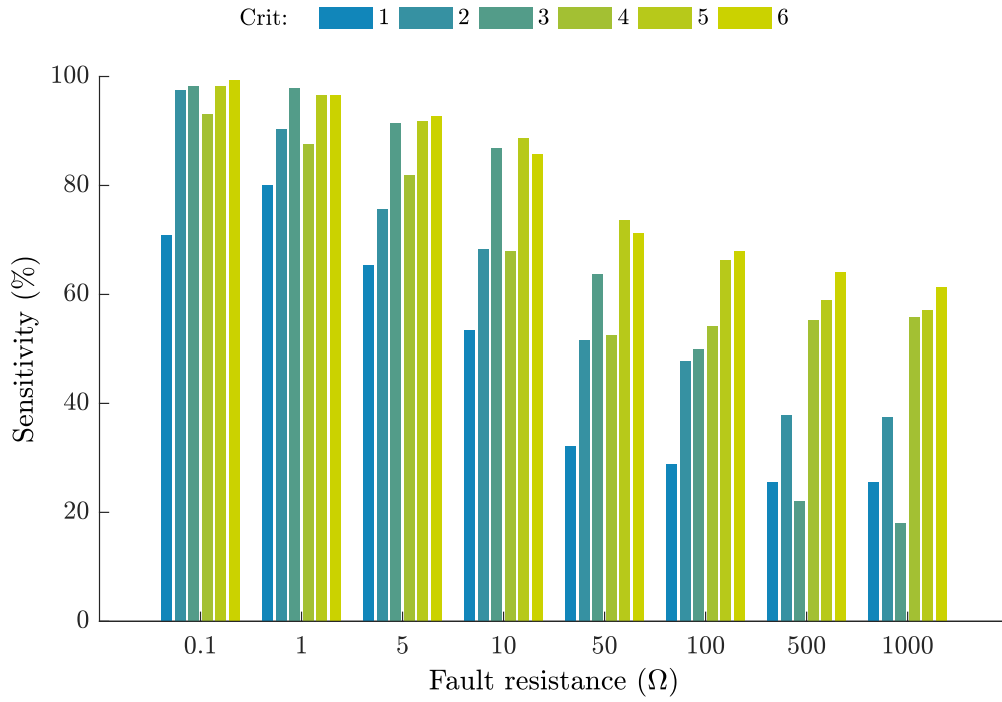


Figure 3.9 Sensitivity of fault detection criteria per fault resistance for single phase to ground faults.

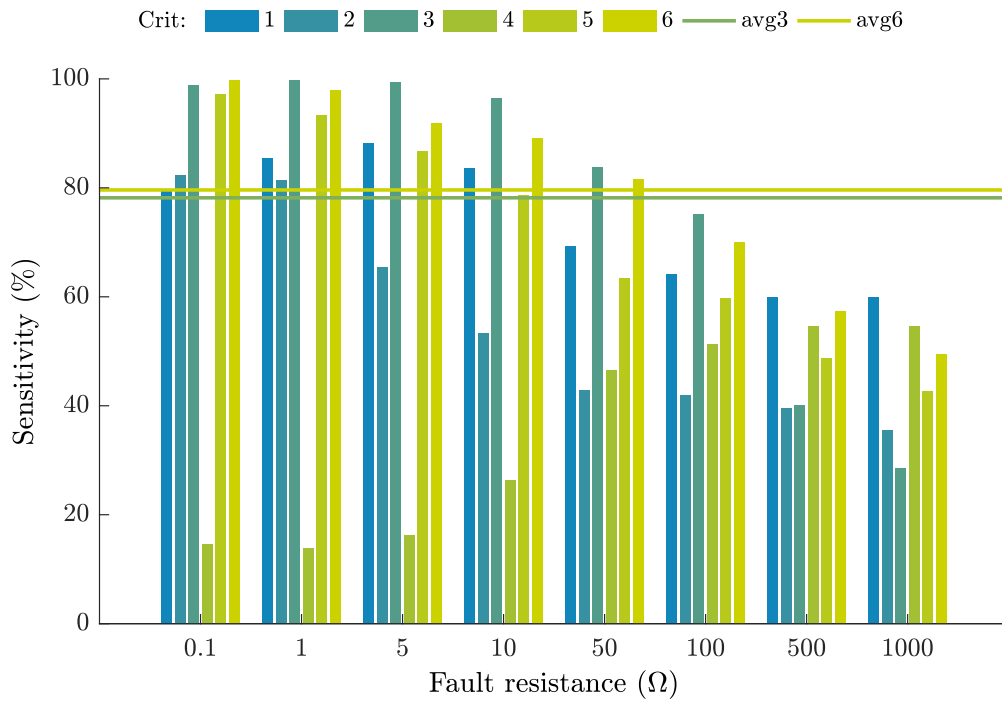


Figure 3.10 Sensitivity of fault detection criteria per fault resistance for three phase faults

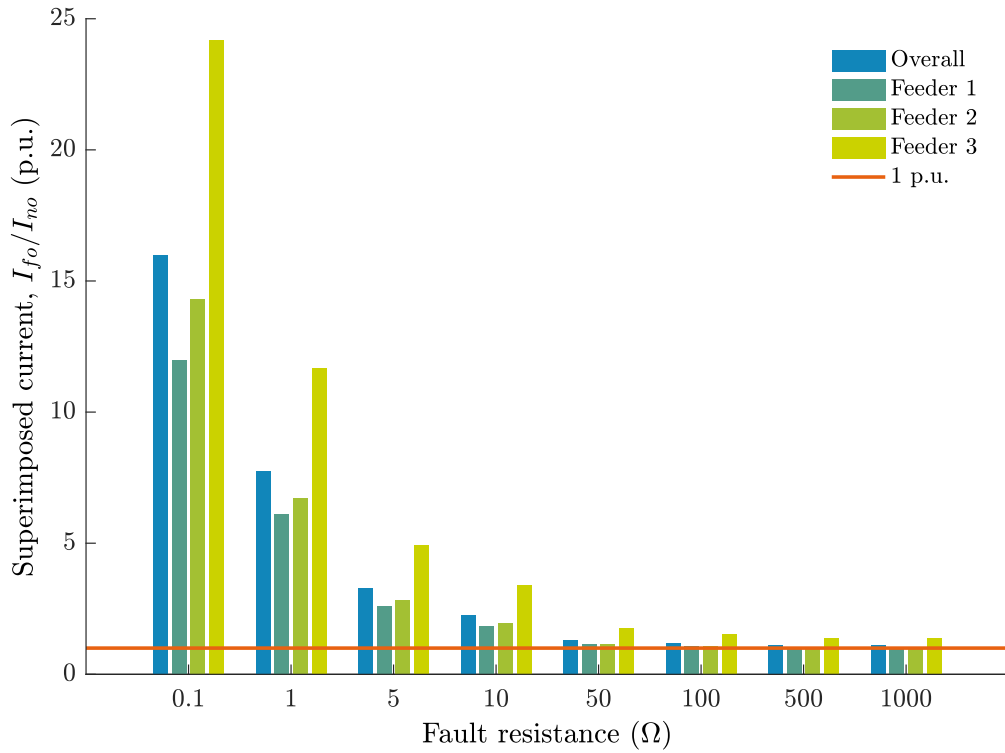


Figure 3.11 Superimposed phase rms current increase per fault resistance in each grid feeder.

it harder to distinguish a faulty from a healthy branch. As presented in the fault detection analysis, currents flowing in the faulty feeder, for a fault resistance higher than  $10 \Omega$ , are very close to the normal operating condition currents. The same holds for voltages. The minimum voltage within a faulty branch (the basic faulty branch identification criterion), with an increase of the fault resistance, would approach the normal operating values. This last effect, would make the method prone to misidentification since it would be more vulnerable to voltage variations due to the unbalanced and heterogeneous nature of the grid. A general conclusion of Fig. 3.12 is that the increase of fault resistance decreases the faulty branch identification accuracy with a maximum difference of 42.9 % being recorded between  $0.1 \Omega$  (94.14 %) and  $1000 \Omega$  (51.24 %), for the case of transformed voltages.

2. *Voltage measurements:* In Fig. 3.12a the superiority of the positive sequence component is demonstrated over all the other voltage components and the phase voltage measurements. Consequently, the rest of the analysis was based on rms positive sequence voltage measurements.

3. *Branch identification method:* Both methods, vertical and horizontal, present similar performance with the vertical one being more successful in identifying the faulty branch, by an average of 4.73 % overall (Fig. 3.12b). Consequently, for the rest of the analysis the vertical method will be used.
4. *Fault location:* Two aspects of fault location are analyzed here. First of all, the location of the sector inside a branch (Fig. 3.12c) and secondly, the location of the fault inside a faulty sector (Fig. 3.12d). Regarding the location of the sector within a faulty branch, two categories of sectors are formed, those in the beginning of each feeder belonging to multiple branches (shared) and those in the middle or towards the end of each feeder that belong only to one branch (unshared). As expected, Fig. 3.12c shows that faults in the beginning of the feeder are easier to identify even at high fault resistance values. As far as the location of the fault within the sector is concerned, Fig. 3.12d shows that faults located at the beginning or the end of each sector present higher chances to influence the faulty branch identification process.
5. *Fault type:* The almost identical performance of the method in all three phases in the case of single phase to ground faults, demonstrated in Fig. 3.12e, leads to the conclusion that the method is immune to the unbalanced per phase distribution of loads and PV units. The higher performance for the case of the three phase faults is attributed to the use of the positive sequence component of the voltage as a basis. In single phase to ground faults, with the use of the positive sequence component, information from all three phases, both faulty and healthy, are taken into account whereas in three phase faults, since all three phases are under fault, the final rms values are not corrupted with healthy input data.
6. *Time of the day:* From Fig. 3.12f, the only safe conclusion is that the effect of the load demand is stronger than this of PV generation since the worst performance is recorded at 20 h, with 90 % of load demand and 0 % of PV generation.

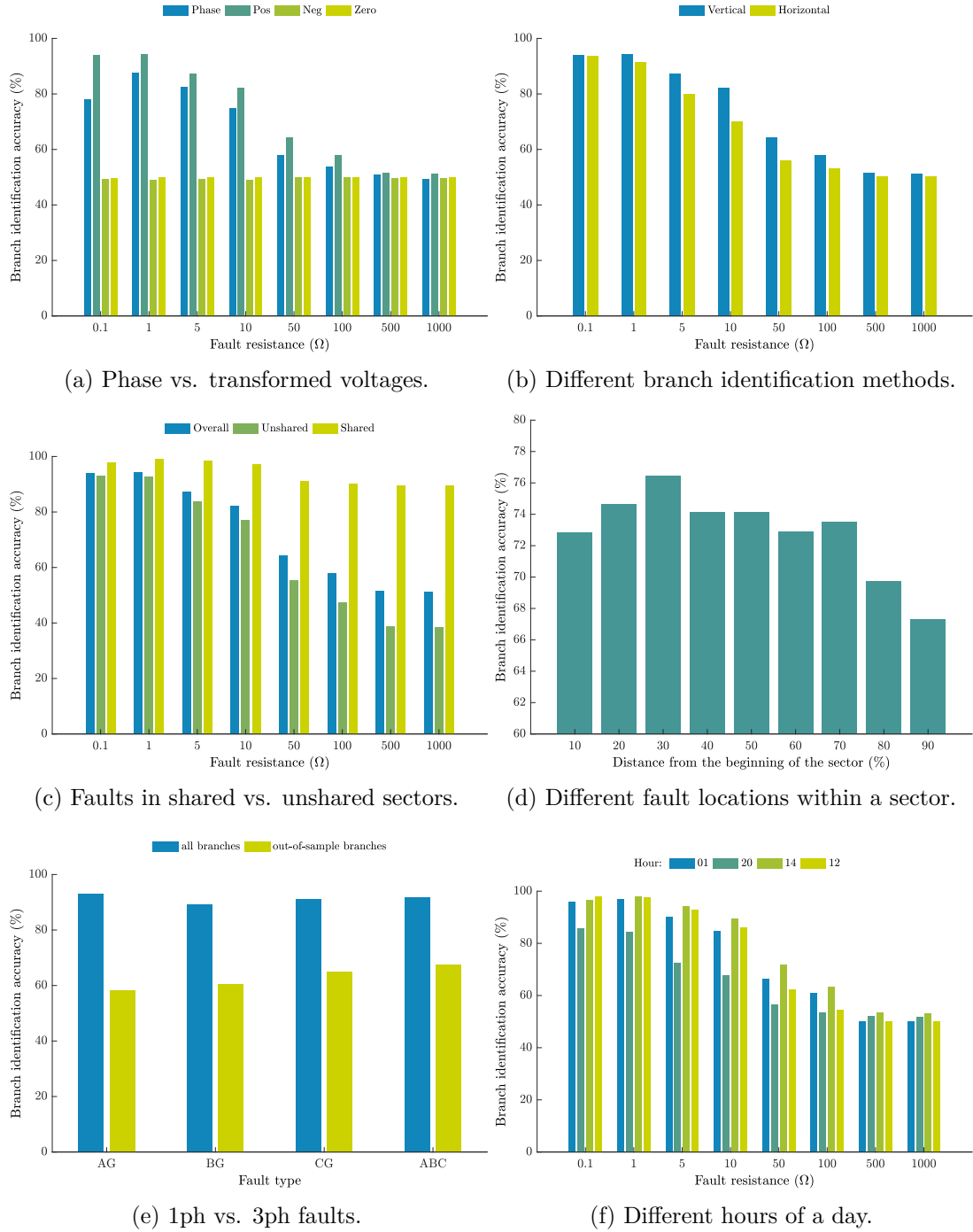


Figure 3.12 Branch identification accuracy.

### 3.2.3.2 Sector localization

The influence of the following parameters is analyzed:

1. *Fault resistance*: Since the increase of fault resistance will decrease the voltage differences between adjacent nodes, the sector localization method is expected to be quite vulnerable to the increase of fault resistance. This

is validated in Fig. 3.13 where it is demonstrated that the increase of fault resistance decreases the faulty sector localization accuracy with a maximum difference of 36.68 % being recorded between 0.1  $\Omega$  (61.61 %) and 1000  $\Omega$  (24.93 %), for the case of transformed voltages.

2. *Voltage measurements:* As with branch identification, the use of the positive sequence component of the voltage is proven to be more accurate compared to phase measurements by 8.48 % in average as presented in Fig. 3.13a.
3. *Fault location:* In contrast to the effect on the branch identification process, the location of a fault in a sector in the beginning of the feeder (shared) hinders its isolability (Fig. 3.13b). The problem that arises is the selection of one branch, among the different branches where the sector belongs to, for the sector localization process as different branches might provide different results in that case and that can possibly lead to the mislocalization of the sector. For example, if a fault between nodes two and five occurs (Fig. 3.6), any of the four branches of the first feeder could be identified as the faulty one from the branch identification step since the specific sector is shared by all of them. Moreover, the slope of the voltage profile curve is more likely to be altered (less likely to stabilize) if more sectors, thus more elements interfering, remain after the one under fault. There, the presence of loads is decreasing the voltage (negative slope to the curve) while the presence of PVs is increasing the voltage (positive slope to the curve). With regards to the location of a fault within a faulty sector, the same tendency as with branch identification process is observed in Fig. 3.13c: faults located at the beginning or the end of each sector present smaller accuracy.
4. *Fault type:* Once more, the performance of the method in all three phases in the case of single phase to ground faults is almost identical (Fig. 3.13d). Three phase faults present, in general, a higher accuracy.
5. *Time of the day:* The conclusion that the effect of the load demand is stronger than this of PV generation is again validated since the worst performance for the sector localization method is recorded at 20  $h$  (Fig. 3.13e). An interesting observation, is that the increased PV generation at 12  $h$  (100 %) seems to partially mitigate the effect of the load demand (60 %). Another example of this mitigation, is derived from the comparison

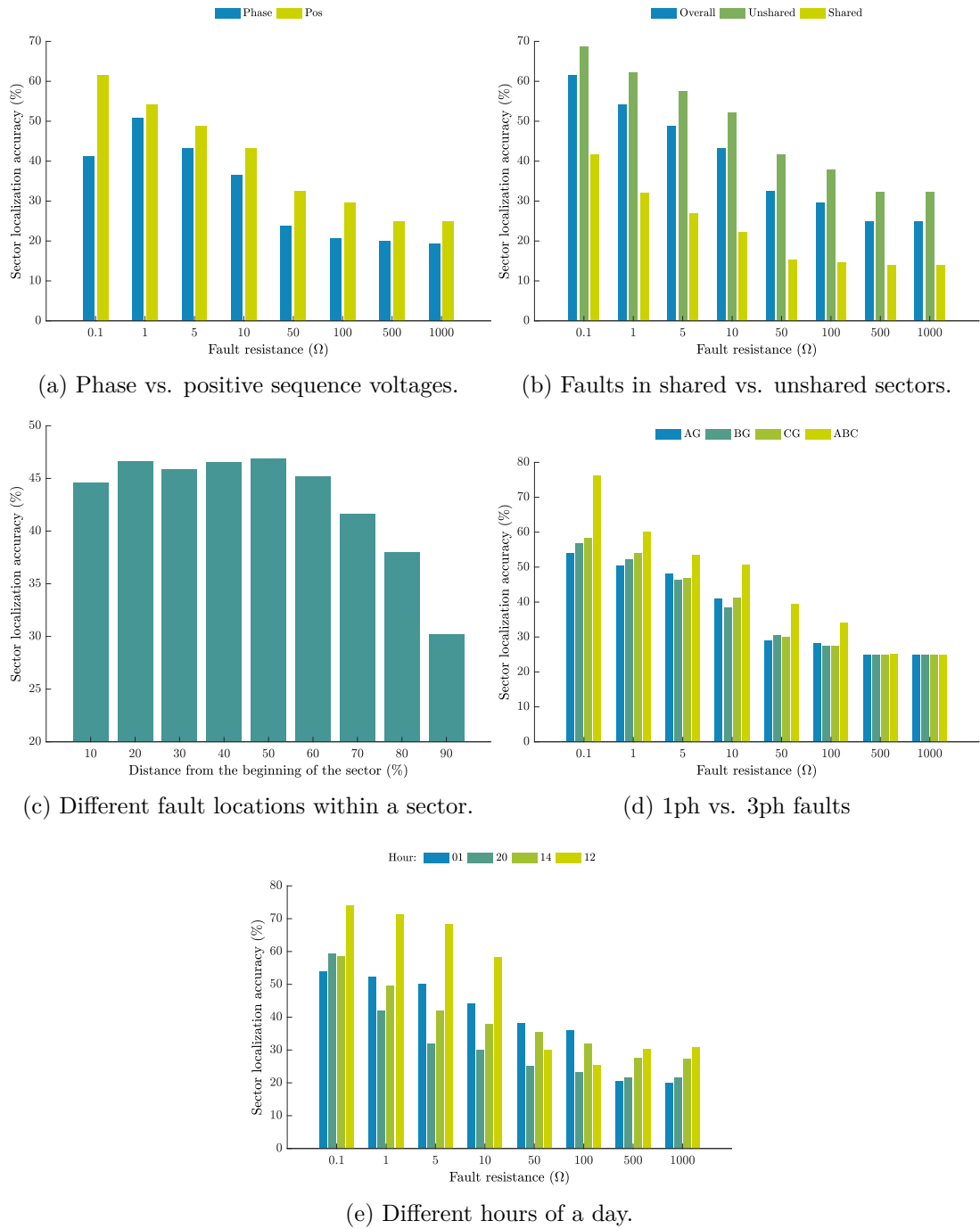


Figure 3.13 Sector localization accuracy.

of the results at 1 *h* with those at 14 *h*. There, the increase of load was accompanied by an increase of PV generation which diminished the accuracy decrease if compared to the one observed at 20 *h* where only the load increased. Finally, the PV influence is more noticeable at fault resistances higher than 500 Ω where the fault current is very small.

### 3.2.3.3 Distance estimation

The influence of the following parameters is analyzed:

1. *Fault resistance*: In the “ideal” case of having zero influence from the connected loads or the PVs on the voltage profile along a faulty branch, the increase of the fault resistance would only transpose the voltage profile to higher values as shown in Fig. 3.2 and Fig. 3.14. Figure 3.14 presents an example of the distance estimation graphic method for the case of a single phase to ground fault with fault resistance of 0.1, 1 and 5  $\Omega$ . It is shown that the distance estimation graphic method is immune against fault resistance variations in the ideal case since the absence of loads and PVs permits the perfect stabilization of the slope of the voltage profile curve to zero, after the faulty sector, in all cases. However, in reality, from Fig. 3.15, it is clear that the increase of the fault resistance affects the fault distance estimation as well, for the same reasons that were analyzed in the previous tasks. This influence is limited with the use of the correct sector criterion as shown in Fig. 3.15b.
2. *Voltage measurements*: Again, the use of phase voltages was compared with the use of the positive sequence components of voltages. Figure 3.15a shows a slightly better performance of the phase voltages in estimating the fault distance by an average of 2.19%. However, since the distance estimation process follows the branch identification and sector localization tasks, and from the fact that the positive sequence component demonstrated a higher accuracy in both of the two previous tasks, the accuracy error difference of 2.19% is not considered important enough to reconsider the choice between phase and symmetrical components. Hence, for the rest of the analysis the positive sequence voltage components have been used.
3. *Distance estimation method*: The use of two different methods is analyzed in Fig. 3.15a and 3.15b as explained in the theoretical analysis of the method. The two options are: a) to use all the available sectors before and after the one under fault for the linear extrapolation step or b) to use only the ones adjacent to the one under fault (Prev.-Next). At the same time, another aspect is investigated: in case the distance estimation fell outside

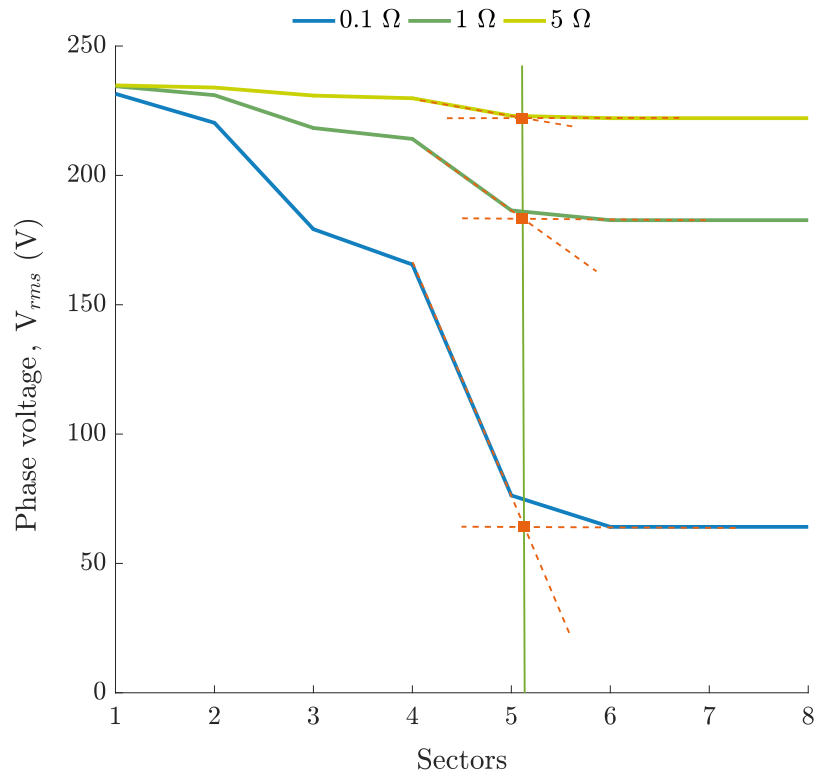


Figure 3.14 Effect of fault resistance under 0 % load and 0 % PV penetration on the distance estimation method for single phase to ground fault located in branch 5 at 270  $m$  from the beginning of feeder 2.

the limits of the identified as faulty sector, a correction process was initiated and restrained the estimation to the limits of the faulty sector; at this point the error sign can be used as an indication of the fault location, upstream or downstream. Although the use of all the available measurements seems to be a better option according to Fig. 3.15a, when the correct sector criterion is activated the use of the adjacent sectors is becoming slightly better as shown in Fig. 3.15b. Moreover, from a comparison between the results demonstrated in Fig. 3.15a and 3.15b, the use of the correct sector criterion is shown to improve the estimation accuracy by an average of 8.82 % for the case of the adjacent sectors. Furthermore, to further justify the use of the adjacent sectors over all the available ones, in Fig. 3.15c the sensitivity percentage of the sector correction criterion is demonstrated; the adjacent sectors option triggers less often the criterion meaning that its estimation is more accurate and less based on other influencing parameters such as the effect of loads or PV units connected further down the faulty branch. As a result, for the rest of the analysis, the use of the correct sector criterion



was used as a default option. Finally, a drawback of the proposed graphic method is that even under the adjacent sectors option, the sector criterion sensitivity reaches an average of 51.65 % meaning that in one out of two cases the graphic method fails to estimate a fault location distance within the identified, as faulty, sector.

4. *Fault location*: Figure 3.15d shows that the influence of the location of the fault within a faulty sector on the distance estimation accuracy is eliminated with the use of the correct sector criterion as the distance estimation error remains almost the same for all cases.
5. *Fault type*: Very slight, almost negligible differences are noticed between the different fault types in Fig. 3.15e.
6. *Time of the day*: As for the previous tasks, the increased number of activated loads hinders the fault distance estimation process with 20 *h* being overall the worst case of the studied scenarios as shown in Fig. 3.15f. On the other hand, at 12 *h*, where the maximum of PV generation is attained, the effect of voltage drop caused by loads is mitigated by the PV units thus decreasing the estimation error.
7. *Grid heterogeneity*: A study was done without any load or PVs present in the grid in order to investigate the effect of the heterogeneity of the grid on the method; as demonstrated before, this removes the influence of fault resistance on the distance estimation method. One parameter was set: the distance of the fault from the beginning of each feeder. Two fault cases were studied: single phase to ground faults at a distance of a) 270 *m* and b) 400 *m* from the beginning of the feeder. One branch from the second feeder and one from the third were selected; branch six from feeder two with a total length of 455 *m* and branch nine which is the longest of the grid with a length of 640 *m* (Fig. 3.6).

In order to distinguish the influence of the different conductor types in terms of resistance and reactance from the potential influence of the branch length itself, a comparative analysis is presented in Table 3.4 between heterogeneous and completely homogeneous lines of equally distanced sectors and a fixed  $R/X$  ratio of 4.76.

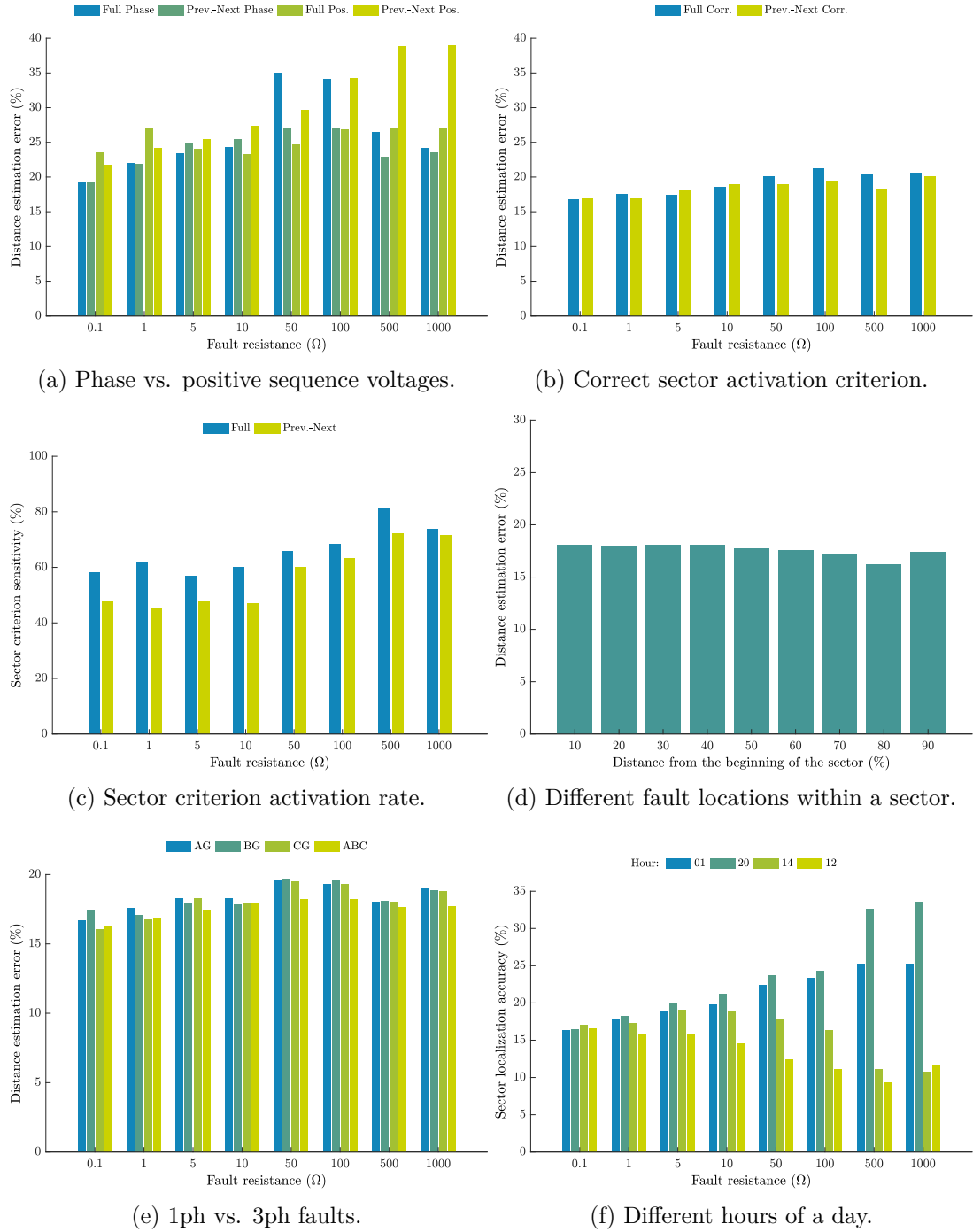


Figure 3.15 Distance estimation error.

A first observation from Table 3.4 is that there is no difference in the estimation accuracy between homogeneous and heterogeneous lines at a distance of 270 m. This however, is attributed to the fact that both branches are composed of sectors with quite similar  $R/X$  ratios as shown in Table A.4 in the Appendix. On the other hand, at 400 m a difference

of 4.33 % and 4.74 % is noticed for branch six and nine respectively. The stable performance between the distance difference of 270 and 400  $m$  under homogeneous lines versus the increased error that this distance increase brings at heterogeneous lines, leads to the safe conclusion that the ones responsible for this error increase are the different types of conductors (various  $R/X$  ratios and lengths) and not the increase of the distance itself or the branch length.

In general, the increase of the fault distance from the beginning of the feeder means a bigger variety of conductors before the faulty sector and thus a more significant influence on the form of the voltage profile which increases the distance estimation error.

Table 3.4 Distance estimation error (%) in homogeneous and heterogeneous lines.

	Homogeneous		Heterogeneous	
	270 $m$	400 $m$	270 $m$	400 $m$
<b>branch 6</b>	0.01	0.07	0.02	4.40
<b>branch 9</b>	0.11	0.05	0.01	4.79

### 3.2.4 Measurement uncertainty

As stated in [237], electricity meter accuracy ranges from  $\pm 0.5\%$  to  $\pm 2.5\%$ . The accuracy of smart meters is usually classified as *class 1* ( $\pm 1\%$ ) for active power measurements and usually less (0.5) for current or voltage. Such an example is the MBox smart meter developed by Efacec [234]. In order to test the accuracy of the developed method versus measurement uncertainty conditions, three types of errors were introduced in each phase current and voltage measurement:

- a) *2 % underestimation (U2)*: values between 0.98 - 1 p.u.
- b) *class 1 accuracy (C1)*: values between 0.99 - 1.01 p.u.
- c) *class 2 accuracy (C2)*: values between 0.98 - 1.02 p.u.

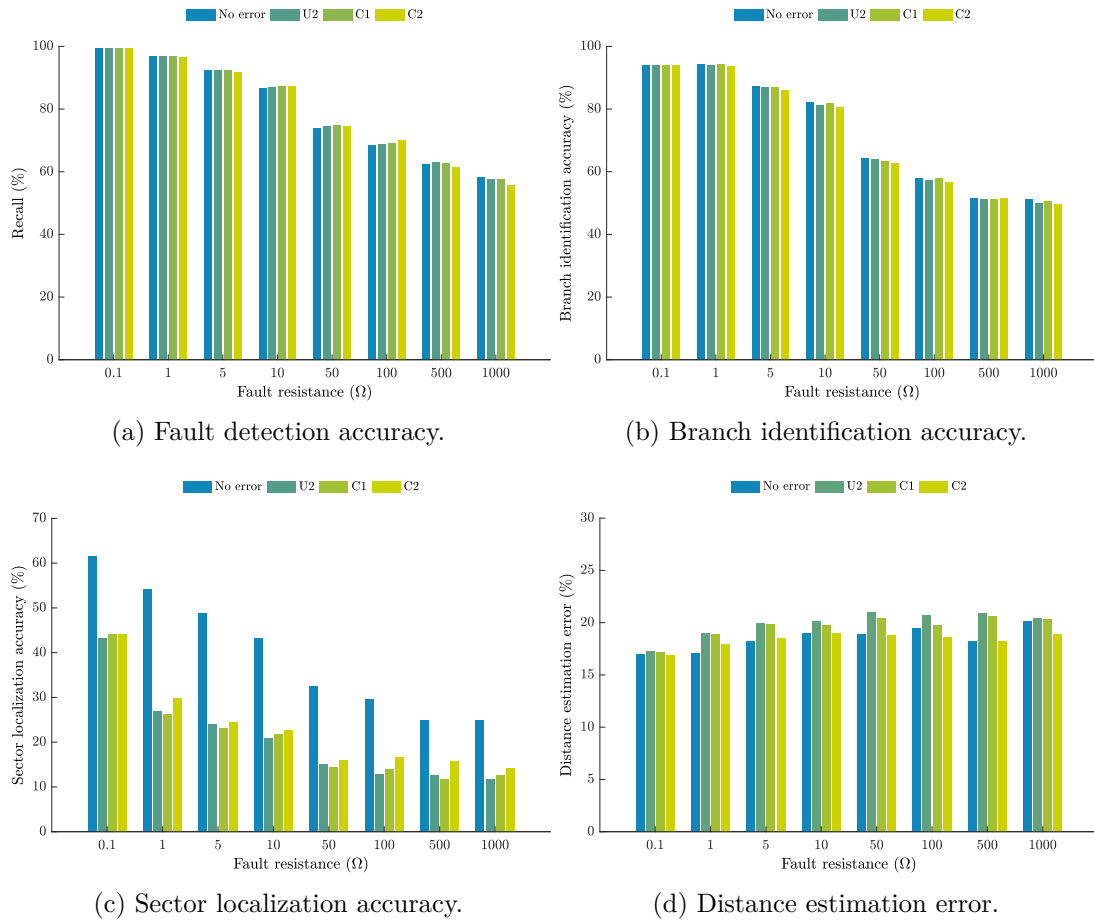


Figure 3.16 Measurement error effect on fault detection and fault location methods.

Figure 3.16 shows that fault detection and faulty branch identification tasks are not affected by the introduction of error in the current and voltage measurements respectively. On the other hand, the passing from accurate to erroneous measurements affected severely the sector localization method; its reduction in accuracy reached an average of 20 % (Fig. 3.16c). This result was expected as the sector localization part is the most sensitive of the proposed fault location method. Erroneous measurements can easily alter the form of the voltage profile curve leading to a mislocalization of the faulty sector. Additionally, the fault resistance increase amplifies the effect of erroneous measurements as nodal voltages reach similar amplitudes. Moreover, the distance estimation error is increased by a maximum of 2 % as shown in Fig. 3.16d. Finally, although no conclusive result can be obtained, the 2 % underestimation error seems to be the most challenging for the sector localization method and the class 2 accuracy the most challenging for distance estimation.

### 3.2.5 Less available measurements

A limitation of the localization method is that at least three distinct sectors are required, meaning a minimum of four measurements per branch for the sector localization and distance estimation steps. Based on an analysis made in [238], a strategy of sensor placement was developed according to the following topological criteria:

- a) voltage sensors should be spread throughout the branch so that the voltage profile curve would be a good approximation of the ideal case, and
- b) nodes with big loads and/or PV units connected to them should be prioritized taking into account all three phases.

Through a preliminary analysis [238], the bare minimum of four measurements per branch was deemed insufficient for the method. The following strategy was applied to reduce the sensors:

Step 1: First and last measurements in each branch are mandatory. This means that nodes 1, 29, 10, 17, 25, 30, 33, 21, 22 and 32 will be equipped with sensors.

Step 2: Removed sensors from nodes with zero loads and PVs. Those nodes are: 3, 4, 14, and 15. However a choice between 4 and 14 is necessary to maintain four measurements in branch eight. Between the two, node 4 is more important as it is shared by branch nine as well. That means that nodes 3, 14 and 15 were removed.

After these two first steps the candidate for removal nodes are: 9, 16, 24, 11 and 18 from the first feeder, 6, 12, 19, 26, 20, 27, 31 from the second and 23 and 28 from the third feeder (fourteen in total). To achieve a 30 % reduction of sensors, ten in total need to be removed. Since three were already removed during Step 2, seven out of the fourteen listed above remain. An extra step is necessary to choose one scenario among the 3,432 possible ones.

Step 3: The top three nodes in terms of connected loads (in  $kVA$ ) in each branch were chosen to remain (Table A.5 to A.7 in the Appendix). This removed exactly seven nodes from the remaining fourteen.

The final choice of nodes to be left unmonitored was: 16, 18, 3, 6, 19, 20, 31, 14, 15, 23 and is presented in Fig. 3.17. Given this sensor availability constraint, the complete set of simulations was repeated without considering any error in the measurements.

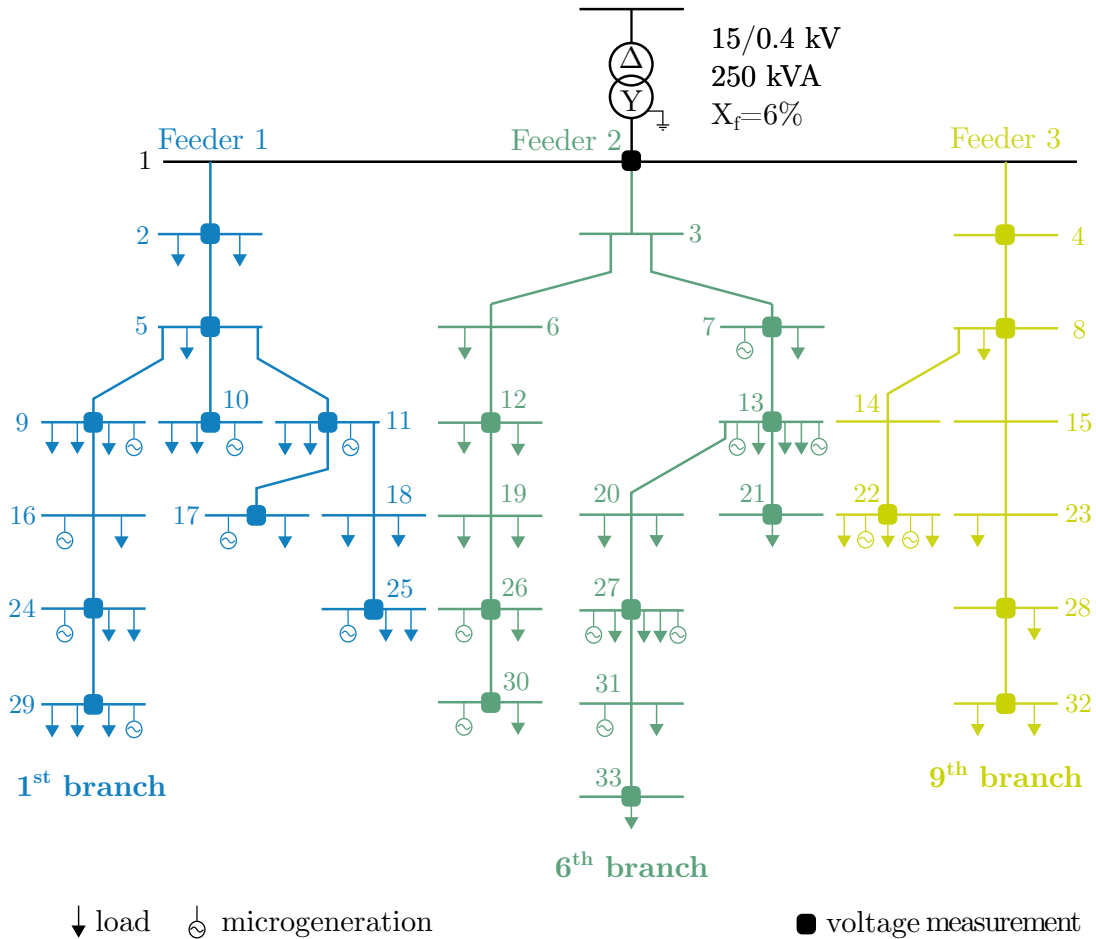


Figure 3.17 Grid schematic with the reduced number of available nodal voltage measurements.

Figure 3.18 presents the performance of the fault location method under the scenario of less available measurements described above. The faulty branch identification process is not affected in this case (Fig. 3.18a). This was expected since the most important measurements for branch identification are the ones towards the end of the branch where the higher voltage drop is expected; a condition that is met in this case since all the last nodes in each branch are equipped with voltage sensors. Again, the sector localization is where the method performance is expected to decrease. However, the availability of less sensors leads unavoidably to longer sectors (as seen from the extracted voltage profile

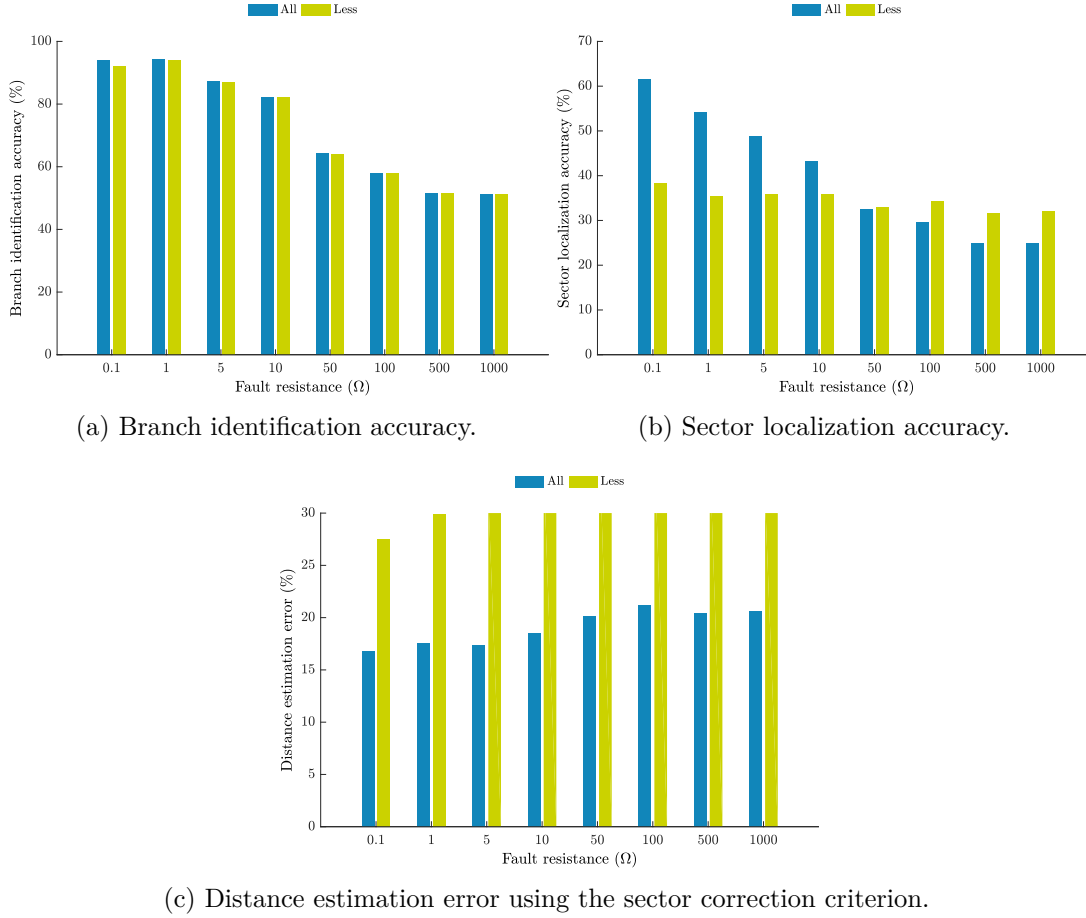


Figure 3.18 Less available measurements effect on the fault location method.

point of view) which makes the method less sensitive to fault resistance as the number of possible sectors within a branch is significantly reduced; a decrease of only 6.4 % between 0.1  $\Omega$  and 1000  $\Omega$  was observed. This effect is presented in Fig. 3.18b. Regarding distance estimation performance, an average error increase of 14.72 % is noticed between the case of full and limited available measurements as the longer sectors affect significantly the extracted voltage curve .

Given the graphic nature of the method and its sensitivity when it comes to localizing the faulty sector and estimating the fault distance, a reduction of the available measurements is not advised even though the faulty branch identification process is unaffected. Finally, there was no need to examine the fault detection performance as from the beginning, the bare minimum of sensors was used to monitor the operation of the grid in terms of current (one sensor per feeder).

### 3.3 Conclusion

In this chapter, a fault detection and fault location method for the LV distribution grid were presented. For the validation of the proposed methods a simulated real semi-rural LV distribution grid of Portugal was used. A total of five influencing parameters were considered: a) fault resistance, b) type of measurements, c) fault location in the grid, d) fault type and e) time of the day. In general, the increase of fault resistance hinders both the fault detection and fault location methods. While for the fault detection phase current measurements are the most suitable tool, for the fault location process, consisting of three distinct steps: a) branch identification, b) sector localization and c) distance estimation, the positive sequence component of the measured rms voltage is a better choice. Moreover, faults located towards the end of the feeder are more difficult to tackle. More specifically, faults near the limits of a sector can impede the fault location method in its attempt to locate them. Furthermore, three phase faults, being the most severe, are the easiest ones to detect and localize. Finally, an increased load demand during peak hours decreases the accuracy of the method, with PVs, though still connected to the grid, not being able to mitigate this effect.

The fault detection method, based solely on phase current rms measurements located in the beginning of each of the grid feeders, is considered as a reliable fault detection tool only for fault resistance values up to  $10\ \Omega$ . The average accuracy of the fault location method is: 72.85 % for the faulty branch identification, 43 % for the sector localization and 82 % for the distance estimation task. Additionally, under measurement uncertainty, the sector localization method is severely affected with an average decrease of its accuracy by 20 % (compared to uncorrupted measurement data) while all the other parts of the methods are unaffected. Finally, for the case of less available measurements, the faulty branch identification method was the only one to remain unaffected under the proposed scenario of sensor placement.





# Chapter 4

## Gradient boosting trees

In order to overcome the disadvantages of the conventional methods and improve the chances of correctly detecting and locating a fault, a new artificial intelligence method is proposed in this Chapter based on *gradient boosting trees (GBT)*. The proposed method can detect and identify both single phase to ground faults, and three phase faults. The reasons for selecting this algorithm as a first approach and not other, e.g. a neural network, were: a) this algorithm has been shown to outperform other regression tree methods and has recently become the winner of several challenges in Kaggle, a site that hosts machine learning competitions, b) it has been successfully used in other energy-based applications, e.g. forecasting electricity prices [239] or solar irradiance forecasting [240], and c) it is a very fast model to train which allows real-time applications.

Being a knowledge based method, its performance is tested against out-of-sample data. The case study of Chapter 3 is extended to include more fault resistance values and more hours of the day. The contribution of this method is threefold and is summarized below:

- a) fault detection and faulty feeder identification: the occurrence of the fault is detected with a simultaneous identification of the feeder under fault.
- b) faulty branch identification: following the feeder and phase identification, the faulty branch within a faulty feeder is also identified.
- c) fault type identification: a distinction of the faulty and non-faulty phases is achieved thus identifying the fault type, single phase to ground (AG, BG or CG) or three phase fault.

## 4.1 Method description

### 4.1.1 Model definition

The GBT algorithm [241] is a prediction model based on the principle of combining several regression trees. In particular, regression trees are models characterized by either having high bias and low variance errors if the tree is shallow, or low bias and high variance errors if the tree is deep. To solve this issue, there are two families of algorithms that combine several regression trees to reduce high errors. The first family are random forests and it is based on the principle of bagging [242], i.e. combining models with low bias and high variance error in order to reduce the variance while keeping a low bias. The second family are gradient boosting trees and it is based on the principle of boosting [242], i.e. combining models with high bias and low variance error in order to reduce the bias while keeping a low variance error.

### 4.1.2 Algorithm functionality

The proposed algorithm has three distinct functionalities:

- **Fault detection:** the first functionality of this algorithm is the detection of a fault occurrence with a simultaneous identification of the feeder under fault.
- **Faulty branch identification:** the last functionality of this proposed method is the faulty branch identification, which is the faulty branch within a faulty feeder.
- **Fault type identification:** an extra element which is often omitted by fault location algorithms is the fault type identification process. In this study, the GBT algorithm can also differentiate faulty from non-faulty phases.

### 4.1.3 Working principle

The main idea of the algorithm is to make use of its prediction capabilities to diagnose the grid faults. In particular, the algorithm uses a training dataset  $\mathbb{S} = \{X_i, Y_i\}_{i=1}^N$ , where  $X$  are the inputs of the GBT model and  $Y$  are the desired

predicted output. For all the identification tasks, the inputs  $X$  are the same: specific data corresponding to a specific branch, e.g. voltage on that branch. The outputs  $Y$  however depend on the specific task. Particularly, the output  $Y$  changes with the task as each algorithm has a slightly different working principle in each of the three tasks:

- **Fault detection:** to identify a faulty feeder, the algorithm considers data from healthy branches in healthy feeders and data from faulty branches. Then, it labels the healthy branches with a 0 and faulty branches with a 1 and the algorithm is trained to predict 0 or 1 to indicate the existence of a fault in a branch. In real time, to identify a faulty feeder, the algorithm is simply tested on all the branches of a feeder.
- **Faulty branch identification:** to identify the faulty branch within a faulty feeder, the algorithm considers data from healthy branches in a faulty feeder and data from faulty branches. Then, the algorithm is trained to distinguish between the two cases using two labels, i.e. 0 and 1. In real time, to identify the branch, the algorithm is tested on the branches of a faulty feeder.
- **Fault type identification:** to identify the type of fault, the algorithm considers only data from faulty branches. Then, it labels each branch datapoint with 1, 2, or 3 to respectively denote single phase fault in phase A, B and C, and uses a label 4 to denote three-phase faults. In real time, to identify the fault, the algorithm is simply tested on the faulty branch.

#### 4.1.4 Training, validation & test

Independently of the task, as they are all classification tasks, the algorithm is trained to minimize the cross-entropy loss of the training dataset, i.e. the price paid for inaccurate predictions. Moreover, to optimize the structure of the algorithm, all the boosting tree hyperparameters, e.g. number of branches or tree depth, are optimally selected using the Bayesian optimization [243]. In particular, the dataset is divided in three subsets: a training dataset, a validation dataset and a test dataset. The training dataset is used to estimate the algorithm parameters, the validation dataset is used to estimate the algorithm hyperparameters and finally the test dataset is used to evaluate the quality of the algorithm.

### 4.1.5 Input features

In terms of the inputs of the model several design choices were made. In particular, to make the model general enough, i.e. to make the model applicable to different grid topologies with various number of branches and available measurements, two design choices were made. First, the use of branch-specific features was avoided, e.g. the branch length or the branch resistances and reactances. Second, all branch-specific measurements were substituted with a fixed number of interpolated values so that each branch could have the exact same number of features. For instance, independently of the number of voltage measurements in a branch, five equally spaced points within the branch were selected and the voltage values from the voltage measurements were interpolated to these five locations. This is also necessary to ensure the uniformity of the dataset.

With that motivation, in order to identify if a fault occurs at time  $t$  the following input features were considered:

1. *Time*: the hour of the day corresponding to  $t$ . This is important because the load and PV penetration in the grid change along the day.
2. *Load*: the load in the grid at time  $t$ .
3. *Generation*: the PV generation in the grid at time  $t$ .
4. *Current at time  $t$* : the current at the beginning of each feeder at time  $t$  was considered as shown in Fig. 3.6. In particular, the current through the three phases and the neutral.
5. *Current 5 min before  $t$* : the current at the beginning of each feeder five minutes before  $t$  was also considered. As before, current through the three phases and the neutral was considered. These features are important to have a comparison between two points close in time so that if a fault occurs at time  $t$ , the method can compare the current at time  $t$  with the values of the current during normal operation.
6. *Voltages at time  $t$* : voltage values across each branch at time  $t$  were considered. More specifically, as mentioned before, five virtual/interpolated equally spaced measurements that were obtained from the real measurements in the branch were considered. Moreover, the voltages for each phase were considered, i.e. in total fifteen voltage points per branch.

7. *Voltage 5 min before  $t$* : voltage values across the branch five minutes before  $t$  were also considered. The same fifteen voltage points as in time  $t$  were used. As with the current, the motivation behind these input features is to provide the method with voltage measurements during normal operation.

In order to further explain the choice of 5 *min* before the fault occurrence as a reference for normal operating conditions, a more thorough analysis is provided below. Smart meters are able to provide measurements with a frequency of 1-10 s [244]. However, they cannot yet be used to their full potential due to telecommunication restrictions in the *supervisory control and data acquisition (SCADA)* system. As mentioned in [245], the French smart meter Linky, is providing measurements of injected and consumed power every 30 *min*. In other cases, collection of smart meter data from the service provider, has been reduced to 15 *min* as reported in [246]; such an example is the MBox smart meter of Efacec [234]. The need of a close in time reference point of normal operation resides in the intermittent and rapidly changing behaviour of the PV generators, e.g. in the presence of clouds passing over the arrays production from the PV cells will vary. However, as explained in Chapter 3, the contribution of PVs to the faults is insignificant if compared to that of the loads and since 30 *min* were considered a quite long time to track changes in the energy consumption behavior, the trade-off of measurements every 5 *min* is proposed here, looking into the future of improved telecommunication infrastructure that 5G technologies are about to bring in smart grids.

## 4.2 Application

### 4.2.1 Case study

#### 4.2.1.1 Grid characteristics

The same grid presented in Fig. 3.6 was used again for this case study.

#### 4.2.1.2 Influencing parameters

Six different influencing parameters were identified for this study:

1. *Fault resistance*: Extending the fault resistance values considered for the conventional method case study, presented in Chapter 3, eighteen different fault resistances were investigated and divided in two sets: (a) training set (included at the training phase): 0.1, 0.5, 1, 3, 5, 7.5, 10, 30, 50, 75, 100, 300, 500, 750 and 1000  $\Omega$  and (b) out-of-sample set (excluded from the training phase): 4, 40 and 400  $\Omega$ , covering the full spectrum of faults, both low and high resistance ones.
2. *Fault location*: Similarly with Chapter 3, in every sector, nine possible locations of fault occurrence were considered for distances of 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 % and 90 % from the beginning of the sector.
3. *Fault types*: Again, single phase to ground faults and three phase faults were selected.
4. *Simultaneity factor*: The assumption of a global simultaneity factor of 0.3 was made for the loads.
5. *Time of the day*: The same PV generation and load profiles with the case study of the previous chapter were used as a basis. Only this time, during the simulations, in order to create multiple healthy operation scenarios, a uniformly distributed noise of 20 % and 2 % was introduced in the PV generation and load respectively as presented in Fig. 4.1.
6. *Measurement noise*: As stated before, this study considered phase rms voltage measurements at every node and phase rms current measurements

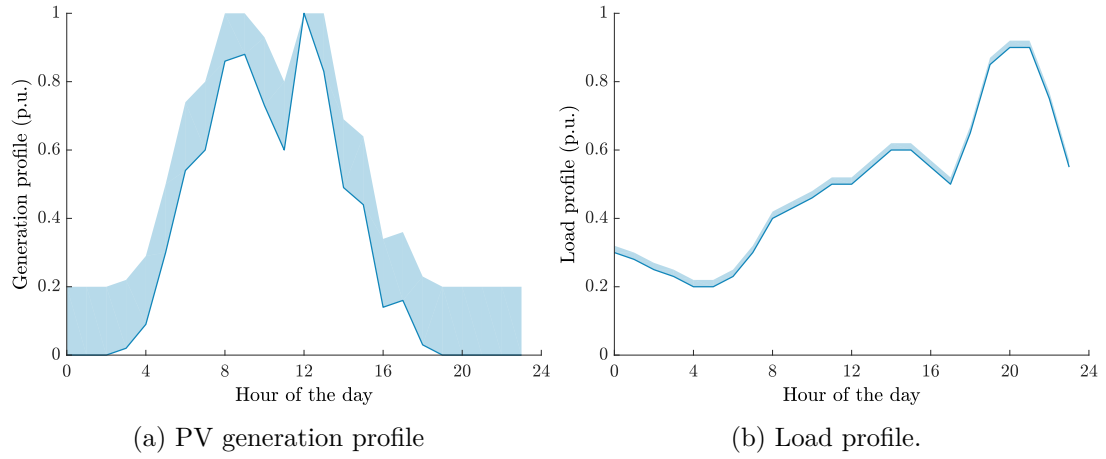


Figure 4.1 Generation and load profiles for one day. The shaded areas show the pool from which values were drawn to create multiple healthy operation scenarios.

(including the neutral) at the beginning of each feeder. In order to approximate real measurement conditions as much as possible, a 2% of underestimation error (values from 0.98 to 1 p.u.) was introduced to each phase measurement independently.

Faulty operation measurements were taken 150 *ms* after the fault occurrence. This choice was made for the fault to be as close to the steady state as possible if not already there and to avoid corruption of the data by the activation of any protective element. As mentioned in Chapter 3, this choice also ensures that the PV units are connected to the grid as the inverter has not yet isolated them [4].

For every fault resistance value, all nine of the possible fault locations in each of the thirty two sectors were studied; for every fault location all four possible fault types were considered (AG, BG, CG and ABC); and finally, for every fault type, eight hours within a day were selected randomly in intervals of three hours, in order to gather samples from every hour of the day. These cascaded *for loops* lead to a total of 165,888 simulation datapoints covering all twenty four hours of the day.

In order for the algorithm to be able to detect the occurrence of the fault, i.e. distinguish faulty from normal operation, a healthy dataset was also created. For that purpose, for every hour of the day (twenty four cases), three hundred scenarios of PV generation and load combinations were simulated for each branch leading to a total of 64,800 datapoints.



### 4.2.2 Implementation

The algorithm was implemented in `python` using the `XGBoost` [241] library for the GBT model, and the `hyperopt` [243] library to perform the hyperparameter optimization based on Bayesian optimization. The case study with the simulations were implemented in `Matlab/Simulink`.

### 4.2.3 Model training and evaluation

The model was repeatedly trained with the training dataset and the algorithm was evaluated in the validation dataset for guiding the Bayesian optimization algorithm to find the optimal parameters. Then, after the optimal hyperparameters were found, the algorithm was evaluated in the test dataset. Table 4.1 presents the size of each dataset; the test dataset consisted of 28,800 healthy data and 38,880 faulty data.

In addition, with the same motivation that was described when defining the model and its input features, i.e. to have a model that generalizes to different grid topologies or faults, the algorithm was evaluated in out-of-sample fault resistances and out-of-sample branches. In particular, while the training and validation datasets considered the first set of fault resistances, set (a), the test dataset comprised of fault resistances of 4, 40, and 400  $\Omega$  (set (b)). Similarly, while for the training and validation datasets only data from the first and third feeder were taken into account, branches 1-4 and 8-9 respectively, for the test dataset the branches 5-7 belonging to the second feeder were employed (Fig. 3.6). The last choice is justified as feeder one and three have the maximum and minimum number of branches respectively. In that way, as it will be shown in the next section, the algorithm was able to provide promising results not only on fault resistances and branches belonging to the training dataset, but also in out-of-sample fault resistances and branches.

For the results that follow, the accuracy formula described in eq. 3.3 was used.

### 4.2.4 Fault detection

The first functionality of this algorithm is the detection of a fault occurrence with a simultaneous identification of the faulty feeder. The results for out-of-

Table 4.1 Dataset sizes.

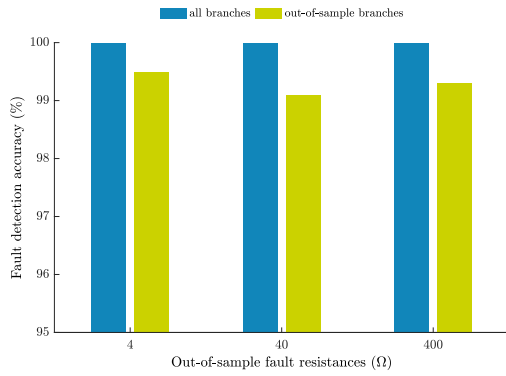
Dataset type	Size
Train	69,993
Validation	23,332
Test	67,680

sample fault resistances for the cases of: a) all branches participating in the training of the algorithm and b) out-of-sample branches, are presented in Fig. 4.2. In the first case, an accuracy of 100 % is achieved. For the second one, a minimum accuracy in fault detection of 99.15 % was noticed (Fig. 4.2a). More specifically, for the case of out-of-sample branches, a slightly different behavior is presented in the three branches of the second feeder. As shown in Fig. 4.2b, branch five presents the best accuracy, followed by branch seven and branch six. This difference in accuracy is attributed to the fact that branch six is the one with the highest contracted and installed power (higher number of connected loads and PVs) as shown in Table A.6. The increased number of existing loads and PVs in branch six is affecting the voltage measurements across the faulty branch thus leading to a slightly increased error of 1.5 % in fault detection.

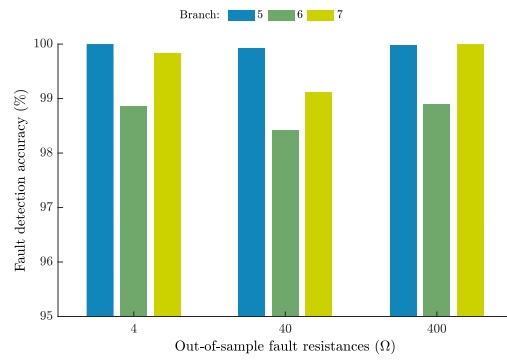
With a minimum accuracy of 99.15 %, when all branches are included in the training of the algorithm, the fault detection and feeder identification results are considered excellent. The excellent performance of the GBT algorithm, with a minimum accuracy in fault detection of 98.5 % even when tested with out-of-sample branches, makes it a very reliable and generalizable method against different grid topologies as well. Since the maximum error was 1.5 % for both single phase to ground and three phase faults, it is also safe to assume that the algorithm will also detect out-of-sample two phase to ground and phase to phase faults.

#### 4.2.5 Fault location

The second functionality of this proposed method is the faulty branch identification which in LV grids with many laterals (branches) is the most important aspect of the fault location process. Four different cases were considered to

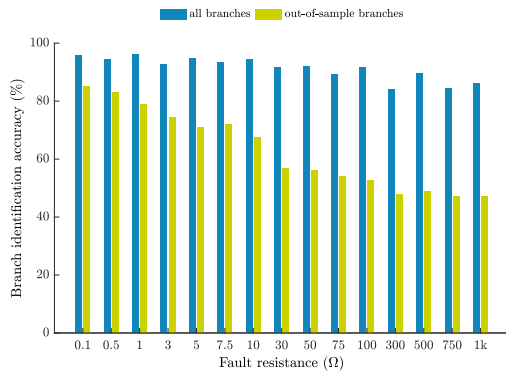


(a) All branches vs. out-of-sample branches.

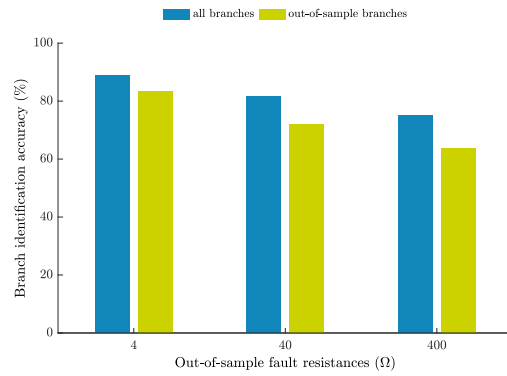


(b) Out-of-sample branches (Fig. refgrid).

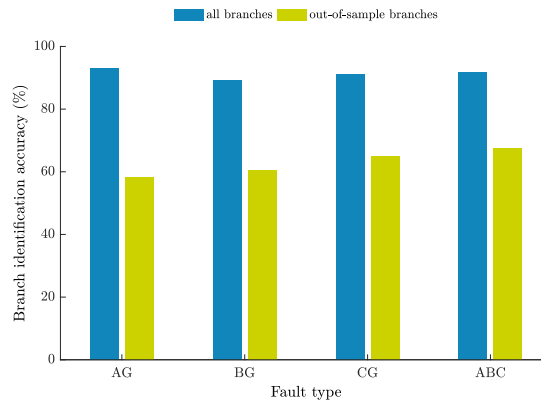
Figure 4.2 Fault detection accuracy.



(a) Fault resistances included in training.



(b) Out-of-sample fault resistance values.



(c) 1ph vs. 3ph faults.

Figure 4.3 Branch identification accuracy: all branches as part of the training set vs. out-of-sample branches.

measure the accuracy of the algorithm. First of all, the performance of the GBT algorithm was tested against out-of-sample data but for resistance values that were part of its training (Fig. 4.3a). Secondly, its accuracy was measured against out-of-sample fault resistance values that the algorithm was not trained for (Fig.

4.3b). For both cases both scenarios of a) all the available branches being part of the training set and b) out-of-sample branches were explored.

In Fig. 4.3a, a higher accuracy of the GBT algorithm is demonstrated in case the tested fault resistance is part of the training set, with a maximum accuracy of 95.8 % for 0.1  $\Omega$  and a minimum of 84.1 % for 300  $\Omega$ . Similarly, for out-of-sample branches, the accuracy decreases from 85.1 % for 0.1  $\Omega$  to 47.1 % for 750 and 1000  $\Omega$ . A general tendency of a decrease of the accuracy with the increase of the fault resistance is observed for both cases. According to the above results, the attempt to identify a faulty branch in case the grid topology changes, is not considered very successful for high fault resistance values. In that case, a retraining of the GBT is advised. The same is also advised in the case of another microgrid which needs to operate in isolated mode; this is not the case of this grid as the installed generation is not sufficient to serve the existing loads as shown in Table A.4. The algorithm rapid training time facilitates that process and makes it ideal for real-time applications.

Furthermore, in Fig. 4.3b, the accuracy of the faulty branch identification process is presented for the out-of-sample fault resistances of 4, 40 and 400  $\Omega$ . The same conclusions that were drawn before are validated here again.

Moreover, in Fig. 4.3c, the accuracy of the faulty branch identification process is presented for each fault type. The GBT algorithm, shows similar behavior to every fault type with one exception: it performs better in the case of the three phase fault for the out-of-sample branches case compared to the single phase to ground fault cases. The latter, is attributed to the fact that in the case of a three phase fault, the voltage drop across the faulty branch is more symmetrical and severe and thus more likely to be differentiated from other healthy branches even for high fault resistance values.

#### 4.2.6 Fault identification

An extra element which is often omitted by fault location algorithms is the fault type identification process. In this study, the GBT algorithm was also implemented to differentiate faulty from non-faulty phases. The results for this functionality for the various fault resistances included in the training phase are provided in Fig. 4.4a while in Fig. 4.4b the accuracy against the different fault types is provided. Although the results presented in Fig. 4.4 concern the

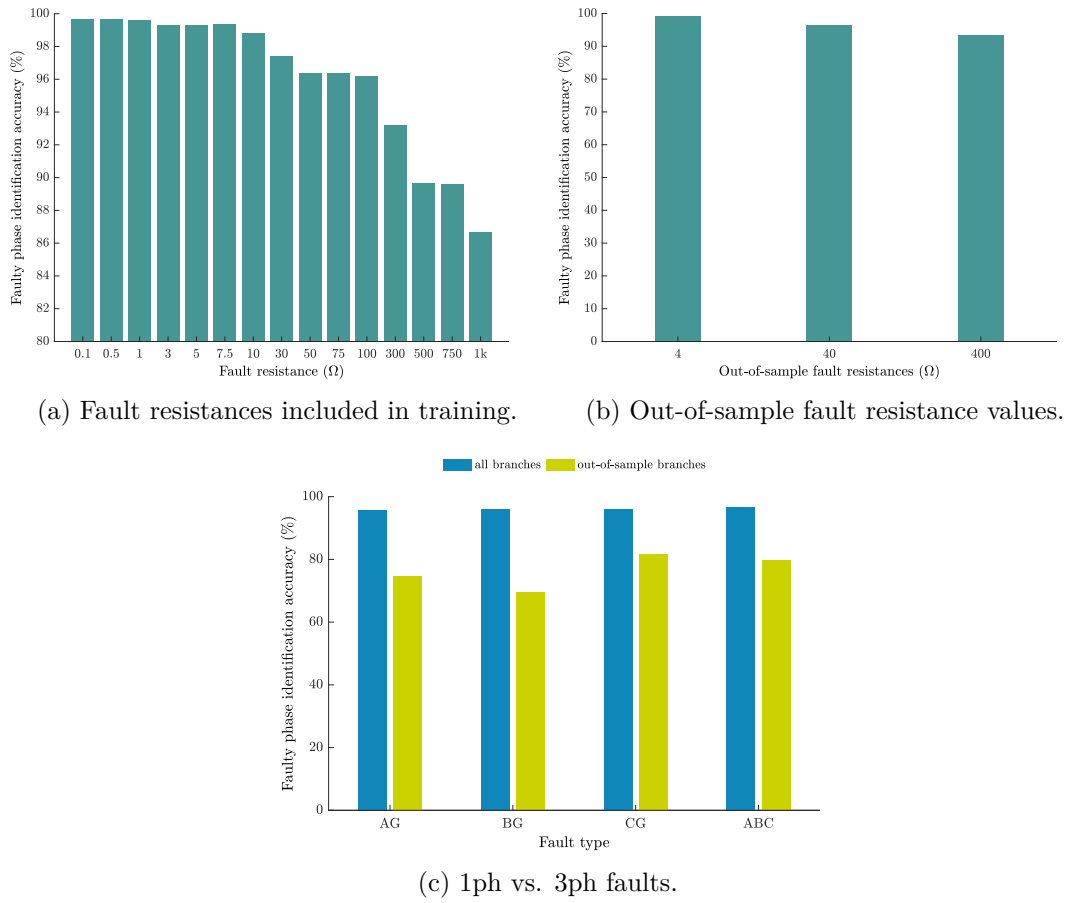


Figure 4.4 Phase identification accuracy.

fault resistance values that were used to train the algorithm, the accuracy was measured against out-of-sample data of the same dataset.

In Fig. 4.4a, the first effect of the increase of fault resistance is noticed. For low resistance faults (below 10  $\Omega$ ), the accuracy of faulty phase identification is maintained at a level higher than 98.8%. After that, a more and more significant decrease of accuracy is noticed with the increase of the fault resistance down to a minimum of 86.7% for 1000  $\Omega$ . This was an expected result as the increase of fault resistance will decrease the voltage drop during a fault and thus bring the voltages across a faulty branch closer to the values of faulty operation, as stated in Chapter 3.

Furthermore, in Fig. 4.4b the accuracy of the faulty phase identification process is presented for the out-of-sample fault resistances of 4, 40 and 400  $\Omega$ . The same conclusions that were drawn before are validated here again.

Finally, in Fig. 4.4c, a similar performance was noticed in all four types of faults: single phase to ground (AG, BG, CG) and three phase faults (ABC) with

---

a maximum deviation of 0.9 % between single phase to ground and three phase faults. This is an indication that besides the different combinations of generation and load penetration in the grid, the unbalanced nature of the grid, i.e. the per phase asymmetry in the distribution of PVs and loads in the grid, did not affect the GBT algorithm.

### 4.3 Conclusion

In this chapter, a gradient boosting tree model was proposed to detect, identify and locate faults in *low voltage (LV)* smart grids. To estimate the model, a set of non branch-specific input features was employed to ensure the robustness of the algorithm against different grid topologies and available number of voltage measurements per branch. The proposed method was evaluated in a case study of a real case semi-rural LV distribution grid of Portugal. In detail, the case study comprised: a) fault resistances between 0.1 to 1000  $\Omega$ , b) different fault locations inside each sector, c) different fault types, d) different hours of the day, and e) a 2 % of underestimation error in the phase rms current and voltage measurements.

To test the accuracy of the proposed algorithm, the method was tested in an out-of-sample dataset. In addition, to analyze the robustness and generalization capabilities of the algorithm, the method was also tested against out-of-sample fault resistances and branches (resistance values and grid branches not included in the training dataset).

An excellent accuracy for fault detection was achieved. Faulty phase and branch identification showed promising results. A great feature of the algorithm is that, as can be seen in the symmetrical performance in all the phases, the asymmetrical distribution of loads and PVs across the phases and branches does not really affect the algorithm performance. In addition, as it was expected, the increase of the fault resistance decreased the accuracy across all three tasks.

In detail, the algorithm achieved an accuracy of 99.15 % when identifying the faulty feeder, an accuracy between 95.8–84.1 % when identifying the faulty branch, and an accuracy between 98.8–86.7 % (the higher the fault resistance the lower the accuracy) when identifying the fault type. However, the attempt to identify a faulty branch in case the topology or the operational mode of the grid change (e.g. installation of additional loads, PVs or batteries, isolated mode operation or voltage support capabilities), is not considered very successful for high fault resistance values. In that case, a retraining of the GBT is advised.

In general, the results demonstrated in this chapter show a clear superiority of the proposed method with regards to the conventional method for LV grids.

# Chapter 5

## Deep Neural Networks

In recent years, the research on neural networks has achieved several breakthroughs that have lead to what is now known as deep learning. In particular, due to these breakthroughs, the usage of neural networks whose depth is no longer limited to a single hidden layer is now possible. These deeper neural networks, as proven by several applications, have systematically proven to be better at estimation problems due to their better generalization properties [247].

In this chapter, the use of *deep neural networks (DNN)* is proposed as a solution to fault detection and location problems in *low voltage (LV)* distribution grids. In detail, the contribution of the developed method is fourfold:

1. A fault detection method that detects the occurrence of a single phase to ground or three phase fault with a simultaneous identification of the faulty feeder.
2. A faulty branch identification algorithm that identifies the faulty branch within a faulty feeder.
3. A fault distance estimation method that determines the location of the fault.
4. An average accuracy decrease of only 4.5% during the faulty branch identification and distance estimation tasks in the case of extremely limited measurements, i.e. when having only one available measurement in the terminal node of each branch and one at the substation level (common for all feeders).



## 5.1 Method description

### 5.1.1 Deep neural networks

In general, a neural network is nothing else than a model  $F(\mathbf{X}, \mathbf{W})$ , with parameters  $\mathbf{W}$ , that uses some input features  $\mathbf{X}$  in order to predict some variable of interest  $\mathbf{Y}$ . Thus, to use the neural network for a given task, i.e. predict  $\mathbf{Y}$ , one only needs to gather a dataset  $\mathcal{S}_{\mathcal{T}} = \{(\mathbf{X}_k, \mathbf{Y}_k)\}_{k=1}^N$  and use this dataset to estimate the optimal parameters  $\mathbf{W}^*$  that best fit the dataset. Let us define the input of a neural network by  $\mathbf{X} = [x_1, \dots, x_n]$  and the output by  $\mathbf{Y} = [y_1, \dots, y_m]$ . Let us also define the number of neurons of the  $k^{\text{th}}$  hidden layer by  $n_k$  and by  $\mathbf{z}_k = [z_{k1}, \dots, z_{kn_k}]$  the state vector in the same layer. Using these definitions, a general DNN with two hidden layers can be represented by Fig. 5.1.

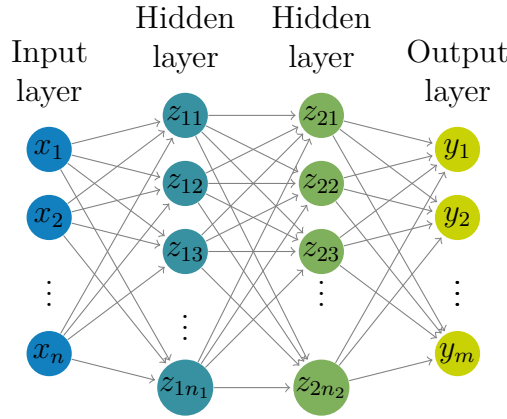


Figure 5.1 Example of a DNN.

In this model, the parameters  $\mathbf{W}$  are the weights establishing the mapping connections between the different neurons of the network. In detail, the mapping equation of a general neuron  $i$  in the  $k^{\text{th}}$  layer is given by:

$$z_{ki} = f_{ki} \left( \mathbf{W}_{ki}^{\top} \cdot \mathbf{z}_{k-1} + b_{ki} \right) \quad (5.1)$$

where  $f_{ki}$  represents the activation function of the neuron,  $\mathbf{z}_{k-1}$  the values of the neurons of the previous layer, i.e.  $k - 1$ ,  $\mathbf{W}_{ki}$  the matrix of weights establishing the connection between all the neurons of layer  $k - 1$  and the neuron  $i$  in the  $k^{\text{th}}$ , and where  $b_{ki}$  is the so-called bias parameter of the neuron. Typical activation

functions are the sigmoid function, the hyperbolic tangent function, or the rectified linear unit [242].

The process of estimating the model weights is usually called training. Given the previously defined  $N$ -dimensional set  $\mathcal{S}_{\mathcal{T}} = \{(\mathbf{X}_k, \mathbf{Y}_k)\}_{k=1}^N$ , the network training is done by solving a general optimization problem with the following structure:

$$\underset{\mathbf{W}}{\text{minimize}} \quad \sum_{k=1}^N g_k(\mathbf{Y}_k, F(\mathbf{X}_k, \mathbf{W})), \quad (5.2)$$

where  $g_k$  is the problem-specific cost function. For grid fault diagnosis, this cost function varies depending on the specific task (more details on this later).

### 5.1.2 Network inputs

While the output of the neural network will vary depending on the specific fault diagnosis task, the possible inputs of the network stay constant across all tasks. In particular, while a feature selection method is performed for each specific task to obtain the most representative inputs, the set of possible input features that the feature selection selects from, is kept constant. In detail, to detect a fault at time  $t$ , this feature set is defined by several branch-related information:

- Five branch voltages at time  $t$  obtained by interpolation of the real branch measurements into five equally spaced measurements.
- The current at time  $t$  at the beginning of the feeder where the branch is located.
- The generation and load in the distribution grid.
- The same three values but 5 minutes before  $t$ . The idea of using the same values some time before is to have some measurements during a similar operation to compare with.

An important thing to note in this feature subset is that the inclusion of branch specific features, e.g. branch length or branch resistance, is avoided. Similar, independently of the number of voltage measurements in a branch, five equally spaced measurements were always considered by interpolating the real measurements. The reason behind these two design choices is to have a model

that can be generalized to new grids. In particular, for the model to be applied to a different grid, the input of the network cannot be branch-specific. Similarly, if the number of voltage measurements depend on the branch, any proposed method would be, not only grid-specific, but branch-specific.

As a remark, to test that excluding branch-specific features did not harm the performance of the model, the DNN performance was compared with and without these features. While these results are out of the scope of this thesis, it is worth mentioning that it was observed that both approaches lead to similar performance.

### 5.1.3 Feature and hyperparameters selection

As indicated above, the inputs of the network were selected for each specific task based on a large input set. To do this selection, motivated by the success in other energy-related studies [240, 248], the *Tree-Structured Parzen Estimator (TPE)* was used [243]. In addition, together with the feature selection, this algorithm is also used for selecting the optimal hyperparameters of the DNN, e.g. number of neurons of the hidden layers, the number of hidden layers, the type of activation functions, etc.

### 5.1.4 DNN estimation

For each fault diagnosis task, in order to estimate the optimal network, i.e. optimal DNN weights  $\mathbf{W}$ , optimal input features and optimal hyperparameters, the same procedure is repeated:

1. Divide the dataset  $\mathcal{S}_{\mathcal{T}} = \{(\mathbf{X}_k, \mathbf{Y}_k)\}_{k=1}^N$  in three subsets: training (60 %) + validation (20 %) + test (20 %).
2. Perform the feature and hyperparameter optimization using the training and validation dataset:
  - The training dataset is used to solve (5.2) and estimate the weights  $\mathbf{W}$ .
  - The validation dataset is used as an out-of-sample dataset to select the optimal features and hyperparameters.

3. Using the optimal network, evaluate its performance on the test dataset, i.e. an out-of-sample dataset never used during training.

To solve (5.2), the Adam optimizer was used [249]. In addition, to avoid overfitting, i.e. estimate a DNN that fits perfectly to the training dataset but it cannot generalize to new data, the network is trained in combination with early stopping and out-of-sample data to evaluate the performance.

### 5.1.5 Network outputs and cost function

As mentioned before, the only DNN differences between the fault diagnosis tasks are the loss function and the network output:

- **Fault detection and feeder identification:** detecting whether a feeder has a fault can be modeled with a classification network. In particular, for each branch-related input  $X$ , the output  $Y$  is defined to be 1 or 0 to respectively denote that the branch has a fault or that the branch has no fault and is in a healthy feeder (the data from those branches who are healthy but in a faulty feeder were excluded). Then, for training, the neural network can simply minimize the binary cross entropy loss (standard loss for classifying between two classes). In real time, this network can simply be applied to all the branches of a feeder to identify if there is a fault on the feeder.
- **Branch identification:** identifying the faulty branch in a faulty feeder is a very similar task to identifying a faulty feeder. However, instead of labeling 1 and 0 the branches in and out of faulty feeders, only branches within a faulty feeder were considered. For all these branches, the output was defined as 1 for the faulty branch and as 0 for the healthy branch in a faulty feeder (data from healthy branches in healthy feeders were excluded). In real time, once the feeder is identified with the previous method, this network is run to identify the faulty branch. As it is a binary classification problem, the network considers again the binary cross entropy loss.
- **Fault distance estimation:** unlike the previous two tasks, estimating the distance at which the fault occurs is no longer a classification problem but a regression problem as the output is now quantitative; fault distance

is estimated in a range between 0 and 1 where 0 marks the beginning of the feeder and 1 the end of the branch. To solve it, the subset of faulty branches was considered. Then, each output  $Y$  was defined as the distance at which the fault occurred. For the loss function, the standard mean squared error was considered. In real time, once the branch has been identified by the previous methods, this network indicates the fault distance.

## 5.2 Application

### 5.2.1 Case study

#### 5.2.1.1 Grid characteristics

The same grid presented in Fig. 3.6 was used again for this case study.

#### 5.2.1.2 Influencing parameters

Eight different influencing parameters were identified for this study:

1. *Fault resistance*: To sample fault resistances, a log-uniform distribution between 0.1 and 1000  $\Omega$  was implemented. The studied fault resistances were divided into six groups according to their order of magnitude in Ohms( $\Omega$ ):
  - (a) very low [0.1-1)
  - (b) low [1-10)
  - (c) common [10-50)
  - (d) high [50-100)
  - (e) very high [100-500)
  - (f) extra high [500-1000]
2. *Voltage measurement*: The choice between using phase rms voltages or the positive, negative and zero components after a symmetrical analysis was analyzed.
3. *Number of layers*: DNN with up to seven layers were used.
4. *Fault location*: Every sector of the grid (section between two consecutive nodes), was divided in ten subsections and the following nine locations were chosen at a distance of 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 % and 90 % from the beginning of each sector.
5. *Fault types*: As explained in the Chapter 2, single phase to ground faults (AG, BG and CG) were chosen because they are the most frequent and three phase faults (ABC) because they are the most severe.

6. *Simultaneity factor (SF)*: Since not all the loads of every consumer are going to be activated simultaneously, three coincident-simultaneity factors (described in (3.5)) of 0.3, 0.5 and 0.8 were considered .
7. *Time of the day*: The generation and load profiles that were considered in this study are provided in Fig. 5.2a and Fig. 5.2b respectively. The shaded areas portray the Gaussian noise of a 0 mean and 20 % and 2 % standard deviation, for the generation and load respectively, that was introduced in the original profiles in order to create a broader spectrum of generation and load combinations.
8. *Measurement noise*: As a first step, a 2 % of underestimation error (values from 0.98 to 1 p.u.) was introduced to each phase measurement independently.

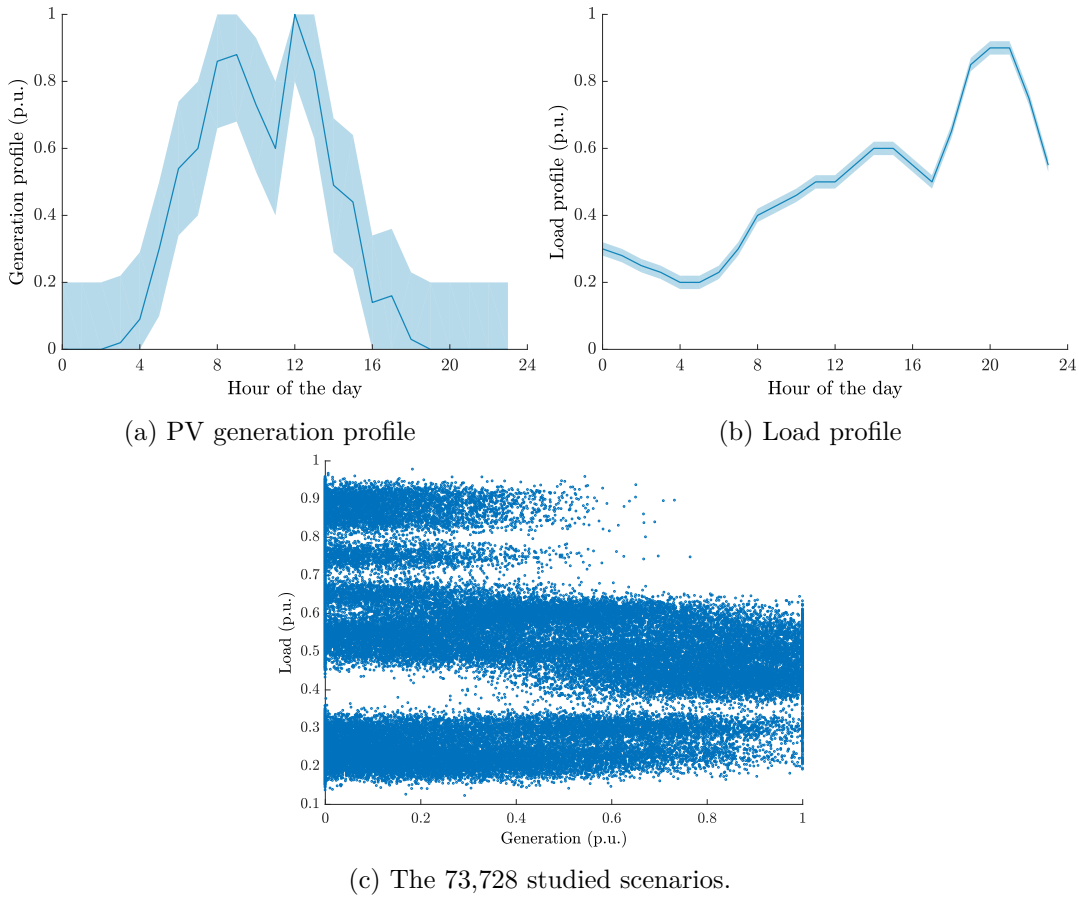


Figure 5.2 Generation and load profiles for one day. The shaded areas show the pool from which values were drawn to create multiple healthy and faulty operation scenarios. Figure (c) presents the scenarios that were considered in this study.

Taking into account all the above parameters, a total of 73,728 fault scenarios were generated for each simulation dataset. Similarly, 64,800 scenarios of healthy operation simulations (under different generation and load profiles) were also generated. The process is presented analytically in Algorithm 4. As a result, a total of 70,334 combinations of generation and load values, and a total of 63,772 fault resistance values were studied. As explained in the previous section, the data were divided into training, validation and test datasets. In addition, faults were studied 150 *ms* after their occurrence, as with all the previous cases.

Before moving to the analysis of the results, three theoretical remarks that were validated in this study should be underlined regarding the effect of fault resistance, PV penetration and increased load demand (higher *SF*).

First, it is important to note that the higher the fault resistance, the more difficult it is to detect and localize a fault due to the very small currents flowing through the faulty branch. Those very small currents will in their turn decrease the effect of voltage drop across the faulty branch, bringing the voltages to a level very close to that of normal operating conditions.

Secondly, the contribution of the PV units, being an inverter controlled source, to the fault current is expected to be somewhat limited (current limited at 1.1 to 1.4 p.u. of normal operation during a fault) [230].

Finally, the higher the load demand, the higher the voltage drop during faulty operation will be across a faulty branch. Nodes further away from the beginning of the feeder will also experience even higher voltage drops than those located at the beginning [250].

### 5.2.2 Fault detection

The DNN, with the use of just two layers, achieved an 100 % accuracy in detecting the fault occurrence and identifying the feeder under fault in all the studied scenarios. With such a level of accuracy, it is safe to assume that the DNN will also manage to distinguish faulty from normal operation if exposed to other fault types, besides the ones it was trained to detect, e.g. phase to phase, or double phase to ground faults.



**Algorithm 4** Simulations

---

```

1: procedure GENERATE NORMAL OPERATION DATA
2:   for  $t=1:24$  do                                 $\triangleright t$ : time of the day
3:     define measurement error
4:     define  $SF$ 
5:     for  $i=1:300$  do                                 $\triangleright i$ : iteration number
6:       for  $f_{br}=1:9$  do                             $\triangleright f_{br}$ : faulty branch
7:         create random generation and load
8:         simulate normal operation
9: procedure GENERATE FAULTY OPERATION DATA
10:  for  $f_s=1:32$  do                                 $\triangleright f_s$ : faulty sector
11:    define measurement error
12:    define  $SF$ 
13:    for  $t=1:3:24$  do
14:       $t \leftarrow t + randi([0, 2])$      $\triangleright$  random selection in intervals of 3 hours
15:      for  $f_d=0.1:0.9$  do                     $\triangleright f_d$ : location in the sector
16:        for  $f_t=1:4$  do                         $\triangleright f_t$ : fault type
17:          for  $R_f=1:8$  do                     $\triangleright R_f$ : fault resistance
18:            create a random  $R_f$  between 0.1 and 1000  $\Omega$ 
19:            create random generation and load
20:            simulate faults

```

---

**5.2.3 Fault location**

The fault location part of the method is divided in two steps as described above: a) branch identification and b) distance estimation. The intermediate step of the conventional method, the localization of the faulty sector, was omitted in this case for two reasons: a) in order to make the method applicable to different grid topologies, a fix number of virtual sectors per branch was created after linear interpolation of the voltages from the available measurement points and b) the fix number of sectors was also necessary for the uniformity of the input features per branch of the DNN. However, in the case of a grid-specific method, the use of sector localization could prove as a useful step to improve the distance estimation accuracy as it will be shown in the next chapter.

**5.2.3.1 Branch identification**

The influence of the following parameters is analyzed:

1. *Fault resistance*: As with every other case study, Fig. 5.3a and Fig. 5.3e show that with the increase of fault resistance the accuracy decreases. More

specifically, from the best case scenario with very low fault resistance value (between 0.1 and 1  $\Omega$ ) and low  $SF$  ( $SF = 0.3$ ) with an accuracy of 97.4 %, the accuracy drops to 60.9 % for extra high fault resistance values (between 500 and 1000  $\Omega$ ) and very high  $SF$  ( $SF = 0.8$ ).

2. *Voltage measurements:* As stated before, an analysis of which type of voltages are more suitable was performed. In Fig. 5.3a, the obtained results are presented between phase measurements and symmetrical components for a DNN of three layers and a  $SF$  of 0.5. The symmetrical components proved to be more accurate by an average of 4.1 % (presented by the avg. lines in Fig. 5.3a). For that reason, for the following analysis only the symmetrical components were used.
3. *Number of layers:* The advantage of deep learning is demonstrated in Fig. 5.3b. By increasing the number of layers, a better accuracy in identifying the faulty branch is achieved. This is more noticeable for an increase from one to two layers than from two to three or three to four. An important observation is that the increased number of layers increases the accuracy of the DNN for higher fault resistances, as shown in Fig. 5.3c, i.e. for fault cases where it is more difficult to draw conclusions of whether a branch is under fault or not. However, a number of layers higher than four, might provide worse results due to overfitting. Overall, for three layers and  $SF = 0.5$ , the DNN presented an average branch identification accuracy of 84.6 %
4. *Fault type:* With an average accuracy over all the studied cases of 86 %, 85.2 % and 85.8 % for phase A, B and C to ground faults respectively (Fig. 5.3d), the DNN was an homogeneous response to single phase faults regardless of the unbalanced nature of the studied grid. On the other hand, for three phase faults, the DNN presents an accuracy of 91.2 % in identifying the faulty branch. The improvement of 5.5 % in the case of three phase faults is attributed to the use of symmetrical components and more specifically to the information the negative and zero sequence carry compared to the phase measurements; during a fault, while the positive sequence component of the voltage is expected to drop, negative and zero components are going to rise.

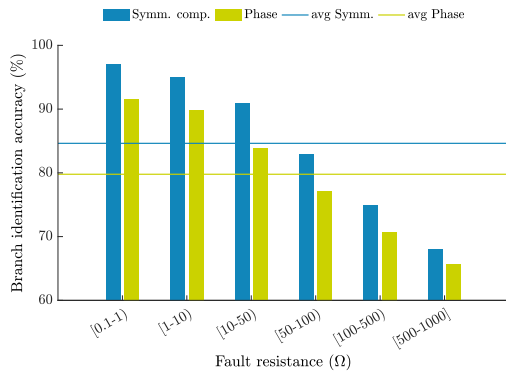
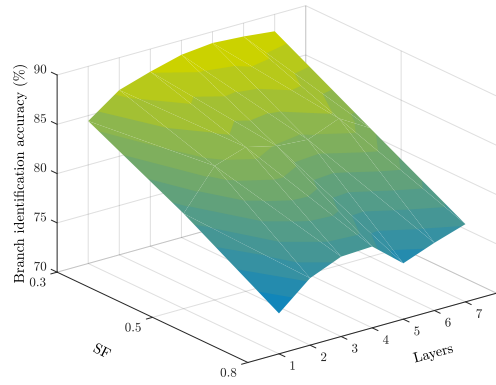
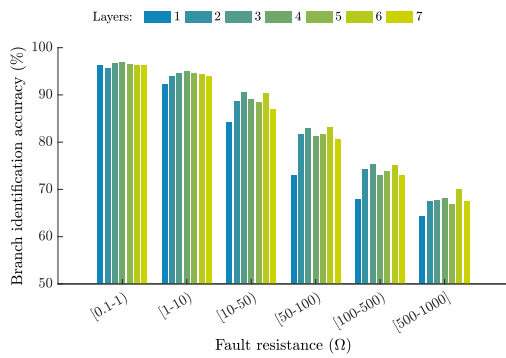
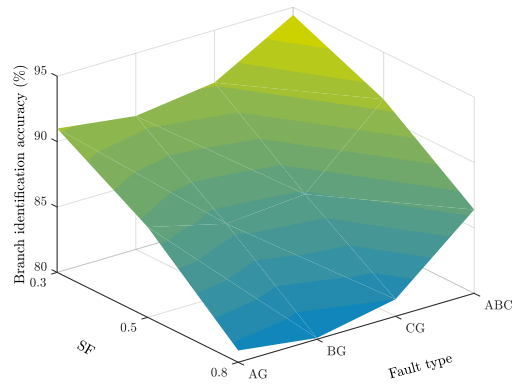
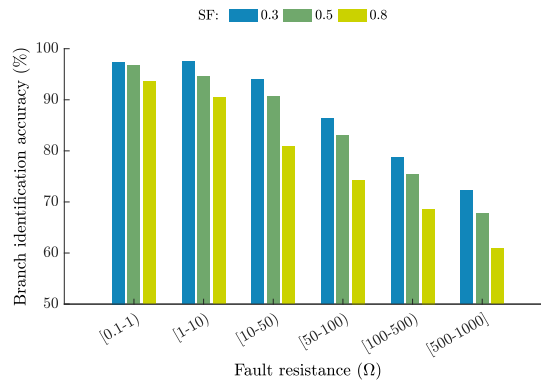
(a)  $SF = 0.5$  and 3 layers.(b) Average accuracy per layer and  $SF$ .(c) Different number of layers and  $SF = 0.5$ (d) Average accuracy per fault type and  $SF$ .(e) Different  $SF$  and 3 layers.

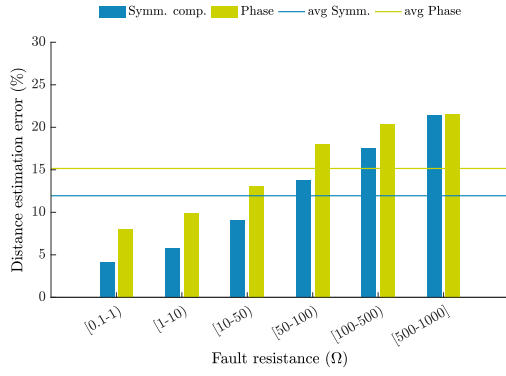
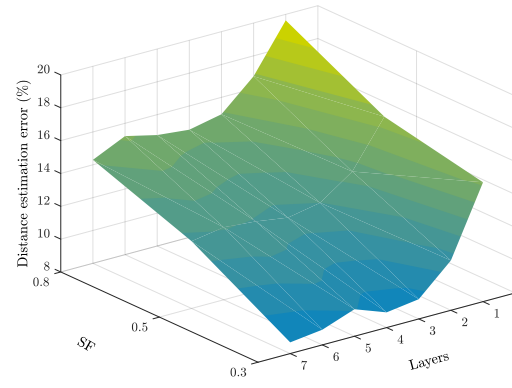
Figure 5.3 Branch identification accuracy.

5. *Simultaneity factor*: The results of the simultaneity factor analysis are presented in Fig. 5.3b, 5.3d and 5.3e. The higher the number of active loads in the grid, the more difficult it becomes to identify the faulty branch as another branch might exist with a higher voltage drop due to the presence of loads. This phenomenon is amplified with the increase of the fault resistance.

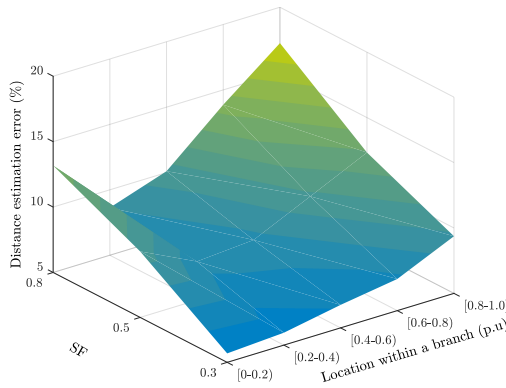
### 5.2.3.2 Distance estimation

The same metric as in Chapter 3, was used to measure the error in distance estimation accuracy (3.4). The influence of the following parameters is analyzed:

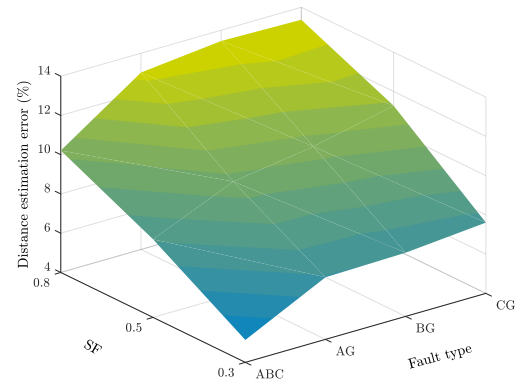
1. *Fault resistance*: Once again the effect of the fault resistance on the method is validated: as the fault resistance increases so does the estimation error (Fig. 5.4). For the analysis that follows, the symmetrical components were used.
2. *Voltage measurements*: The choice of symmetrical components is once more proven to be a better choice presenting a better accuracy by 3.21 % as presented in Fig. 5.4a.
3. *Number of layers*: In Fig. 5.4b, the advantage of deep learning through the use of multiple layers is once more demonstrated. The distance estimation error gets reduced by an average of 4.6 % when the number of layers is increased from one to four. A slight error increase is noticed again for a higher number of layers (5, 6 and 7) which is again attributed to unavoidable overfitting of the DNN.
4. *Fault location*: An average error increase of 4.37 % is noticed in distance estimation for faults located at the end of the feeder compared to those located in the beginning or the middle of the feeder as presented in Fig. 5.4c.
5. *Fault type*: With an average estimation error over all the studied cases of 9.98 %, 10.35 % and 10.75 % for phase A, B and C to ground faults respectively (Fig. 5.4d), the DNN exhibits an homogeneous response to single phase faults regardless of the unbalanced nature of the studied grid, as already noted during the branch identification step. On the other hand, for three phase faults, the DNN presents a distance estimation error of 7.78 %. The improvement of 2.58 %, on average, in the case of three phase faults is again attributed to the use of symmetrical components and the extra information that the transformation itself offers.
6. *Simultaneity factor*: Similarly to the branch identification task, the average estimation error increases with the increase of the  $SF$  as shown in Fig. 5.4b, 5.4c, 5.4d and 5.5.

(a)  $SF = 0.5$  and 4 layers.

(b) Number of layers.



(c) Fault location within a branch.



(d) Fault type

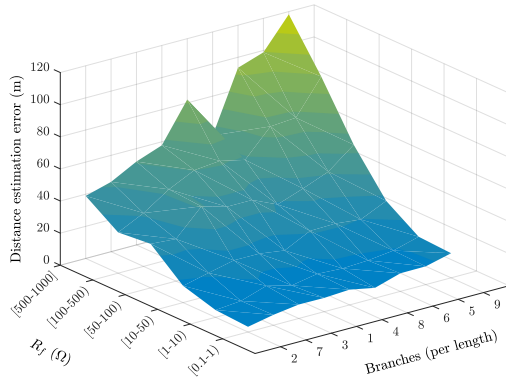
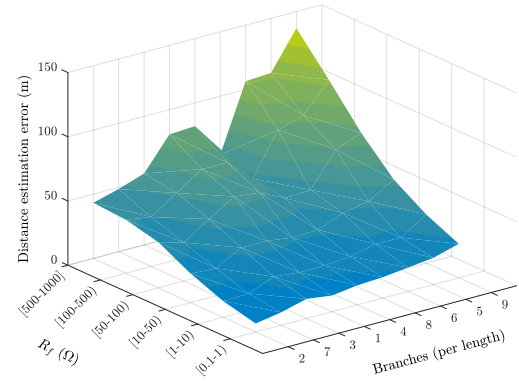
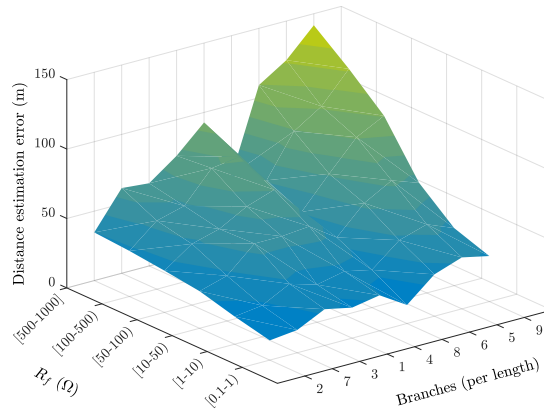
(e)  $SF = 0.3$  and 4 layers(f)  $SF = 0.5$  and 4 layers(g)  $SF = 0.8$  and 4 layers

Figure 5.4 Distance estimation error.

7. *Time of the day:* Under a per hour analysis of the distance estimation, more precise conclusions about the effect of generation and load profiles in the distance estimation can be extracted. In Fig. 5.5 three regions are identified: I) low load demand and low PV penetration, II) medium load demand and high PV penetration and III) high load demand and low PV penetration. The more elements of the grid are activated the more difficult it becomes to locate a fault. Region (I) is the one where the minimum average errors of all  $SF$ s are encountered since load and generation also present a minimum level of penetration. As the load demand increases so does the estimation error. In Fig. 5.5, as we move from region (I) to region (II) and (III), the estimation error increases. Region (II) presents special interest because of the increased penetration of PV units. There, the presence of PV generation is partially mitigating the error increase that comes with the increased load demand. In general, however, the effect of load increase is deemed more serious than the increase of PV generation since all three of the  $SF$  curves follow the same trend with that of the curve of the load profile.
8. *Grid heterogeneity:* The effect of the length of each branch in combination with the fault resistance on the estimation of the fault distance is presented in Fig. 5.4e, 5.4f, 5.4g for an  $SF$  of 0.3, 0.5 and 0.8 respectively. Branches are sorted in an ascending order from the shortest, 2 (185  $m$ ), to the longest, 9 (640  $m$ ). For very low and low fault resistances the distance estimation error does not exceed the 38  $m$ . Additionally, with the increase of the fault resistance the effect of the heterogeneity of the grid is noticeable. Longer branches demonstrate bigger errors with a maximum of 137  $m$  (in the case of an  $SF = 0.5$  for example) for branch nine (the longest).
9. *Grid imbalance:* An exception to the rule that wants the distance estimation error to increase with the length of the branch, is branch eight where the dip in the 3D plots of Fig. 5.4e, 5.4f, 5.4g is noticed. This is attributed to the particularity of branch eight where almost all the loads and PVs are connected in its terminal node (Fig. 3.6) thus reducing the effect of imbalance due to the topological distribution of loads and PVs. Hence, although the method is immune to the per phase distribution of loads and PVs, their position in the grid plays a significant role. If all the connected

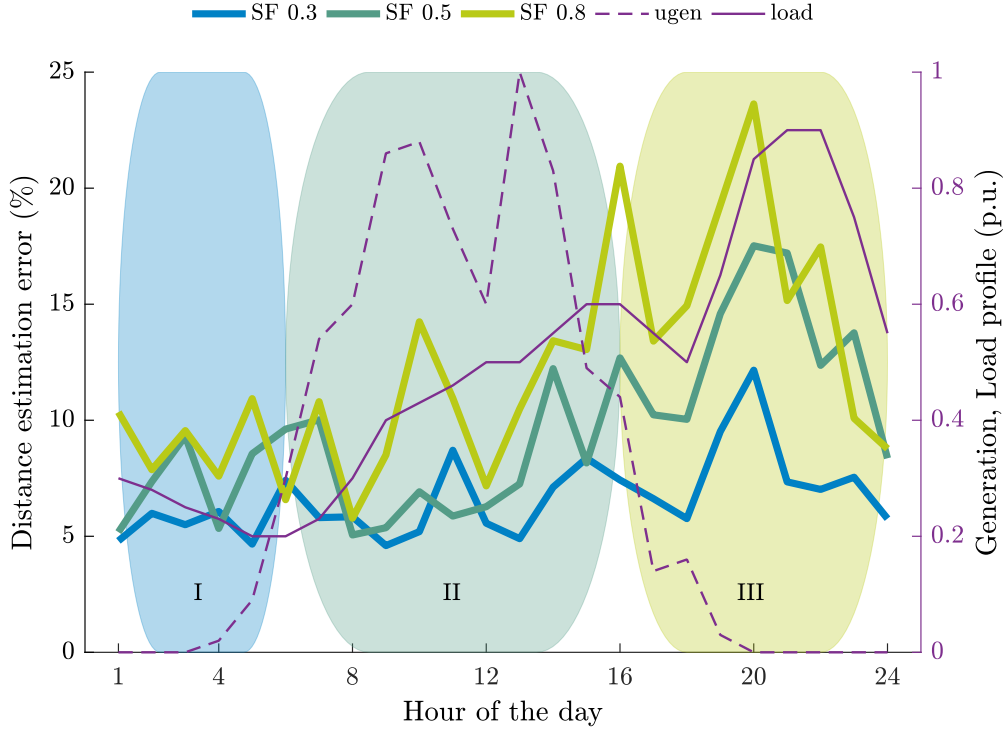


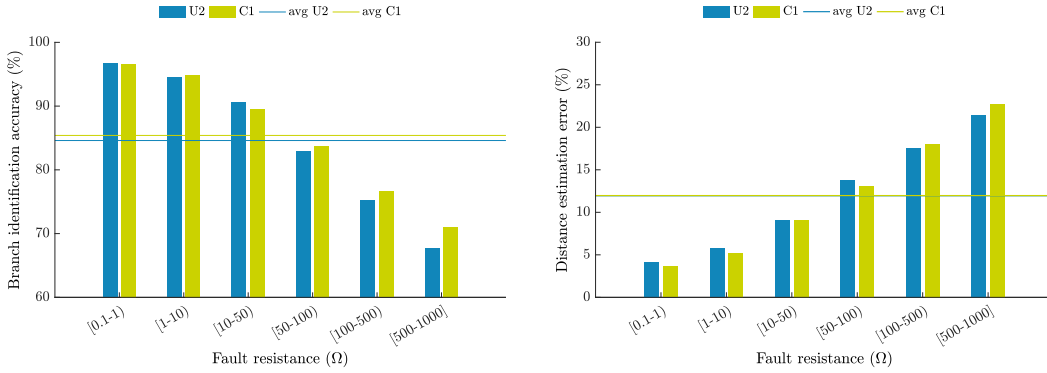
Figure 5.5 Mean absolute percentage error for distance estimation for different  $SF$  throughout the day. The main generation and load profiles are provided in the secondary y axis. Three distinct regions (I, II and III) of different performance are marked. The dominant effect of the load demand over the PV generation is shown as the  $SF$  curves follow the trend of the load curve and not so much that of generation.

elements in one branch are concentrated in one or two nodes, which is the case of branch eight, then the distance estimation process is simplified substantially.

#### 5.2.4 Measurement uncertainty

Two types of errors in measurements were considered: a) an error of 2% (between 0.98 and 1 p.u.), for each phase measurement for both current and voltage and b) a class 1 accuracy of the measurements (values between 0.99 and 1.01 p.u.).

The different error types: 2% underestimation and class 1 accuracy in the voltage and current measurements did not seem to affect the performance of the DNN as the observed differences were negligible; on average (avg lines in Fig. 5.6a) 0.77% for the branch identification task and 0.03% for the distance estimation (Fig. 5.6b).



(a) Branch identification for  $SF = 0.5$  and 3 layers. (b) Distance estimation for  $SF = 0.5$  and 4 layers.

Figure 5.6 Class 1 accuracy vs. 2% underestimation error in measurements for branch identification (a) and the distance estimation method (b).

### 5.2.5 Less available measurements

The two extreme scenarios were considered in this study regarding the available voltage measurements: a) available in every node and b) only two available for each branch, one at the transformer level (node 1 in Fig. 3.6) and one at each terminal node. In the second case, with the bare minimum of available measurements, a reduction of 70% in the available measurements is achieved.

For this experiment, a DNN of three layers was used for the branch identification task and of four layers for the distance estimation task (best options according to previous analysis). For an  $SF = 0.5$ , an average decrease of 4.58% of the DNN accuracy was observed in identifying the faulty branch and a decrease of 4.37% in estimating the fault distance as shown in Fig. 5.7a and 5.7b. An interesting observation is that the accuracy of the DNN during the limited measurements case study, was very similar to that with full availability of measurements for very low and extra high values of fault resistance in the branch identification task. The case of very low fault resistance is considered the easiest to treat while the one of extra high values is the most difficult one and the DNN has a hard time identifying the faulty branch either way.

The above results are considered extremely encouraging since they render DNN as an implementable solution for many LV distribution grids of today. Using only the bare minimum of voltage and current measurements with a decrease of only 4.5% in their accuracy, the biggest problem being the absence of measurement devices in the grid due to their increased cost, is tackled.



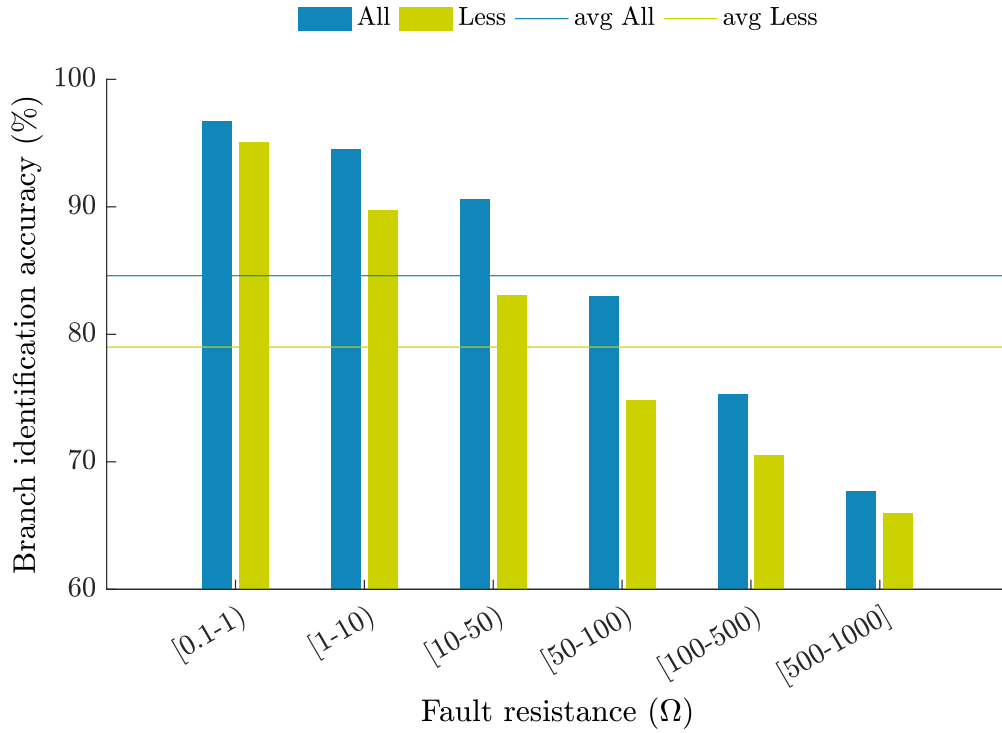
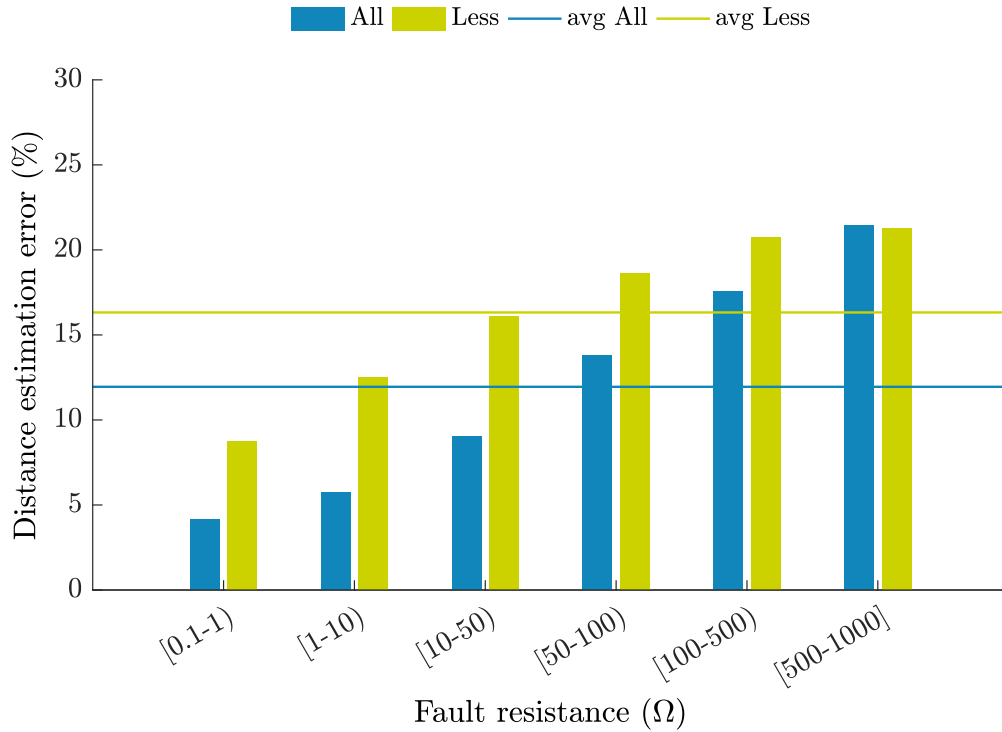
(a) Branch identification for  $SF = 0.5$  and 3 layers.(b) Distance estimation for  $SF = 0.5$  and 4 layers.

Figure 5.7 Less available measurements effect on the faulty branch identification (a) and the distance estimation method (b).

## 5.3 Conclusion

In this chapter, *deep neural networks* (*DNN*) were proposed as a solution to the fault detection and location processes in *low voltage* (*LV*) smart distribution grids. To test the accuracy of the DNN, its performance was evaluated in a simulated real LV distribution grid of Portugal. The DNN was designed to: a) detect the fault occurrence and identify the feeder under fault, b) identify the faulty branch within a faulty feeder and c) estimate the fault distance from the beginning of the feeder. Branch-specific parameters were excluded from the input features so that the method would be applicable to other grid topologies as well.

The case study was further extended by: a) including a bigger variety of fault resistance values reaching the number of 63,772 values between 1 and 1000  $\Omega$ , b) testing different simultaneity factors (0.3, 0.5 and 0.8), c) covering a big spectrum of PV generation and load demand scenarios with 70,334 studied combinations, d) studying the effect of different types of measurement errors (2% underestimation error and class 1 accuracy) and e) exploring the possibility of limiting the available voltage and current measurements.

The DNN excels in detecting a fault occurrence with an accuracy of 100%. For the branch identification task, an average accuracy of 84.6% is achieved while for the distance estimation task the error does not exceed on average the 12%. Moreover, the use of symmetrical components is more efficient than using phase voltage measurements. Furthermore, the increase of fault resistance hinders the fault location process in general. Another important remark is that the DNN is immune to the per phase distribution of loads and PV generation but not to their topological distribution which can affect the fault location process. Moreover, for the extreme case of only two available measurements per branch (one common for all at the substation level and one at each terminal node), the method accuracy is decreased by only 4.58% for branch identification and 4.37% for distance estimation. Finally, being a knowledge based method and although for fault detection it might not encounter many difficulties, for a reliable fault location estimation, a retraining of the DNN is advised in case the topology or the operational mode of the grid change.

In general, the results demonstrated in this chapter show that DNN is a reliable solution for fault detection and location in LV grids with high accuracy even under high resistance faults.



# Chapter 6

## Comparative analysis

In this chapter, a comparative analysis is being performed between the three proposed methods under a common case study. Moreover, four methods from the literature are used to compare the accuracy of the methods of this thesis, two for the branch identification task [79, 251] and two for the fault distance estimation task [93, 90]. However, as it will be explained later, due to the absence of other similar studies in *low voltage (LV)* distribution grids, the employed methods are based on *medium voltage (MV)* grids and therefore the comparison is not exactly fair since the LV grid is a more complex case study. As noted in Chapter 2, only very limited studies of fault location have been performed in the LV grid case and most of them are conceptual proposals not offering numerical results of distance estimation accuracy to make a direct comparison with the methods of this thesis.

## 6.1 Common case study

The case study presented in the DNN chapter was used as a basis to compare the three different methods among themselves and to other studies in the literature. More specifically:

1. *Fault resistance*: The same dataset of 63,772 of fault resistance values between 0.1 and 1000  $\Omega$  was used.
2. *Voltage measurement*: The positive, negative and zero components of the measured voltage were used.
3. *Fault location*: The nine possible locations of a fault within a sector, that were used before were considered again considered here at a distance of 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 % and 90 % from the beginning of each sector.
4. *Fault types*: Both single phase to ground and three phase faults were again considered.
5. *Simultaneity factor*: The  $SF$  of 0.5, being neither the highest nor the lowest of the studied values, was used for the comparative analysis.
6. *Time of the day*: The same 70,334 combinations of PV generation and load created during the DNN case study were used again here.
7. *Measurement noise*: The 2 % of underestimation error (values from 0.98 to 1 p.u.) was selected as the noise type to be introduced to each phase measurement independently.

## 6.2 Fault detection

While the criteria for fault detection presented in Chapter 3 did not prove reliable for high fault resistance values, the use of artificial intelligence tools such as the *gradient boosting trees (GBT)* and the *deep neural networks (DNN)*, solved this problem demonstrating excellent accuracy (100 % for the DNN).

## 6.3 Fault location

### 6.3.1 Branch identification

Due to a lack of available research papers for LV distribution grids, two methods designed for MV distribution grids [79, 251] were employed to compare the proposed methods of the three previous chapters. These references were used to compare the faulty branch identification results.

In the first case [79], the authors developed a general fault location method based on voltage and current measurements at the *point of common coupling (PCC)* of *distributed generators (DG)*. They considered all the different types of faults, i.e. single phase to ground, double phase to ground, phase to phase and three phase. Moreover, they studied faults in three possible locations within a sector at distances of 5 %, 50 % and 95 % from the beginning of each sector. However, the maximum fault resistance value for phase to phase and three phase faults was 5  $\Omega$  and 50  $\Omega$  for the rest.

In the second study that was used as reference [251], the authors developed a method based on *phase measurement unit (PMU)* real time state estimation that detects faults and identifies faulted lines. The authors considered single phase to ground, double phase to ground and three phase faults. Furthermore, they considered only two possible fault locations within a faulty sector, at the middle of the line and at a distance equal to 25 % of the sector's length. Although they investigated high impedance faults of up to 1000  $\Omega$ , the data they presented for such high fault resistances were applicable only to a single fault case of an unearthed neutral. For the rest of the cases, the maximum fault resistance they tested was that of 100  $\Omega$ .

All these information are gathered in Table 6.1 where the considerably bigger number of fault scenarios that were considered in this study is demonstrated.

Figure 6.1 presents the performance of each of the three methods developed in this work, the conventional one, the GBT and the DNN. Each one is proved to be better than the previous with the difference in their accuracy being more significant with the increase of the fault resistance. DNN outperform the conventional method by 4.01 % for very low fault resistances and by 16.96 % for extra high ones. DNN present an accuracy over 90 % in detecting the faulty branch for fault resistances until 50  $\Omega$ . For extra high fault resistances, their

Table 6.1 Comparison of studied parameters between branch identification methods.

Parameters	Brahma [79]	Pignati [251]	This study
<b>Grid</b>	12.4 kV (MV), U.S.A.	10 kV (MV), The Netherlands	400 V (LV), Portugal
<b>Fault types</b>	1ph-G, 2ph-G, ph-ph, 3ph	1ph-G, 2ph-G, 3ph	1ph-G, 3ph
<b>Fault resistance (<math>\Omega</math>)</b>	1-5 (ph-ph, 3ph), 1-50 (1ph-G, 2ph-G)	1, 100, 1000 <sup>a</sup>	0.1-1000 (63,772 different values)
<b>Fault location within the sector</b>	0.05, 0.5, 0.95	0.25, 0.5	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9
<b>Measurements, inputs</b>	synchronized $I, V$ at PCC of DG	$I, V$ from PMUs on every node	rms $V$ at every node and rms $I$ at each feeder
<b>Noise in measurements</b>	-	0.016 % for voltage, 1.2 % for current	2 % underestimation

<sup>a</sup>1000  $\Omega$  only for one case

accuracy decreases to a bit less than 70 %. This last observation increases the reliability of DNN as they maintain decent levels of accuracy even under the most challenging scenarios. Another advantage of the DNN is that, if more datapoints were to be created, the accuracy of the DNN would be improved to an extend as the chances of overfitting would be further reduced. Regarding the comparison with other studies of the literature, DNN seem to match and outperform the Pignati [251] method while they are a bit less accurate than Brahma [79]. However, it needs to be underlined that this is not a fair comparison since these methods were tested against considerably less influencing parameters

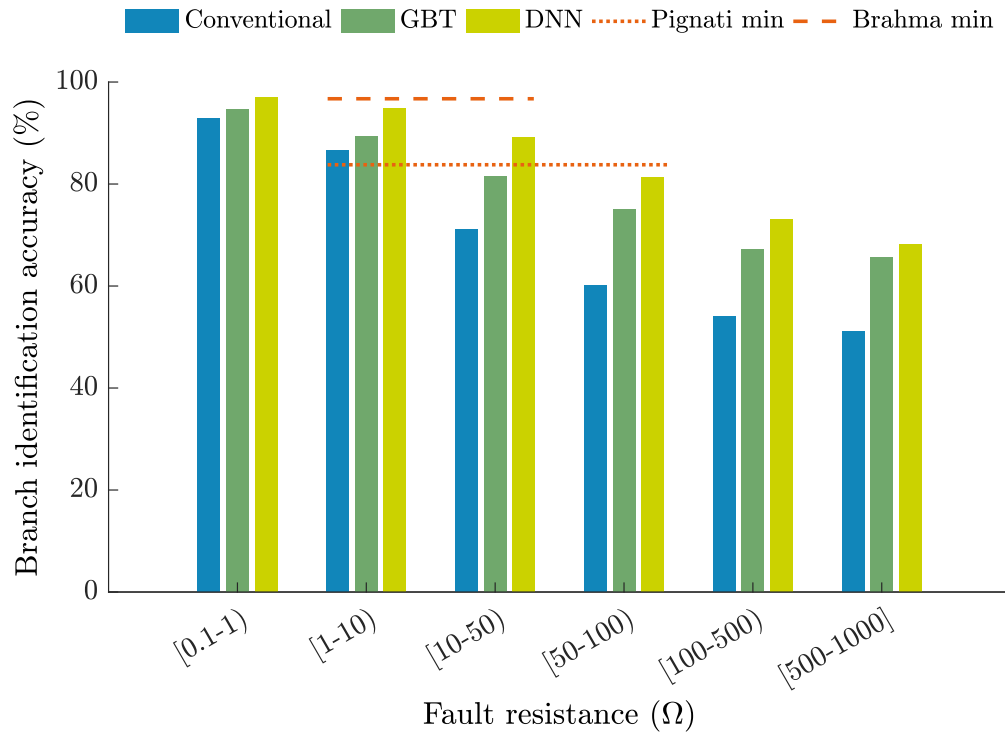


Figure 6.1 Branch identification accuracy per fault resistance values for the three developed methods compared to two from the literature.

and were designed for MV grids. It is probable that if applied to a LV grid, their accuracy would decrease.

### 6.3.2 Sector localization

The step of sector localization is the most crucial one of the conventional method. Since the following step, distance estimation, is based on a graphic method, localizing the correct faulty sector is of paramount importance to the method. However, it is a step that could, and was, omitted during the DNN design choices. In the case of the DNN for inter-branch consistency, uniformity of the dataset and generalization reasons, five virtual sectors were created in each branch using linear interpolation from the available measurement points. Figure 6.2 presents the accuracy of the conventional method in localizing a real sector as these were originally defined (line connecting two adjacent measurement points - in the ideal case adjacent nodes) or a virtual sector (interpolated from branch voltage measurements). The accuracy of both is compared to the accuracy of the DNN in localizing which of the five sectors is under fault. It is obvious that the DNN outperforms the conventional method. However, the accuracy of



the DNN remains low still, with an average accuracy of 45.43%. It should be underlined that the DNN, as noted in Chapter 5, was not designed to localize a faulty sector as it is capable of bypassing this step. If this functionality was included in the training of the DNN, its accuracy would increase if the sector localization task was to demonstrate a high accuracy itself. Finally, it is shown that the conventional method fails to localize the faulty sector in the case of virtual sectors.

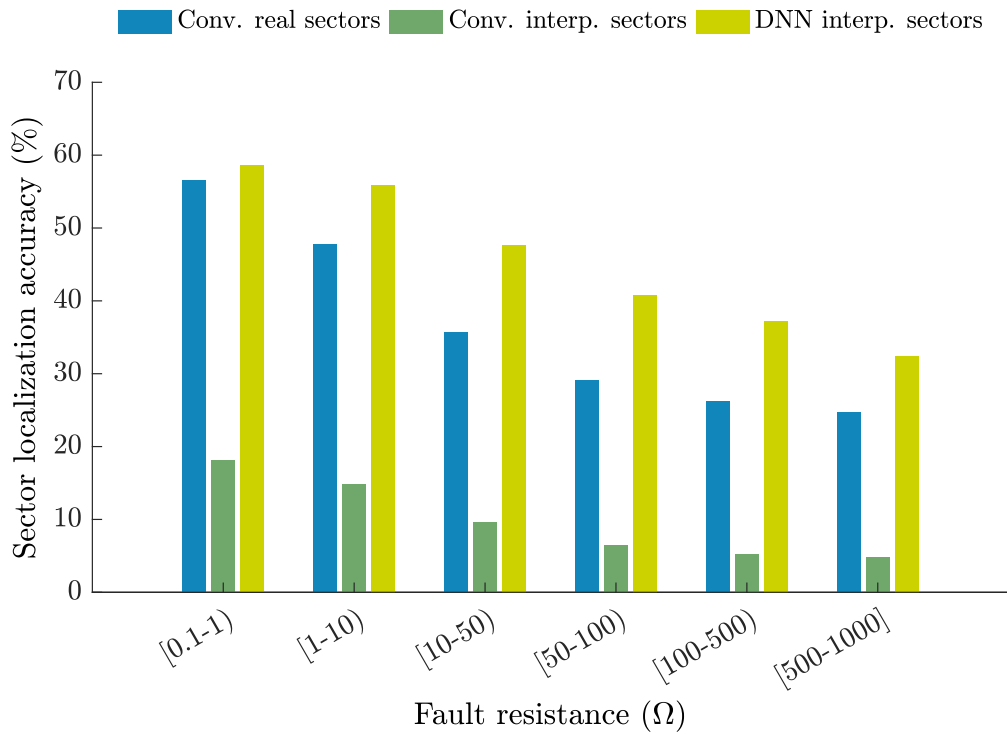


Figure 6.2 Faulty sector localization accuracy per fault resistance and method.

### 6.3.3 Distance estimation

As with the branch identification case, three more studies were employed from the literature to compare the conventional and the DNN methods with them for the fault distance estimation task. Two of the cited studies concern once again MV distribution grids [90, 93] while the third one studies arc faults in LV grids [135].

In the first case [93], the authors used NN in combination with *support vector machines (SVM)*, proposing two fault location schemes using measurements of current and voltage at the substation level; their method also requires a boolean

input of the status of all the available circuit breakers and relays. Moreover, they considered all the different types of faults and they chose three possible fault locations within each branch at 20 %, 50 % and 80 % from the beginning of the feeder. However, the studied fault resistance did not exceed 100  $\Omega$  and no noise was considered in the measurements.

In the second study that was used as a reference [90], the authors developed a fault locator based on NN. Their method requires current and voltage input from a single point in the distribution line. They too, considered all the possible fault types. Furthermore, they chose multiple fault locations within the single feeder they studied, one every 2.5 *km* with a maximum line length of 40 *km*. Similarly to the previous case, the maximum studied fault resistance was once again set at 100  $\Omega$

All these information are gathered in Table 6.2 where the considerably bigger number of fault scenarios that were considered in this study is demonstrated.

Finally, in [135], a single-end impedance-based method based on time-domain formulation, was proposed for LV grids and arc faults. The maximum arc resistance was 6  $\Omega$ .

Figure 6.3 presents the performance of each of the two methods developed in this work, the conventional one (with and without using the sector correction criterion) and the DNN. Each one is proved to be better than the previous one. It is clear that the average error of the DNN method is way below from the maximum estimation error of the Thukaram method and slightly lower than maximum error of the Aslan method. In the Thukaram method [93], the maximum studied fault resistance was 100  $\Omega$  and the fault location accuracy was 300 *m* for Scheme I and 55 *m* for Scheme II. At the same time, for the Aslan method [90] the maximum distance estimation error was 90 *m* for fault resistances between 2 and 100  $\Omega$ . Finally, in [135], the maximum error of the algorithm for 6  $\Omega$  of fault resistance was 23 *m*.

With a fault distance estimation error of less than 38, 78 and 137 *m* for fault resistances under 10, 100 and 1000  $\Omega$  respectively for an  $SF = 0.5$ , and with the average estimation errors for the same fault resistance margins and  $SF$  being 18, 30 and 44 *m* respectively, the DNN method is considered superior to the considered references taking into consideration the increased complexity of the case study and the number of influencing parameters.

Table 6.2 Comparison of studied parameters between fault location methods.

Parameters	Thukaram [93]	Aslan [90]	This study
<b>Grid</b>	11 kV (MV), 52-bus distribution system	34.5 kV (MV), Simplified distribution feeder	400 V (LV), Portugal
<b>Fault types</b>	1ph-G, 2ph-G, ph-ph, 3ph	1ph-G, 2ph-G, ph-ph, 3ph	1ph-G, 3ph
<b>Fault resistance (<math>\Omega</math>)</b>	50, 60, 70, 80, 90, 100	2, 5, 10, 15, 20, 30, 40, 50, 60, 80, 100	0.1-1000 (63,772 different values)
<b>Fault location</b>	0.2, 0.5, 0.8 (within each branch)	every 2.5 <i>km</i> (total length of 40 <i>km</i> )	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 (within each sector)
<b>Measurements, inputs</b>	$I, V$ at substation, status of circuit breakers, relays	$I, V$ at one end of the line	rms $V$ at every node and rms $I$ at each feede
<b>Noise in measurements</b>	-	-	2 % underestimation

Finally, Fig. 6.4 shows the advantage of knowing the faulty sector in the fault distance estimation method accuracy. For the conventional method, limiting the estimation to the limits of the faulty sector reduces the estimation error by 6.42 % on average (activation of the sector correction criterion). At the same time, the DNN accuracy in distance estimation if the fault is located in a correctly identified sector is improved by 6.91 %. Hence, it is obvious that the information of the faulty sector is a powerful tool for the distance estimation task. The problem is, as stated in the previous section, the accuracy of the sector localization methods.

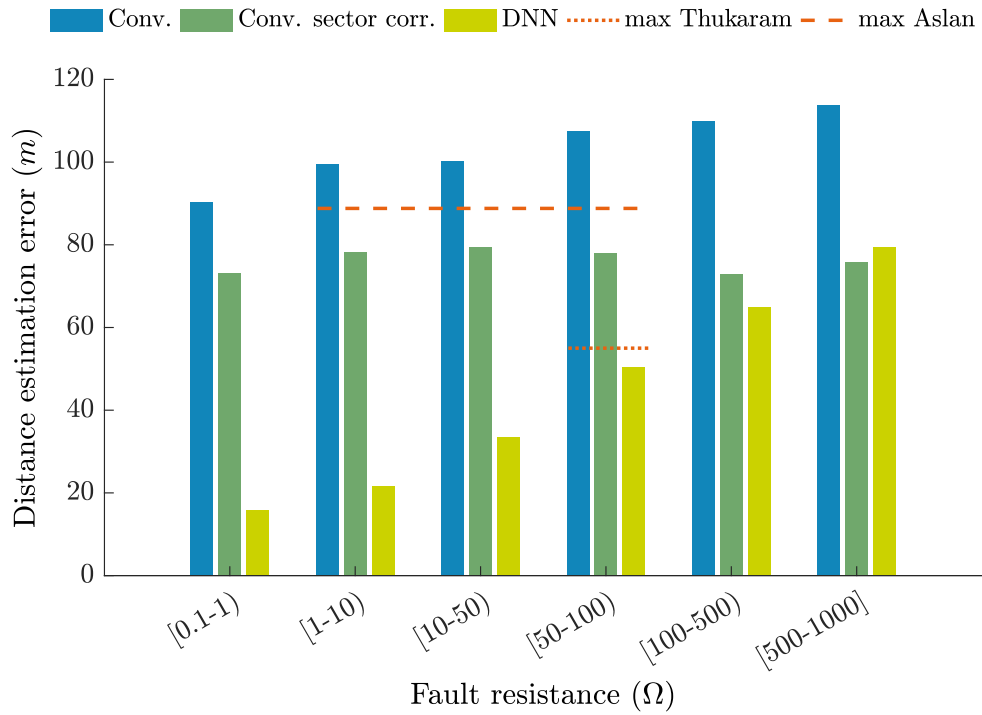


Figure 6.3 Fault distance estimation error in meters per fault resistance values for the two developed methods compared to two from the literature.

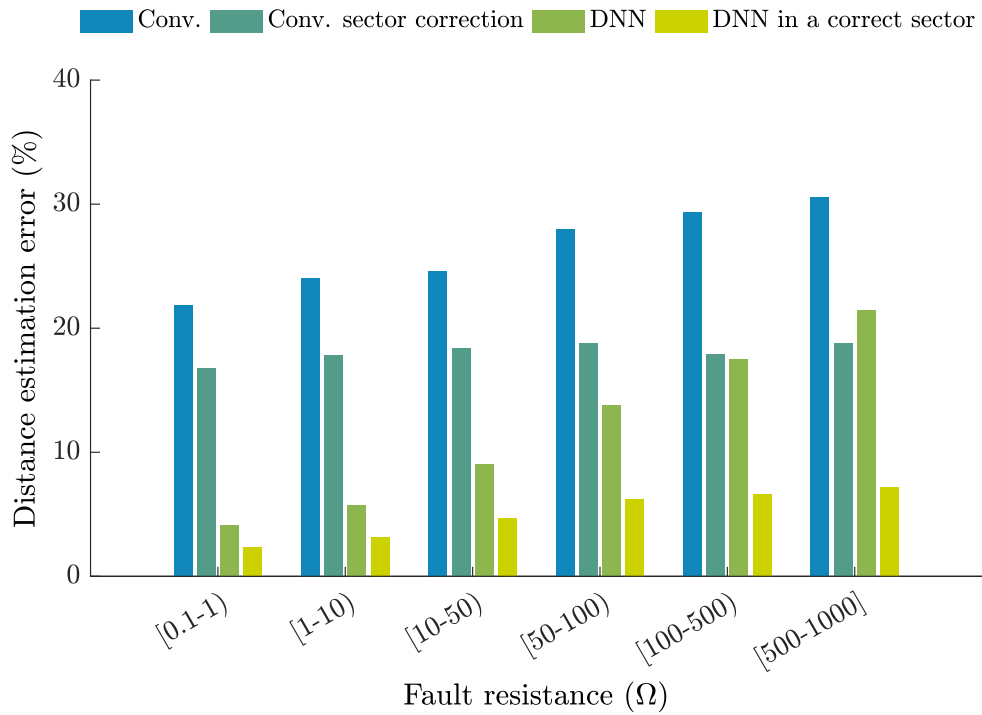


Figure 6.4 Fault distance estimation error per fault resistance values for the conventional method and the DNN method (in general and within a correctly localized sector).

## 6.4 Conclusion

In this chapter, a common case study, was used to compare the three proposed methods of this first part among themselves but also with some representative studies from the literature. Due to the absence of studies in *low voltage (LV)* distribution grids, studies from the *medium voltage (MV)* distribution grids were employed. Although, as explained in the Introduction, the LV grid is a more complex system than the MV one the *deep neural networks (DNN)* seem to either outperform or in the worst case scenario match the performance of the other existing methods in the literature.

The case study included: a) a big variety of fault resistance values reaching the number of 63,772 values between 1 and 1000  $\Omega$ , b) nine different fault locations within each sector, c) two fault types (single phase to ground and three phase faults), d) a simultaneity factor of 0.5, e) a big spectrum of PV generation and load demand scenarios with 70,334 studied combinations and f) a 2 % underestimation error in measurements.

Regarding fault detection, with an accuracy of 100 % using DNN there is a guarantee that a fault occurrence will not remain undetected, something that could cause a lot of problems in the grid.

For the fault location steps: a) branch identification, b) sector localization and c) distance estimation the following comparative conclusions were drawn. First of all, DNN is found to be the best choice among the developed methods for the faulty branch identification task. With regards to other studies in the literature [251, 79], the DNN outperformed in some cases the existing methods [251] or matched their accuracy [79].

Moreover, the importance of a reliable sector localization method is underlined as it can lead to a decrease of the distance estimation error by an average of 6.5 % in both the conventional and the DNN methods. However, both methods present an average accuracy of less than 50 % in localizing the faulty sector.

Furthermore, in the distance estimation analysis the use of DNN proved again to be the most reliable solution. With a distance estimation error that on average does not exceed 12 % it outperformed both the conventional and other methods of the literature.

In general, the results demonstrated in this chapter show a clear superiority of the DNN with regards to other methods.

## Part II

# Fault detection and isolation in distributed generators



## Chapter 7

# Fault detection and isolation in grid-connected photovoltaic systems

As indicated in the Introduction, *grid-connected photovoltaics (GCPVs)* is a rapidly growing field. Their monitoring to ensure safe and reliable operation is thus of paramount importance. In this chapter, two algorithms are proposed for fault diagnosis of faults occurring both in the dc and the ac side through the monitoring of the voltage and current at the output of the inverter. A total of nineteen fault cases are investigated covering faults inside the PV array, in the dc bus, the power electronic devices and the grid side of the system. Furthermore, the isolation speed of the algorithms is presented and its importance is analyzed. Moreover, a sensitivity analysis is also presented against different irradiance levels (1000 to 200  $W/m^2$ ) and erroneous measurements. Finally, the results are validated through simulations on a 43.2 kW GCPV.



## 7.1 Method description

At this point, before passing on to the fault detection process, it should be underlined that the single-fault assumption was made. In other words, it was assumed that only one fault was occurring at a time.

### 7.1.1 Fault detection

The first step of every *fault detection and isolation (FDI)* method is to detect the occurrence of a fault. As stated in the Introduction, the main objective of this study is to determine whether or not a fault is occurring while using the minimum amount of measurements and more specifically through the monitoring of electrical variables at the output of the inverter. In order to complete the set of the monitored electrical variables, through the use of three phase current ( $I$ ) and voltage ( $V$ ) measurements, the active ( $P$ ) and reactive ( $Q$ ) power are calculated.

The main tool used to detect a change in the operation of a GCPV system is the monitoring of the output electrical power. A lower value of the active power than the one expected for a certain irradiation level is a first indicator of a fault occurrence. As it was mentioned in Chapter 2, the calculation of residuals and their check against thresholds is the basis of the comparative methods that compare the actual operating conditions with the ones from modeling and simulations. To achieve that however, knowledge of the actual irradiance is necessary. For that purpose, a sensor that measures irradiance at all times could be used. The best technology for that application is the calibrated PV reference device which can provide irradiance measurements with an accuracy of 97.6% [252]. As described in [252], the equivalent irradiance according to the IEC 60904-3 reference solar spectrum is used to compare the electrical response of the system with the one under the actual solar irradiation.

Another important parameter is the threshold crossing method and the configuration of its settings as it is the most common way to track an anomaly in fault detection theory. In order for an alarm signal to be created during a fault occurrence, certain thresholds need to be crossed. Let us assume that the parameter we want to track is the decrease to zero of the active power during a fault. Actually, it will never be exactly zero due to various factors one of which

could be noise in the measurement. As a result, a safety margin around this value is necessary in order to ensure an accurate detection of the fault and avoid false alarms or the fault going unnoticed. The decision making about the size of this safety margin is the process of configuring the threshold settings.

So, conceptually, under the assumption that a difference in the output active power of the PV system is related to the occurrence of the fault, an alarm signal is created and triggers the initiation of the fault isolation algorithm.

Furthermore, a crucial element in the fault detection process is the evolution of faults in time from the moment of their occurrence until the point where they are established in the system; what is also called steady-state of a fault. The different fault types presented in Chapter 2 can be grouped in two major categories: a) faults on the dc side of the plant and b) faults on the ac side. The occurrence of a fault is generally divided in three periods of time: the sub-transient, the transient and the steady-state. For the ac faults specifically, the sub-transient period is defined as the first cycle (period) after the fault occurrence and the transient period as the next five to ten cycles [253]. An example is given in Fig. 7.1 where the *short circuit (SC)* fault between two phases and the ground (on the ac side) is presented. The fault appears at  $t_0 = 1$  s. The sub-transient period of the fault begins with the fault occurrence at  $t_0$  and lasts until  $t_1$  with a duration of one cycle ( $T = 1/f = 1/50 = 20$  ms). The transient period begins at  $t_1$  and lasts until  $t_2$  -in this case 140 ms- where the fault current stabilizes in a final steady state. In this work, faults are studied close to their steady-state.

### 7.1.2 Fault isolation

The next step in an FDI method is the localization of the detected faults. This process is known as fault isolation and as mentioned before, it is the core of this study. The fault isolation, in fault diagnosis theory, is usually broken down to four following steps [231, 254]: 1) symptom generation, 2) fault signature table construction, 3) isolation algorithm design and 4) threshold crossing configuration.

For the first step (symptom generation), a list of symptoms is created based on some of the monitored variables according to the way those variables are affected by different types of faults. For example, a symptom can be an increase or a decrease of the current or the voltage after the occurrence of a fault. The objective in this case is to select the minimum number of symptoms that guarantee the

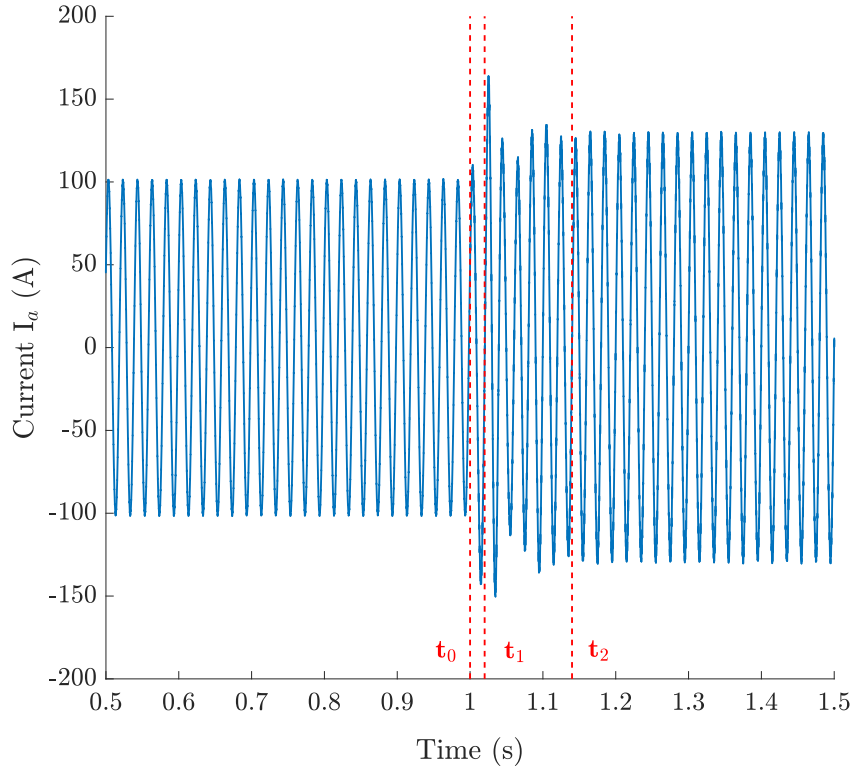


Figure 7.1 Evolution over time of the phase current  $I_a$  under a SC fault between phases A, B and the ground.

maximum number of isolable faults. For the GCPV system case, three families of symptoms were investigated in order to extract the maximum amount of necessary features from the monitored variables:

- a) voltage and current response
- b) harmonic content of the voltage
- c) symmetrical component analysis of the voltage

As it will be presented in the next section, the first two: (a) and (b), were used for the fault isolation *algorithm A* which covers only dc faults while (a) and (c) were used for the fault isolation *algorithm B* that included grid side faults as well.

Furthermore, the fault signature is defined as the set of alterations that symptoms are subjected to, in the presence of a single fault. Usually, a fault signature is a series of “0” and “1” where “1” indicates a change in the associated symptom while “0” the fact that symptom remains unaffected. In some cases however, more indices can be used to mark a specific alteration to a symptom as

we will see later in this chapter. All the fault signatures are gathered in a single table, the fault signature table, where the rows signify the different fault types and the columns the selected symptoms. For a fault to be completely isolable it would require a unique fault signature, meaning that each fault causes a unique combination of alterations to the complete set of symptoms. Following the simplified example above, a table full of “0” would indicate a normal operation of the system.

Once the two first steps are completed, the design of the isolation algorithm should be possible. Very often, multiple algorithms can be used based in a single fault signature table. This is the case of the current study as well, as presented in the next section, where one algorithm is proposed per case study but they certainly are not the only algorithms that can be derived from the signature tables.

Finally, as for the fault detection case, the configuration of the threshold settings is also necessary in the isolation algorithms. It is the factor that will define the regions of each symptom that are affected by a fault and will ensure the robustness of the final algorithm against uncertainties, noise or external disturbances. The configuration of the threshold settings is presented for each case in the following section.

## 7.2 Application

In this section, the case study is presented along with the complete list of all the studied faults. Based on those faults, a fault detection analysis is presented followed by two fault isolation methods; *algorithm A* studies faults on the dc side of the system while *algorithm B* takes into account faults on the ac side as well.

### 7.2.1 Case study

As explained in Chapter 2, the selected topology for the *grid-connected photovoltaic system (GCPV)* system is a three-phase direct connection to the grid through a dc-dc boost converter and an inverter as shown in Fig. 7.2. In Fig. 7.2 the considered locations of faults are also presented: faults in the PV array (1); faults in the power electronic devices (2); faults in the dc bus (3); faults in the grid side, close to the plant (4) and far away (5). The PV array is connected to grid with a three phase connection. The eight parameters of Bishop's model of the PV cells and the general characteristics of the system are provided in Table 7.1. At this point, it should be mentioned that the grid side was simulated as an ideal voltage source.

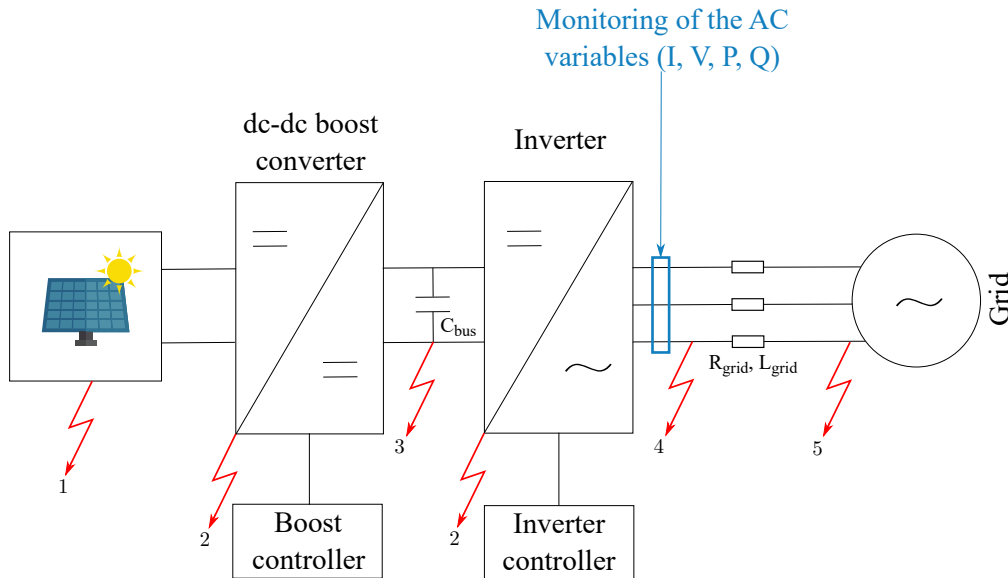


Figure 7.2 GCPV system schematic.

Table 7.1 PV system characteristics.

Parameter	Symbol	Value	Unit
Photo current	$I_{ph}$	4.838	A
Reverse diode saturation current	$I_o$	$10^{-7}$	A
Thermal voltage	$V_t$	0.0257	V
Series resistance	$R_s$	0.005	$\Omega$
Shunt resistance	$R_{sh}$	9	$\Omega$
Bishop's coefficient	$k$	0.1	—
Bishop's coefficient	$n$	3.4	—
Breakdown voltage	$V_b$	−30	V
Number of cells per module	$N_{cell}$	72	—
Number of modules per string	$N_{mod}$	10	—
Number of strings	$N_{str}$	30	—
Number of bypass diodes per module	$N_{bpd}$	4	—
Short circuit current	$I_{sc}$	145.3	A
Open circuit voltage	$V_{oc}$	407.67	V
Output power	$P_{pv}$	43.2	kW
Dc bus capacitance	$C_{bus}$	0.6	mF
Dc bus voltage	$V_{bus}$	800	V
Grid inductance	$L_{grid}$	5	mH
Grid resistance	$R_{grid}$	5	m $\Omega$
Grid frequency	$f_{grid}$	50	Hz
Grid voltage	$V_{LL_{grid_{rms}}}$	400	V

### 7.2.2 Studied fault types

As explained in Chapter 2, from the excessive list of possible faults that was briefly described in the previous section, the following nineteen were selected according to their severity and occurrence frequency:

a) faults inside the PV array:

[f01] shading of a number of PV modules

[f02] inverse bypass diode

[f03] short-circuited bypass diode

[f04] bypass diode breakdown

b) faults between the dc bus and the ground:

[f05] SC between the positive pole and the ground

[f06] SC between the negative pole and the ground

c) faults in the power electronic devices:

[f07] open-circuited converter IGBT

[f08] short-circuited converter IGBT

[f09] open-circuited inverter IGBT (1 IGBT in 1 leg)

[f10] open-circuited inverter leg (both IGBTs in 1 leg)

[f11] short-circuited inverter IGBT (1 IGBT in 1 leg)

[f12] short-circuited inverter leg (both IGBTs in 1 leg)

d) faults in the grid:

[f13] SC between one phase and the ground

[f14] SC between two phases and the ground

[f15] SC between two phases

[f16] SC between three phases

[f17] voltage sag in one phase

[f18] voltage sag in two phases

[f19] voltage sag in three phases

More specifically, in the shading fault case (01), two scenarios were examined: a) 50 % and b) 30 % of the surface of the first five modules in the first twenty out of a total of thirty strings was covered by shade; the shade was simulated by a proportional decrease of the photocurrent. Moreover, in the bypass diode breakdown case (f04), the diode was replaced by a resistance of 5  $\Omega$ . Furthermore, in all SC faults, the fault resistance was set to 1  $\Omega$ . Finally, voltage sags (f17-f19) were simulated as a 50 % sharp decrease of the initial voltage.

### 7.2.3 Fault detection

The values of the monitored variables ( $I, V, P, Q$ ) during *normal operation* (*NO*) of the system were compared to the values of the variables of *faulty operation* (*FO*). A difference between the two suggested the existence of a fault under the condition that the irradiance has not changed. The impact of the occurrence of

Table 7.2 Fault detectability table of the monitored electrical variables.

Faults	$\Delta I$	Changes in variables		$\Delta Q$
		$\Delta V$	$\Delta P$	
<b>f01</b>	1	0	1	1
<b>f02</b>	1	0	1	1
<b>f03</b>	1	0	1	1
<b>f04</b>	1	0	1	1
<b>f05</b>	1	1	1	1
<b>f06</b>	1	1	1	1
<b>f07</b>	1	1	1	1
<b>f08</b>	1	1	1	1
<b>f09</b>	1	1	1	1
<b>f10</b>	1	1	1	1
<b>f11</b>	1	1	1	1
<b>f12</b>	1	1	1	1
<b>f13</b>	1	1	1	1
<b>f14</b>	1	1	1	1
<b>f15</b>	1	1	1	1
<b>f16</b>	1	1	1	1
<b>f17</b>	1	1	1	1
<b>f18</b>	1	1	1	1
<b>f19</b>	1	1	1	1
<b>f20</b>	1	1	1	1

the different types of faults on the monitored electrical variables is presented in Table 7.2 where “1” signifies a change in the value of the variable after the fault occurrence while “0” the fact that the variable remains unaffected. An obvious observation from Table 7.2, is that three of the four variables were affected in every fault case: the current, the active and the reactive power. Hence, a change in the active power is a strong indicator of a fault occurrence given the fact that the irradiance has not changed and the maximum power point tracker is working properly.

#### 7.2.4 Fault isolation algorithm A: dc side faults

As a first step in the development of a fault isolation method for the whole system, only faults on the dc side (f01-f12) were considered. From the faults presented in Table 7.2, only fault categories can be discriminated and not individ-



ual fault cases. Following the four-step approach in developing a fault isolation method, as described in the previous section and in order to obtain a unique fault signature for each fault case, six symptoms were generated.

### **Step 1: Symptom generation**

First of all, the equality to zero of the sum of the phase currents was examined (s1). Secondly, the number of the phases where the current during faulty operation was zero, was also recorded (s2). Then the behavior of the current and the voltage in each phase was monitored and the results were compared to their rms values during normal operation. Thus, the indices of  $\Delta I = I_{FO} - I_{NO}$  and  $\Delta V = V_{FO} - V_{NO}$  were created. From those indices, two additional symptoms were created: the per phase increase or decrease of the current (s3) and the number of phases where the voltage decreases after the fault occurrence (s4). Finally, two more symptoms were generated by monitoring the frequency components of the measured voltage. The two indicators of a fault occurrence in this case were the absence of the 10 kHz (s5) and 50 Hz (s6) frequency components of the voltage. The 10 kHz and 50 Hz frequencies represent the inverter PWM frequency and the grid frequency respectively. The complete list of symptoms is provided below:

[s1] the sum of phase currents is equal to zero,  $\sum i = 0$

[s2] in how many of the three phases is  $I_{FO} = 0$

[s3] current increase  $\Delta I > 0$  or current decrease  $\Delta I < 0$

[s4] in how many of the three phases is  $\Delta V < 0$

[s5] absence of the 10 [kHz] frequency component of the voltage

[s6] absence of the 50 [Hz] frequency component of the voltage

### **Step 2: Fault signature table construction**

Based on these symptoms the fault signature table, Table 7.3, was created. For symptoms s1, s5 and s6 the “√” symbol is used to mark an affirmative answer to the posed question while “x” is used to mark a negative response. For s1 specifically, “+” and “−” indicate the sign of the sum of the phase currents.

Table 7.3 Fault signature table.

Faults	Symptoms						Isolable Faults
	s1	s2	s3	s4	s5	s6	
f01	✓	0	—	0	x	x	1
f02	✓	0	—	0	x	x	
f03	✓	0	—	0	x	x	
f04	✓	0	—	0	x	x	
f07	✓	3	—	3	x	x	2
f08	✓	3	—	3	x	x	
f05	—	0	—	3	✓	✓	3
f06	+	0	—	3	✓	✓	4
f09	✓	0	+—	0	x	x	5
f10	✓	1	+—	1	✓	x	6
f11	✓	0	+—	3	x	x	7
f12	✓	0	+—	3	✓	✓	8

Similarly, for s3, “—” means that the current during faulty operation had a smaller value than normal operation and “+—” that the current was in some phases greater (+) than the one in normal operation and smaller (—) in the rest. Finally, for symptoms s2 and s4 the numbers “0-3” signify the number of the affected phases.

A first conclusion from Table 7.3 is that 50 % of the faults are completely isolable (six cases) while the rest are identifiable as two different groups of faults. The first group of faults consists of all the faults inside the PV array (f01-f04), an expected result since the point of origin of the fault is far from the measurement point (at the output of the inverter). For the same reason, faults inside the

converter (f07, f08) can only be identified as a group of faults since they present an identical fault signature as well. The rest of the fault cases, faults on the dc bus (f05, f06) and faults inside the inverter (f09-f12), can all be uniquely identified as they all have different fault signatures.

### **Step 3: Fault isolation algorithm**

As mentioned before, many algorithms can be designed out of a fault signature table. One of them is presented in Fig. 7.3 where the flowchart of the algorithm is provided. The algorithm initiates when the measured active power at the inverter output is different from the estimated one for the current irradiance level. The first symptoms to be checked were the s1 and s2. If the sum of the phase currents was not equal to zero (s1) and if on top of that the sum was positive, then the negative pole to ground SC fault (f06) on the dc bus was identified. If however, the sum of the currents was negative, then the positive pole to ground SC fault (f05) was identified.

Following the right branch of the flowchart, the symptoms s2, s3 and s4 were used consecutively to identify a series of faults. The first one to be checked was s2 and in how many phases the current was zero. If it was zero in all three phases then the category of boost faults (f07-f08) was identified. On the other hand, if the current was not zero in any of the phases the next criterion, s3, was used. There, if the current during faulty operation was smaller than the one in normal operation in all three phases, then the group of PV faults (f01-f04) was found. Furthermore, if the current difference was positive in some phases and negative in others, then the algorithm proceeded to next step where the s4 was checked. Moreover, if no voltage drop was noticed during faulty operation in any of the phases, then the identified fault was the OC of one IGBT in one of the legs of the inverter (f09). Finally, if the frequency component of the voltage at 50  $Hz$  was not zero and if at the same time a voltage drop was noticed in all three phases, then the SC of one of the IGBTs in one of the inverter legs was identified (f11). In this case, it was also possible to identify the faulty IGBT by checking the behavior of the voltage in each phase (close to zero with some spikes on the faulty phase).

Returning back to the case where the sum of phase currents was equal to zero (s1) and following the “yes” branch of this decision block, if both the frequency

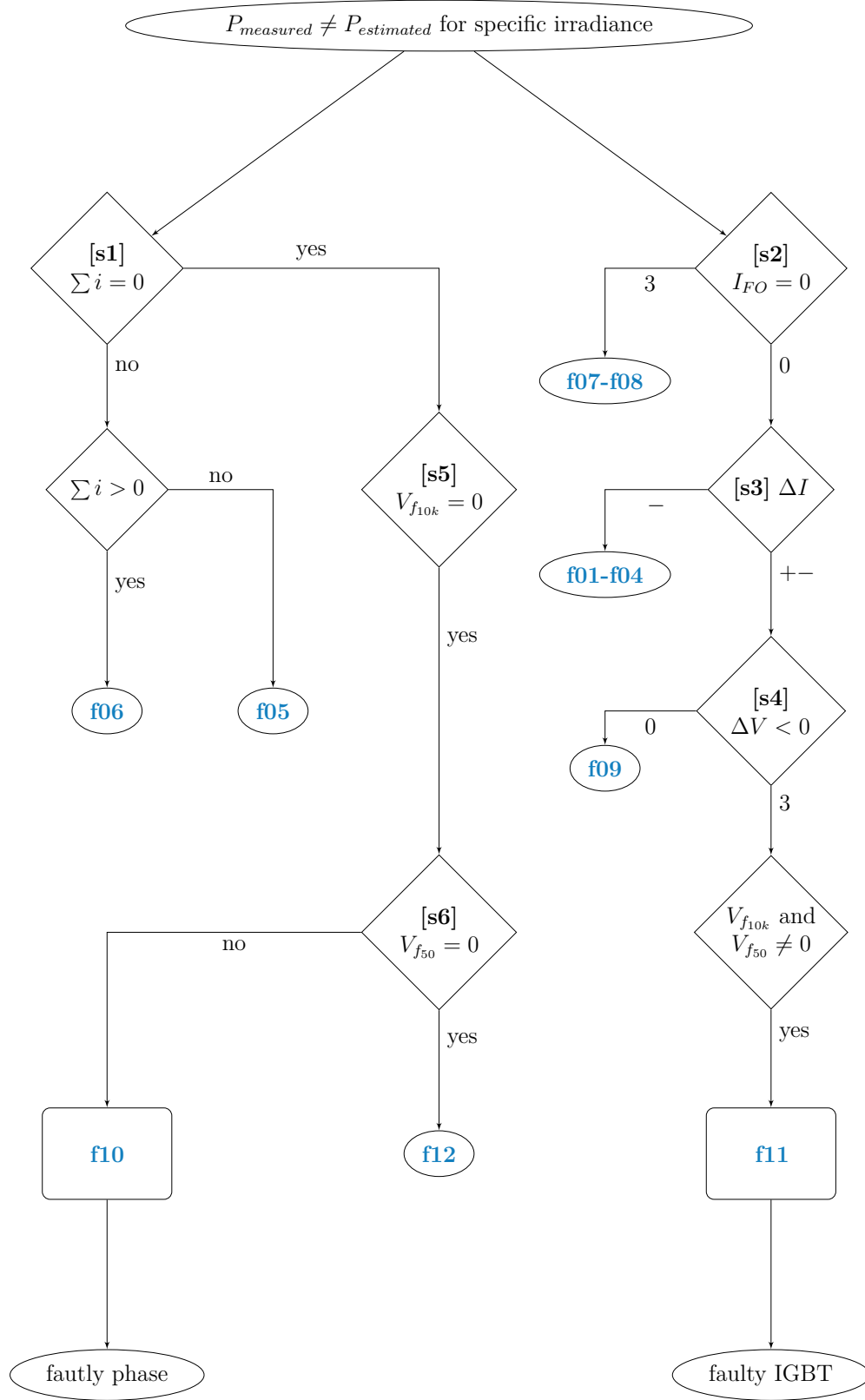


Figure 7.3 Flowchart of the fault isolation algorithm A.

components of the voltage at 10  $kHz$  and 50  $Hz$  disappeared after the fault occurrence, then the identified fault case was the SC of both IGBTs in one of the inverter legs. On the contrary if the frequency component at 50  $Hz$  was still present, the OC of both of the IGBTs in one of the inverter legs (f10) was identified. Again, the detection of the faulty phase was possible since the current of the phase under fault was equal to zero.

#### **Step 4: Threshold settings configuration**

For the implementation of the isolation algorithm, the environment of **Matlab/Simulink** was used. Since many methods face problems at low irradiance levels, as mentioned in Chapter 2, a step by step decrease of the irradiance (from 1000 to 800, 600, 400 and 200  $W/m^2$ ) was applied to verify that the thresholds set for 1000  $W/m^2$  of irradiance would still hold.

Beginning with s1, the instantaneous values of the currents were added in order to trace their sum. Continuing with s2, the rms value of the current was monitored. For both s1 and s2, the equality to zero was verified through the implementation of a threshold of 2  $A$  around zero ( $-2A < I < 2A$ ) which corresponded to 2% of the nominal value of the current at STC.

Passing on to the s3 and s4 symptoms the difference of the monitored variables ( $I, V$ ) between faulty and normal operation was created based on their moving maximum values over a sliding window of 40  $ms$  (equal to twice the period of the signal, 20  $ms$ ) in order to attain the maximum values of the sinusoidal signals at all times. The negative value of  $\Delta V$  was defined as lower than 5  $V$  (1% of  $V_{NO}$  at STC) while the current settings, in order to avoid any false detection at low irradiance levels under 400  $W/m^2$ , were designed around 2  $A$  (2% of  $I_{NO}$ ).

For the threshold settings of the frequency components of the voltage, a bandpass filter and a lowpass filter were used for 10  $kHz$  and 50  $Hz$  respectively. The equality to zero of their output was checked with a threshold crossing around 5  $V$ ,  $-5V < V_{f50}, V_{f10k} < 5V$ , which again corresponded to 1% of  $V_{NO}$  at STC. This disappearance of the voltage frequency components is visible in the spectrograms presented in Fig. 7.4a and Fig. 7.4b for the 10  $kHz$  and 50  $Hz$  cases respectively; the example of f12 is used to demonstrate the changes in the frequency components. In both cases after the fault occurrence at 0.5  $s$ , the disappearance of the frequency components is marked by a color change in the dBm scale.

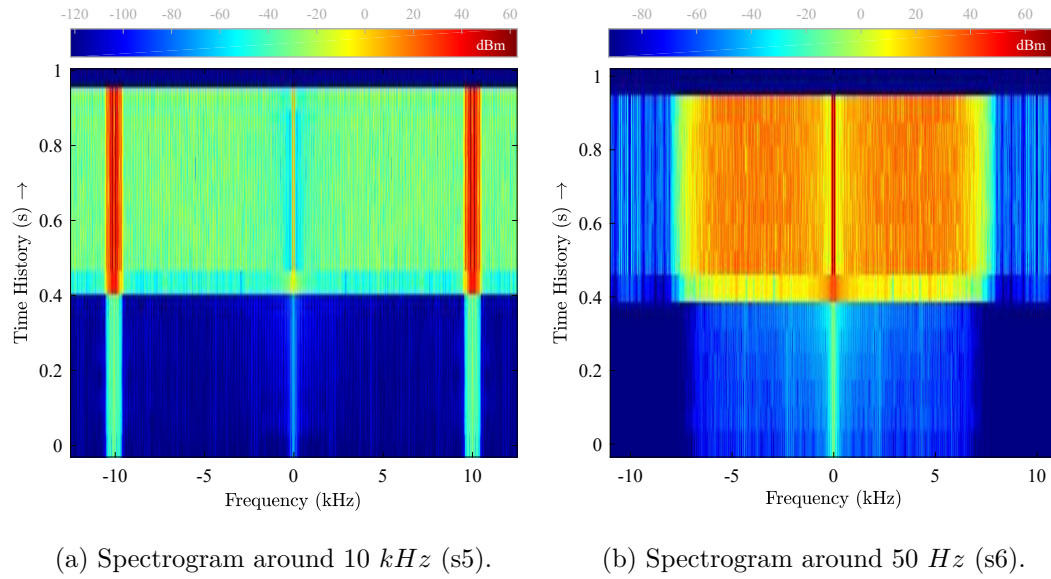


Figure 7.4 Spectrograms of the lowpass filters for the SC inverter leg fault case (f12) around a) 10 kHz and b) 50 Hz show the disappearance of the pre-fault existing frequency components depicted in red.

### 7.2.5 Fault isolation algorithm B: dc and ac side faults

#### Step 1: Symptom generation

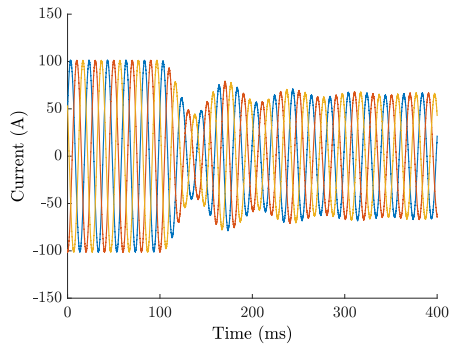
In this second method, more detailed criteria around the current and the voltage were developed in order to generate a list of symptoms to include grid side faults. A first set of criteria included comparisons between faulty and normal operation values of voltage and current similar to those used in *algorithm A* and conclusions were drawn based on whether  $\Delta I$  was negative or positive (s1) or  $V_{FO}$  was equal to zero (s3). Another aspect of changes in  $I$  and  $\Delta V$  was whether they appeared symmetrically in all three phases or not (s2, s4). Furthermore, the sum of all three phase currents being equal to zero was considered as an extra symptom (s5). In addition, the method of symmetrical components was used to further analyze the measured voltage. Moreover, the existence of negative and zero sequence component at the monitored voltage (s6) was considered as another fault symptom. Finally, based on observations from the behavior of the phase of the positive, negative and zero sequence components of the voltage during simulations, the rest of the symptoms (s07-s09) were created. A list summarizing all those symptoms is presented below:

[s1] current increase  $\Delta I > 0$  or current decrease  $\Delta I < 0$

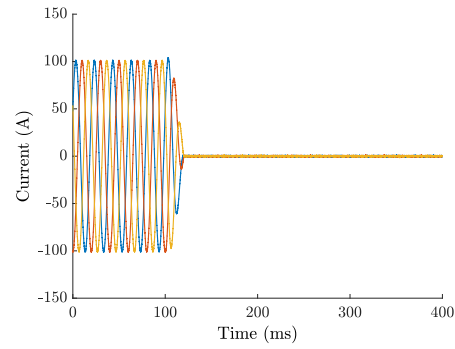
[s2] in how many of the three phases is  $I_{FO} = 0$

- [s3] voltage is equal to zero,  $V_{FO} = 0$
- [s4] in how many of the three phases is  $\Delta V < 0$
- [s5] the sum of phase currents is equal to zero,  $\sum i = 0$
- [s6] both the negative and the zero components exist
- [s7] phase of the positive component,  $\phi_{pos} < -25^\circ$
- [s8] phase of the negative component,  $\phi_{neg} < -105^\circ$ ,  $-105^\circ < \phi_{neg} < -20^\circ$  or  $\phi_{neg} > -20^\circ$
- [s9] phase of the zero component,  $\phi_0 > 110^\circ$

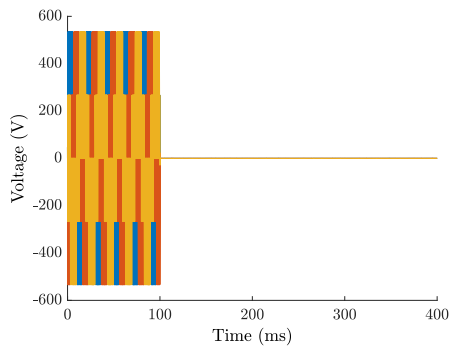
Examples of the monitored variables used to generate the symptoms are provided in Fig. 7.5. In Fig. 7.5a the current decrease can be observed during a SC on the bypass diode (f03). Additionally, Fig. 7.5b shows an example of a decrease of the current to zero after a fault in the boost converter (f07 or f08). Similarly, Fig. 7.5c shows the decrease to zero of the voltage after a short-circuited inverter leg (f12). On the other hand, the distortions on the voltage waveforms noticed under a SC fault between two phases and the ground (on the ac side) are depicted in Fig. 7.5d. Furthermore, the non equality of the sum of the instantaneous values of currents to zero under a dc bus fault (f06), after the fault inception at 100 *ms*, is presented in Fig. 7.5e. Moreover, as shown in Fig. 7.5f, the negative sequence component of the voltage under normal operation is zero (or very close to it) and gains a positive value after the inception of a double phase fault (f18) on the grid side. Regarding the monitoring of the phase of the negative sequence component of the voltage (Fig. 7.5g), its value under normal operation is oscillating between  $200^\circ$  and  $-200^\circ$  degrees as the voltage itself is not exactly equal zero but oscillating around it with a negligible amplitude. However, after the fault occurrence it was observed in simulations that the value of the phase of the negative sequence component of the voltage is stabilizing to negative values as shown in Fig. 7.5g for a SC fault between two phases (f15). Finally, changes in the phase of the zero sequence component of the voltage were also recorded. An example is provided in Fig. 7.5h where the phase of the zero sequence is increasing during a single phase fault (f17).



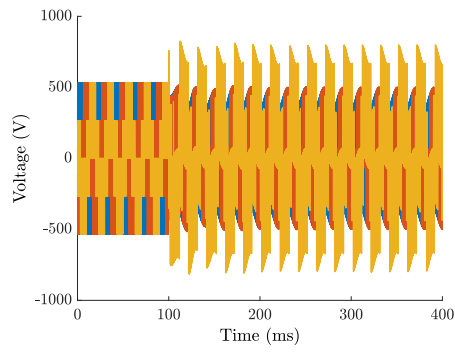
(a) Instantaneous 3ph. current (f03).



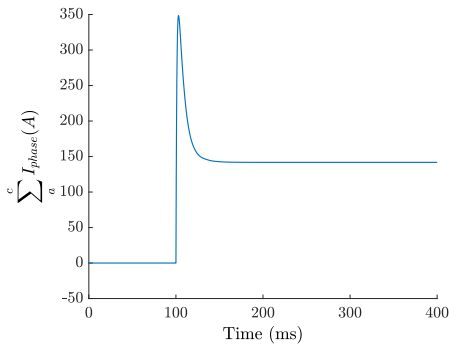
(b) Instantaneous 3ph. current (f07-f08).



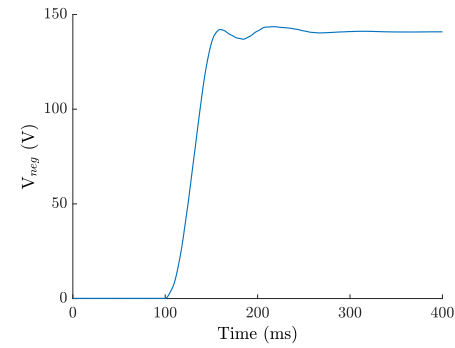
(c) Instantaneous 3ph. voltage (f12).



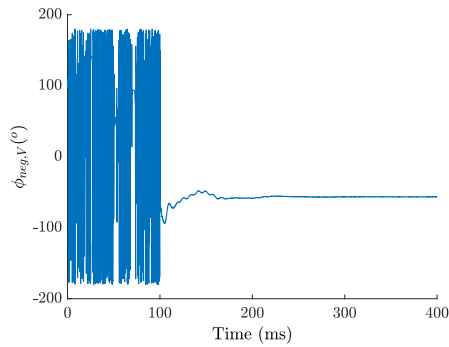
(d) Instantaneous 3ph. voltage (f14).



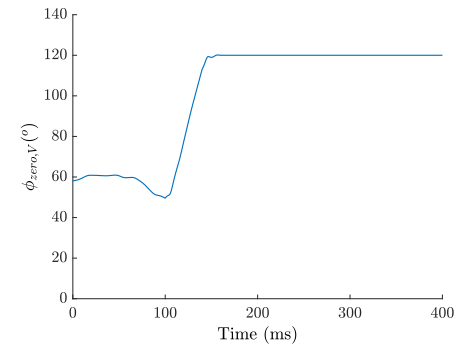
(e) Sum of currents (f06).



(f) Negative voltage component (f18).



(g) Phase of negative voltage component (f15).



(h) Phase of zero voltage component (f17).

Figure 7.5 Monitored variables under various faults that were used for the symptom generation. Fault inception is at 100 ms.



### **Step 2: Fault signature table construction**

The symptoms described above along with the studied fault cases, were used to construct the fault signature table, Table 7.4. In the fault signature table, the symbol “√” is used to verify that the criteria described in the associated symptom were met while “x” marks the opposite. The symbol “/” is used as indicator that the specific symptom was of no interest to the associated fault. Another set of symbols was necessary to describe the changes in the current in s1; “+” is used when  $\Delta I > 0$  and “-” when  $\Delta I < 0$  while “+-” to indicate different behavior in different phases. For s2 and s4 the numbers “0 – 3” indicate how many of the three phases are affected; “0” is for none of the phases, “1” for one phase etc. Finally, for s8, the following symbols were assigned to the phase value ranges: “--” for  $\phi_{neg} < -105^\circ$ , “+-” for  $-105^\circ < \phi_{neg} < -20^\circ$  and “+” for  $\phi_{neg} > -20^\circ$ .

From the fault signature table, Table 7.4, 70 % of the faults are completely isolated (thirteen out of the total of nineteen faults cases). All faults inside the PV array (f01-f04), constitute altogether a group of faults that present exactly the same fault signature, making them impossible to discriminate from each other. This can be easily explained since the measurement point was located far from the source of the fault and by the time the signal reached the sensors, it had already been altered by the intermediary elements. The second and last group of indiscriminable faults are the faults inside the boost converter (f07, f08) since they too have the same fault signature. All the other faults can be completely isolated. Finally, in the cases of f10 and f11 inverter faults, as with *algorithm A*, the faulty IGBT or inverter leg were also identified through monitoring of the current and voltage behavior.

### **Step 3: Fault isolation algorithm**

An algorithm was developed in order to isolate each fault or group of faults. In Fig. 7.6, a flowchart with all the decision levels of the proposed algorithm is presented. When a fault is identified, the algorithm is terminated.

As derived from Table 7.4, the only faults that had a sum of phase currents different from zero were the SC faults between the dc bus and the ground (f05, f06). As a result, the first symptom that was checked was the symptom s5 and whether or not the sum of phase currents was equal to zero. In case the sum was different from zero and positive then the fault occurred between the negative

Table 7.4 Fault signature table

Faults	Symptoms								
	<i>s1</i>	<i>s2</i>	<i>s3</i>	<i>s4</i>	<i>s5</i>	<i>s6</i>	<i>s7</i>	<i>s8</i>	<i>s9</i>
<b>f01</b>	—	0	x	0	✓	/	/	/	/
<b>f02</b>	—	0	x	0	✓	/	/	/	/
<b>f03</b>	—	0	x	0	✓	/	/	/	/
<b>f04</b>	—	0	x	0	✓	/	/	/	/
<b>f07</b>	—	3	x	3	✓	/	/	/	/
<b>f08</b>	—	3	x	3	✓	/	/	/	/
<b>f05</b>	—	0	✓	3	—	/	/	/	/
<b>f06</b>	—	0	✓	3	+	/	/	/	/
<b>f09</b>	+—	0	x	0	✓	/	/	/	/
<b>f10</b>	+—	1	x	1	✓	/	/	/	/
<b>f11</b>	+—	0	x	3	✓	/	/	/	/
<b>f12</b>	+—	0	✓	3	✓	/	/	/	/
<b>f14</b>	+	0	x	2	✓	✓	/	/	/
<b>f16</b>	+	0	x	0	✓	x	✓	/	/
<b>f19</b>	+	0	x	0	✓	x	x	+	/
<b>f15</b>	+	0	x	0	✓	x	x	—	/
<b>f13</b>	+	0	x	1	✓	✓	/	—	/
<b>f17</b>	+	0	x	0,1	✓	✓	x	+	x
<b>f18</b>	+	0	x	0	✓	✓	x	/	✓

pole and the ground whereas if the sum was negative then the fault occurred between the positive pole and the ground. On the other hand, if the sum was equal to zero then the algorithm proceeded to the next symptom.

Besides the dc bus to ground SC faults, the only other case where the voltage was equal to zero, was the case of both short-circuited IGBTs in one of the inverter legs (f12). Consequently, the next symptom to be checked was s3. If  $V_{FO}$  was indeed equal to zero, f12 was identified and if not the algorithm advanced to the next step.

Symptom s2 describing in how many phases the fault current is equal to zero, followed. If  $I_{FO} = 0$  in all three phases, then the group of boost converter faults (f07-f08) was isolated. Moreover, if  $I_{FO} = 0$  in only one phase then the case of open-circuited inverter leg (f10) was identified. On the contrary, if the fault current was not zero in any of the phases the algorithm continued with symptom s4.

Symptom s4 checks in how many of the three phases, the fault voltage drops to smaller rms values compared to the normal operation. If  $\Delta V_{FO} < 0$  in all three phases then f11, a short-circuited inverter IGBT fault, was identified. Furthermore, if  $\Delta V_{FO} < 0$  in two of the three phases, the SC between two phases and the ground fault (f14) was isolated. But if  $\Delta V_{FO} < 0$  in only one phase, then there were two possibilities: a) one phase voltage sag (f17) or b) SC between one phase and the ground (f13). As shown in Table 7.4, the latter can take two values when it comes to symptom s4, either 0 or 1. This happens because in the case of a very low irradiance ( $200W/m^2$ ), small changes were noticed in the measured voltage during the single phase voltage sag. To be able to distinguish faults f17 and f13, the monitoring of the phase of the negative sequence component of the voltage (symptom s8) was employed and if  $\phi_{neg} < -105^\circ$  then the case of a SC between one phase and the ground (f13) was identified.

The most important indicator in order to discriminate the remaining faults was the one described in s1: differences in the phase current. If  $\Delta I < 0$  then the group of PV faults (f01-f04) was isolated; else if the current was strictly negative in all three phases then the open-circuited inverter IGBT (f09) was isolated. The remaining faults are all faults in the grid side for which a current increase was observed ( $\Delta I > 0$ ).

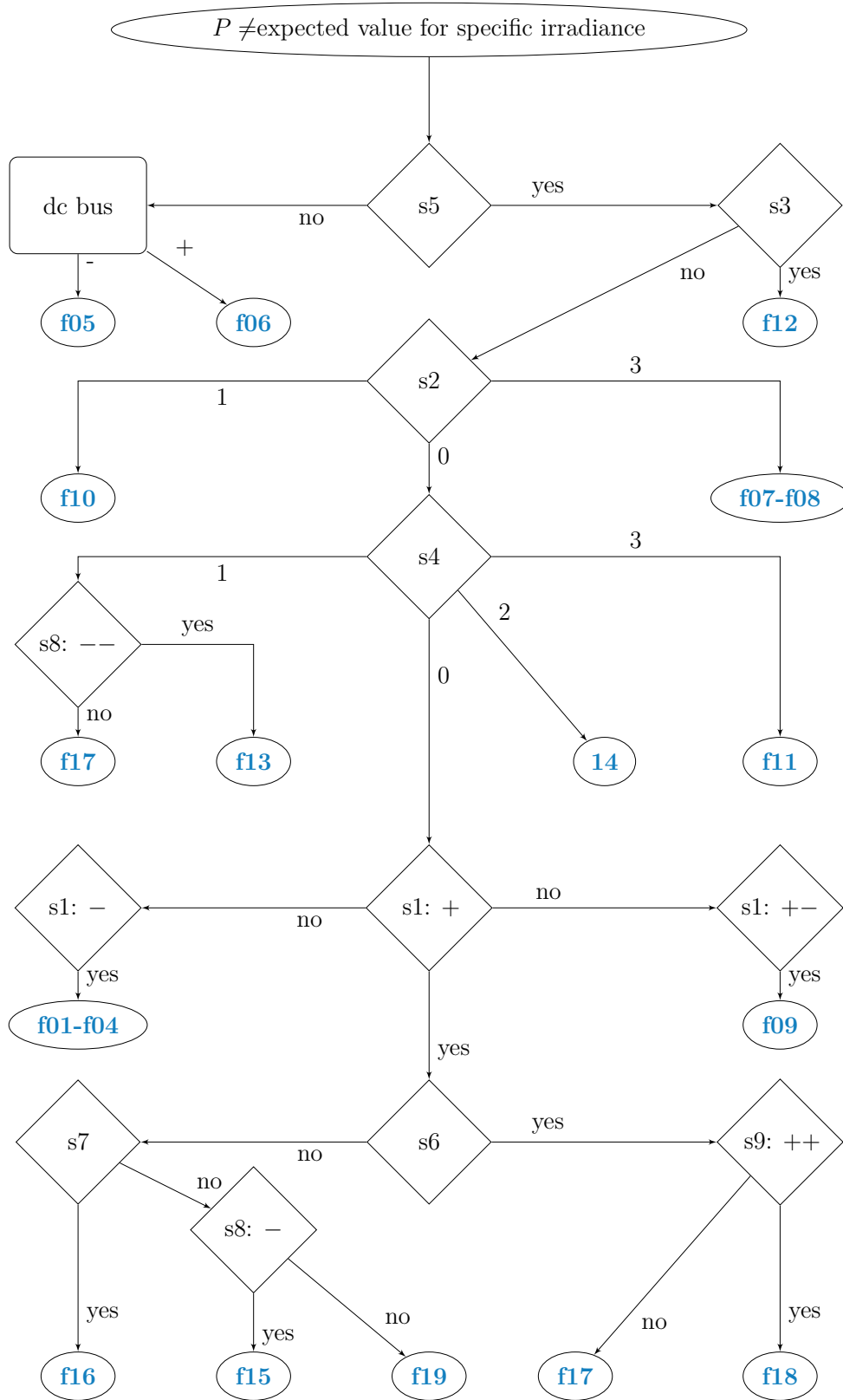


Figure 7.6 Fault isolation algorithm B flowchart.

As described in the previous steps, for the distinction of the grid faults, an analysis of the symmetrical components of the measured voltage was necessary. The existence of both the negative and the zero components of the voltage, symptom s6, was used to separate one phase and two phase voltage sags (f17 and f18 respectively) from the rest of the faults. To discriminate those two faults the phase of the zero component was used (s9) as it was observed during simulations that in the case of two phase voltage sags, the phase of the zero component of the voltage was increasing and stabilizing in a value higher than  $105^\circ$ .

Moving to the other branch of the flowchart, in the case where the existence of both the negative and the zero sequence components of the voltage was not confirmed, faults f15, f16 and f19 were left to isolate. As a first step, s7, the phase of the positive component,  $\phi_{pos}$ , was checked. If it was smaller than  $-25^\circ$ , f16 was isolated. If not, s8, was employed again to distinguish f15 and f19. If  $-105^\circ < \phi < -20^\circ$  then the double phase SC fault would get isolated (f15), leaving the last fault f19, three phase voltage sag, isolated as well as the only remaining possibility.

#### **Step 4: Threshold settings configuration**

In order to implement the diagnostic strategy presented in Fig. 7.6, further specification of the threshold crossing settings was deemed necessary. The configuration of the implemented threshold crossing settings was made in order for them to be compliant with the various levels of irradiation. For the common symptoms with *algorithm A* (s1, s2, s4 and s5) the same thresholds were applied.

For s3, for the voltage equality to zero, the thresholds were set again at 1 % of the peak value at normal operation and STC but this time the rms values were used to monitor the decrease to zero, for accuracy reasons.

Regarding the symmetrical component analysis, their moving average was traced, once more over a sliding window of 40 ms. The existence of the negative and the zero sequence components was verified by applying a threshold of 5.53 V around zero for consistency reasons. For the rest of the symptoms (s7 to s9) the implemented threshold settings were the angles described in Step 1: symptom generation, including a safety margin of at least  $5^\circ$ . Finally, in Fig. 7.7 and Fig. 7.8 two examples of threshold crossing settings application for symptoms s7 and s8 are given. More can be found in the Appendix B.

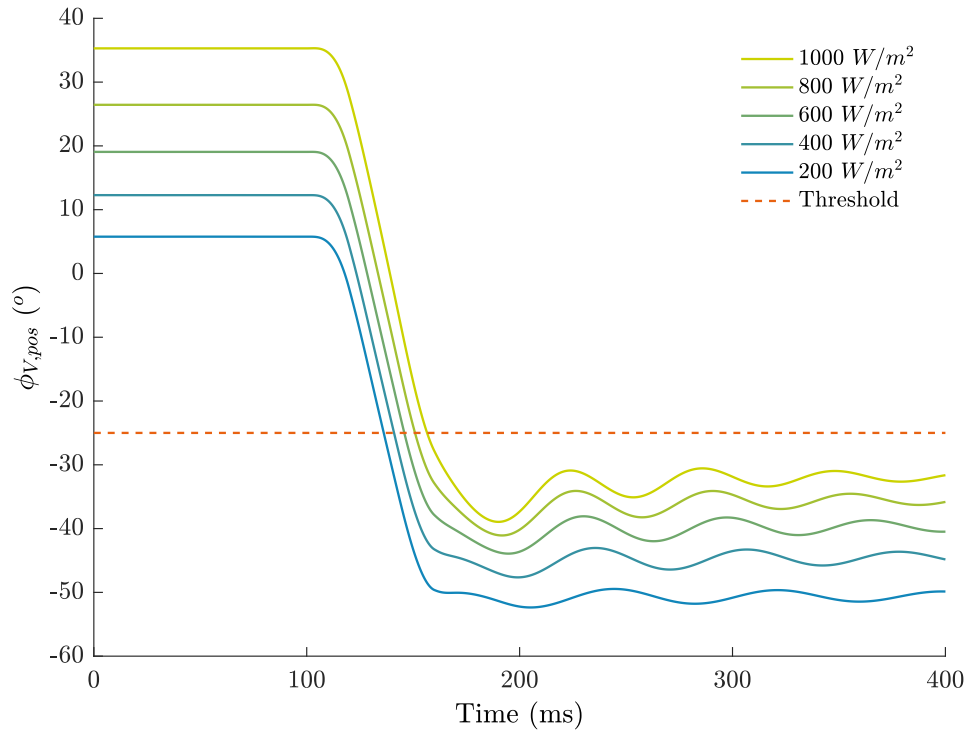


Figure 7.7 Moving average of the phase angle of the positive sequence component of the voltage (s7) for a three phase SC fault (f16) at different irradiance levels. Fault inception at 100  $ms$ .

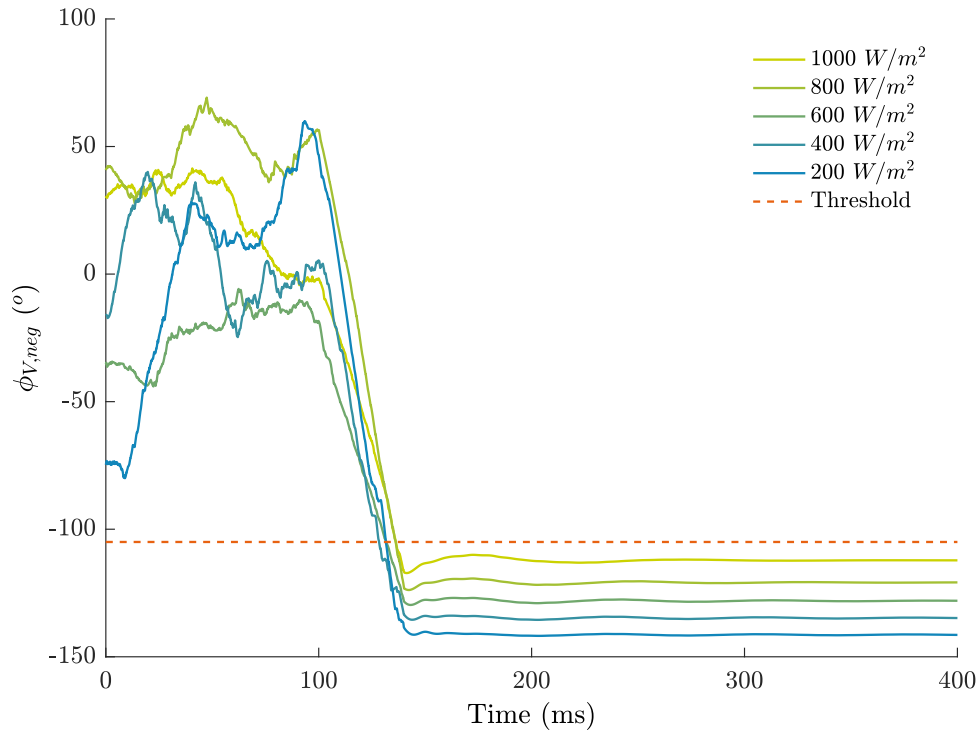


Figure 7.8 Moving average of the phase angle of the negative sequence component of the voltage (s8) for a single phase to ground SC fault (f13) at different irradiance levels. For NO curve shape refer to Fig. 7.5g. Fault inception at 100  $ms$ .

### 7.2.6 Algorithm performance

The fault diagnosis speed is an important aspect of the FDI process since a fast diagnosis will leave some time to plan if possible the necessary course of action. As mentioned before the inverter will act at around 200 *ms* after the fault occurrence [255]. The required time for the methods to create the alarm signal for the different irradiance levels and faults was monitored. The maximum and minimum diagnosis time for each fault or group of faults were also recorded and they are presented in Table 7.5. The algorithms can isolate a fault in some cases as fast as 1 *ms* after the fault occurrence and in general in less than 60 *ms*. However, there is one exception to that general rule for both algorithms. The open-circuited inverter leg fault (f10) for *algorithm A* requires 97 *ms* from fault occurrence to be isolated. At the same time, in *algorithm B* the extra feature of the faulty phase detection for the one short-circuited IGBT inside the inverter fault (f11), which is not depicted on the table, presents an extra latency of 30 *ms* linked to the reaction of the inverter controller.

The advantage of isolating a fault in such a short time is the fact that it happens before the protective systems are triggered. The minimum time that the protective elements require to act in order to isolate a GCPV system after a fault occurrence, is 120 *ms* [256]. In the case of a fault in the grid side, as time passes, more and more PVs will gradually be disconnected, beginning with those located towards the end of the feeder where the voltage drop will be higher [253]. It is clear that in this case, the designated symptoms would certainly be affected and would possibly cause misslocalizations of faults. Hence, the fact that the algorithms are faster than the minimum trigger time of protection elements (120 *ms*), ensures the correct isolation of the fault.

As mentioned before, *algorithm A* completely isolates 50 % of the faults while the remaining six fault cases are localized as two groups of faults: faults inside the PV (f01-f04) and faults inside the boost converter (f07-f08). Similarly, *algorithm B* completely isolates 70 % of the faults and the remaining in the same groups of faults as in *algorithm A*.

Both methods were validated by simulations means for different irradiance levels and more specifically for 1000, 800, 600, 400 and 200  $W/m^2$ . More figures can be found in the Appendix B.

Table 7.5 Algorithms isolation speed in *ms*.

Faults	algorithm A		algorithm B	
	$t_{min}$	$t_{max}$	$t_{min}$	$t_{max}$
f01-f04	25	37	37	45
f05	14	20	1	1
f06	14	20	3	3
f07-f08	14	20	32	39
f09	10	13	10	13
f10	43	97	18	22
f11	56	60	56	60
f12	24	25	20	20
f13	/	/	48	56
f14	/	/	55	56
f15	/	/	40	59
f16	/	/	40	40
f17	/	/	26	44
f18	/	/	20	23
f19	/	/	38	39
f20	/	/	17	44
global	10	97	1	60



### 7.3 Measurement uncertainty

After the first stress test of the algorithm against various irradiance levels, another parameter was examined to verify its robustness: erroneous measurements. Although a thorough analysis and evaluation of measurement data is provided by the Joint Committee for Guides in Metrology in [257], describing the full details of modeling measurement uncertainty is out of the scope of this thesis. In this case, the model described in [258] was used. This model of transforming an ideal measurement to a distorted signal closer to what a sensor would give in reality, is presented in Fig. 7.9. In this model the ideal signal is submitted to a  $20 \mu s$  delay then passed through a  $10 kHz$  second-order low pass filter and finally multiplied with a gain of 0.98 in order to acquire the corrupted measurement with an amplitude deviation of 2%.

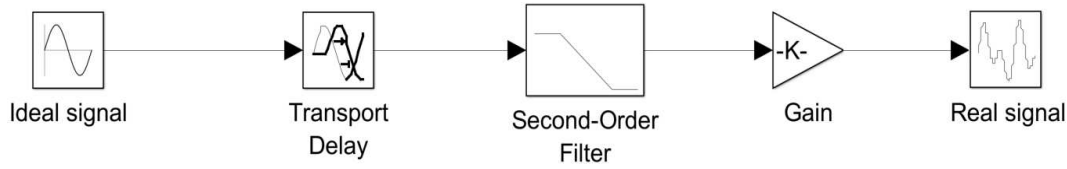


Figure 7.9 Signal distortion modeling of the measured  $I$  or  $V$  signal

The above model is described by the following equation:

$$y(t) = k \cdot H(s) \cdot x(t - t_d) \quad (7.1)$$

where:

$t_d$  is the time delay and

$H(s)$  is the low pass filter transfer frequency given by:

$$H(s) = \frac{\omega_n^2}{s^2 + 2 \cdot \zeta \cdot \omega_n \cdot s + \omega_n^2} \quad (7.2)$$

with:

$s$ : the Laplace operator

$\omega_n = 62832 \text{ rad/s}$  ( $= 2 \cdot \pi \cdot 10^4$ ): the natural frequency

$\zeta = 0.707$ : the damping ratio

When exposed to the distorted current and voltage measurements both algorithms produced exactly the same results; neither any of the symptoms nor any of the threshold crossing settings nor the algorithm speed were affected by the noise in the measured current and voltage. A few examples are provided in Fig. 7.10 to 7.12 where erroneous measurements under the two extreme irradiance cases of 1000 and 200  $W/m^2$  are presented for the symptoms s7, s8 and s9.

In Fig. 7.10, the threshold of  $-25^\circ$  for the phase angle of the positive sequence component of the voltage (s7) was proved sufficient under the three phase SC fault (f16). Moreover, in Fig. 7.11 the angle of the phase of the negative sequence component of the voltage (s8) also remains within the threshold limits under a single phase to ground SC fault (f13). Finally, the changes in measurements did not seem to affect the case of the phase angle of the zero sequence component of the voltage (s9) under a double phase voltage sag error (f18) either.

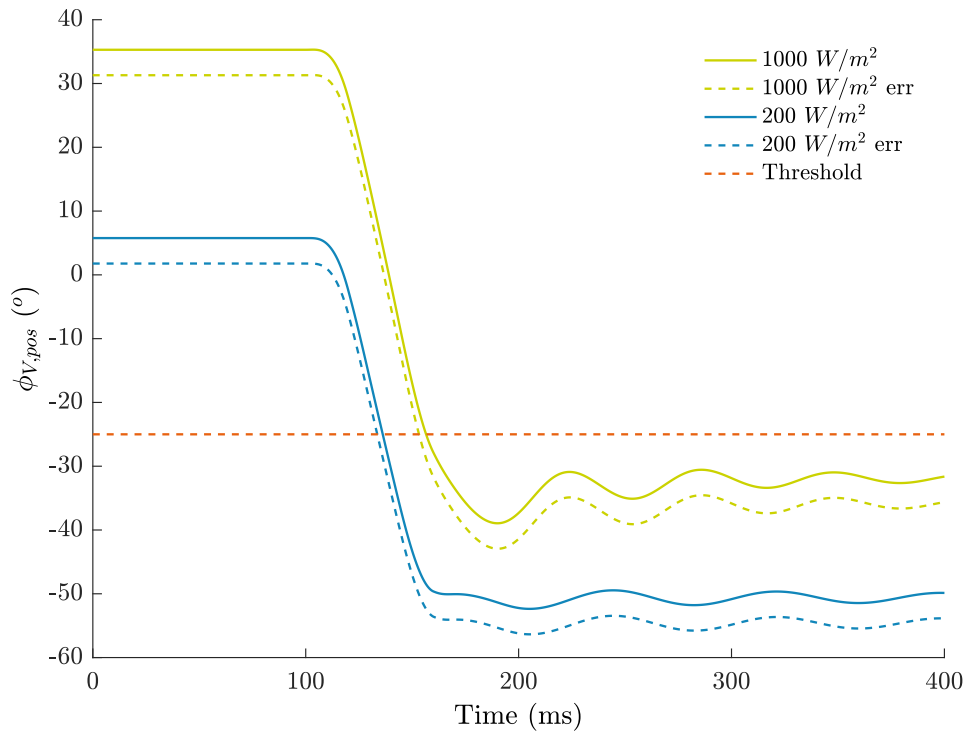


Figure 7.10 Erroneous (dashed line) vs. ideal measurements (solid line) for symptom s7 under a f16 fault at 1000 and 200  $W/m^2$ . The chosen threshold at  $-25^\circ$  appears in a red dashed line. Fault inception at 100 ms.

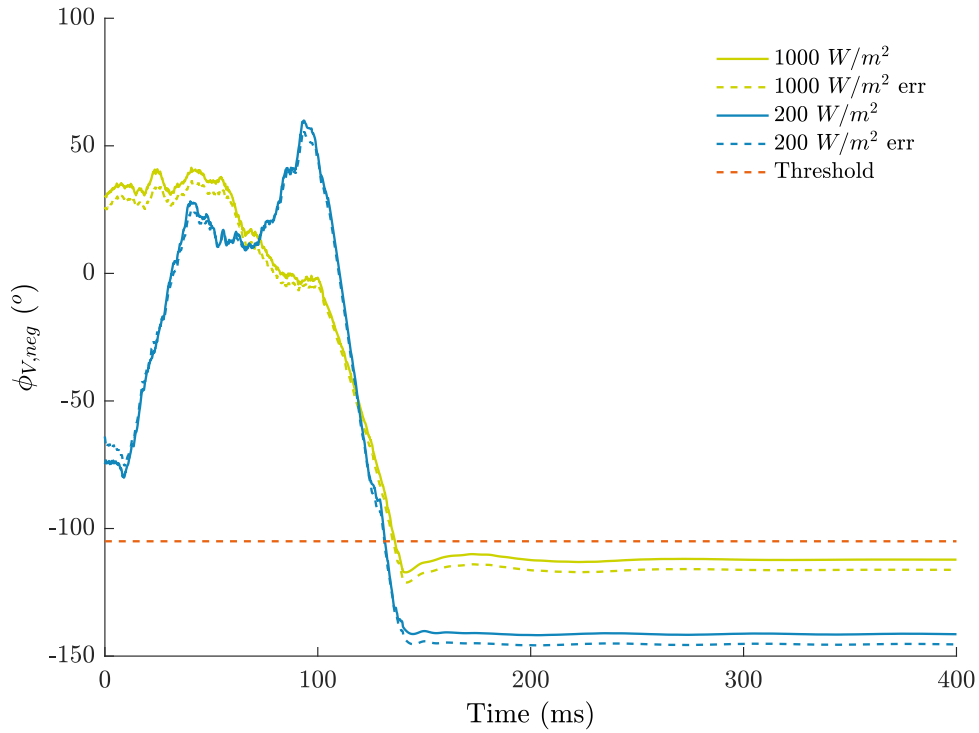


Figure 7.11 Erroneous (dashed line) vs. ideal measurements (solid line) for symptom s8 under a f13 fault at 1000 and 200  $W/m^2$ . The chosen threshold at  $-105^\circ$  appears in a red dashed line. Fault inception at 100  $ms$ .

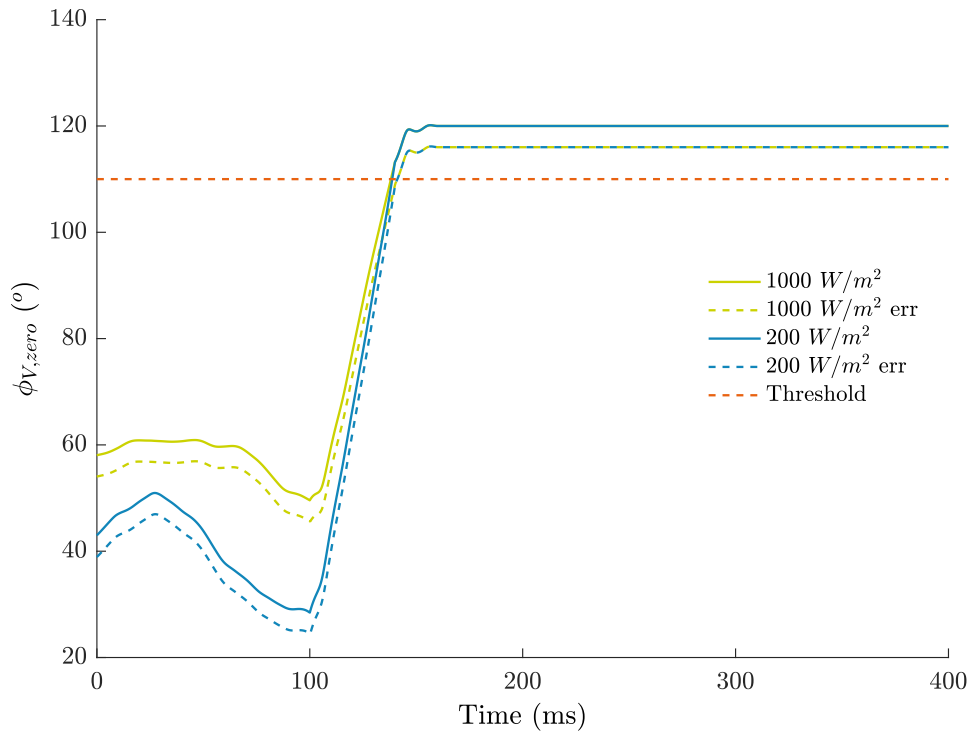


Figure 7.12 Erroneous (dashed line) vs. ideal measurements (solid line) for symptom s9 under a f18 fault at 1000 and 200  $W/m^2$ . The chosen threshold at  $110^\circ$  appears in a red dashed line. Fault inception at 100  $ms$ .

## 7.4 Conclusion

In this chapter, the developed fault diagnosis methods for *grid-connected photovoltaic (GCPV)* systems were analyzed. The developed methods rely solely on voltage and current measurements taken at the output of the inverter. The theoretical steps of a successful fault detection and isolation method were described: a) symptom generation, b) fault signature table construction, c) fault isolation algorithm design and d) threshold crossing settings configuration. Then, a list of the studied faults was provided and the fault detection process was explained.

Under the assumption that a calibrated PV reference device was available to monitor the actual irradiance of the plant and that the maximum power point tracking was working correctly, the expected active power was compared to the measured active power. In that case, a decrease of the active power was considered as a strong indicator of a fault occurrence and marked the initiation of the fault isolation algorithms.

Two fault isolation algorithms were developed, the first, *algorithm A*, focused only on dc side faults while the second, *algorithm B*, included ac side faults as well thus covering the whole system. *Algorithm B* is capable of isolating thirteen faults and two groups of faults out of the nineteen fault cases. Moreover, both algorithms are able to isolate faults quite fast (during their transient or close to steady-state phase) within 100 ms and 60 ms for *A* and *B* respectively. This level of speed ensured an isolation before the protective schemes are triggered and thus alter the fault behaviors to the associated symptoms.

A two-level robustness verification was used to test the resilience of the algorithm against various conditions. The first parameter that was modified was the irradiance level with the algorithms performing accurately in all five of the studied levels (1000, 800, 600, 400 and 200 W/m<sup>2</sup>). The second parameter was the measurement uncertainty. Noise was introduced in the voltage and current measurements and the algorithm proved resilient against that as well. To conclude with, no impact of the erroneous measurements was noticed on the symptoms, threshold crossing settings or isolation performance and speed.



# Chapter 8

## General conclusion

Changes in policies under the light of the rapid climate change impose the installation of renewable energy sources in *low voltage grids (LV)*. A rapid growth of *grid-connected PV systems (GCPVs)* is being recorded over the last years. In 2017 alone, the global PV installed capacity increased by 99 GW reaching a total of 403 GW which represents the 2.5% of the global energy demand and 55 % of the installed capacity of new renewable energy sources in 2017 [3]. The LV distribution grid, which was traditionally designed on the basis of the “fit and forget” principle allowing only a unidirectional power flow, is now facing several problems such as congestion, voltage rises and decrease of power quality [21], with the interconnection of renewable energy sources.

So far, the focus of operators was to improve the transmission and *medium voltage (MV)* distribution parts of the grid. However, the installation of *photovoltaics (PV)* and other green types of microgeneration units in the LV grid, obliges operators to shift their attention to the monitoring and control of the LV grid increasing its functionalities [22], i.e. installing smart meters and implementing self-healing strategies. In this context, three fault diagnosis methods for LV distribution grids and two fault isolation algorithms for GCPVs have been proposed in this thesis.

In this chapter, a review of the contributions of this thesis is presented and a discussion on future work and perspectives is held.

## 8.1 Results & Contributions

### 8.1.1 Part I: Fault diagnosis in low voltage smart distribution grids

While several fault location methods for distribution grids exist in the literature, the majority of them focuses on MV grids. Only a very limited number of studies have considered the LV grid; just eight studies have been performed (to the author's knowledge) with the oldest one being published in 2012 (Fig. 2.10). The lack of studies in LV grids is attributed to their increased complexity compared to the MV ones. A LV grid presents five basic characteristics that hinder the fault location methods: a) an increased number of laterals, b) multi-phase and unbalanced operation, c) unbalanced distribution of loads, d) various types of conductors connecting the nodes and e) up to now, limited number of available measurements (a fact that is now changing with the transition to smart grids). Moreover, fault location methods for MV grids usually consider only low fault resistance values that rarely surpass the  $100\ \Omega$ . Taking into account that: a) faults of a higher fault resistance can occur, e.g. when a downed conductor touches the earth, fault resistances vary from  $90\ \Omega$  (concrete as ground) to  $1500\ \Omega$  (wet sand as ground), and b) that distribution system operators usually rely on customer phone calls to detect and locate faults in LV grids, the need for fault detection and location techniques that cover these cases, i.e. large fault resistances and LV distribution grids, is evident.

The three fault detection and location methods for LV smart distribution grids proposed in this thesis are:

- A conventional fault detection method based on overcurrent monitoring in combination with a method that uses sparse voltage measurements to build the voltage profile across the faulty branch for fault location.
- Gradient boosting trees (GBT), a method that has been proven to excel in many applications the last few years.
- Deep neural networks (DNN), a method that improved the traditional neural network architecture by taking advantage of an increased number of hidden layers.

Simulations on a real semi-rural LV distribution grid of Portugal were performed to validate the results. A common case study was used to compare the three methods. The influencing parameters were: a) a big variety of fault resistance values (63,772 values between 1 and 1000  $\Omega$ ), b) nine different fault locations within each sector, c) two fault types (single phase to ground and three phase faults), d) a simultaneity factor of 0.5, e) a big spectrum of PV generation and load demand scenarios with 70,334 studied combinations and f) a 2 % underestimation error in measurements.

In general, the increase of fault resistance hinders the fault location methods as it decreases fault current and increases voltages, thus bringing the monitored variables closer to their normal operating conditions, i.e. making it harder to discriminate faulty from normal operation. In addition, the symmetrical components of the monitored variables are proven to be a better option than the phase measurements. Moreover, faults towards the end of the feeder are harder to locate due to the heterogeneity of the grid and the length of the branches. Furthermore, three phase faults are shown to be easier to locate when a symmetrical component analysis is used. As expected, an increase of the load demand decreases the accuracy of the fault location methods while on the other hand the contribution of the PV units to the faults seems limited. Finally, a decreased number of available measurements seems to seriously affect the conventional method but only slightly the DNN.

The contributions of this first part of the thesis are summarized below:

1. Three different methods of fault detection and fault location were proposed for the quite unexplored case of LV grids: one that was initially conceived for the MV case and is now extended to the LV grid [2] and two artificial intelligence methods. It is, to the author's knowledge, the first application of artificial intelligence for fault detection and fault location in the LV grid case.
2. With the use of GBT and DNN the fault detection problem is tackled with an accuracy of 99.15 % and 100 % respectively.
3. The use of DNN for fault location is the most reliable solution of the three with an average error of only 12 % considering erroneous measurements.



4. All three of the fault location methods are immune to the per phase distribution of loads and microgeneration units in the LV grid.
5. Under the case of extremely limited measurements, one in the beginning of the feeder and one at each terminal node, the fault location accuracy of the DNN is decreased by only 4.5 % while the fault detection accuracy remains intact.

### 8.1.2 Part II: Fault detection and isolation in distributed generators

The second part of the thesis, Chapter 7, studies faults in GCPV, a field that presents an increased interest the last decade with at least 76 different methods since 2010.

A big variety of faults can disrupt the operation of a GCPV power plant. Based on their location faults can appear: a) in the PV array, b) in the power converters, c) on the dc bus and d) in the grid side. The development of fast, efficient and reliable fault detection and isolation methods for GCPVs, capable of dealing with the different types of faults, is a recognized necessity from the scientific community and a prerequisite for their integration in the smart grids.

So far, to the author's knowledge, and although different approaches have been tested, no research has been found to monitor the GCPV as a complete system, i.e. isolating faults in every part of the plant with a single method. Moreover, several methods tend to fail in very low irradiance levels. For this reason, two algorithms based on a signal approach, one considering faults only on the dc side of the plant and a second one considering faults on both the dc and the ac side, that consider various irradiance levels, are proposed as a fault isolation strategy.

The algorithms use current and voltage measurements at the output of the inverter. The choice of the output of the inverter, i.e. the point of common coupling, as the monitoring source of the status of the GCPV system is in accordance with the location of voltage sensors used in the previous case of fault location methods in the LV distribution grid. Both algorithms work under the single-fault assumption, i.e. that only one fault is occurring at a time. A case study with a PV array of 43.2 kW connected to the grid through a dc-dc boost

converter and an inverter was used to study faults occurring on all four of the aforementioned possible locations. Different irradiance levels were considered ranging from 1000 to 200  $W/m^2$ . Furthermore, a sensitivity analysis was conducted against erroneous measurements. Finally, It should be noted that faults inside the PV array and the boost converter are identified as group of faults by the algorithms as they are located very far from the measurement point; all the other faults are completely isolable.

As part of this second part of the thesis, a preliminary study of fault isolation in grid-connected *proton exchange membrane fuel cells (PEMFC)* was also conducted. In detail, the same fault types with the case study of photovoltaics were studied; in the place of faults at the source of the renewable energy system, in this case, only the drying fault is considered. The isolation of the drying fault is achieved without interfering with the isolation of the faults from the other parts of the system. The complete analysis is presented in Appendix D.

The contributions of this second part of the thesis are summarized below:

1. For the first time, to the author's knowledge, a study considers faults on both the dc and the ac side with measurements taken only from the ac side. This limitation minimizes the amount of necessary sensors and hence decreases the monitoring cost.
2. The second algorithm, the one that considers faults in both the dc and the ac side, achieves an isolation of 15 out of the 19 studied fault cases.
3. The algorithms are unaffected by different levels of solar irradiance. There is no need to readjust threshold settings even for 200  $W/m^2$  of solar irradiance.
4. Robust algorithms, without the need to readjust threshold settings against distorted measurement signals.
5. The isolation of the faults is achieved in the first 100  $ms$  after their occurrence, time inferior to the reaction margin of the inverter which is expected to trip at approximately 200  $ms$  [4].

## 8.2 Future work & Perspectives

The possible extensions of this work are divided in short, mid and long term objectives.

### 8.2.1 Short term

A few extensions of this work that could be implemented within a short period of time are:

For the conventional method proposed for fault location in the LV grid:

- Alternative ways to the linear interpolation could be investigated to obtain the voltage profile across a faulty branch; the least squares method could be an option.
- A technique to create virtual sectors in the beginning and the end of the feeder in order to waive the limitation of the method of not being able to include the first and the last sector in its fault location estimation.

For the GBT method:

- The algorithm could be extended to include fault distance estimation as well.

For the DNN method:

- Include sector localization as well.

For all the LV grid fault location methods:

- Include phase to phase and double phase to ground faults in the analysis.

For the PV isolation algorithms:

- Include more irradiance levels in the study.
- Include faults in the ac side of the plant of various fault resistance values.

A common objective for all the developed methods would be to test them directly to experimental set ups.

### 8.2.2 Mid term

For all the LV grid fault location methods:

- Test the methods against changes in the grid topology, e.g. installation of batteries or new PVs.
- Test the methods on different grids.

For the PV isolation algorithms:

- Study different modes of connection of the PV to the grid.
- Develop a model that takes into account temperature variations.
- Test the same algorithms against renewable sources of the same type of connection, e.g. fuel cells, to validate the isolation of non PV-specific faults.

### 8.2.3 Long term

A long term perspective of this study would be to integrate both methods in a common approach and test them under the same case study (grid). This would provide a complete solution of fault detection and isolation for LV smart distribution grids with renewable energy sources to the distribution system operators.

To conclude with, further research is needed in order to protect the largest and most complex machine ever built: our electrical power network. The scientific community has underestimated some threats such as the geomagnetic storms. Events like the blackout of 1989 in Quebec, of 2003 in Italy and 2012 in India serve as alarms for utilities and consumers. Let's not ignore them.



# Bibliography

- [1] Turan Gönen. *Electric Power Distribution Engineering*. Taylor & Francis, Boca Raton, third edition edition, 2014.
- [2] A. Teninge, C. Pajot, B. Raison, and D. Picault. Voltage profile analysis for fault distance estimation in distribution network. In *2015 IEEE Eindhoven PowerTech*, pages 1–5, June 2015.
- [3] Trends 2018 in Photovoltaic Applications. Technical Report Report IEA PVPS T1-34:2018, 2018.
- [4] J. Yang and Y. Wang. Review on protection issues of low-voltage distribution network with multiple power-electronic-converter-interfaced distribution energy resources. In *International Conference on Renewable Power Generation (RPG 2015)*, pages 1–6, October 2015.
- [5] Rolf Isermann. *Fault-Diagnosis Applications: Model-Based Condition Monitoring: Actuators, Drives, Machinery, Plants, Sensors, and Fault-Tolerant Systems*. Springer, Berlin, 2011. OCLC: 740883705.
- [6] Pedro Linares and Luis Rey. The costs of electricity interruptions in Spain. Are we sending the right signals? *Energy Policy*, 61:751–760, October 2013.
- [7] Ettore Bompard, Tao Huang, Yingjun Wu, and Mihai Cremenescu. Classification and trend analysis of threats origins to the security of power systems. *International Journal of Electrical Power & Energy Systems*, 50:50–64, September 2013.
- [8] Z. Lu, X. Lu, W. Wang, and C. Wang. Review and evaluation of security threats on the communication networks in the smart grid. In *2010 - MILCOM 2010 MILITARY COMMUNICATIONS CONFERENCE*, pages 1830–1835, October 2010.
- [9] Henrik Bjørnebye. *Investing in EU Energy Security: Exploring the Regulatory Approach to Tomorrow’s Electricity Production*. Kluwer Law International B.V., August 2010.
- [10] Leora Lawton, Michael Sullivan, Kent Van Liere, Aaron Katz, and Joseph Eto. A framework and review of customer outage costs: Integration and analysis of electric utility outage cost surveys. Technical Report LBNL-54365, Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States), November 2003.

- [11] Abhishek Shivakumar, Manuel Welsch, Constantinos Taliotis, Dražen Jakšić, Tomislav Baričević, Mark Howells, Sunay Gupta, and Holger Rogner. Valuing blackouts and lost leisure: Estimating electricity interruption costs for households across the European Union. *Energy Research & Social Science*, 34:39–48, December 2017.
- [12] Lloyd's. Solar Storm Risk to the North American Electric Grid. Technical report, 2013.
- [13] Nikolaos Sapountzoglou. The Last Question: Does Space Affect Electrical Power Networks?, February 2019.
- [14] Robin Harding. Japan taxpayers foot \$100bn bill for Fukushima disaster. *Financial Times*, March 2016.
- [15] Justin McCurry. Japanese government held liable for first time for negligence in Fukushima. *The Guardian*, March 2017.
- [16] Helen Pidd. India blackouts leave 700 million without power. *The Guardian*, July 2012.
- [17] CEER. 6th Benchmarking Report on the Continuity of Electricity and Gas Supply. Technical Report C 18 - EQS - 86 - 03, Brussels, 2016.
- [18] ENEDIS. Qualité de fourniture | Enedis. <https://www.enedis.fr/qualite-de-fourniture>.
- [19] S. S. Gururajapathy, H. Mokhlis, and H. A. Illias. Fault location and detection techniques in power distribution systems with distributed generation: A review. *Renewable and Sustainable Energy Reviews*, 74:949–958, July 2017.
- [20] André D. Filomena, Mariana Resener, Rodrigo H. Salim, and Arturo S. Bretas. Distribution systems fault analysis considering fault resistance estimation. *International Journal of Electrical Power & Energy Systems*, 33(7):1326–1335, September 2011.
- [21] Sergio Bruno and Massimo La Scala. Unbalanced Three-Phase Optimal Power Flow for the Optimization of MV and LV Distribution Grids. In *From Smart Grids to Smart Cities*, pages 1–42. John Wiley & Sons, Ltd, January 2017.
- [22] James Northcote-Green and Robert Wilson. *Control and Automation of Electric Power Distribution Systems*. Number 28 in Power Engineering. Taylor & Francis, Boca Raton, 2007.
- [23] Jen-Hao Teng. A direct approach for distribution system load flow solutions. *IEEE Transactions on Power Delivery*, 18(3):882–887, July 2003.
- [24] Morteza Shabanzadeh and Mohsen Parsa Moghaddam. What is the Smart Grid? Definitions, Perspectives, and Ultimate Goals. In *Power System Conference*, November 2013.

- [25] Europäische Kommission, editor. *European Technology Platform Smart-Grids: Vision and Strategy for Europe's Electricity Networks of the Future*. Number 22040 in EUR. Office for Official Publications of the European Communities, Luxembourg, 2006. OCLC: 255229219.
- [26] A. Bahmanyar, S. Jamali, A. Estebsari, and E. Bompard. A comparison framework for distribution system outage and fault location methods. *Electric Power Systems Research*, 145:19–34, April 2017.
- [27] Mini S Thomas and John D McDonald. *Power System SCADA and Smart Grids*. CRC Press, 2015. OCLC: 942514123.
- [28] <https://www.freepik.com/s/macrovector/> / Freepik.
- [29] A. Zidan, M. Khairalla, A. M. Abdrabou, T. Khalifa, K. Shaban, A. Abdrabou, R. El Shatshat, and A. M. Gaouda. Fault Detection, Isolation, and Service Restoration in Distribution Systems: State-of-the-Art and Future Trends. *IEEE Transactions on Smart Grid*, 8(5):2170–2185, September 2017.
- [30] Economic assessment of smart grids solutions: Analysis carried out by the distribution network operators. Executive Summary 2017, ENEDIS, ADEeF, May 2017.
- [31] T. Kupila, T. Ihonen, T. Keränen, and L. Anttila. Efficient coordination in major power disruption. *CIREN - Open Access Proceedings Journal*, 2017(1):1064–1067, 2017.
- [32] A. A. Girgis, C. M. Fallon, and D. L. Lubkeman. A fault location technique for rural distribution feeders. *IEEE Transactions on Industry Applications*, 29(6):1170–1175, November 1993.
- [33] K. Srinivasan and A. St.-Jacques. A new fault location algorithm for radial transmission lines with loads. *IEEE Transactions on Power Delivery*, 4(3):1676–1682, July 1989.
- [34] Myeon-Song Choi, Seung-Jae Lee, Duck-Su Lee, and Bo-Gun Jin. A new fault location algorithm using direct circuit analysis for distribution systems. *IEEE Transactions on Power Delivery*, 19(1):35–41, January 2004.
- [35] A. A. Girgis, D. G. Hart, and W. L. Peterson. A new fault location technique for two- and three-terminal lines. *IEEE Transactions on Power Delivery*, 7(1):98–107, January 1992.
- [36] Mohamed A. Gabr, Doaa K. Ibrahim, Eman S. Ahmed, and Mahmoud I. Gilany. A new impedance-based fault location scheme for overhead unbalanced radial distribution networks. *Electric Power Systems Research*, 142:153–162, January 2017.
- [37] F. M. Aboshady, D. W. P. Thomas, and Mark Sumner. A new single end wideband impedance based fault location scheme for distribution systems. *Electric Power Systems Research*, 173:263–270, August 2019.
- [38] Yuan Liao. A novel method for locating faults on distribution systems. *Electric Power Systems Research*, 117:21–26, December 2014.



- [39] R. Das, M. S. Sachdev, and T. S. Sidhu. A technique for estimating locations of shunt faults on distribution lines. In *IEEE WESCANEX 95. Communications, Power, and Computing. Conference Proceedings*, volume 1, pages 6–11 vol.1, May 1995.
- [40] Rahman Dashti and Javad Sadeh. Accuracy improvement of impedance-based fault location method for power distribution network using distributed-parameter line model. *International Transactions on Electrical Energy Systems*, 24(3):318–334, 2014.
- [41] Janne Altonen and Ari Wahlroos. Advancements in Fundamental Frequency Impedance Based Earth-Fault Location in Unearthed Distribution Networks. In *19th International Conference on Electricity Distribution*, page 4, Vienna, May 2007.
- [42] Seung-Jae Lee, Myeon-Song Choi, Sang-Hee Kang, Bo-Gun Jin, Duck-Su Lee, Bok-Shin Ahn, Nam-Seon Yoon, Ho-Yong Kim, and Sang-Bong Wee. An intelligent and efficient fault location and diagnosis scheme for radial distribution systems. *IEEE Transactions on Power Delivery*, 19(2):524–532, April 2004.
- [43] Rajarshi Dutta and S. R. Samantaray. Assessment of impedance based fault locator for AC micro-grid. *Renewable Energy Focus*, 26:1–10, September 2018.
- [44] Jun Zhu, D. L. Lubkeman, and A. A. Girgis. Automated fault location and diagnosis on electric power distribution feeders. *IEEE Transactions on Power Delivery*, 12(2):801–809, April 1997.
- [45] E. C. Senger, G. Manassero, C. Goldemberg, and E. L. Pellini. Automated fault location system for primary distribution networks. *IEEE Transactions on Power Delivery*, 20(2):1332–1340, April 2005.
- [46] S. Das, N. Karnik, and S. Santoso. Distribution Fault-Locating Algorithms Using Current Only. *IEEE Transactions on Power Delivery*, 27(3):1144–1153, July 2012.
- [47] R. H. Salim, M. Resener, A. D. Filomena, K. R. Caino de Oliveira, and A. S. Bretas. Extended Fault-Location Formulation for Power Distribution Systems. *IEEE Transactions on Power Delivery*, 24(2):508–516, April 2009.
- [48] André D. Filomena, Mariana Resener, Rodrigo H. Salim, and Arturo S. Bretas. Fault location for underground distribution feeders: An extended impedance-based formulation with capacitive current compensation. *International Journal of Electrical Power & Energy Systems*, 31(9):489–496, October 2009.
- [49] X. Yang, M. Choi, S. Lee, C. Ten, and S. Lim. Fault Location for Underground Power Cable Using Distributed Parameter Approach. *IEEE Transactions on Power Systems*, 23(4):1809–1816, November 2008.
- [50] Alireza Bahmanyar and Sadegh Jamali. Fault location in active distribution networks using non-synchronized measurements. *International Journal of Electrical Power & Energy Systems*, 93:451–458, December 2017.

- [51] S. F. Alwash, V. K. Ramachandaramurthy, and N. Mithulananthan. Fault-Location Scheme for Power Distribution System with Distributed Generation. *IEEE Transactions on Power Delivery*, 30(3):1187–1195, June 2015.
- [52] R. H. Salim, K. C. O. Salim, and A. S. Bretas. Further improvements on impedance-based fault location for power distribution systems. *Transmission Distribution IET Generation*, 5(4):467–478, April 2011.
- [53] Y. Liao. Generalized Fault-Location Methods for Overhead Electric Distribution Systems. *IEEE Transactions on Power Delivery*, 26(1):53–64, January 2011.
- [54] F. C. L. Trindade and W. Freitas. Low Voltage Zones to Support Fault Location in Distribution Systems With Smart Meters. *IEEE Transactions on Smart Grid*, 8(6):2765–2774, November 2017.
- [55] Wanjing Xiu and Yuan Liao. Novel fault location methods for ungrounded radial distribution systems using measurements at substation. *Electric Power Systems Research*, 106:95–100, January 2014.
- [56] Lifeng Yang. One-terminal fault location system that corrects for fault resistance effects, June 1998.
- [57] Damir Novosel, David Hart, Yi Hu, and Jorma Myllymaki. System for locating faults and estimating fault resistance in distribution networks with tapped loads, November 1998.
- [58] Mario Tremblay, Bruno Fazio, and Denis Valiquette. Using voltage sag measurements for advanced fault location and condition-based maintenance. *CIREN - Open Access Proceedings Journal*, 2017(1):893–896, October 2017.
- [59] Javad Sadeh, Ehsan Bakhshizadeh, and Rasoul Kazemzadeh. A new fault location algorithm for radial distribution systems using modal analysis. *International Journal of Electrical Power & Energy Systems*, 45(1):271–278, February 2013.
- [60] U. D. Dwivedi, S. N. Singh, and S. C. Srivastava. A wavelet based approach for classification and location of faults in distribution systems. In *2008 Annual IEEE India Conference*, volume 2, pages 488–493, December 2008.
- [61] Z. Q. Bo, G. Weller, and M. A. Redfern. Accurate fault location technique for distribution system using fault-generated high-frequency transient voltage signals. *Transmission and Distribution IEE Proceedings - Generation*, 146(1):73–79, January 1999.
- [62] H. Nouri, Chun Wang, and T. Davies. An accurate fault location technique for distribution lines with tapped loads using wavelet transform. In *2001 IEEE Porto Power Tech Proceedings (Cat. No.01EX502)*, volume 3, pages 4 pp. vol.3–, September 2001.
- [63] A. Borghetti, M. Bosetti, M. Di Silvestro, C. A. Nucci, and M. Paolone. Continuous-Wavelet Transform for Fault Location in Distribution Power Networks: Definition of Mother Wavelets Inferred From Fault Originated Transients. *IEEE Transactions on Power Systems*, 23(2):380–388, May 2008.

- [64] N. I. Elkalashy, N. A. Sabiha, and M. Lehtonen. Earth Fault Distance Estimation Using Active Traveling Waves in Energized-Compensated MV Networks. *IEEE Transactions on Power Delivery*, 30(2):836–843, April 2015.
- [65] Amany M. El-Zonkoly. Fault Diagnosis in Distribution Networks with Distributed Generation. *Smart Grid and Renewable Energy*, 02(01):1–11, 2011.
- [66] Rui Liang, Guoqing Fu, Xueyuan Zhu, and Xue Xue. Fault location based on single terminal travelling wave analysis in radial distribution network. *International Journal of Electrical Power & Energy Systems*, 66:160–165, March 2015.
- [67] D. W. P. Thomas, R. J. O. Carvalho, and E. T. Pereira. Fault location in distribution systems based on traveling waves. In *2003 IEEE Bologna Power Tech Conference Proceedings*, volume 2, pages 5 pp. Vol.2–, June 2003.
- [68] S. Robson, A. Haddad, and H. Griffiths. Fault Location on Branched Networks Using a Multiended Approach. *IEEE Transactions on Power Delivery*, 29(4):1955–1963, August 2014.
- [69] H. Hizman, P. A. Crossley, P. F. Gale, and G. Bryson. Fault section identification and location on a distribution feeder using travelling waves. In *IEEE Power Engineering Society Summer Meeting*, volume 3, pages 1107–1112 vol.3, July 2002.
- [70] D. W. P. Thomas, R. J. O. Carvalho, E. T. Pereira, and C. Christopoulos. Field trial of fault location on a distribution system using high frequency transients. In *2005 IEEE Russia Power Tech*, pages 1–7, June 2005.
- [71] W. C. Santos, F. V. Lopes, N. S. D. Brito, and B. A. Souza. High-Impedance Fault Identification on Distribution Networks. *IEEE Transactions on Power Delivery*, 32(1):23–32, February 2017.
- [72] A. Borghetti, M. Bosetti, C. A. Nucci, M. Paolone, and A. Abur. Integrated Use of Time-Frequency Wavelet Decompositions for Fault Location in Distribution Networks: Theory and Experimental Validation. *IEEE Transactions on Power Delivery*, 25(4):3139–3146, October 2010.
- [73] A. Borghetti, S. Corsi, C. A. Nucci, M. Paolone, L. Peretto, and R. Tinarelli. On the use of continuous-wavelet transform for fault location in distribution power systems. *International Journal of Electrical Power & Energy Systems*, 28(9):608–617, November 2006.
- [74] M. Goudarzi, B. Vahidi, R. A. Naghizadeh, and S. H. Hosseini. Improved fault location algorithm for radial distribution systems with discrete and continuous wavelet analysis. *International Journal of Electrical Power & Energy Systems*, 67:423–430, May 2015.
- [75] S. Jamali and A. Bahmanyar. A new fault location method for distribution networks using sparse measurements. *International Journal of Electrical Power & Energy Systems*, 81:459–468, October 2016.

- [76] Cristian Grajales-Espinal, Juan Mora-Flórez, and Sandra Pérez-Londoño. Advanced fault location strategy for modern power distribution systems based on phase and sequence components and the minimum fault reactance concept. *Electric Power Systems Research*, 140:933–941, November 2016.
- [77] A. Bahmanyar, A. Estebarsari, E. Pons, E. Patti, S. Jamali, E. Bompard, and A. Acquaviva. Fast fault location for fast restoration of smart electrical distribution grids. In *2016 IEEE International Smart Cities Conference (ISC2)*, pages 1–6, September 2016.
- [78] F. C. L. Trindade, W. Freitas, and J. C. M. Vieira. Fault Location in Distribution Systems Based on Smart Feeder Meters. *IEEE Transactions on Power Delivery*, 29(1):251–260, February 2014.
- [79] S. M. Brahma. Fault Location in Power Distribution System With Penetration of Distributed Generation. *IEEE Transactions on Power Delivery*, 26(3):1545–1553, July 2011.
- [80] R. A. F. Pereira, L. G. W. da Silva, M. Kezunovic, and J. R. S. Mantovani. Improved Fault Location on Distribution Feeders Based on Matching During-Fault Voltage Sags. *IEEE Transactions on Power Delivery*, 24(2):852–862, April 2009.
- [81] Hazlie Mokhlis and Haiyu Li. Non-linear representation of voltage sag profiles for fault location in distribution networks. *International Journal of Electrical Power & Energy Systems*, 33(1):124–130, January 2011.
- [82] S. Lotfifard, M. Kezunovic, and M. J. Mousavi. Voltage Sag Data Utilization for Distribution Fault Location. *IEEE Transactions on Power Delivery*, 26(2):1239–1246, April 2011.
- [83] Arturo Suman Bretas, Luciano Pires, Miguel Moreto, and Rodrigo Hartstein Salim. A BP Neural Network Based Technique for HIF Detection and Location on Distribution Systems with Distributed Generation. In De-Shuang Huang, Kang Li, and George William Irwin, editors, *Computational Intelligence*, Lecture Notes in Computer Science, pages 608–613. Springer Berlin Heidelberg, 2006.
- [84] L. S. Martins, J. F. Martins, C. M. Alegria, and V. F. Pires. A network distribution power system fault location based on neural eigenvalue algorithm. In *2003 IEEE Bologna Power Tech Conference Proceedings*, volume 2, pages 6 pp. Vol.2–, June 2003.
- [85] F. Dehghani and H. Nezami. A new fault location technique on radial distribution systems using artificial neural network. In *22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013)*, pages 1–4, June 2013.
- [86] Zhongjian Kang, Aina Tian, and Yanyan Feng. A New Method for Fault Type Identification Based on HHT and Neural Network in Distribution Network. In Wenjiang Du, editor, *Informatics and Management Science IV*, Lecture Notes in Electrical Engineering, pages 187–194. Springer London, 2013.

- [87] M. Majidi, M. Etezadi-Amoli, and M. Sami Fadali. A Novel Method for Single and Simultaneous Fault Location in Distribution Networks. *IEEE Transactions on Power Systems*, 30(6):3368–3376, November 2015.
- [88] Hadi Zayandehroodi, Azah Mohamed, Hussain Shareef, and Masoud Farhoodnea. A novel neural network and backtracking based protection coordination scheme for distribution system with distributed generation. *International Journal of Electrical Power & Energy Systems*, 43(1):868–879, December 2012.
- [89] H. A. Darwish. A precise fault locator algorithm with a novel realization for MV distribution feeders. In *2006 IEEE Power Engineering Society General Meeting*, pages 8 pp.–, June 2006.
- [90] Yilmaz Aslan. An alternative approach to fault location on power distribution feeders with embedded remote-end power generation using artificial neural networks. *Electr Eng*, 94(3):125–134, September 2012.
- [91] J. Zhang, Z. Y. He, S. Lin, Y. B. Zhang, and Q. Q. Qian. An ANFIS-based fault classification approach in power distribution system. *International Journal of Electrical Power & Energy Systems*, 49:243–252, July 2013.
- [92] Ramón Perez, Carmen Vásquez, and Amelec Vilorio. An intelligent strategy for faults location in distribution networks with distributed generation. *IFS*, 36(2):1627–1637, March 2019.
- [93] D. Thukaram, H. P. Khincha, and H. P. Vijaynarasimha. Artificial neural network and support vector Machine approach for locating faults in radial distribution systems. *IEEE Transactions on Power Delivery*, 20(2):710–721, April 2005.
- [94] P. Janik and T. Lobos. Automated classification of power-quality disturbances using SVM and RBF networks. *IEEE Transactions on Power Delivery*, 21(3):1663–1669, July 2006.
- [95] J. Cosier, D. T. do Vale, and J. G. Rolim. Design and Training of Artificial Neural Networks for Locating Low Current Faults in Distribution Systems. In *2007 International Conference on Intelligent Systems Applications to Power Systems*, pages 1–6, November 2007.
- [96] S. A. M. Javadian, A. M. Nasrabadi, M. Haghifam, and J. Rezvantalab. Determining fault’s type and accurate location in distribution systems with DG using MLP Neural networks. In *2009 International Conference on Clean Electrical Power*, pages 284–289, June 2009.
- [97] Y. Dong, C. Zheng, and M. Kezunovic. Enhancing Accuracy While Reducing Computation Complexity for Voltage-Sag-Based Distribution Fault Location. *IEEE Transactions on Power Delivery*, 28(2):1202–1212, April 2013.
- [98] J. J. Mora, J. C. Bedoya, and J. Melendez. Extensive Events Database Development using ATP and Matlab to Fault Location in Power Distribution Systems. In *2006 IEEE/PES Transmission Distribution Conference and Exposition: Latin America*, pages 1–6, August 2006.

- [99] M. U. Usman, J. Ospina, and M. O. Faruque. Fault Classification and Location Identification in a Smart Distribution Network Using ANN. In *2018 IEEE Power Energy Society General Meeting (PESGM)*, pages 1–6, August 2018.
- [100] Dabit Sonoda, A. C. Zambroni de Souza, and Paulo Márcio da Silveira. Fault identification based on artificial immunological systems. *Electric Power Systems Research*, 156:24–34, March 2018.
- [101] S. M. Torabi. Fault location and classification in distribution systems using clark transformation and neural network. In *16th Electrical Power Distribution Conference*, pages 1–8, April 2011.
- [102] Q. Jin and R. Ju. Fault Location for Distribution Network Based on Genetic Algorithm and Stage Treatment. In *2012 Spring Congress on Engineering and Technology*, pages 1–4, May 2012.
- [103] J. C. S. Souza, M. A. P. Rodrigues, M. T. Schilling, and M. B. Do Coutto Filho. Fault location in electrical power systems using intelligent systems techniques. *IEEE Transactions on Power Delivery*, 16(1):59–67, January 2001.
- [104] Meshal A. Al-shaher, Manar M. Sabry, and Ahmad S. Saleh. Fault location in multi-ring distribution network using artificial neural network. *Electric Power Systems Research*, 64(2):87–92, February 2003.
- [105] A. Zapata-Tapasco, J. Mora-Flórez, and M. C. de Almeida. Fault location in power distribution systems using a learning approach based on decision trees. In *2014 IEEE PES Transmission Distribution Conference and Exposition - Latin America (PES T D-LA)*, pages 1–6, September 2014.
- [106] J. J. Mora, G. Carrillo, and L. Perez. Fault Location in Power Distribution Systems using ANFIS Nets and Current Patterns. In *2006 IEEE/PES Transmission Distribution Conference and Exposition: Latin America*, pages 1–6, August 2006.
- [107] Farzad Dehghani, Fereydoun Khodnia, and Esfandiar Dehghan. Fault location of unbalanced power distribution feeder with distributed generation using neural networks. In *CIREN - Open Access Proceedings Journal*, volume 2017, pages 1134–1137, October 2017.
- [108] Z. Galijasevic and A. Abur. Fault location using voltage measurements. *IEEE Transactions on Power Delivery*, 17(2):441–445, April 2002.
- [109] Prashant P. Bedekar, Sudhir R. Bhide, and Vijay S. Kale. Fault section estimation in power system using Hebb’s rule and continuous genetic algorithm. *International Journal of Electrical Power & Energy Systems*, 33(3):457–465, March 2011.
- [110] B. Das. Fuzzy logic-based fault-type identification in unbalanced radial power distribution system. *IEEE Transactions on Power Delivery*, 21(1):278–285, January 2006.

- [111] Patrick E. Farias, Adriano Peres de Moraes, Jean Pereira Rossini, and Ghendy Cardoso. Non-linear high impedance fault distance estimation in power distribution systems: A continually online-trained neural network approach. *Electric Power Systems Research*, 157:20–28, April 2018.
- [112] L. Sousa Martins, J. F. Martins, V. Fernao Pires, and C. M. Alegria. The application of neural networks and Clarke-Concordia transformation in fault location on distribution power systems. In *IEEE/PES Transmission and Distribution Conference and Exhibition*, volume 3, pages 2091–2095 vol.3, October 2002.
- [113] Shu Hongchun, Gongfa Li, Xia Qi, Wu Qinjin, and Tian Xincui. A fault location method of traveling wave for distribution network with only two-phase current transformer using artificial neural network. *2010 3rd International Congress on Image and Signal Processing*, 6:2942–2945, 2010.
- [114] A. A. Girgis and M. B. Johns. A hybrid expert system for faulted section identification, fault type classification and selection of fault location algorithms. *IEEE Transactions on Power Delivery*, 4(2):978–985, April 1989.
- [115] Ali Rafinia and Jamal Moshtagh. A new approach to fault location in three-phase underground distribution system using combination of wavelet analysis with ANN and FLS. *International Journal of Electrical Power & Energy Systems*, 55:261–274, February 2014.
- [116] J. A. Momoh, L. G. Dias, and D. N. Laird. An implementation of a hybrid intelligent tool for distribution system fault diagnosis. *IEEE Transactions on Power Delivery*, 12(2):1035–1040, April 1997.
- [117] A. Estebarsari, E. Pons, E. Bompard, A. Bahmanyar, and S. Jamali. An improved fault location method for distribution networks exploiting emerging LV smart meters. In *2016 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS)*, pages 1–6, June 2016.
- [118] Lei Ye, Dahai You, Xianggen Yin, Ke Wang, and Junchun Wu. An improved fault-location method for distribution system using wavelets and support vector regression. *International Journal of Electrical Power & Energy Systems*, 55:467–472, February 2014.
- [119] D. S. Gazzana, G. D. Ferreira, A. S. Bretas, A. L. Bettiol, A. Carniato, L. F. N. Passos, A. H. Ferreira, and J. E. M. Silva. An integrated technique for fault location and section identification in distribution systems. *Electric Power Systems Research*, 115:65–73, October 2014.
- [120] Fan Chunju, K. K. Li, W. L. Chan, Yu Weiyong, and Zhang Zhaoning. Application of wavelet fuzzy neural network in locating single line to ground fault (SLG) in distribution lines. *International Journal of Electrical Power & Energy Systems*, 29(6):497–503, July 2007.
- [121] A. Ngaopitakkul, C. Pothisarn, S. Bunjongjit, and B. Suechoey. DWT and RBF neural networks algorithm for identifying the fault types in underground cable. In *TENCON 2011 - 2011 IEEE Region 10 Conference*, pages 1379–1382, November 2011.

- [122] Teke Gush, Syed Basit Ali Bukhari, Raza Haider, Samuel Admasie, Yun-Sik Oh, Gyu-Jung Cho, and Chul-Hwan Kim. Fault detection and location in a microgrid using mathematical morphology and recursive least square methods. *International Journal of Electrical Power & Energy Systems*, 102:324–331, November 2018.
- [123] Fault identification in electrical power distribution system using combined discrete wavelet transform and fuzzy logic. *Journal of Electrical Systems and Information Technology*, 2(2):257–267, September 2015.
- [124] J. Mora-Florez, V. Barrera-Nunez, and G. Carrillo-Caicedo. Fault Location in Power Distribution Systems Using a Learning Algorithm for Multivariable Data Analysis. *IEEE Transactions on Power Delivery*, 22(3):1715–1721, July 2007.
- [125] L. Peretto, R. Sasdelli, E. Scala, and R. Tinarelli. Fault Location Method Integrating a Distributed Measurement System and Wavelet Analysis. In *2007 IEEE Instrumentation Measurement Technology Conference IMTC 2007*, pages 1–6, May 2007.
- [126] R. H. Salim, K. R. C. de Oliveira, A. D. Filomena, M. Resener, and A. S. Bretas. Hybrid Fault Diagnosis Scheme Implementation for Power Distribution Systems Automation. *IEEE Transactions on Power Delivery*, 23(4):1846–1856, October 2008.
- [127] Wen-Hui Chen, Chih-Wen Liu, and Men-Shen Tsai. On-line fault diagnosis of distribution substations using hybrid cause-effect network and fuzzy rule-based method. *IEEE Transactions on Power Delivery*, 15(2):710–717, April 2000.
- [128] M. Pourahmadi-Nakhli and A. A. Safavi. Path Characteristic Frequency-Based Fault Locating in Radial Distribution Systems Using Wavelets and Neural Networks. *IEEE Transactions on Power Delivery*, 26(2):772–781, April 2011.
- [129] G. Niu, L. Zhou, W. Pei, and Z. Qi. A novel fault location and recognition method for low voltage active distribution network. In *2015 5th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*, pages 876–881, November 2015.
- [130] K. Sun, Q. Chen, and Z. Gao. An Automatic Faulted Line Section Location Method for Electric Power Distribution Systems Based on Multisource Information. *IEEE Transactions on Power Delivery*, 31(4):1542–1551, August 2016.
- [131] A. M. Pasdar, Y. Sozer, and I. Husain. Detecting and Locating Faulty Nodes in Smart Grids Based on High Frequency Signal Injection. *IEEE Transactions on Smart Grid*, 4(2):1067–1075, June 2013.
- [132] L. Marques, N. Silva, I. Miranda, E. Rodrigues, and H. Leite. Detection and localisation of non-technical losses in low voltage distribution networks. In *Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MedPower 2016)*, pages 1–8, November 2016.



- [133] G. A. Orcajo, J. M. Cano, M. G. Melero, M. F. Cabanas, C. H. Rojas, J. F. Pedrayes, and J. G. Norniella. Diagnosis of Electrical Distribution Network Short Circuits Based on Voltage Park's Vector. *IEEE Transactions on Power Delivery*, 27(4):1964–1972, October 2012.
- [134] N. Silva, F. Basadre, P. Rodrigues, M. S. Nunes, A. Grilo, A. Casaca, F. Melo, and L. Gaspar. Fault detection and location in Low Voltage grids based on distributed monitoring. In *2016 IEEE International Energy Conference (ENERGYCON)*, pages 1–6, April 2016.
- [135] M. M. Alamuti, H. Nouri, R. M. Ciric, and V. Terzija. Intermittent Fault Location in Distribution Feeders. *IEEE Transactions on Power Delivery*, 27(1):96–103, January 2012.
- [136] Antonio Luque López and Steven Hegedus, editors. *Handbook of Photovoltaic Science and Engineering*. Wiley, Chichester, 2. ed., [fully rev. and updated] edition, 2011. OCLC: 845812578.
- [137] Efstratios Theodosiou, E Danezis, Vassilios Manimanis, and E.-M Kalyva. From Phytagoreans to Kepler: The dispute between the geocentric and the heliocentric systems. *Journal of Astronomical History and Heritage*, January 2002.
- [138] U.S. Naval research Labolatory. Vanguard | Space Systems Development Department. <https://www.nrl.navy.mil/ssdd/heritage/vanguard>.
- [139] Damien Picault. *Reduction of Mismatch Losses in Grid-Connected Photovoltaic Systems Using Alternative Topologies*. PhD Thesis, Institut National Polytechnique de Grenoble - INPG, October 2010.
- [140] W. De Soto, S. A. Klein, and W. A. Beckman. Improvement and validation of a model for photovoltaic array performance. *Solar Energy*, 80(1):78–88, January 2006.
- [141] Long Bun. *Détection et localisation de défauts pour un système PV*. PhD Thesis, Université Grenoble Alpes, November 2011.
- [142] J. W. Bishop. Computer simulation of the effects of electrical mismatches in photovoltaic cell interconnection circuits. *Solar Cells*, 25(1):73–89, October 1988.
- [143] E. I. Batzelis and S. A. Papathanassiou. A Method for the Analytical Extraction of the Single-Diode PV Model Parameters. *IEEE Transactions on Sustainable Energy*, 7(2):504–512, April 2016.
- [144] Albert Yaw Appiah, Xinghua Zhang, Ben Beklisi Kwame Ayawli, and Frimpong Kyeremeh. Review and Performance Evaluation of Photovoltaic Array Fault Detection and Diagnosis Techniques. *International Journal of Photoenergy*, 2019:19, February 2019.
- [145] R. Hariharan, M. Chakkarapani, G. Saravana Ilango, and C. Nagamani. A Method to Detect Photovoltaic Array Faults and Partial Shading in PV Systems. *IEEE Journal of Photovoltaics*, 6(5):1278–1285, September 2016.

- [146] M. Bressan, Y. El-Basri, and C. Alonso. A new method for fault detection and identification of shadows based on electrical signature of defects. In *2015 17th European Conference on Power Electronics and Applications (EPE'15 ECCE-Europe)*, pages 1–8, September 2015.
- [147] Y H Chen, R Liang, Y Tian, and F Wang. A novel fault diagnosis method of PV based-on power loss and I-V characteristics. *IOP Conference Series: Earth and Environmental Science*, 40:012022, August 2016.
- [148] K. A. Klise and J. S. Stein. Automated performance monitoring for PV systems using pecos. In *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*, pages 3431–3435, June 2016.
- [149] Santiago Silvestre, Aissa Chouder, and Engin Karatepe. Automatic fault detection in grid connected PV systems. *Solar Energy*, 94:119–127, August 2013.
- [150] A. Chouder and S. Silvestre. Automatic supervision and fault detection of PV systems based on power losses analysis. *Energy Conversion and Management*, 51(10):1929–1937, October 2010.
- [151] J. Chen, C. Kuo, S. Chen, C. Kao, T. Zhan, C. Lin, and Y. Chen. DC-side fault detection for photovoltaic energy conversion system using fractional-order dynamic-error-based fuzzy Petri net integrated with intelligent meters. *IET Renewable Power Generation*, 10(9):1318–1327, 2016.
- [152] A. Umana and A. P. S. Meliopoulos. Detection of cell-level fault conditions within a photovoltaic array system. In *2016 IEEE/PES Transmission and Distribution Conference and Exposition (T D)*, pages 1–5, May 2016.
- [153] T. Shimakage, K. Nishioka, H. Yamane, M. Nagura, and M. Kudo. Development of fault detection system in PV system. In *2011 IEEE 33rd International Telecommunications Energy Conference (INTELEC)*, pages 1–5, October 2011.
- [154] M. Miwa, S. Yamanaka, H. Kawamura, H. Ohno, and H. Kawamura. Diagnosis of a Power Output Lowering of PV Array with a  $(-dI/dV)$ -V Characteristic. In *2006 IEEE 4th World Conference on Photovoltaic Energy Conference*, volume 2, pages 2442–2445, May 2006.
- [155] E. Garoudja, K. Kara, A. Chouder, S. Silvestre, and S. Kichou. Efficient fault detection and diagnosis procedure for photovoltaic systems. In *2016 8th International Conference on Modelling, Identification and Control (ICMIC)*, pages 851–856, November 2016.
- [156] Santiago Silvestre Bergés and Aissa Chouder. Fault detection and automatic supervision methodology for PV systems. 2010.
- [157] W. Chine, A. Mellit, A. Massi Pavan, and S.A. Kalogirou. Fault detection method for grid-connected photovoltaic plants. *Renewable Energy*, 66:99–110, June 2014.
- [158] P. Jain, J. X. Xu, S. K. Panda, J. Poon, C. Spanos, and S. R. Sanders. Fault diagnosis via PV panel-integrated power electronics. In *2016 IEEE 17th Workshop on Control and Modeling for Power Electronics (COMPEL)*, pages 1–6, June 2016.

- [159] J. E. Quiroz, J. S. Stein, C. K. Carmignani, and K. Gillispie. In-situ module-level I-V tracers for novel PV monitoring. In *2015 IEEE 42nd Photovoltaic Specialist Conference (PVSC)*, pages 1–6, June 2015.
- [160] Aissa Chouder, Santiago Silvestre, Nawel Sadaoui, and Lazhar Rahmani. Modeling and simulation of a grid connected PV system based on the evaluation of main PV module parameters. *Simulation Modelling Practice and Theory*, 20(1):46–58, January 2012.
- [161] Aissa Chouder, Santiago Silvestre, Bilal Taghezouit, and Engin Karatepe. Monitoring, modelling and simulation of PV systems using LabVIEW. *Solar Energy*, 91:337–349, May 2013.
- [162] M. Davarifar, A. Rabhi, A. El Hajjaji, and M. Dahmane. New method for fault detection of PV panels in domestic applications. In *3rd International Conference on Systems and Control*, pages 727–732, October 2013.
- [163] Santiago Silvestre, Mário Aires da Silva, Aissa Chouder, Daniel Guasch, and Engin Karatepe. New procedure for fault detection in grid connected PV systems based on the evaluation of current and voltage indicators. *Energy Conversion and Management*, 86:241–249, October 2014.
- [164] S. J. Devi, B. Umamaheswari, and R. Parthiban. Novel parametric tolerance based fault identification technique for PV system. In *2016 IEEE 25th International Symposium on Industrial Electronics (ISIE)*, pages 568–573, June 2016.
- [165] X. Lin, Y. Wang, D. Zhu, N. Chang, and M. Pedram. Online fault detection and tolerance for photovoltaic energy harvesting systems. In *2012 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pages 1–6, November 2012.
- [166] R. Platon, J. Martel, N. Woodruff, and T. Y. Chau. Online Fault Detection in PV Systems. *IEEE Transactions on Sustainable Energy*, 6(4):1200–1207, October 2015.
- [167] C. Kuo, J. Chen, S. Chen, C. Kao, H. Yau, and C. Lin. Photovoltaic Energy Conversion System Fault Detection Using Fractional-Order Color Relation Classifier in Microdistribution Systems. *IEEE Transactions on Smart Grid*, 8(3):1163–1172, May 2017.
- [168] Mohamed Hassan Ali, Abdelhamid Rabhi, Ahmed El Hajjaji, and Giuseppe M. Tina. Real Time Fault Detection in Photovoltaic Systems. *Energy Procedia*, 111:914–923, March 2017.
- [169] B. Andò, S. Baglio, A. Pistorio, G. M. Tina, and C. Ventura. Sentinella: Smart Monitoring of Photovoltaic Systems at Panel Level. *IEEE Transactions on Instrumentation and Measurement*, 64(8):2188–2199, August 2015.
- [170] Nuri Gokmen, Engin Karatepe, Berk Celik, and Santiago Silvestre. Simple diagnostic approach for determining of faulted PV modules in string based PV arrays. *Solar Energy*, 86(11):3364–3377, November 2012.

- [171] G. Liu and W. Yu. A fault detection and diagnosis technique for solar system based on Elman neural network. In *2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pages 473–480, December 2017.
- [172] Zhehan Yi and A. H. Etemadi. A novel detection algorithm for Line-to-Line faults in Photovoltaic (PV) arrays based on support vector machine (SVM). In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–4, July 2016.
- [173] W. Rezgui, L. Mouss, N. K. Mouss, M. D. Mouss, and M. Benbouzid. A smart algorithm for the diagnosis of short-circuit faults in a photovoltaic generator. In *2014 First International Conference on Green Energy ICGE 2014*, pages 139–143, March 2014.
- [174] Elyes Garoudja, Aissa Chouder, Kamel Kara, and Santiago Silvestre. An enhanced machine learning based approach for failures detection and diagnosis of PV systems. *Energy Conversion and Management*, 151:496–513, November 2017.
- [175] Yue Wu, Zhicong Chen, Lijun Wu, Peijie Lin, Shuying Cheng, and Peimin Lu. An Intelligent Fault Diagnosis Approach for PV Array Based on SA-RBF Kernel Extreme Learning Machine. *Energy Procedia*, 105:1070–1076, May 2017.
- [176] P. Ducange, M. Fazzolari, B. Lazzerini, and F. Marcelloni. An intelligent system for detecting faults in photovoltaic fields. In *2011 11th International Conference on Intelligent Systems Design and Applications*, pages 1341–1346, November 2011.
- [177] Z. Wang and R. S. Balog. Arc fault and flash detection in photovoltaic systems using wavelet transform and support vector machines. In *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*, pages 3275–3280, June 2016.
- [178] H. Mekki, A. Mellit, and H. Salhi. Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules. *Simulation Modelling Practice and Theory*, 67:1–13, September 2016.
- [179] C. B. Jones, M. Martínez-Ramón, R. Smith, C. K. Carmignani, O. Lavrova, C. Robinson, and J. S. Stein. Automatic fault classification of photovoltaic strings based on an in situ IV characterization system and a Gaussian process algorithm. In *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*, pages 1708–1713, June 2016.
- [180] Syafaruddin, E. Karatepe, and T. Hiyama. Controlling of artificial neural network for fault diagnosis of photovoltaic array. In *2011 16th International Conference on Intelligent System Applications to Power Systems*, pages 1–6, September 2011.
- [181] Y. Zhao, L. Yang, B. Lehman, J. F. de Palma, J. Mosesian, and R. Lyons. Decision tree-based fault detection and classification in solar photovoltaic arrays. In *2012 Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*, pages 93–99, February 2012.

- [182] Xiao Li, Pu Yang, Jiangfan Ni, and Jing Zhao. Fault diagnostic method for PV array based on improved wavelet neural network algorithm. In *Proceeding of the 11th World Congress on Intelligent Control and Automation*, pages 1171–1175, June 2014.
- [183] Y. Zhao, F. Balboni, T. Arnaud, J. Mosesian, R. Ball, and B. Lehman. Fault experiments in a commercial-scale PV laboratory and fault detection using local outlier factor. In *2014 IEEE 40th Photovoltaic Specialist Conference (PVSC)*, pages 3398–3403, June 2014.
- [184] Y. Zhao, R. Ball, J. Mosesian, J. F. de Palma, and B. Lehman. Graph-Based Semi-supervised Learning for Fault Detection and Classification in Solar Photovoltaic Arrays. *IEEE Transactions on Power Electronics*, 30(5):2848–2858, May 2015.
- [185] Zhicong Chen, Lijun Wu, Shuying Cheng, Peijie Lin, Yue Wu, and Wencheng Lin. Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and I-V characteristics. *Applied Energy*, 204:912–931, October 2017.
- [186] M. N. Akram and S. Lotfifard. Modeling and Health Monitoring of DC Side of Photovoltaic Array. *IEEE Transactions on Sustainable Energy*, 6(4):1245–1253, October 2015.
- [187] A. Belaout, F. Krim, A. Sahli, and A. Mellit. Multi-class neuro-fuzzy classifier for photovoltaic array faults diagnosis. In *2017 5th International Conference on Electrical Engineering - Boumerdes (ICEE-B)*, pages 1–4, October 2017.
- [188] A. Belaout, F. Krim, A. Mellit, B. Talbi, and A. Arabi. Multiclass adaptive neuro-fuzzy classifier and feature selection techniques for photovoltaic array fault detection and classification. *Renewable Energy*, 127:548–558, November 2018.
- [189] A. Mellit and S. A. Kalogirou. Neuro-Fuzzy Based Modeling for Photovoltaic Power Supply System. In *2006 IEEE International Power and Energy Conference*, pages 88–93, Putra Jaya, Malaysia, November 2006. IEEE.
- [190] C. B. Jones, J. S. Stein, S. Gonzalez, and B. H. King. Photovoltaic system fault detection and diagnostics using Laterally Primed Adaptive Resonance Theory neural network. In *2015 IEEE 42nd Photovoltaic Specialist Conference (PVSC)*, pages 1–6, June 2015.
- [191] Z. Wang, S. McConnell, R. S. Balog, and J. Johnson. Arc fault signal detection - Fourier transformation vs. wavelet decomposition techniques using synthesized data. In *2014 IEEE 40th Photovoltaic Specialist Conference (PVSC)*, pages 3239–3244, June 2014.
- [192] I. S. Kim. Fault detection algorithm of the photovoltaic system using wavelet transform. In *India International Conference on Power Electronics 2010 (IICPE2010)*, pages 1–6, January 2011.

- [193] T. Takashima, J. Yamaguchi, and M. Ishida. Fault detection by signal response in PV module strings. In *2008 33rd IEEE Photovoltaic Specialists Conference*, pages 1–5, May 2008.
- [194] M. González, B. Raison, S. Bacha, and Long Bun. Fault diagnosis in a grid-connected photovoltaic system by applying a signal approach. In *IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society*, pages 1354–1359, November 2011.
- [195] M. S. Agamy, M. Harfman-Todorovic, and A. Elasser. Ground fault and insulation degradation detection and localization in PV plants. In *2013 IEEE 39th Photovoltaic Specialists Conference (PVSC)*, pages 2840–2844, June 2013.
- [196] M. Davarifar, A. Rabhi, A. Hajjaji, E. Kamal, and Z. Daneshifar. Partial shading fault diagnosis in PV system with discrete wavelet transform (DWT). In *2014 International Conference on Renewable Energy Research and Application (ICRERA)*, pages 810–814, October 2014.
- [197] E. Jamshidpour, P. Poure, and S. Saadate. Photovoltaic Systems Reliability Improvement by Real-Time FPGA-Based Switch Failure Diagnosis and Fault-Tolerant DC–DC Converter. *IEEE Transactions on Industrial Electronics*, 62(11):7247–7255, November 2015.
- [198] S. Chen and X. Li. PV series arc fault recognition under different working conditions with joint detection method. In *2016 IEEE 62nd Holm Conference on Electrical Contacts (Holm)*, pages 25–32, October 2016.
- [199] L. Chen, S. Li, and X. Wang. Quickest Fault Detection in Photovoltaic Systems. *IEEE Transactions on Smart Grid*, 9(3):1835–1847, May 2018.
- [200] M. Davarifar, A. Rabhi, A. El-Hajjaji, and M. Dahmane. Real-time model base fault diagnosis of PV panels using statistical signal processing. In *2013 International Conference on Renewable Energy Research and Applications (ICRERA)*, pages 599–604, October 2013.
- [201] K. Satpathi, Y. M. Yeap, A. Ukil, and N. Gedddada. Short-Time Fourier Transform Based Transient Analysis of VSC Interfaced Point-to-Point DC System. *IEEE Transactions on Industrial Electronics*, PP(99):1–1, 2017.
- [202] S. Buddha, H. Braun, V. Krishnan, C. Tepedelenlioglu, A. Spanias, T. Yeider, and T. Takehara. Signal processing for photovoltaic applications. In *2012 IEEE International Conference on Emerging Signal Processing Applications*, pages 115–118, January 2012.
- [203] N. L. Georgijevic, M. V. Jankovic, S. Srdic, and Z. Radakovic. The Detection of Series Arc Fault in Photovoltaic Systems Based on the Arc Current Entropy. *IEEE Transactions on Power Electronics*, 31(8):5917–5930, August 2016.
- [204] S. Roy, M. K. Alam, F. Khan, J. Johnson, and J. Flicker. An Irradiance-Independent, Robust Ground-Fault Detection Scheme for PV Arrays Based on Spread Spectrum Time-Domain Reflectometry (SSTDTR). *IEEE Transactions on Power Electronics*, 33(8):7046–7057, August 2018.

- [205] T. Takashima, J. Yamaguchi, K. Otani, K. Kato, and M. Ishida. Experimental Studies of Failure Detection Methods in PV Module Strings. In *2006 IEEE 4th World Conference on Photovoltaic Energy Conference*, volume 2, pages 2227–2230, May 2006.
- [206] J. Johnson, B. Pahl, C. Luebke, T. Pier, T. Miller, J. Strauch, S. Kuszmaul, and W. Bower. Photovoltaic DC Arc Fault Detector testing at Sandia National Laboratories. In *2011 37th IEEE Photovoltaic Specialists Conference*, pages 003614–003619, June 2011.
- [207] M. K. Alam, F. H. Khan, J. Johnson, and J. Flicker. PV arc-fault detection using spread spectrum time domain reflectometry (SSTD). In *2014 IEEE Energy Conversion Congress and Exposition (ECCE)*, pages 3294–3300, September 2014.
- [208] M. K. Alam, F. Khan, J. Johnson, and J. Flicker. PV ground-fault detection using spread spectrum time domain reflectometry (SSTD). In *2013 IEEE Energy Conversion Congress and Exposition*, pages 1015–102, September 2013.
- [209] E. Garoudja, F. Harrou, Y. Sun, K. Kara, A. Chouder, and S. Silvestre. A statistical-based approach for fault detection and diagnosis in a photovoltaic system. In *2017 6th International Conference on Systems and Control (ICSC)*, pages 75–80, May 2017.
- [210] S. Vergura, G. Acciani, V. Amoroso, and G. Patrono. Inferential statistics for monitoring and fault forecasting of PV plants. In *2008 IEEE International Symposium on Industrial Electronics*, pages 2414–2419, June 2008.
- [211] Mahmoud Dhimish, Violeta Holmes, and Mark Dales. Parallel fault detection algorithm for grid-connected photovoltaic plants. *Renewable Energy*, 113(Supplement C):94–111, December 2017.
- [212] M. Davarifar, A. Rabhi, A. Hajjaji, and Z. Daneshifar. Real-time diagnosis of PV system by using the Sequential Probability Ratio Test (SPRT). In *2014 16th International Power Electronics and Motion Control Conference and Exposition*, pages 508–513, September 2014.
- [213] Fouzi Harrou, Ying Sun, Bilal Taghezouit, Ahmed Saidi, and Mohamed-Elkarim Hamlati. Reliable fault detection and diagnosis of photovoltaic systems based on statistical monitoring approaches. *Renewable Energy*, 116:22–37, February 2018.
- [214] M. Dhimish, V. Holmes, B. Mehrdadi, and M. Dales. Simultaneous fault detection algorithm for grid-connected photovoltaic plants. *IET Renewable Power Generation*, 11(12):1565–1575, 2017.
- [215] Elyes Garoudja, Fouzi Harrou, Ying Sun, Kamel Kara, Aissa Chouder, and Santiago Silvestre. Statistical fault detection in photovoltaic systems. *Solar Energy*, 150:485–499, July 2017.
- [216] S Wang, X N Zhang, D D Gao, H X Liu, J Ye, and L R Li. A fault diagnosis system for PV power station based on global partitioned gradually

- approximation method. *IOP Conference Series: Earth and Environmental Science*, 40:012030, August 2016.
- [217] S.K. Firth, K.J. Lomas, and S.J. Rees. A simple model of PV system performance and its use in fault detection. *Solar Energy*, 84(4):624–635, April 2010.
- [218] S. Dhar and P. K. Dash. Differential current-based fault protection with adaptive threshold for multiple PV-based DC microgrid. *IET Renewable Power Generation*, 11(6):778–790, 2017.
- [219] F. Ancuta and C. Cepisca. Fault analysis possibilities for PV panels. In *Proceedings of the 2011 3rd International Youth Conference on Energetics (IYCE)*, pages 1–5, July 2011.
- [220] Y. Hu, W. Cao, J. Ma, S. J. Finney, and D. Li. Identifying PV Module Mismatch Faults by a Thermography-Based Temperature Distribution Analysis. *IEEE Transactions on Device and Materials Reliability*, 14(4):951–960, December 2014.
- [221] Mohammed Tadj, Khalil Benmouiza, Ali Cheknane, and Santiago Silvestre. Improving the performance of PV systems by faults detection using GISTEL approach. *Energy Conversion and Management*, 80:298–304, April 2014.
- [222] L. Tang and B. T. Ooi. Locating and Isolating DC Faults in Multi-Terminal DC Systems. *IEEE Transactions on Power Delivery*, 22(3):1877–1884, July 2007.
- [223] A. Drews, A.C. de Keizer, H.G. Beyer, E. Lorenz, J. Betteke, W.G.J.H.M. van Sark, W. Heydenreich, E. Wiemken, S. Stettler, P. Toggweiler, S. Bofinger, M. Schneider, G. Heilscher, and D. Heinemann. Monitoring and remote failure detection of grid-connected PV systems based on satellite observations. *Solar Energy*, 81(4):548–564, April 2007.
- [224] Y. Hu, J. Zhang, W. Cao, J. Wu, G. Y. Tian, S. J. Finney, and J. L. Kirtley. Online Two-Section PV Array Fault Diagnosis With Optimized Voltage Sensor Locations. *IEEE Transactions on Industrial Electronics*, 62(11):7237–7246, November 2015.
- [225] Yanli Liu, Bingfeng Li, and Ze Cheng. Research on PV module structure based on fault detection. In *2010 Chinese Control and Decision Conference*, pages 3891–3895, May 2010.
- [226] Moath Alsafasfeh, Ikhlas Abdel-Qader, Bradley Bazuin, Qais Alsafasfeh, and Wencong Su. Unsupervised Fault Detection and Analysis for Large Photovoltaic Systems Using Drones and Machine Vision. *Energies*, 11(9):2252, August 2018.
- [227] R. A. Walling, R. Saint, R. C. Dugan, J. Burke, and L. A. Kojovic. Summary of Distributed Resources Impact on Power Delivery Systems. *IEEE Transactions on Power Delivery*, 23(3):1636–1644, July 2008.
- [228] Voltage characteristics of electricity supplied by public electricity networks. Technical Report EN 50160:2010 F, European Committee for Electrotechnical Standardization [CENELEC], Brussels, Belgium, July 2010.



- [229] IEEE Recommended Practice for Monitoring Electric Power Quality. *IEEE Std 1159-2009 (Revision of IEEE Std 1159-1995)*, pages c1–81, June 2009.
- [230] Mohamed El Khatib, Javier Hernandez Alvidrez, and Abraham Ellis. Fault Analysis and Detection in Microgrids with High PV Penetration. Technical Report SAND2017-5472, 1367437, May 2017.
- [231] Teresa Escobet, Anibal Bregon, Belarmino Pulido, and Vicenç Puig, editors. *Fault Diagnosis of Dynamic Systems: Quantitative and Qualitative Approaches*. Springer International Publishing, 2019.
- [232] Yu Qin, Minghao Wen, Junchao Zheng, and Yu Bai. A novel distance protection scheme for HVDC lines based on R-L model. *International Journal of Electrical Power & Energy Systems*, 100:167–177, September 2018.
- [233] Murari Mohan Saha, Jan Izykowski, and Eugeniusz Rosołowski. *Fault Location on Power Networks*. Power Systems. Springer, London, 2010. OCLC: ocn495597133.
- [234] Luís Marques. *Detection and Location of Non-Technical Losses in Low Voltage Distribution Networks*. Master Thesis, Faculty of Engineering of the University of Porto, Porto, July 2016.
- [235] C.R. Bayliss and B.J. Hardy. *Transmission and Distribution Electrical Engineering*. Newnes, Oxford, 4th edition, 2012. OCLC: 793076688.
- [236] C. J. Ziser, Z. Dong, and T. K. Saha. Probabilistic Modelling of Demand Diversity and its Relationship with Electricity Market Outcomes. In *2007 IEEE Power Engineering Society General Meeting*, pages 1–6, June 2007.
- [237] Q. Cetina, R. A. J. Roscoe, and P. S. Wright. Challenges for Smart Electricity Meters due to Dynamic Power Quality Conditions of the Grid: A Review. In *2017 IEEE International Workshop on Applied Measurements for Power Systems (AMPS)*, pages 1–6, September 2017.
- [238] N. Sapountzoglou, B. Raison, and N. Silva. Fault Detection and Localization in LV Smart Grids. In *2019 IEEE Milan PowerTech*, pages 1–6, June 2019.
- [239] Jesus Lago, Fjo De Ridder, and Bart De Schutter. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Applied Energy*, 221:386–405, July 2018.
- [240] Jesus Lago, Karel De Brabandere, Fjo De Ridder, and Bart De Schutter. Short-term forecasting of solar irradiance without local telemetry: A generalized model using satellite data. *Solar Energy*, 173:566–577, October 2018.
- [241] Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, pages 785–794, San Francisco, California, USA, 2016. ACM Press.

- [242] Trevor Hastie, Jerome Friedman, and Robert Tibshirani. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York, New York, NY, 2001.
- [243] James S. Bergstra, Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. Algorithms for Hyper-Parameter Optimization. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24*, pages 2546–2554. Curran Associates, Inc., 2011.
- [244] Mihai Sanduleac, Gianluca Lipari, Antonello Monti, Artemis Voulkidis, Gianluca Zanetto, Antonello Corsi, Lucian Toma, Giampaolo Fiorentino, and Dumitru Federenciuc. Next Generation Real-Time Smart Meters for ICT Based Assessment of Grid Data Inconsistencies. *Energies*, 10(7):857, June 2017.
- [245] J. Duplex, S. Gossweiler, and S. Fagnoni. A better knowledge of electricity consumption for residential customers through the Linky smart meter. In *22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013)*, pages 1–4, June 2013.
- [246] Imtiaz Parvez, Arif I. Sarwat, Longfei Wei, and Aditya Sundararajan. Securing Metering Infrastructure of Smart Grid: A Machine Learning and Localization Based Key Management Approach. *Energies*, 9(9):691, September 2016.
- [247] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.
- [248] Jesus Lago, Fjo De Ridder, Peter Vrancx, and Bart De Schutter. Forecasting day-ahead electricity prices in Europe: The importance of considering market integration. *Applied Energy*, 211:890–903, February 2018.
- [249] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. *arXiv:1412.6980 [cs]*, December 2014.
- [250] I. Kim. The effect of load current on a three-phase fault. In *2016 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–4, September 2016.
- [251] M. Pignati, L. Zanni, P. Romano, R. Cherkaoui, and M. Paolone. Fault Detection and Faulted Line Identification in Active Distribution Networks Using Synchrophasors-Based Real-Time State Estimation. *IEEE Transactions on Power Delivery*, 32(1):381–392, February 2017.
- [252] L. Dunn, M. Gostein, and K. Emery. Comparison of pyranometers vs. PV reference cells for evaluation of PV array performance. In *2012 38th IEEE Photovoltaic Specialists Conference*, pages 002899–002904, June 2012.
- [253] M. E. Baran and I. El-Markaby. Fault analysis on distribution feeders with distributed generators. *IEEE Transactions on Power Systems*, 20(4):1757–1764, November 2005.

- [254] Andreas Varga. *Solving Fault Diagnosis Problems*, volume 84 of *Studies in Systems, Decision and Control*. Springer International Publishing, Cham, 2017.
- [255] T. S. Sidhu and D. Bejmert. Short-circuit current contribution from large scale PV power plant in the context of distribution power system protection performance. In *IET Conference on Renewable Power Generation (RPG 2011)*, pages 1–6, September 2011.
- [256] H. Hooshyar and M. E. Baran. Fault Analysis on Distribution Feeders With High Penetration of PV Systems. *IEEE Transactions on Power Systems*, 28(3):2890–2896, August 2013.
- [257] Evaluation of measurement data - Guide to the expression of uncertainty in measurement. Technical Report JCGM 100:2008, Joint Committee for Guides in Metrology, September 2008.
- [258] Geoffrey Auran. *Full Selective Protection Strategy for Multi-Terminal Cable HVDC Grids Based on HB-MMC Converters*. PhD thesis, Université Grenoble Alpes, October 2017.
- [259] Frano Barbir. *PEM Fuel Cells: Theory and Practice*. Elsevier/Academic Press, Amsterdam ; Boston, 2nd ed edition, 2013.
- [260] Z. Zheng, R. Petrone, M. C. Péra, D. Hissel, M. Becherif, C. Pianese, N. Yousfi Steiner, and M. Sorrentino. A review on non-model based diagnosis methodologies for PEM fuel cell stacks and systems. *International Journal of Hydrogen Energy*, 38(21):8914–8926, July 2013.
- [261] R. I. Salim, H. Noura, and A. Fardoun. A review on fault diagnosis tools of the proton exchange Membrane Fuel Cell. In *2013 Conference on Control and Fault-Tolerant Systems (SysTol)*, pages 686–693, October 2013.
- [262] N. Sapountzoglou and B. Raison. A Grid Connected PV System Fault Diagnosis Method. In *2019 IEEE International Conference on Industrial Technology (ICIT)*, pages 977–982, February 2019.
- [263] A. Gebregergis, P. Pillay, and R. Rengaswemy. PEMFC Fault Diagnosis, Modeling, and Mitigation. In *2008 IEEE Industry Applications Society Annual Meeting*, pages 1–8, October 2008.
- [264] N. Yousfi-Steiner, Ph. Moçotéguy, D. Candusso, D. Hissel, A. Hernandez, and A. Aslanides. A review on PEM voltage degradation associated with water management: Impacts, influent factors and characterization. *Journal of Power Sources*, 183(1):260–274, August 2008.

# Appendix A

## LV grid characteristics

Table A.1 Characteristics of the eleven different types of conductors connecting the nodes of the grid.

From Node	To Node	R ( $\Omega$ )	X ( $\Omega$ )	R/X	$l$ (m)	Cable type
3	6	0.641	0.15	4.273	120	1
1	3	0.476	0.10	4.760	40	2
2	5	0.476	0.10	4.760	65	
12	19	0.476	0.10	4.760	80	
19	26	0.476	0.10	4.760	50	
1	2	0.667	0.10	6.670	85	3
1	4	0.667	0.10	6.670	55	
3	7	0.667	0.10	6.670	105	
4	8	0.667	0.10	6.670	100	
5	9	0.667	0.10	6.670	70	
7	13	0.667	0.10	6.670	35	
9	16	0.667	0.10	6.670	35	
8	15	1.910	0.15	12.733	65	4
11	18	1.910	0.15	12.733	50	
13	20	1.910	0.15	12.733	80	
15	23	1.910	0.15	12.733	140	

Table A.2 Cont. characteristics of the eleven different types of conductors connecting the nodes of the grid.

From Node	To Node	R ( $\Omega$ )	X ( $\Omega$ )	R/X	$l$ (m)	Cable type
16	24	1.330	0.10	13.300	35	5
26	30	1.330	0.10	13.300	40	
5	10	2.970	0.15	19.800	35	6
8	14	3.060	0.15	20.400	65	7
27	31	3.060	0.15	20.400	70	
5	11	2.080	0.10	20.800	105	8
6	12	2.080	0.10	20.800	140	
20	27	2.080	0.10	20.800	90	
13	21	4.610	0.15	30.733	105	9
18	25	4.610	0.15	30.733	35	
24	29	4.610	0.15	30.733	40	
28	32	4.610	0.15	30.733	70	
31	33	4.610	0.15	30.733	35	
23	28	4.450	0.10	44.500	35	10
11	17	7.130	0.15	47.533	170	11
14	22	7.130	0.15	47.533	210	

Table A.3 Characteristics of the different types of conductors per branch.

Branch	From Node	To Node	R ( $\Omega/km$ )	X ( $\Omega/km$ )	$l$ (m)	R ( $\Omega$ )	X ( $\Omega$ )	Total dist.
1	1	2	0.667	0.10	85	0.0567	0.0085	85
	2	5	0.476	0.10	65	0.0309	0.0065	150
	5	9	0.667	0.10	70	0.0467	0.0070	220
	9	16	0.667	0.10	35	0.0233	0.0035	255
	16	24	1.330	0.10	35	0.0466	0.0035	290
	24	29	4.610	0.15	40	0.1844	0.0060	330
2	1	2	0.667	0.10	85	0.0567	0.0085	85
	2	5	0.476	0.10	65	0.0309	0.0065	150
	5	10	2.970	0.15	35	0.1040	0.0053	185
3	1	2	0.667	0.10	85	0.0567	0.0085	85
	2	5	0.476	0.10	65	0.0309	0.0065	150
	5	11	2.080	0.10	105	0.2184	0.0105	255
	11	17	7.130	0.15	35	0.2496	0.0053	290
4	1	2	0.667	0.10	85	0.0567	0.0085	85
	2	5	0.476	0.10	65	0.0309	0.0065	150
	5	11	2.080	0.10	105	0.2184	0.0105	255
	11	18	1.910	0.15	50	0.0955	0.0075	305
	18	25	4.610	0.15	35	0.1614	0.0053	340
5	1	3	0.476	0.10	40	0.0190	0.0040	40
	3	6	0.641	0.15	120	0.0769	0.0180	160
	6	12	2.080	0.10	140	0.2912	0.0140	300
	12	19	0.476	0.10	80	0.0381	0.0080	380
	19	26	0.476	0.10	50	0.0238	0.0050	430
	26	30	1.330	0.10	40	0.0532	0.0040	470

Table A.4 Cont. characteristics of the different types of conductors per branch.

Branch	From Node	To Node	R ( $\Omega/km$ )	X ( $\Omega/km$ )	$l$ (m)	R ( $\Omega$ )	X ( $\Omega$ )	Total dist.
6	1	3	0.476	0.10	40	0.0190	0.0040	40
	3	7	0.667	0.10	105	0.0700	0.0105	145
	7	13	0.667	0.10	35	0.0233	0.0035	180
	13	20	1.910	0.15	80	0.1528	0.0120	260
	20	27	2.080	0.10	90	0.1872	0.0090	350
	27	31	3.060	0.15	70	0.2142	0.0105	420
	31	33	4.610	0.15	35	0.1614	0.0053	455
7	10	3	0.476	0.10	40	0.0190	0.0040	40
	3	7	0.667	0.10	105	0.0700	0.0105	145
	7	13	0.667	0.10	35	0.0233	0.0035	180
	13	21	4.610	0.15	105	0.4841	0.0158	285
8	1	4	0.667	0.10	55	0.0367	0.0055	55
	4	8	0.667	0.10	100	0.0667	0.0100	155
	8	14	3.060	0.15	65	0.1989	0.0098	220
	14	22	7.130	0.15	170	1.2121	0.0255	390
9	10	4	0.667	0.10	55	0.0367	0.0055	55
	4	8	0.667	0.10	100	0.0667	0.0100	155
	8	15	1.910	0.15	65	0.1242	0.0098	220
	15	23	1.910	0.15	140	0.2674	0.0210	360
	23	28	4.450	0.10	210	0.9345	0.0210	570
	28	32	4.610	0.15	70	0.3227	0.0105	640



Table A.5 Total per branch and phase ( $a, b, c$ ) contracted (loads) and installed (PVs) power in the first feeder.

Branch	Node	Contracted Power ( $kVA$ )			Installed Power ( $kW$ )		
		$S_a$	$S_b$	$S_c$	$P_a$	$P_b$	$P_c$
1	2	3,450	3,450	0	0	0	0
	5	0 0	3,450	0	0	0	
	9	6,900	3,450	3,450	0	1,700	0
	16	0	6,900	0	0	3,450	0
	24	0	6,900	10,350	0	0	3,680
	29	3,450	3,450	10,350	0	0	3,680
Sum		13,800	24,150	27,600	0	5,150	7,360
2	2	3,450	3,450	0	0	0	0
	5	0	0	3,450	0	0	0
	10	3,450	10,350	0	0	3,680	0
Sum		6,900	13,800	3,450	0	3,680	0
3	2	3,450	3,450	0	0	0	0
	5	0	0	3,450	0	0	0
	11	3,450	0	6,900	0	0	3,450
	17	10350	0	0	3,680 0	0	
Sum		17,250	3,450	10,350	3,680	0	3,450
4	2	3,450	3,450	0	0	0	0
	5	0	0	3,450	0	0	0
	11	3,450	0	6,900	0	0	3,450
	18	0	3,450	3,450	0	0	0
	25	0	3,450	10,350	0	0	3,680
Sum		6,900	10,350	24,150	0	0	7,130

Table A.6 Total per branch and phase ( $a, b, c$ ) contracted (loads) and installed (PVs) power in the second feeder.

Branch	Node	Contracted Power ( $kVA$ )			Installed Power ( $kW$ )		
		$S_a$	$S_b$	$S_c$	$P_a$	$P_b$	$P_c$
5	3	0	0	0	0	0	0
	6	3,450	0	0	0	0	0
	12	3,450	3,450	0	0	0	0
	19	3,450	3,450	0	0	0	0
	26	0	10,350	0	0	3,680	0
	30	0	0	10,350	0	0	3,680
	Sum	10,350	17,250	10,350	0	3,680	3,680
6	3	0	0	0	0	0	0
	7	10,350	0	0	3,680	0	0
	13	6,900	3,450	3,450	3,450	0	1,700
	20	0	3,450	3,450	0	0	0
	27	6,900	3,450	3,450	3,450	0	1,700
	31	0	6,900	0	0	3,450	0
	33	0	0	3,450	0	0	0
	Sum	24,150	17,250	13,800	10,580	3,450	3,400
7	3	0	0	0	0	0	0
	7	10,350	0	0	3,680	0	
	13	6,900	3,450	3,450	3,450	0	1,700
	21	0	6,900	0	0	0	0
	Sum	17,250	10,350	3,450	7,130	0	1,700

Table A.7 Total per branch and phase ( $a, b, c$ ) contracted (loads) and installed (PVs) power in the third feeder.

Branch	Node	Contracted Power ( $kVA$ )			Installed Power ( $kW$ )		
		$S_a$	$S_b$	$S_c$	$P_a$	$P_b$	$P_c$
8	4	0	0	0	0	0	0
	8	0	0	6,900	0	0	0
	14	0	0	0	0	0	0
	22	10,350	3,450	3,450	3,680	0	1,700
	Sum	10,350	3,450	10,350	3,680	0	1,700
9	4	0	0	0	0	0	0
	8	0	0	6,900	0	0	0
	15	0	0	0	0	0	0
	23	0	0	3,450	0	0	0
	28	0	6,900	0	0	0	0
	32	0	3,450	3,450	0	0	0
	Sum	0	10,350	13,800	0	0	0

Table A.8 Total per feeder and phase ( $a, b, c$ ) contracted (loads) and installed (PVs) power.

Feeder	Node	Contracted Power ( $kVA$ )			Installed Power ( $kW$ )		
		$S_a$	$S_b$	$S_c$	$P_a$	$P_b$	$P_c$
1	2	3,450	3,450	0	0	0	0
	5	0	0	3,450	0	0	0
	9	6,900	3,450	3,450	0	1,700	0
	16	0	6,900	0	0	3,450	0
	24	0	6,900	10,350	0	0	3,680
	29	3,450	3,450	10,350	0	0	3,680
	10	3,450	10,350	0	0	3,680	0
	11	3,450	0	6,900	0	0	3,450
	17	10,350	0	0	3,680	0	0
	18	0	3,450	3,450	0	0	0
	25	0	3,450	10,350	0	0	3,680
	Sum	31,050	41,400	48,300	3,680	8,830	14,490
2	3	0	0	0	0	0	0
	6	3,450	0	0	0	0	0
	12	3,450	3,450	0	0	0	0
	19	3,450	3,450	0	0	0	0
	26	0	10,350	0	0	3,680	0
	30	0	0	10,350	0	0	3,680
	7	10,350	0	0	3,680	0	0
	13	6,900	3,450	3,450	3,450	0	1,700
	20	0	3,450	3,450	0	0	0
	27	6,900	3,450	3,450	3,450	0	1,700
	31	0	6,900	0	0	3,450	0
	33	0	0	3,450	0	0	0
	21	0	6,900	0	0	0	0
	Sum	3,450	41,400	24,150	10,580	7,130	7,080

Table A.9 Cont. Total per feeder and phase ( $a, b, c$ ) contracted (loads) and installed (PVs) power.

Feeder	Node	Contracted Power ( $kVA$ )			Installed Power ( $kW$ )		
		$S_a$	$S_b$	$S_c$	$P_a$	$P_b$	$P_c$
3	4	0	0	0	0	0	0
	8	0	0	6,900	0	0	0
	14	0	0	0	0	0	0
	22	10,350	3,450	3,450	3,680	0	1,700
	15	0	0	0	0	0	0
	23	0	0	3,450	0	0	0
	28	0	6,900	0	0	0	0
	32	0	3,450	3,450	0	0	0
Sum		10,350	13,800	17,250	3,680	0	1,700
Grid	Total	75,900	96,600	89,700	17,940	15,960	23,270

## Appendix B

### Grid-connected PV fault isolation figures

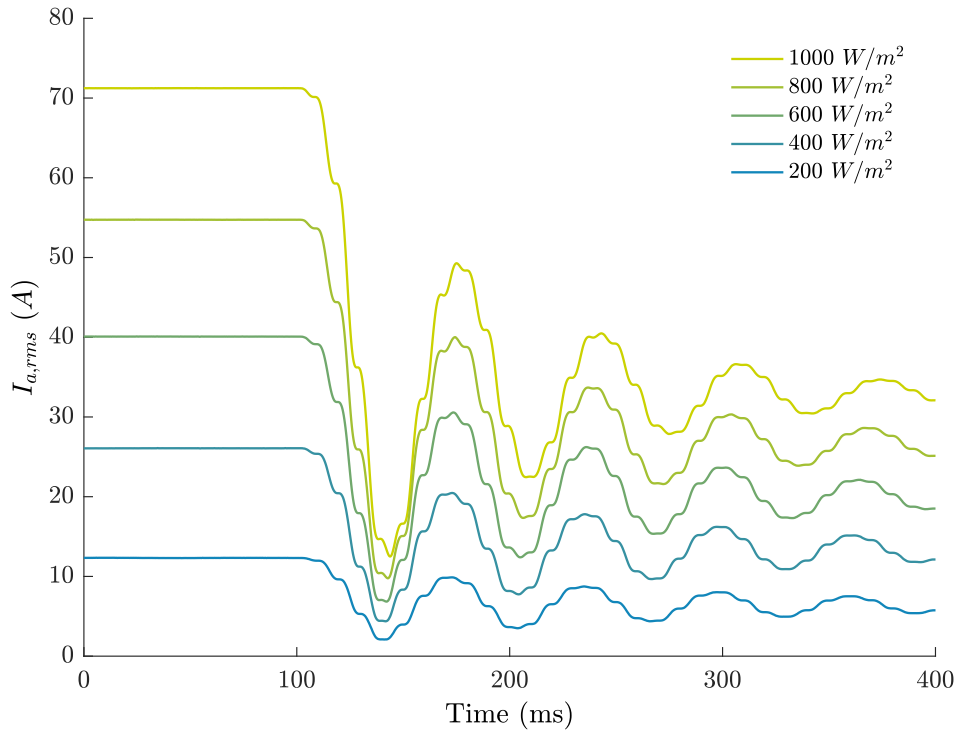


Figure B.1 Decrease of the phase a rms current (s1) for an inverse bypass diode fault (f02) at different irradiance levels. Fault inception at 100 ms. Under this fault all three phases decrease but only phase a is shown here.

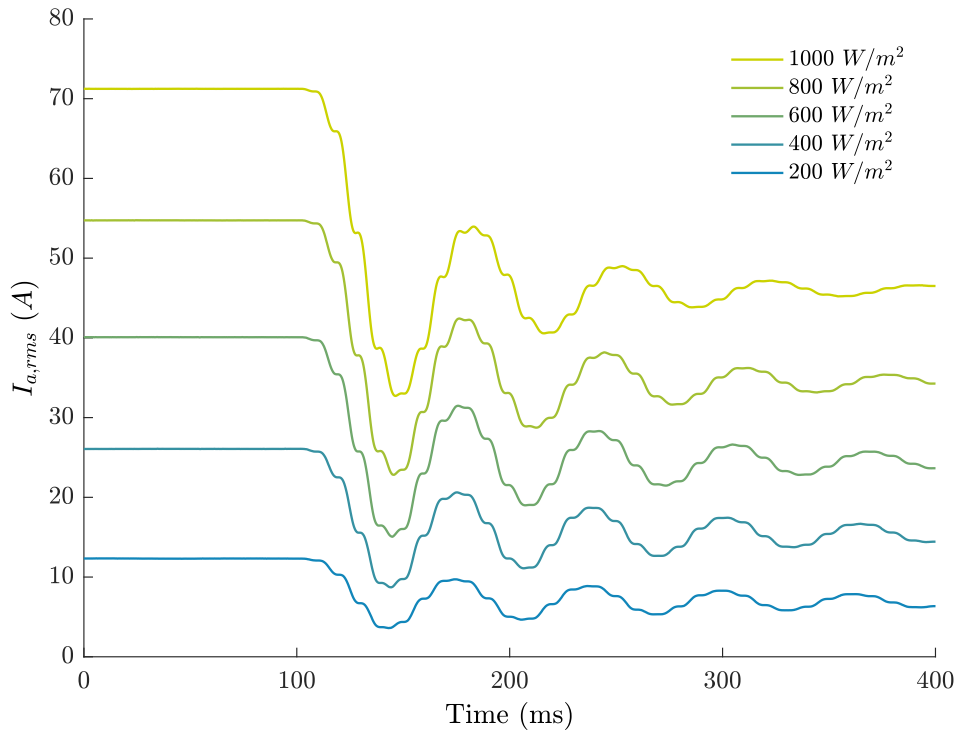


Figure B.2 Decrease of the phase a rms current (s1) for a bypass diode breakdown fault (f04) at different irradiance levels. Fault inception at 100 ms. Under this fault all three phases decrease but only phase a is shown here.

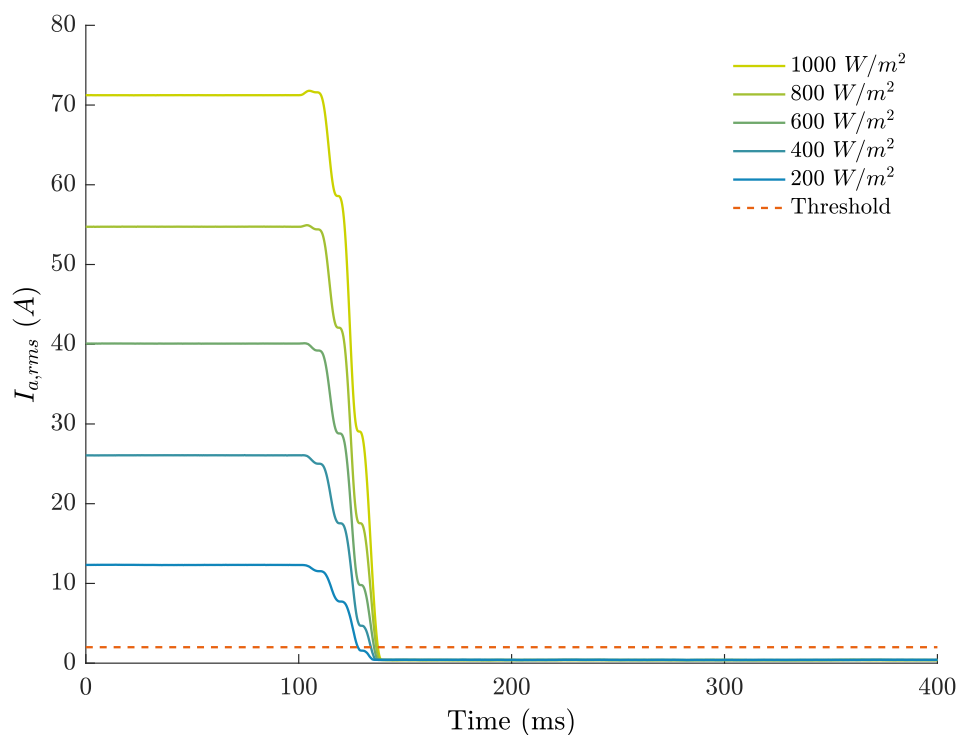


Figure B.3 Decrease of the phase a rms current to zero for an open-circuited boost IGBT fault (f11) at different irradiance levels. Fault inception at 100 ms. Under this fault all three phases drop to zero but only phase a is shown here.

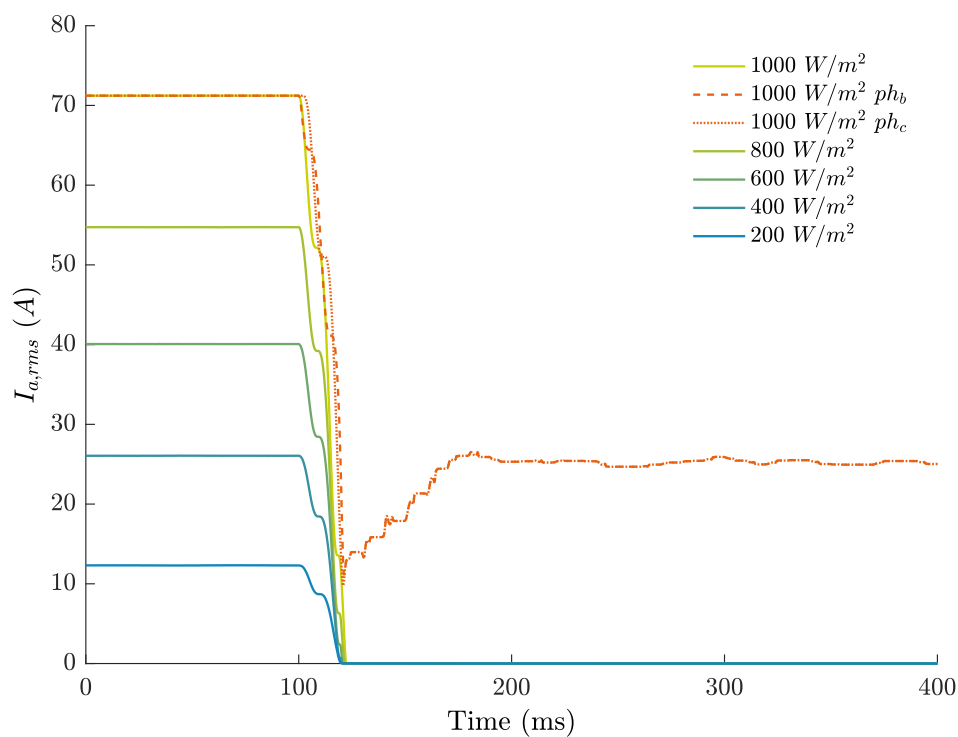


Figure B.4 Decrease of the phase a rms current to zero (s5) for an open-circuited inverter leg fault (f10) at different irradiance levels. Fault inception at 100 ms. The current on the other phases decreases during FO but doesn't drop to zero.



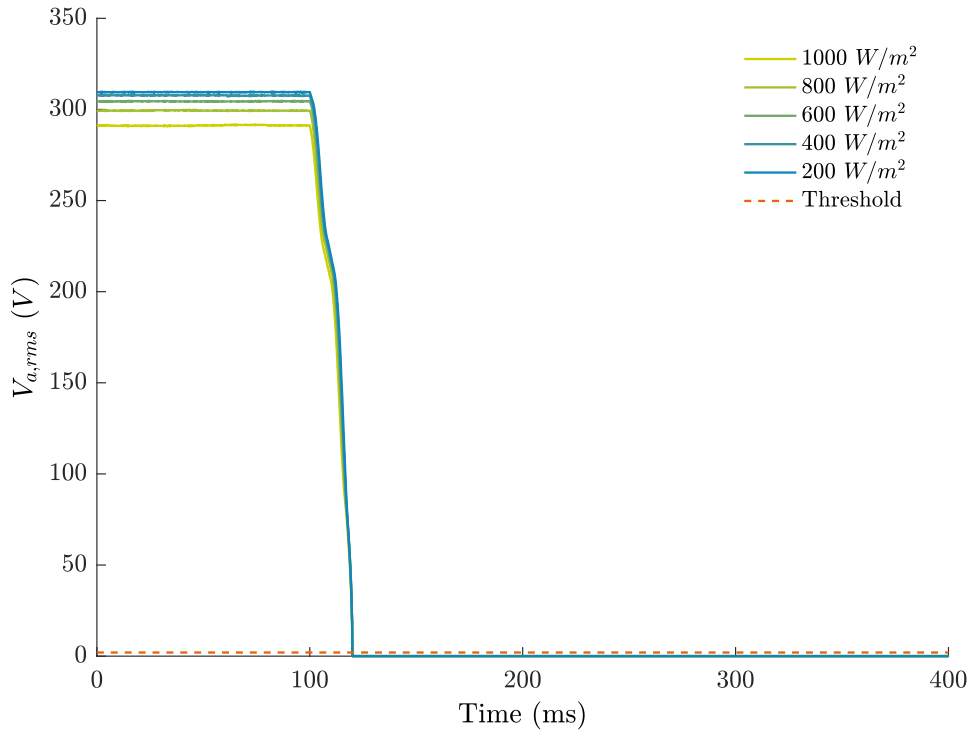


Figure B.5 Decrease of the phase a rms voltage to zero for a short-circuited inverter leg fault (f11) at different irradiance levels. Fault inception at 100 ms. Under this fault all three phases drop to zero but only phase a is shown here.

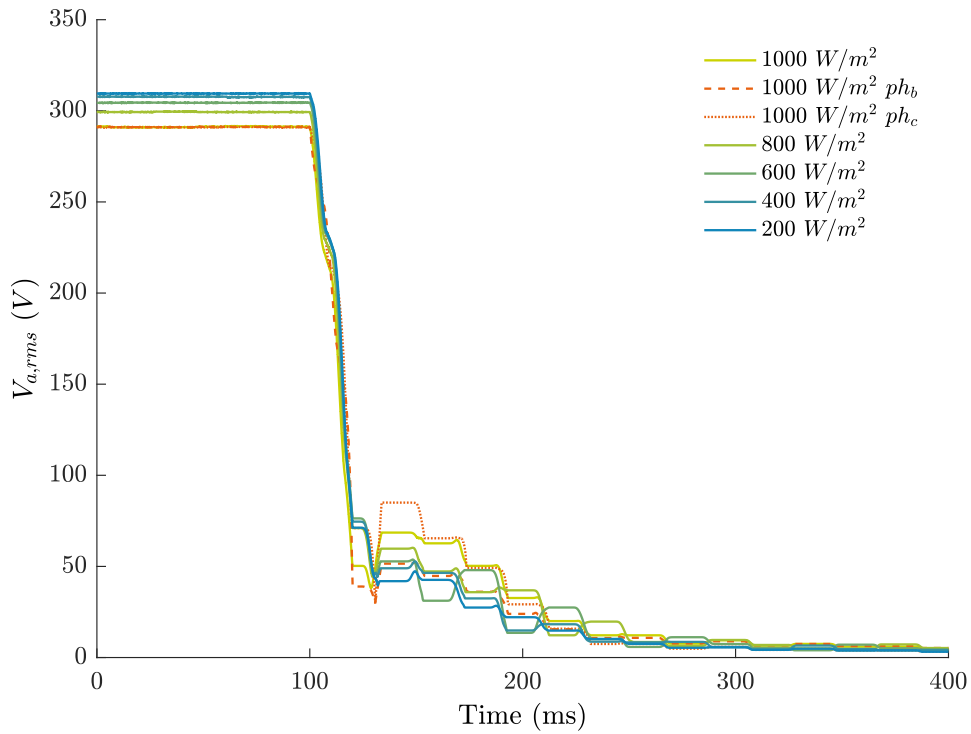


Figure B.6 Decrease of the phase rms voltage in all three phases (s5) for a short-circuited inverter IGBT fault (f11) at different irradiance levels (for phase a only). Fault inception at 100 ms.

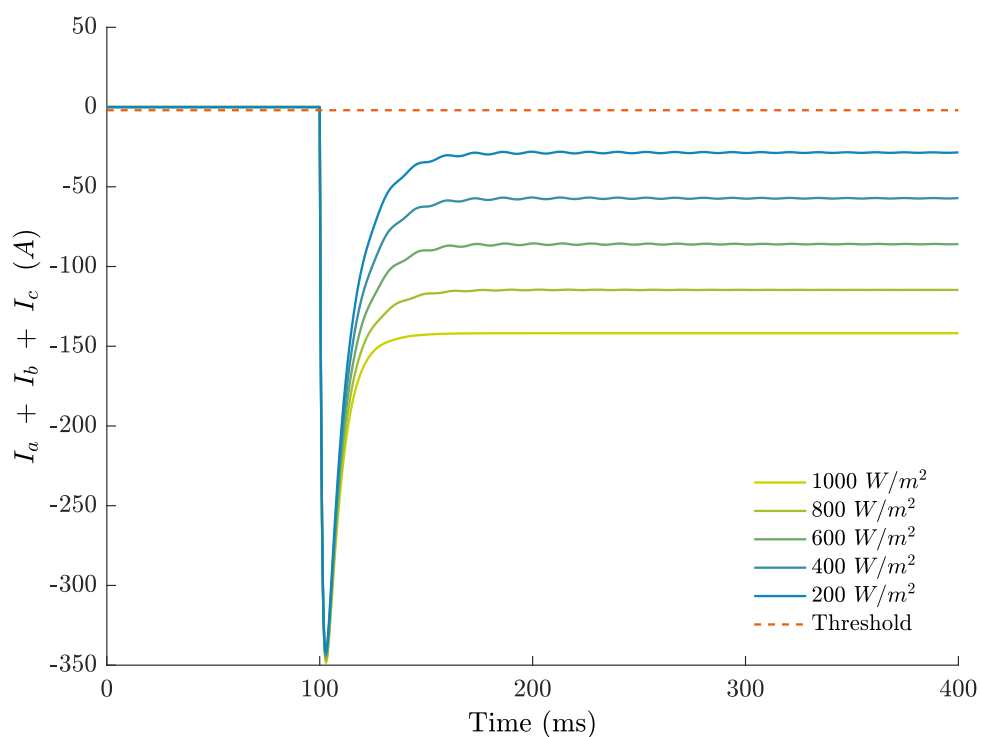


Figure B.7 Sum of phase currents (s5) for a positive pole to ground SC fault (f05) at different irradiance levels. Fault inception at 100 ms.

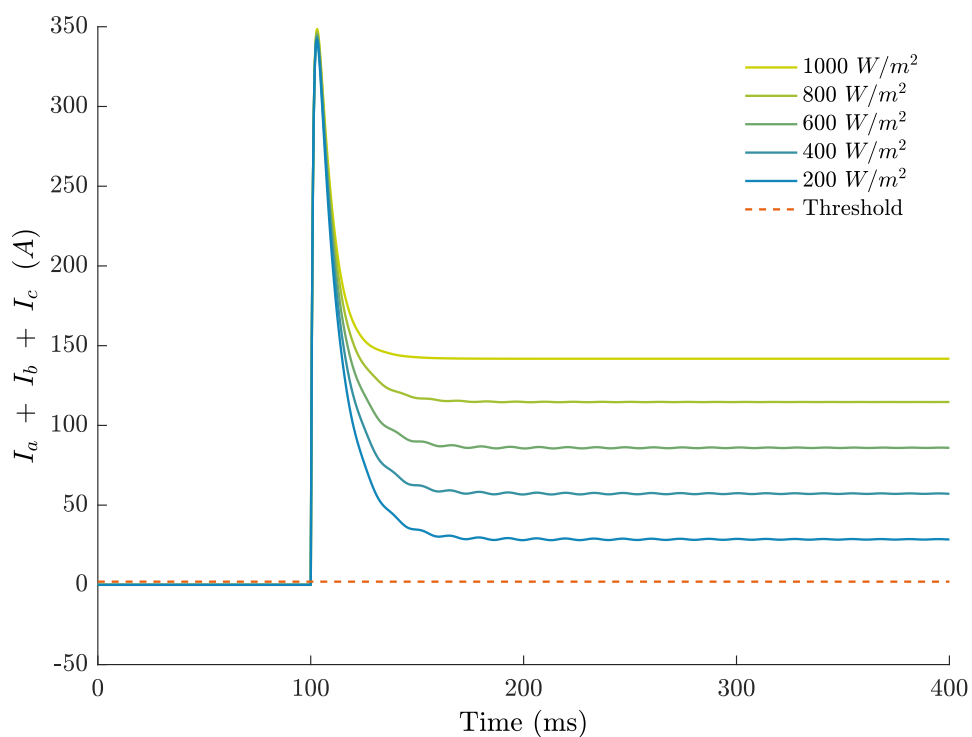


Figure B.8 Sum of phase currents (s5) for a negative pole to ground SC fault (f06) at different irradiance levels. Fault inception at 100 ms.

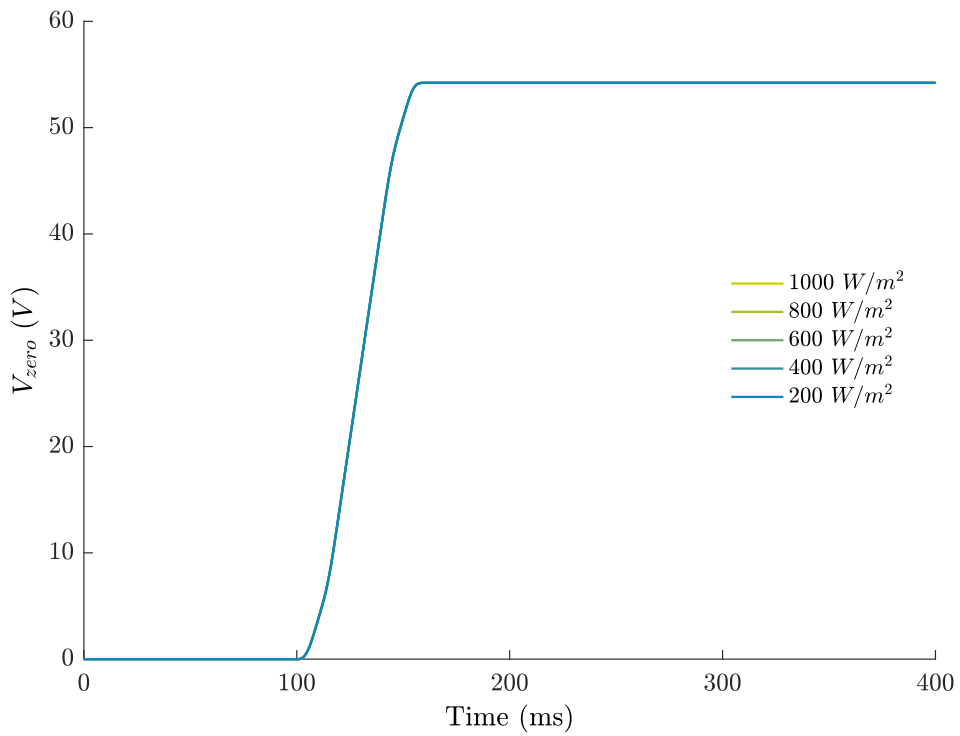


Figure B.9 Moving average of the zero sequence component of the voltage (s6) for a single phase voltage sag fault (f17) at different irradiance levels. Fault inception at 100 ms.

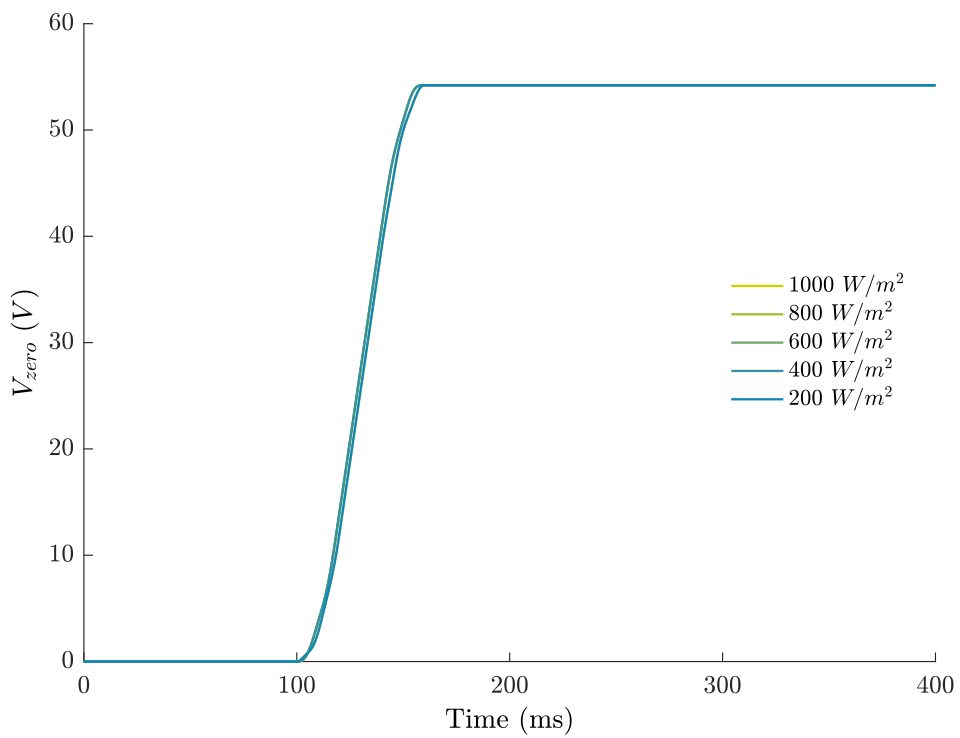


Figure B.10 Moving average of the zero sequence component of the voltage (s6) for a two phase voltage sag fault (f18) at different irradiance levels. Fault inception at 100 ms.

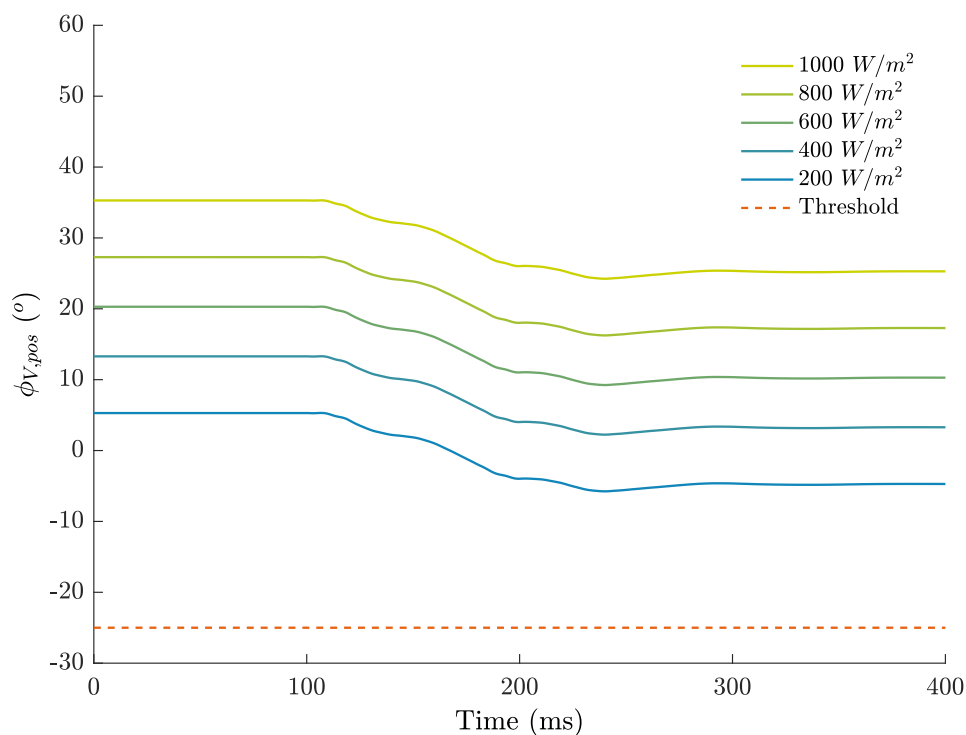


Figure B.11 Moving average of the phase angle of the positive sequence component of the voltage (s7) for a double phase SC fault (f15) at different irradiance levels. Fault inception at 100 *ms*.

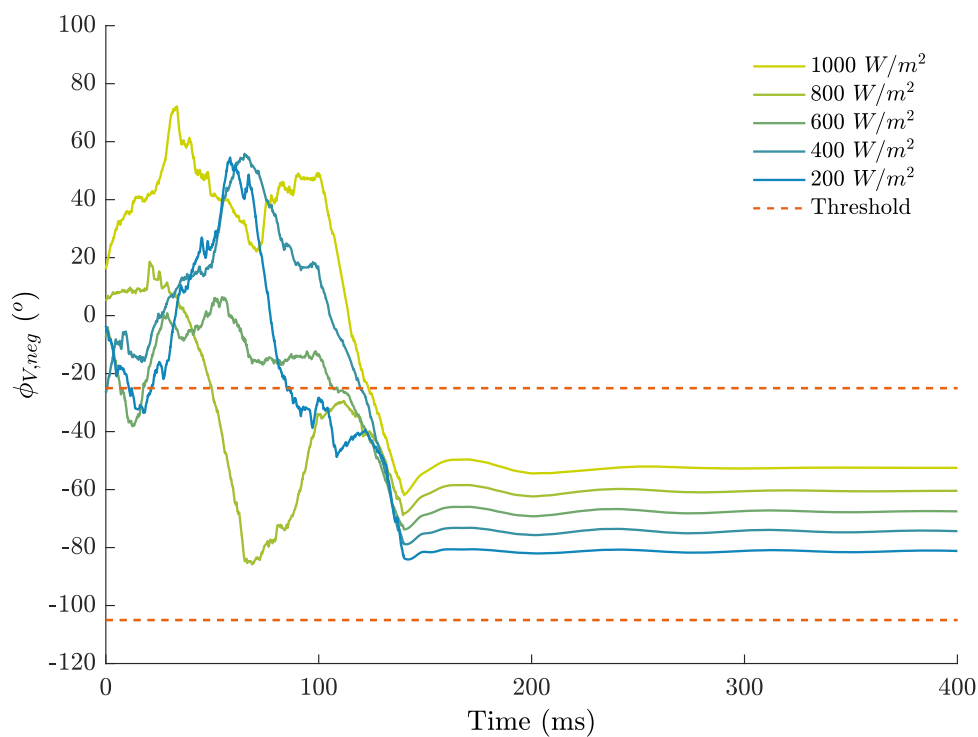


Figure B.12 Moving average of the phase angle of the negative sequence component of the voltage (s8) for a double phase SC fault (f15) at different irradiance levels. For NO curve shape refer to Fig. 7.5g. Fault inception at 100 *ms*.



## Appendix C

### Dependency of the neural network on the dataset size

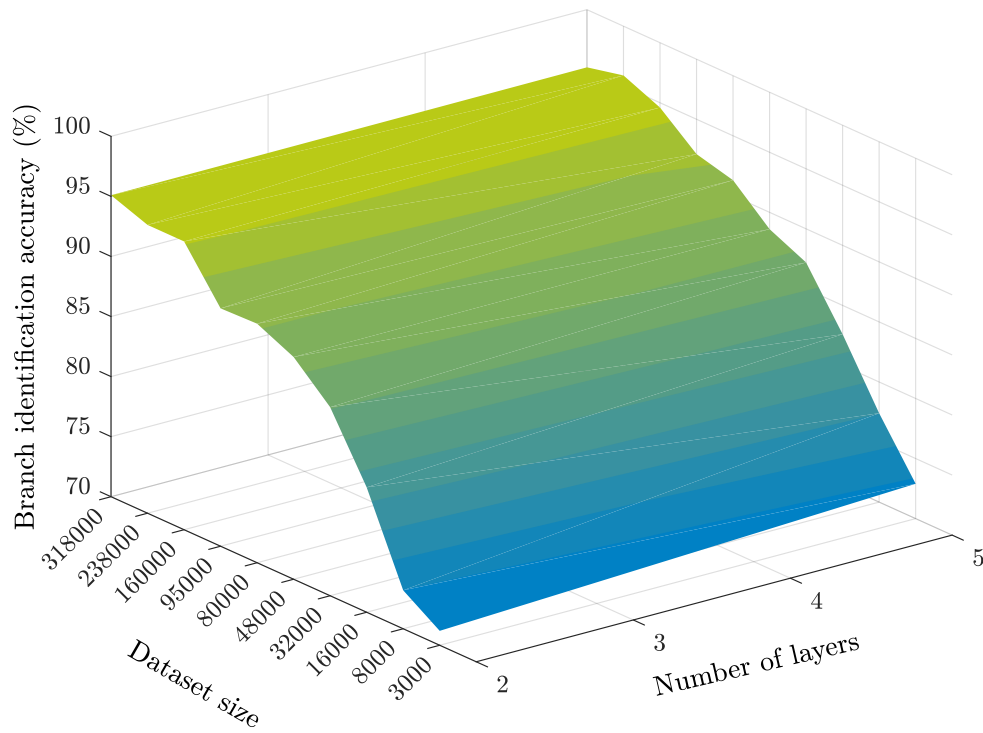


Figure C.1 Branch identification accuracy per dataset size and number of layers.

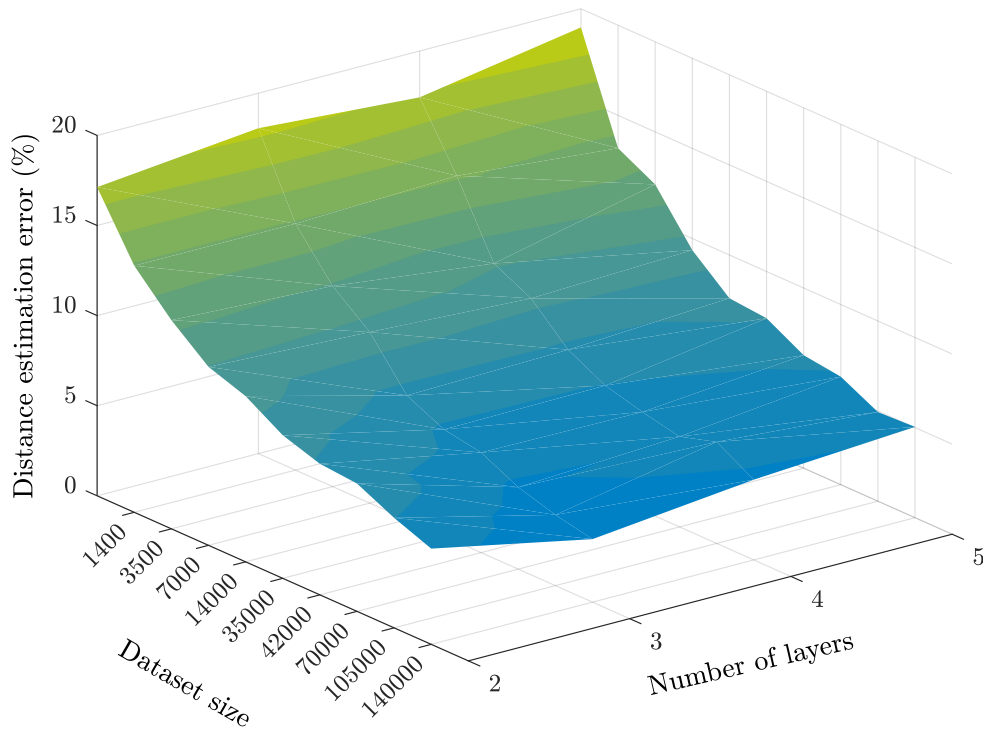


Figure C.2 Distance estimation error per dataset size and number of layers.

# Appendix D

## Fault detection and isolation in grid-connected PEM fuel cells

Polymer electrolyte membrane or proton exchange membrane fuel cells (*PEMFC*) are the most popular among other fuel cell technologies (alkaline, phosphoric acid, molten carbonate and solid oxide) for their simplicity, variability, quick start up capabilities and their suitability for a wide range of applications [259]. As PEMFCs are becoming more and more popular their monitoring and safe operation is of paramount importance.

Monitoring of a PEMFC is a non-linear, multi-fault source with different time-scale problems especially because of all the chemical reactions taking place inside the fuel cell [260]. Its complexity is underlined in the first fuel cell law: *A change of one parameter will affect at least two others of which at least one will have an opposite effect of the expected outcome* [259]. It is thus clear that classification of faults or obtaining unique fault signatures is a challenging task.



## D.1 Faults in fuel cells

Faults in fuel cells can be divided in three big categories: permanent (CO poisoning, fuel cell aging, etc.), transient (flooding and drying) and external faults (power electronics faults, cooling system, etc.) [261]. In this study, the drying fault is studied and its isolation is ensured against other external faults such as power electronics faults as they were described in Chapter 7 [262].

## D.2 Case study

In order to simulate and analyze the drying effect in PEMFC, the detailed model of Matlab/Simulink for fuel cells was used with an output power and voltage of 50 kW and 625 V respectively. The PEMFC was connected to the grid through a dc-dc boost converter and an inverter. The grid was replaced by an ideal voltage source.

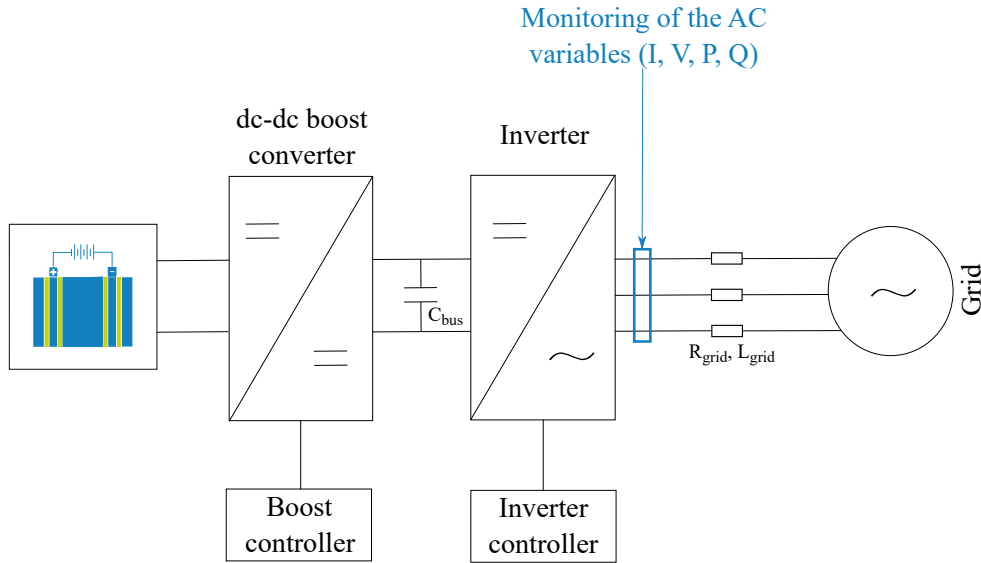


Figure D.1 Schematic of the grid connected PEMFC system.

Water management in PEMFC is a very complicated process affected by various parameters. Equation D.1 relates the saturation pressure of water with the temperature:

$$P_{sat} = P_0 \cdot e^{-(M_{H_2O} \cdot h_{fg}/R) \cdot (1/T - 1/T_0)} \quad (D.1)$$

where  $P_0$  is the pressure at known temperature,  $T_0$  the known temperature,  $M_{H_2O}$  the water molar mass in  $[kg \cdot mol^{-1}]$ ,  $h_{fg}$  the water latent heat vaporization in  $[kJ \cdot kg^{-1}]$ ,  $R$  the gas constant in  $[J \cdot mol^{-1} \cdot K^{-1}]$  and  $T$  the temperature in  $[K]$ . From eq. D.1, it is clear that by increasing the temperature, the saturation pressure will also increase. An increase of the saturation pressure will also mean an increase of the evaporation rate of water which in its turn will decrease the amount of liquid water within the fuel cell [263]. Hence, in order to simulate the drying effect in a fuel cell, the increase of the fuel cell stack operating temperature was implemented. From the nominal operating temperature of  $65^\circ C$  the temperature was linearly increased up to  $115^\circ C$  after 23 s as shown in Fig. D.2.

### D.3 Results

Theoretically, as mentioned in [263], a higher cell temperature leads initially to a higher cell voltage as the voltage losses inside a fuel cell decrease with the increase of temperature which compensates up to a point for the theoretically expected voltage drop. Moreover, an increase of the temperature to values higher than  $100^\circ C$ , according to [264], favors drying effects. After a long time under drying conditions the voltage can drop even to zero [263].

This behavior is observed in Fig. D.2 where while passing from an operating temperature of  $65^\circ C$  to  $75^\circ C$  the output fuel cell voltage reaches its maximum of 625 V (Region I). After that point, and as the temperature keeps increasing, the output voltage decreases (Region II). A more steep voltage drop is observed after the temperature crosses the threshold of  $100^\circ C$  and eventually drops to a value very close to 0 V after 23 s of increasing temperature (Region III).

The further away we move from the fuel cell the harder it gets to detect the drying effect; a phenomenon that was noticed in photovoltaics as well [262]. At the dc bus level, the boost controller is trying to maintain the voltage constant at 800 V (Fig. D.3). A gradual decrease of the dc bus voltage is noticed as the drying of the fuel cell progresses with a maximum voltage drop of 10 V until the point where the boost converter follows the drop of the fuel cell voltage to almost zero with a decrease of at least 300 V.

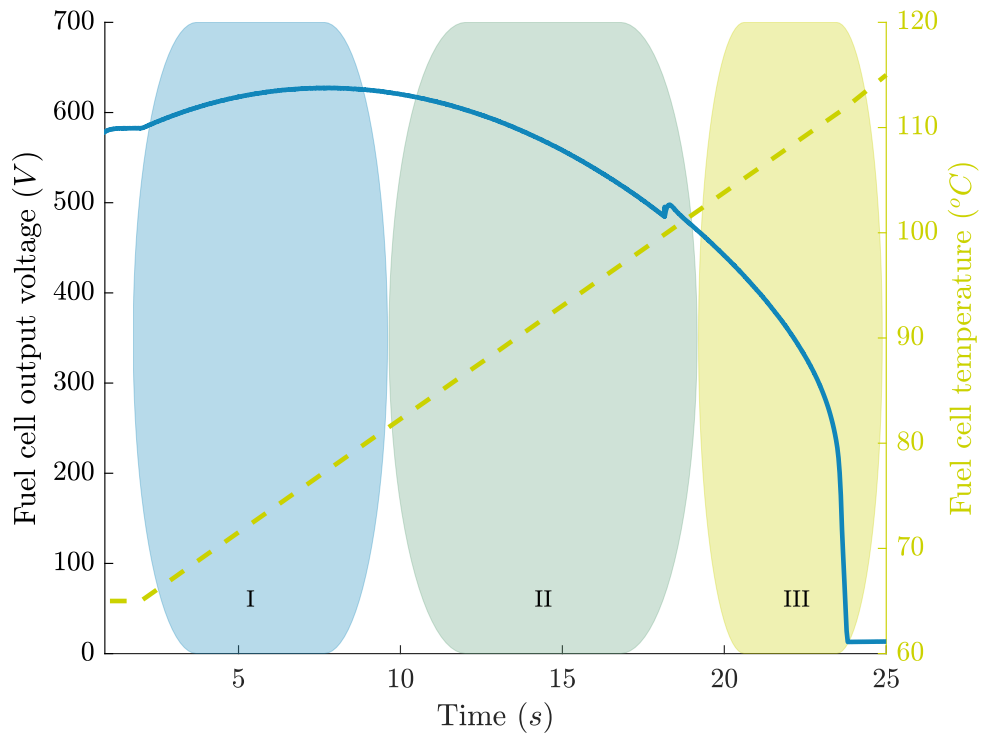


Figure D.2 Fuel cell output voltage profile compared to the increase of temperature inside the fuel cell. Three distinct regions (I, II and III) of output voltage increase or decrease can be identified.

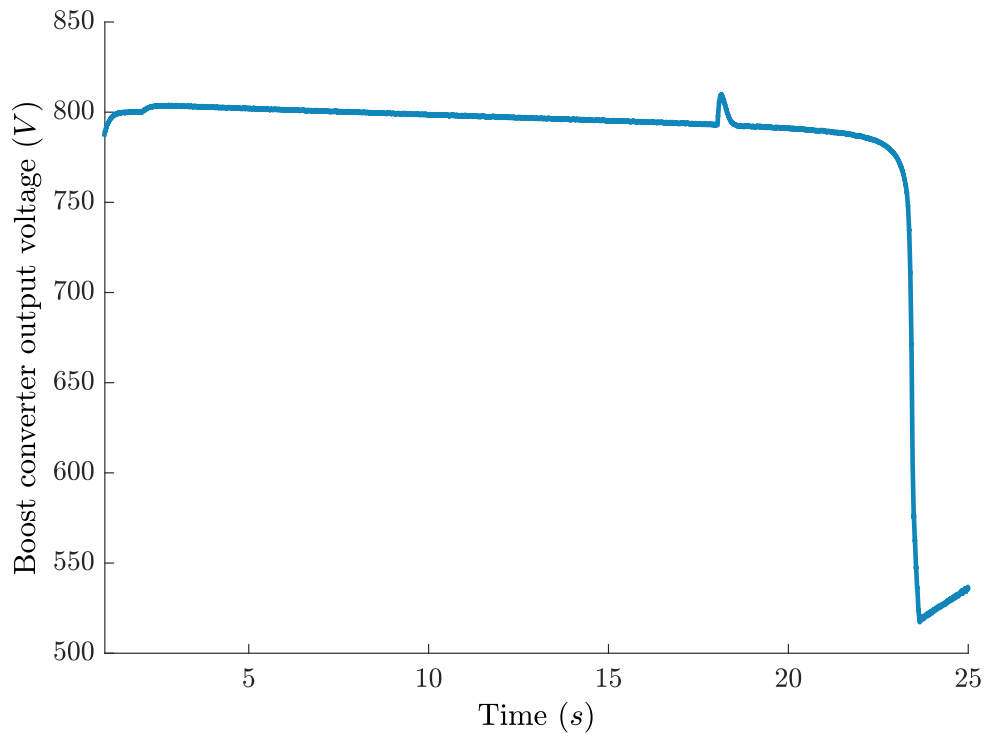


Figure D.3 Boost converter output voltage profile as the drying effect progresses. DC bus voltage was set at 800 V.

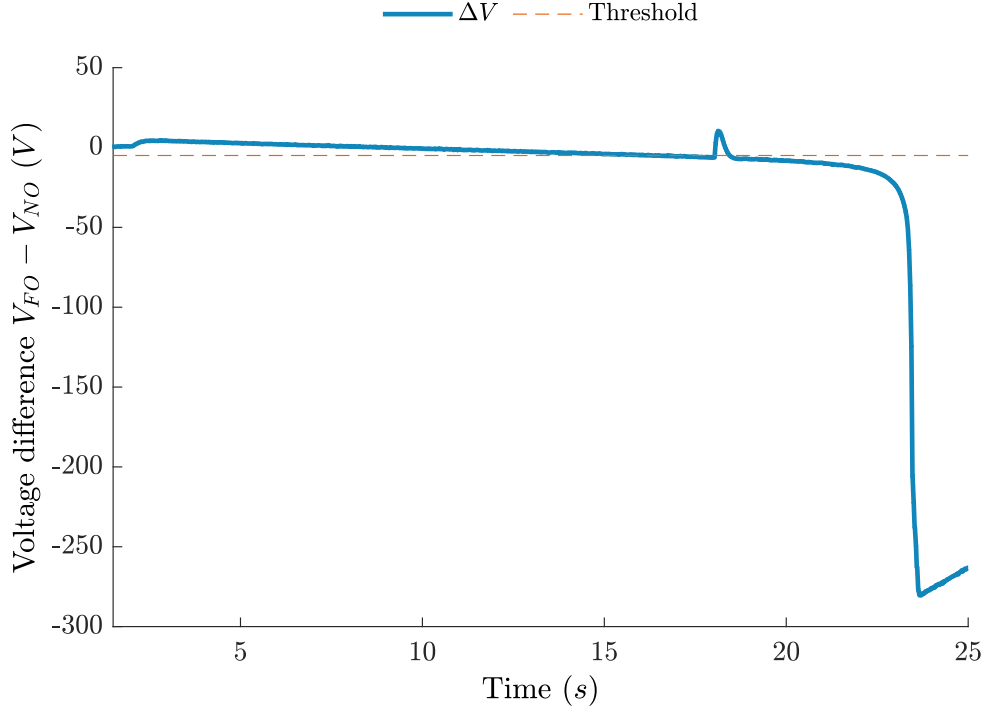


Figure D.4 Inverter output voltage difference of phase A,  $\Delta V$ , between faulty and normal operation as the drying effect progresses. The same behavior is noticed in all three phases.

By moving one step further from the fuel cell and monitoring the voltage at the output of the inverter, an even smaller voltage drop is noticed. The difference,  $\Delta V$ , between faulty (*FO*) and normal (*NO*) operation voltage is presented in Fig. D.4. In concordance with the study in photovoltaics, the voltage drop threshold was set at 1% of the peak voltage under normal operation. This will result in a detection of the drying effect only 15 s after the inception of the fault. However, detection will still occur before the drop of the fuel cell output voltage to zero. As far as isolation is concerned, given the slow dynamics of the fault progression and the affected variables, its complete isolation would still be possible when compared to the other faults inside the boost converter, the dc bus, the inverter and the grid as described in [262]. However, similarly with the photovoltaics case, it is expected that faults inside the fuel cells would still be recognized as a group of faults as a monitoring on the ac side of the system alone, is not enough to provide a discrimination between them.

Finally, since no single method can satisfy all the requirements of a complete fault diagnosis in a complex system like the grid-connected PEMFC system [260], the need of hybrid methods, using different monitoring tools is evident.

## D.4 Conculsion

The drying of a fuel cell was investigated in this study. An increase of the stack temperature was implemented to simulate the drying effect and the voltage was monitored at every conversion stage. Though fault detection and isolation are still possible from the ac side of the system, if a quicker diagnosis is necessary to prevent possible fuel cell degradation then the monitoring of the voltage at the fuel cell level is mandatory. Finally, based on the fault signature table of the photovoltaics case, the isolation of the other listed faults in Chapter 7 will still be possible.

# Résumé de la thèse

Les nouveaux compteurs intelligents, la production décentralisée, les sources d'énergie renouvelable et le changement climatique redéfinissent les réseaux électriques. Pour profiter pleinement des nouveaux réseaux électriques intelligents, nous devons les surveiller et les protéger. La capacité de “self-healing” est donc importante dans les réseaux intelligents afin d'assurer un comportement correct en cas de défaillance et de réduire les temps de panne de courant. À cette fin, cette thèse propose trois méthodes différentes de détection et la localisation des défauts pour les *réseaux de distribution basse tension (BT)* et deux méthodes de localisation des défauts pour les systèmes photovoltaïques raccordés au réseau.

Dans les systèmes de distribution, les pannes de courant sont responsables de 80 % des interruptions des clients. Les défauts les plus courants dans les systèmes de distribution sont les court-circuits entre une phase et la terre qui représentent 70 % du total des cas des défauts. Les plus graves, les défauts triphasés, représentent 5 % du total des cas des défauts. Le reste 25 % est divisé entre les défauts biphasés (15 %) et les défauts biphasés-terre (10 %) [1].

Au départ, les réseaux de distribution BT étaient conçus pour suivre la doctrine traditionnelle “fit and forget”, qui ne permettait qu'un flux unidirectionnel d'énergie du transformateur de distribution aux consommateurs. Avec l'intégration des sources d'énergie renouvelables dans le réseau, le flux d'énergie bidirectionnel devient une réalité et les opérateurs de réseaux de distribution sont confrontés à plusieurs problèmes tels que la congestion, la tension surélevée et la baisse de la qualité de l'alimentation [21]. Contrairement au *réseau moyenne tension (MT)*, celle BT est plus complexe et difficile à gérer en raison de certaines de ses caractéristiques uniques qui lui attribuent une nature déséquilibrée et hétérogène [23]. Ces caractéristiques sont des obstacles importants dans les processus de détection et de localisation des défauts.

Dans ce cadre et par la nécessité d'une information et d'un contrôle accru sur le réseau BT et la disponibilité de nouveaux outils et d'une infrastructure de comptage, le concept de réseau intelligent a vu le jour.

## **Partie I : Diagnostic des erreurs dans les réseaux de distribution intelligente BT**

Pour tenter de résoudre un problème assez inexploré, trois méthodes différentes de détection et de localisation des défauts pour les réseaux de distribution BT basées sur des mesures de courant et de tension rms ont été développées dans cette thèse.

### **Méthode conventionnelle:**

La première approche était une méthode "conventionnell". Une méthode de détection des défauts, basée sur les mesures de courant, et une méthode de localisation des défauts, basée sur les mesures de tension, pour les réseaux de distribution BT sont proposées. La méthode de détection des défauts proposée consiste à surveiller l'augmentation du courant de phase au début de chaque feeder. La méthode de localisation des défauts, repose sur une analyse du profil de tension construite à partir des mesures de tensions nodales. Bien que l'idée de base de la méthode de localisation de défauts ait été initialement conçue pour les réseaux de distribution MT [2], elle est ici testée dans le cas plus complexe des réseaux BT.

En général, l'augmentation de la résistance aux défauts entrave à la fois les méthodes de détection et de localisation des défauts car elle diminue le courant de défaut et augmente les tensions, rapprochant ainsi les variables surveillées de leurs conditions de bon fonctionnement. Alors que pour la détection des défauts, les mesures de courant de phase sont l'outil le plus approprié, pour le processus de localisation des défauts, qui comprend trois étapes distinctes : a) identification des branches, b) localisation des secteurs et c) estimation de la distance, la composante positive de la tension efficace mesurée est un meilleur choix. En outre, les défauts situés vers l'extrémité de la ligne d'alimentation sont plus difficiles à localiser. De plus, les défauts triphasés, étant les plus graves, sont les plus faciles à détecter et à localiser. En outre, une des avantages de

la méthode est que, comme on peut le voir dans la performance symétrique de toutes les phases, la distribution asymétrique des charges et des PV sur les trois phases du réseau n'affecte pas vraiment la performance de l'algorithme. Une augmentation de la demande de charge pendant les heures creuses diminue la précision de la méthode, les PV, bien que toujours connectés au réseau, ne pouvant pas atténuer cet effet. Enfin, bien que la méthode de localisation des défauts ne semble pas être affectée par le bruit des mesures, la diminution du nombre de capteurs disponibles peut sérieusement affecter sa précision.

### **Gradient boosting trees:**

Afin de surmonter les inconvénients de la méthode conventionnelle et d'améliorer les chances de détecter et de localiser correctement un défaut, une nouvelle méthode d'intelligence artificielle est proposée basée sur les "*gradient boosting trees*" (*GBT*). La méthode proposée permet de détecter et d'identifier à la fois les défauts monophasés et les défauts triphasés. Pour estimer le modèle, un ensemble de caractéristiques d'entrée non spécifiques à la branche est utilisé pour assurer la robustesse de l'algorithme par rapport aux différentes topologies du réseau et le nombre disponible de mesures de tension par branche. Comme il s'agit d'une méthode basée sur l'apprentissage, son rendement est évalué par rapport aux données non-échantillonnées. La contribution de cette méthode est triple et est résumée ci-dessous :

- a) Détection des défauts et identification du feeder en défaut: l'occurrence de la défaillance est détectée avec une identification simultanée du feeder en défaut.
- b) Identification du type des défauts: une distinction entre les phases défectueuses et non défectueuses est obtenue, identifiant ainsi le type des défaut, monophasé (AG, BG ou CG) ou triphasé.
- c) Identification de la branche en défaut: suite à l'identification du feeder en défaut et de la phase, la branche défectueuse dans un feeder en défaut est également identifiée.

Une excellente précision pour la détection des défauts est atteinte. L'identification de phase et de branche montre des résultats prometteurs. De même, avec la méthode précédente, cet algorithme est également robuste contre la distribution



déséquilibrée par phase des charges et des unités de production. De plus, comme on s'y attendait, l'augmentation de la résistance aux défauts a réduit la précision des tous les trois tâches. En général, les résultats obtenus grâce à l'utilisation de GBT montrent une nette supériorité de la méthode proposée par rapport à la méthode conventionnelle pour les réseaux BT.

### **Réseaux de neurones profonds:**

Ces dernières années, la recherche sur les réseaux de neurones a permis de réaliser plusieurs percées qui ont abouti à ce qu'on appelle aujourd'hui l'apprentissage profond. En particulier, grâce à ces avancées, l'utilisation de réseaux de neurones dont la profondeur n'est plus limitée à une seule couche cachée est maintenant possible. Ces réseaux de neurones plus profonds, comme l'ont prouvé plusieurs applications, se sont systématiquement révélés meilleurs pour les problèmes d'estimation en raison de leurs meilleures propriétés de généralisation [247]. Dans ce cas, l'utilisation de *réseaux de neurones profonds* (DNN) est proposée comme solution aux problèmes de détection et de localisation des défauts dans les réseaux de distribution BT. En détail, la contribution de la méthode développée est quadruple :

- a) Méthode de détection des défauts qui détecte l'occurrence d'un défaut monophasée ou triphasée avec une identification simultanée du feeder en défaut.
- b) Méthode d'identification de branche en défaut qui identifie la branche en défaut dans un feeder en défaut.
- c) Méthode d'estimation de la distance d'erreur qui détermine l'emplacement exact de l'erreur.
- d) Une diminution moyenne de la précision de seulement 4.5 % lors des tâches d'identification de branche et d'estimation de distance erronées dans le cas de mesures extrêmement limitées, lorsqu'il n'y a qu'une seule mesure disponible dans le nœud terminal de chaque branche et une mesure au niveau de la poste électrique (commune à toutes les lignes d'alimentation).

Le DNN excelle dans la détection d'une occurrence d'erreur avec une précision de 100 %. Pour la tâche d'identification des défaut, une précision moyenne de 84.6 % est atteinte, tandis que pour la tâche d'estimation de distance, l'erreur ne

dépasse pas en moyenne 12 %. De plus, l'utilisation de composants symétriques est plus efficace que l'utilisation de mesures de tension de phase. En outre, l'augmentation de la résistance aux défauts entrave le processus de localisation des défauts en général. Une autre remarque importante est que le DNN est insensible à la nature déséquilibrée du réseau quand il s'agit de la distribution par phase des charges et de la génération PV mais pas à leur distribution topologique qui pourrait soit améliorer ou entraver le processus de localisation des défaut. En outre, pour le cas extrême de seulement deux mesures disponibles par branche (une commune pour tous au niveau de la poste électrique et une à chaque nœud terminal), la précision de la méthode est diminuée de seulement 4.58 % pour l'identification de branche et 4.37 % pour l'estimation de la distance. Enfin, étant une méthode basée sur l'apprentissage et bien que pour la détection des défauts, il pourrait ne pas rencontrer beaucoup de difficultés, pour une estimation fiable de l'emplacement des défauts, un recyclage du DNN est conseillé en cas de changement de topologie ou de mode opérationnel du réseau. En général, les résultats montrent que DNN est une solution fiable pour la détection et localisation des défauts dans les réseaux BT avec une grande précision même en cas des défauts à haute résistance des défaut.

### **Analyse comparative:**

Une étude de cas commune a été utilisée pour comparer les trois méthodes. Les paramètres d'influence étaient: a) une grande variété de valeurs de résistance aux défauts atteignant le nombre de 63,772 valeurs entre 1 et 1000  $\Omega$ , b) neuf emplacements des défauts différents dans chaque secteur, c) deux types des défauts (défauts monophasés et triphasés), d) un facteur de simultanéité de 0.5, e) un large spectre de scénarios de production photovoltaïque et de demande de charge avec 70,334 combinaisons étudiées et f) une erreur de sous-estimation de 2 %.

En ce qui concerne la détection des défauts, avec une précision de 100 % en utilisant DNN, il y a une garantie qu'une occurrence des défaut ne restera pas indétectable, quelque chose qui pourrait causer beaucoup de problèmes dans le réseau, y compris la contribution accrue des PV dans le courant des défaut.

Pour les étapes de localisation des défauts : a) identification de branche, b) localisation de secteur et c) estimation de distance, les conclusions comparatives

suivantes ont été tirées. Tout d'abord, chaque étape qui a été prise dans le passage de l'utilisation d'outils conventionnels à l'utilisation intelligents, a été d'améliorer la précision d'identification de branche avec DNN étant le meilleur choix parmi les méthodes développées. En outre, quelques études de la littérature, conçues pour les réseaux MT en raison de l'indisponibilité d'autres méthodes dans le cas des réseaux BT, ont également été comparées à celles développées dans cette thèse. En ce qui concerne d'autres études dans la littérature [251, 79], le DNN a surperformé dans certains cas les méthodes existantes [251] ou apparié leur précision dans autres [79].

En plus, l'importance d'une méthode de localisation de secteur fiable est soulignée, car elle peut conduire à une diminution de l'erreur d'estimation de distance par une moyenne de 6.5 %, tant dans les méthodes conventionnelles que dans les méthodes DNN. Toutefois, les deux méthodes présentent une précision moyenne inférieure à 50 % pour la localisation du secteur défectueux.

En outre, dans l'analyse de l'estimation de la distance, l'utilisation de DNN s'est avéré à nouveau la solution la plus fiable. Avec une erreur d'estimation de distance qui, en moyenne, ne dépasse pas 12 %, elle a surpassé les méthodes conventionnelles et autres de la littérature. En général, les résultats montrent une nette supériorité du DNN par rapport aux autres méthodes.

## **Partie II : Détection et localisation des défauts pour les générateurs distribués**

Dans cette partie, les méthodes de détection et la localisation des défauts développées pour les systèmes photovoltaïques connectés au réseau ont été analysées. Les méthodes développées reposent uniquement sur les mesures de tension et de courant prises à la sortie de l'onduleur. Les étapes théoriques d'une méthode réussie de détection et de localisation des défauts ont été décrites : a) génération de symptômes, b) construction d'une table de signatures des défauts, c) conception d'un algorithme d'isolation des défauts et d) configuration des paramètres de franchissement de seuil. Ensuite, une liste des défauts étudiés a été fournie et le processus de détection des défauts a été expliqué.

Dans l'hypothèse où un dispositif de référence PV calibré était disponible pour surveiller l'irradiance réelle de la centrale et que le suivi du point de puissance

maximale fonctionnait correctement, la puissance active attendue a été comparée à la puissance active mesurée. Dans ce cas, une diminution de la puissance active a été considérée comme un indicateur fort de l'apparition d'un défaut et a marqué le lancement des algorithmes de localisation des défauts.

Deux algorithmes de localisation des défaut ont été développés, le premier, *algorithme A*, se concentre uniquement sur les défauts côté courant continu tandis que le second, *algorithme B*, inclut les défauts côté courant alternatif et couvre ainsi l'ensemble du système. L'*algorithme B* est capable de localiser treize défauts et deux groupes des défauts sur les dix-neuf cas des défaut. De plus, les deux algorithmes sont capables de localiser les défauts assez rapidement (pendant leur phase transitoire ou proche de l'état stable) dans une fourchette de 100 *ms* et 60 *ms* pour *A* et *B* respectivement. Ce niveau de vitesse a permis d'assurer une localisation avant que les dispositifs de protection ne soient déclenchés et donc de modifier les comportements des défaut en fonction des symptômes associés.

Une vérification de la robustesse à deux niveaux a été utilisée pour tester la résilience de l'algorithme dans diverses conditions. Le premier paramètre qui a été modifié était le niveau d'irradiation, les algorithmes fonctionnant avec précision dans les cinq niveaux étudiés (1000, 800, 600, 400 et 200  $W/m^2$ ). Le deuxième paramètre était l'incertitude de mesure. Du bruit a été introduit dans les mesures de tension et de courant et l'algorithme s'est avéré résistant à cela aussi. En conclusion, aucun impact des mesures erronées n'a été constaté sur les symptômes, les paramètres de franchissement de seuil ou les performances et la vitesse d'isolement.

Dans le cadre de cette deuxième partie de la thèse, une étude préliminaire de détection et localisation de défauts dans les piles à combustible connectées au réseau a également été réalisée. En détail, les mêmes types de défauts ont été étudiés avec l'étude de cas du photovoltaïque; à la place des défauts à la source du système d'énergie renouvelable, dans ce cas, seul le défaut de séchage est pris en compte. La localisation du défaut de séchage est réalisée sans interférer avec la localisation des défauts des autres parties du système. L'analyse complète est présentée à l'Annexe D.



# Publications

## Journals

1. N. Sapountzoglou, J. Lago and B. Raison, “Fault diagnosis in low voltage smart distribution grids using gradient boosting trees”, accepted in Electrical Power Systems Research.
2. N. Sapountzoglou, J. Lago, B. De Schutter and B. Raison, “A generalizable and sensor-independent deep learning method for fault detection and location in low-voltage distribution grids”, under review in Applied Energy.

## Conference Publications

1. N. Sapountzoglou, B. Raison and N. Silva, “Fault Detection and Localization in LV Smart Grids”, in 2019 IEEE Milan PowerTech, Jun. 2019, pp. 1-6. [\[Online\]](#).
2. N. Sapountzoglou, B. Raison and N. Silva, “A Fault Localization Method for Single-phase to Ground Faults in LV Smart Distribution Grids”, in 2019 International Conference ELECTRIMACS, May 2019, Salerno [\[pre-print Online\]](#).
3. N. Sapountzoglou, B. Raison, “A Grid Connected PV System Fault Diagnosis Method”, in 2019 IEEE International Conference on Industrial Technology (ICIT), Feb. 2019, Melbourne [\[Online\]](#).
4. N. Sapountzoglou, B. Raison, “Fault detection through monitoring of the AC variables in Grid Connected PV systems”, in 3<sup>ème</sup> édition de Symposium de Génie Electrique, Jul. 2018, Nancy [\[Online\]](#).

## INCITE project deliverables

7 abstracts at workshop proceedings [\[Online\]](#).



“You can not hope to build a better world without improving the individuals.”

-Marie Skłodowska-Curie







## Abstract

In this thesis, three different methods of fault diagnosis for low voltage (LV) distribution grids and two methods of fault isolation for grid-connected photovoltaic systems (GCPVs) are proposed. The proposed tools for fault diagnosis in LV grids are: a) a conventional method based on overcurrent monitoring and sparse voltage measurements across the faulty branch, b) gradient boosting trees and c) deep neural networks which are the most reliable solution demonstrating a 100 % accuracy in fault detection and an average of 12 % of error in distance estimation. Moreover, under limited available measurements their accuracy is decreased by only 4.5 %. Furthermore, two algorithms based on a signal approach are proposed for fault isolation in GCPVs. They use current and voltage measurements at the output of the inverter, examining faults occurring both in the dc and the ac side. Finally, the proposed algorithms achieve an isolation of 15 out of the 19 studied fault cases in less than 100 *ms*.

**keywords:** fault detection, fault location, fault diagnosis, machine learning, low-voltage distribution grids, distributed generation

## Résumé

Dans cette thèse, trois méthodes différentes de diagnostic des défauts pour les réseaux de distribution basse tension (BT) et deux méthodes de localisation des défauts pour les systèmes photovoltaïques raccordés au réseau (GCPV) sont proposées. Les outils proposés pour le diagnostic des défauts dans les réseaux BT sont: a) une méthode conventionnelle basée sur la surveillance de courant et de tension, b) des gradient boosting trees et c) des réseaux de neurones profonds qui sont la solution la plus fiable démontrant une précision de 100 % dans la détection des défauts et 12 % d'erreur moyenne dans l'estimation de la distance. De plus, sous des mesures disponibles limitées, leur précision n'est réduite que de 4.5 %. En outre, deux algorithmes basés sur une approche signal sont proposés pour la localisation de défauts dans les GCPV. Ils utilisent des mesures de courant et de tension à la sortie de l'onduleur, en examinant les défauts apparaissant aussi bien du côté dc que du côté ac. Enfin, les algorithmes proposés permettent la localisation de 15 sur 19 de cas de défauts étudiés en moins de 100 *ms*.

**mots-clés:** détection des défauts, localisation des défauts, diagnostic des défauts, apprentissage automatique, réseaux de distribution basse tension, production décentralisée