PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTER OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH



YAHIA FARES-MEDEA UNIVERSITY

Faculty of technology

Department of electrical engineering

Advanced Electronic System Laboratory (AESL)

THESIS

PRESENTED FOR GRADUATION FROM

DOCTORAT LMD

Domain: science and technologies

Sector: Automatic

Specialty: Automatic and systems

Presented by

RAI Nawal

Bachelor's degree in industrial maintenance: University of Ghardaïa, Algeria

Master's degree in Automatic and systems: University of Ghardaïa, Algeria

With a view to obtaining the title of Doctorate in Automatic

Theme

Contribution to the detection and localization of faults in the photovoltaic systems

Jury members:

First name and Last name	Grade	
Abdelhalim TLEMÇANI	Professor, University of Médéa	President
Abdelkader MORSLI	Senior lecturer A, University of Médéa	Examiner
Youcef BOT	Professor, University of Khemis Miliana	Examiner
Amel ABBADI	Senior lecturer A, University of Médéa	Supervisor
Fethia HAMIDIA	Professor, University of Médéa	Co-supervisor

Academic year: 2024/2025

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE YAHIA FARES DE MEDEA

Faculté de technologies

Département de Génie électrique

Laboratoire de système électronique avancée (LSEA)

THÈSE

EN VUE DE L'OBTENTION DU DIPLOME DE DOCTORAT

Domaine : Science et technologies

Filière : Automatique

Spécialité : Automatique et systèmes

Présentée par

RAI Nawal

Licence en maintenance industrielle : Université de Ghardaïa, Algérie

Master en Automatique et systèmes : Université de Ghardaïa, Algérie

En vue de l'obtention du titre de Doctorat en Automatique

Thème

Contribution à la détection et localisation des défauts dans les systèmes photovoltaïques

Devant le Jury composé de :

Nom et Prénom	Grade	
TLEMÇANI Abdelhalim	Professeur, Université de Médéa	Président
MORSLI Abdelkader	Maitre de Conférences A, Université de Médéa	Examinateur
BOT Youcef	Professeur, Université de Khemis Miliana	Examinateur
ABBADI Amel	Maitre de Conférences A, Université de Médéa	Directeur de thèse
HAMIDIA Fethia	Professeur, Université de Médéa	Co-Directeur

Année Universitaire : 2024/2025

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE



UNIVERSITE YAHIA FARES DE MEDEA

Faculté de technologies

Département de Génie électrique

Laboratoire de système électronique avancée (LSEA)

THÈSE

EN VUE DE L'OBTENTION DU DIPLOME DE DOCTORAT

Domain : Science et technologies

Filière : Automatique

Spécialité : Automatique et système

Présentée par

RAI Nawal

Licence en maintenance industrielle : Université de Ghardaïa, Algérie

Master en Automatique et systèmes : Université de Ghardaïa, Algérie

En vue de l'obtention du titre de Doctorat en Automatique

Thème

Contribution à la détection et localisation des défauts dans les systèmes photovoltaïques

Contribution to the detection and localization of faults in PV systems

Devant le Jury composé de :

Nom et Prénom	Grade	
TLEMÇANI Abdelhalim	Professeur, Université de Médéa	Président
MORSLI Abdelkader	Maitre de Conférences A, Université de Médéa	Examinateur
BOT Youcef	Proofesseur, Université de Khemis Miliana	Examinateur
ABBADI Amel	Maitre de Conférences A, Université de Médéa	Directeur de thèse
HAMIDIA Fethia	Professeur, Université de Médéa	Co-Directeur

Année Universitaire : 2024/2025

Acknowledgment

From the depths of my heart, I humbly thank Almighty Allah for His grace and guidance, which have empowered me to delve into, understand, and complete my research and thesis journey with clarity and determination.

This thesis would not have been possible without the guidance and help of several individuals who, in one way or another, contributed and extended their valuable assistance in the preparation and completion of this study. It is a pleasure to thank those who made it a possibility.

First and foremost, I would like to express my heartfelt gratitude to my supervisor, Dr. Amel Abbadi. Your unwavering support, insightful feedback, and relentless encouragement have been the cornerstone of my PhD journey. Your belief in my potential and your dedication to my success have been truly inspiring.

I also wish to express my sincere thanks to my co-supervisor, Prof. Dr. Fethia Hamidia. Your expertise, constructive feedback, and continuous support have significantly enriched my research and academic experience.

I am profoundly grateful to all the people I had the pleasure of knowing and collaborating with during my PhD years. Your collective support, insights, and encouragement have been crucial to my development as a researcher and a scholar.

In particular, a special acknowledgment is due to Dr. Khalid Yahya, whose invaluable support and collaboration have been integral to my PhD journey. Dr. Yahya has been a constant source of wisdom, encouragement, and practical advice. His willingness to provide assistance at any time, coupled with his profound insights and dedication to my success, have been truly remarkable. Whether through late-night discussions, meticulous feedback, or simply being available when needed, Dr. Yahya's contributions have profoundly shaped my research and academic development. I am deeply grateful for his mentorship and friendship.

Finally, I would like to express my appreciation to my laboratory research team for their camaraderie, collaborative spirit, and dedication to scientific inquiry. To the professors of the department and the laboratory, your mentorship and expertise have been instrumental in shaping my academic path.

Dedication

In loving memory of my grandmother, Tlemçani Kheira, peace upon her, whose unwavering support and encouragement ignited the spark that led me on this PhD journey. Though you are no longer with us, your belief in my potential and your steadfast encouragement at the outset of this endeavour continue to inspire me every day. This thesis is dedicated to you, with heartfelt gratitude for the foundation of strength and resilience you instilled in me. Your spirit lives on in every page of this work.

To my dear parents, your unwavering love, sacrifices, and unwavering belief in my dreams have been the guiding force behind my pursuit of knowledge and achievement. Your constant support and encouragement have sustained me through challenges and triumphs alike. This thesis is a testament to your enduring love and the values you have instilled in me. I dedicate this accomplishment to you with immense gratitude and love.

To my beloved little sister, Fifi, your infectious enthusiasm and unwavering belief in me have been a source of strength throughout this journey. Your laughter and optimism remind me of what truly matters in life.

To my older brother, Laid, and his wife, Asma, your unwavering support and encouragement have meant the world to me. Your belief in my abilities and sacrifices have inspired me to strive for excellence. To my nephew, Rayan, your joyful presence has brought light into my life.

To my cousins, aunts, and uncles, your love, encouragement, and understanding have been invaluable. Your support and belief in my aspirations have been a constant source of motivation.

Lastly, to my friends who have been with me through thick and thin during these PhD years, especially Nadia and Leila, your unwavering help, encouragement, and companionship have been the pillars that supported me. Whether it was sharing insights over late-night discussions, celebrating small victories, or offering a listening ear during challenges, your friendship has been a cherished gift. Your presence has enriched my journey immeasurably, and I am deeply thankful for your enduring support and camaraderie.

Abstract

With growing concerns about global warming, there is an increasing demand for renewable energy. Among the various types of renewable energies, photovoltaic (PV) energy is the most popular for global electricity production. However, PV power facilities often experience faults that can negatively impact the performance of the PV panels. These faults include hot spots, partial shading, cell aging, cell cracks, short circuits, open circuit faults, and others. An automatic fault detection and diagnosis system is crucial to effectively manage these faults. This thesis addresses two main aspects: PV generator modelling and fault detection and diagnosis. It covers basic principles of photovoltaics, various electrical models for solar cells, and methods of identifying unknown parameters. The research proposes and evaluates a metaheuristic technique for accurately identifying unknown parameters of different three PV models. Additionally, the main contribution of the thesis is a new fault detection method for the PV DC side using Ensemble Learning (EL) techniques. The study achieved advancements in accurate PV generator modelling using metaheuristic approaches, and reliable fault detection using EL techniques.

Keywords: fault detection and diagnosis, PV cell modelling, parameter identification, metaheuristic approaches, Ensemble Learning.

خلاصة:

مع تزايد المخاوف بشأن ظاهرة الاحتباس الحراري، هناك طلب متزايد على الطاقة المتجددة. من بين الأنواع المختلفة للطاقات المتجددة، تعد الطاقة الكهروضوئية هي الأكثر شعبية لإنتاج الكهرباء على مستوى العالم. ومع ذلك، غالبًا ما تواجه مرافق الطاقة الكهروضوئية أخطاء يمكن أن تؤثر سلبًا على أداء الألواح الكهروضوئية. وتشمل هذه العيوب النقاط الساخنة، والتظليل الجزئي، وشيخوخة الخلايا، وتشققات الخلايا، والدوائر القصيرة، وأعطال الدائرة المفتوحة، وغيرها. يعد نظام الكشف عن الأخطاء وتشخيصها تلقائيًا أمرًا ضروريًا لإدارة هذه الأخطاء بشكل فعال. تتناول هذه الأطروحة جانبين رئيسيين: نمذجة المولدات الكهروضوئية واكتشاف الأخطاء وتشخيصها. كما تغطي المبادئ الأساسية للخلايا الكهروضوئية، والنماذج الكهربائية المختلفة للخلايا الشمسية، وطرق تحديد المعلمات غير المعروفة. يقترح هذا البحث ويقيم تقنية أدلة عليا جديدة لتحديد المعلمات غير المعروفة بدقة لثلاثة نماذج كهروضوئية مختلفة. بالإضافة إلى ذلك، فإن المساهمة الرئيسية للأطروحة هي طريقة جديدة للكشف عن الأخطاء في الجانب الكهروضوئي المستمر باستخدام تقنيات التعلم التجميعي. حققت الدراسة تطورات في النمذجة الدقيقة للمولدات الكهروضوئية باستخدام أدلة عليا ، والكشف الموثوق عن الأخطاء باستخدام تقنيات التعلم التجميعي.

الكلمات المفتاحية: اكتشاف الأخطاء وتشخيصها، نمذجة الخلايا الكهروضوئية، تحديد المعلمات، خوارزميات أدلة عليا، التعلم التجميعي.

Résumé

Avec les inquiétudes croissantes concernant le réchauffement climatique, la demande d'énergies renouvelables augmente. Parmi les différents types d'énergies renouvelables, l'énergie photovoltaïque (PV) est la plus populaire pour la production mondiale d'électricité. Cependant, les installations photovoltaïques connaissent souvent des pannes qui peuvent avoir un impact négatif sur les performances des panneaux photovoltaïques. Ces défauts incluent les points chauds, l'ombrage partiel, le vieillissement des cellules, les fissures des cellules, les courts-circuits, les défauts de circuit ouvert et autres. Un système automatique de détection et de diagnostic des défauts est crucial pour gérer efficacement ces défauts. Cette thèse aborde deux aspects principaux : la modélisation des générateurs PV et la détection et le diagnostic des défauts. Il couvre les principes de base du photovoltaïque, divers modèles électriques de cellules solaires et les méthodes d'identification de paramètres inconnus. La recherche propose et évalue une technique métaheuristique pour identifier avec précision les paramètres inconnus de trois modèles PV différents. De plus, la principale contribution de la thèse est une nouvelle méthode de détection de défauts pour le côté CD du PV utilisant des techniques d'ensemble d'apprentissage (EA). L'étude a permis de réaliser des progrès dans la modélisation précise des générateurs photovoltaïques à l'aide d'approches métaheuristiques et dans la détection fiable des défauts à l'aide de techniques EL.

Mots clés : détection et diagnostic de défauts, modélisation de cellules PV, identification de paramètres, approches métaheuristiques, Ensemble d'apprentissage

Table of content

Acknowledgment ……..i

Dedication…….. ii

Abstract…… iii

Table of content…… iv

List of figures……..……..ix

List of tables…….…….. xi

List of abbreviations and symbols xii

Chapter I : General introduction..... 1

I.1 Background and project motivation 1

I.1.1 Background 1

I.1.2 Motivation 2

I.2 Problem statement 3

I.2.1 Manual fault detection methods 3

I.2.2 Semi-automatic fault detection methods 4

I.2.3 Automatic fault detection methods 4

I.3 Aims and objectives 4

I.4 Thesis contributions 5

I.5 Thesis outlines 5

I.5.1 Chapter II 6

I.5.2 Chapter III 6

I.5.3 Chapter IV 6

I.5.4 Chapter V 6

I.5.5 Chapter VI 6

Chapter II : Photovoltaic array and modelling..... 7

II.1 Introduction 7

II.2 Energy production 8

II.2.1 Photovoltaic effect 8

II.2.2 Photovoltaic cell 9

II.2.3 Photovoltaic cell characteristic 11

II.3 Mathematical model 13

II.3.1 One-diode model (1DM) 14

II.3.2 Two-Diode Model (2DM) 15

II.3.3 Three-Diode Model (3DM) 17

II.3.4 PV module Model 18

II.4 PV cell parameters identification methods 19
II.4.1 Analytical approaches 20
II.4.2 Numerical approaches 20
II.4.3 Metaheuristic approaches 20
II.4.4 Hybrid approaches 22
II.5 Conclusion 23
Chapter III : PV parameters identification using BB-TLBO algorithm 24
III.1 Introduction 24
III.2 Related work 24
III.3 Problem formulation 27
III.3.1 Method 01: Conventional RMSE 27
III.3.1.1 1DM of PV cell 27
III.3.1.2 2DM of PV cell 27
III.3.1.3 3DM of PV cell 28
III.3.1.4 1DM of PV module 28
III.3.1.5 2DM of PV module 28
III.3.1.6 3DM of PV module 28
III.3.2 Method 02: RMSE based on Lambert W function 28
III.3.2.1 1DM of PV cell 29
III.3.2.2 2DM of PV cell 30
III.3.2.3 3DM of PV cell 30
III.3.2.4 1DM of PV module 30
III.3.2.5 2DM of PV module 30
III.3.2.6 3DM of PV module 31
III.4 The proposed hybrid BB-TLBO algorithm 31
III.4.1 Biogeography-based optimization 32
III.4.1.1 Migration 32
III.4.1.2 Mutation 32
III.4.2 Teaching–learning-based optimization 33
III.4.2.1 Teacher phase 33
III.4.2.2 Learner phase 34
III.4.3 Proposed BB-TLBO algorithm 34
III.4.3.1 Initialization 35
III.4.3.2 Migration process 35
III.4.3.3 Teaching phase 35
III.4.3.4 Learning phase 35

III.4.3.5 Elitism phase 35
III.5 Results and discussion 37
III.5.1 Accuracy analysis 38
III.5.1.1 RTC France silicon solar cell identification 38
III.5.1.2 Monocrystalline solar module STM6-40/36 identification 38
III.5.2 Statistics analysis 43
III.5.2.1 RTC France silicon solar cell identification 43
III.5.2.2 Monocrystalline solar module STM6-40/36 identification 43
III.5.3 Convergence analysis 46
III.6 Conclusion 49
Chapter IV : Literature review on PV faults 50
IV.1 Introduction 50
IV.2 Factors that affect the efficiency of the PV module 51
IV.2.1 Cell materials 52
IV.2.1.1 Effect of corrosion on PV module efficiency 52
IV.2.1.2 Effect of degradation on PV efficiency 52
IV.2.1.3 The effect of cracked cells on PV module efficiency 53
IV.2.1.4 The effect of contact stability on PV module efficiency 54
IV.2.1.5 Effect of the selected semiconductor on solar panel performance 54
IV.2.2 PV system devices 55
IV.2.2.1 Size of inverters, batteries, and charge controller 55
IV.2.2.2 Effect of Inverter power conversion efficiency on PV performance 55
IV.2.2.3 Effect of incorrect size/ type of conductor 55
IV.2.2.4 Effect of extent of voltage drop on PV performance 56
IV.2.2.5 Odd wiring technique 56
IV.2.2.6 Insufficient use of disconnect 56
IV.2.2.7 Effect of Incorrect grounding on PV performance 57
IV.2.2.8 Underrated component usage 57
IV.2.3 Environmental factors 57
IV.2.3.1 Effect of dust on PV module efficiency 57
IV.2.3.2 The effect of solar radiation on PV module efficiency 58
IV.2.3.3 The effect of humidity on PV modules 58
IV.2.3.4 The effect of ambient temperature on PV module efficiency 58
IV.2.3.5 Effect of shading on the PV module efficiency 60
IV.2.3.6 Effect of wind direction and wind Speed 60
IV.2.3.7 Effect of Tilt angle on PV module efficiency 60

IV.3 Diagnosis terminology definitions 61
IV.3.1 Faults 61
IV.3.2 Failure 61
IV.3.3 Symptom 61
IV.3.4 Residual 61
IV.3.5 Fault detection 62
IV.3.6 Fault isolation 62
IV.3.7 Fault identification (classification) 62
IV.3.8 Fault diagnosis 62
IV.4 Faults in PV systems 62
IV.4.1 DC side faults 63
IV.4.1.1 PV array faults 63
IV.4.1.2 MPPT faults 65
IV.4.2 AC side faults 65
IV.4.2.1 Inverter faults 65
IV.4.2.2 Total blackout 65
IV.5 Fault detection and diagnosis 65
IV.5.1 Visual and thermal methods for detection of faults and diagnosis 66
IV.5.1.1 Ultraviolet (UV) fluorescence (FL) methods 66
IV.5.1.2 Electroluminescence (EL) methods 66
IV.5.1.3 Infrared thermography 67
IV.5.2 Electrical methods for detection of fault and diagnosis 68
IV.5.2.1 Statistical and signal processing approaches 69
IV.5.2.2 Methods-based I-V characteristics analysis 69
IV.5.2.3 Methods based on artificial intelligence techniques 70
IV.6 Conclusion 71
Chapter V : PV fault detection and classification using Voting-Ensemble Learning 72
V.1 Introduction 72
V.2 PV array fault diagnosis using BB-TLBO algorithm 73
V.3 Tested PV array 73
V.4 Examined faulty scenarios 73
V.4.1 Fault scenario n° 01 (Partial shading) 73
V.4.2 Fault scenario n° 02 (Short-circuit fault of module 02) 74
V.4.3 Fault scenario n° 03 (Short-circuit fault of module 02 and 03) 75
V.5 Parameter identification and fault detection and diagnosis: 76

V.6 Artificial intelligence related work 77
V.7 Ensemble learning 79
V.7.1 Voting 79
V.7.2 Bagging 80
V.8 Description of the experimental setup PV array 81
V.9 Dataset description 82
V.10 Considered faulty scenarios 84
V.10.1 Fault scenario n° 01 (string fault) 84
V.10.2 Fault scenario n° 02 (string-to-ground fault) 85
V.10.3 Fault scenario n° 03 (string-to-string fault) 86
V.11 The proposed approach based on ensemble learning 87
V.11.1 Dataset 88
V.11.2 Dataset preparation 88
V.11.2.1 Exploratory Data Analysis (EDA) 88
V.11.2.2 Pre-Processing Data 88
V.11.2.3 Splitting Data 89
V.11.3 Building model 89
V.11.3.1 Base classifier 90
V.11.3.2 Voting-Based Probability 92
V.11.3.3 Ensemble model 92
V.11.3.4 Testing data evaluation 93
V.11.4 PV fault classification 93
V.12 Results and discussion 93
V.12.1 Accuracy performance 94
V.12.2 Precision performance 94
V.12.3 Recall performance 95
V.12.4 F1 score performance 96
V.12.5 Confusion matrix performance 96
V.13 Conclusion 99
Chapter VI : General conclusion and future works 100
VI.1 General conclusion 100
VI.2 Future Works 102
References 106
List of publications and conferences 121
Annex. 122

List of figures

Figure I-1: Global annual solar PV capacity additions, 2019-2023[6].....2
Figure I-2: Different PV faults. 3
Figure II-1: A diagram showing the photovoltaic effect [16].....9
Figure II-2: The basis operation of PV cell [21]. 10
Figure II-3: Hole-electron pairs created by photons in p-n junction of PV cell [20]. 10
Figure II-4: Connecting solar with a load and conducting current [20]. 11
Figure II-5: Complete I-V characteristic of a PV cell [14]. 13
Figure II-6: Equivalent circuit of 1DM solar cell [34]. 15
Figure II-7: Equivalent circuit of 2DM solar cell [34]. 17
Figure II-8: Equivalent circuit of 3DM solar cell [34]. 18
Figure II-9: Equivalent circuit of the PV module [34]. 19
Figure II-10: Different categories of metaheuristic algorithms [49]. 21
Figure III-1: Graphical representation of the Lambert W function, $x = W(xex)$ 29
Figure III-2: Flowchart of BB-TLBO algorithm. 36
Figure III-3: IAE of the current and the power of the three adopted models of the RTC France silicon solar cell. 41
Figure III-4: IAE of the current and the power of the three adopted models of the solar module STM6-40/36 41
Figure III-5: P-V and I-V characteristic curves for the three RTC France silicon solar cell models. 42
Figure III-6: P-V and I-V characteristic curves for the three models of the STM6-40/36 solar modules. 43
Figure IV-1: Factors, which affects the efficiency of photovoltaic module performance [91]. 52
Figure IV-2: Types of PV module degradation [97]. 53
Figure IV-3: Kinds of cracks on PV modules [97]. 54
Figure IV-4: Several ways PV module generates heat [97] 59
Figure IV-5: Types of faults usually occurred in PV systems [130]. 63
Figure IV-6: UV FL failure images [131]. 66
Figure IV-7: EL failure image example [131]. 67
Figure IV-8: Thermography techniques [131]. 67
Figure IV-9: TDR-based method for FD in a disconnected PV string [9] 69
Figure IV-10: I-V curve parameters. Isc, Voc, Pmpp [131]. 70
Figure IV-11: Simplified framework of fault identification and classification based on AI-techniques: a) I–V based machine learning approach and b) thermography based [9]. 71
Figure V-1: Simulink/Matlab PV array simulation setup. 73
Figure V-2: Fault scenario n° 01: Partial shading 74
Figure V-3: Fault scenario n° 02: Short-circuit fault on module # 02. 74
Figure V-4: Fault scenario n° 03: Short-circuit fault on module # 02 and module #03. 75
Figure V-5: Matlab-Simulink PV system simulation setup. 81
Figure V-6: Internal components of the PV array. 82
Figure V-7: Faulty scenario n°01 (string fault). 85
Figure V-8: Fault scenario n° 02 (string-to-ground fault). 86
Figure V-9: Fault scenario n° 03 (string-to-string fault). 87
Figure V-10: Flowchart of $V - ELExT,\ AdB, LR, NB$. 88
Figure V-11: Confusion matrix 96
Figure V-12: The confusion matrix of the propose voting-ensemble model and its base models. 98

List of tables

Table III-1: Upper and lower limits of the 1DM, 2DM and 3DM parameters 37
Table III-2: Best values of the identified parameters for the three adopted models of the RTC France silicon solar cell 39
Table III-3: Best values of the identified parameters for the three adopted models of the solar module STM6-40/36 40
Table III-4: Statistical results obtained during the identification of 1DM, 2DM and 3DM of RTC France 45
Table III-5: Statistical results obtained during the identification of 1DM, 2DM and 3DM of the STM6-40/36 solar module 46
Table III-6: Convergence curves for three models of RTC France silicon solar cell. 47
Table III-7: Convergence curves for three models of STM6-40/36 module. 48
Table V-1: Upper and lower limits of the unknown parameters 75
Table V-2: Presents the outcomes of the PV array 1DM parameter identification process 76
Table V-3: Classification of fault scenarios according to their obtained parameters 76
Table V-4: Module characteristics used in the studied PV array 82
Table V-5: Dataset of the studied PV array 83
Table V-6: Additional features directly measured from the PV array 83
Table V-7: The hyperparameters for the proposed $V - ELExT,\ AdB, LR, NB$ algorithm 93
Table V-8: The accuracy of the proposed ensemble learning and its base models 94
Table V-9: The precision of the proposed ensemble learning and its base models 95
Table V-10: The recall of the proposed ensemble learning and its base models 95
Table V-11: The F1 score of the proposed ensemble learning and its base models 96

List of abbreviations and symbols

Abbreviations:

1DM	One diode model
2DM	Two diode model
3DM	Three diode model
ABSO	Artificial Bee Swarm Optimization
AC	Alternative current
ACO	Ant Colony Optimization
ACGWO	Enhanced form of chaotic grey wolf optimizer
AdB	AdaBoost
ADET	Adaptative differential evolution techniqe
AE	Absolute Error
AI	Artificial intelligence
ATLDE	Adaptive teaching learning and differential evolution
BA	Bat-inspired Algorithm
BB-TLBO	Biogeography-based and teaching learning based optimization
BBO	Biogeography-based Optimization
BBO-M	Modified Biogeography-based Optimization
BHCS	Biogeography-based heterogeneous cuckoo search
CatBoost	Categorical Boosting
CART	Classification and regression trees
CFO	Central Force Optimization
CIS	Copper-indium selenide
CM	Confusion Matrix
CNN	Convolutional neural network
CPSO	Chaos particle swarm optimization
CSC	Charged silicon camera
CS	Cuckoo Search
DOL	Dynamic opposite learning strategy
DOLADE	Hybrid dynamic opposite learning and adaptive differential evolution
DE	Differential Evolution
DL	Deep learning
DSEL	Deep Stack-based Ensemble Learning

DT	Decision tree
DVADE	Difference vector with adaptative mutation differential evolution
ECM	Earth capacitance measurement
EL	Ensemble learning
EMA	Exchange Market Algorithm
ES	Evolutionary Strategy
EVA	Ethylene vinyl acetate
ExT	Extra Trees Classifier
FA	Firefly Algorithm
FDD	Fault detection and diagnosis
FF	Fill Factor
FL	Fluorescence
GA	Genetic Algorithm
GACCC	Genetic Algorithm with Convex Combination Crossover
GOTLBO	Generalized oppositional teaching learning-based optimization
GOBL	Generalized opposition-based learning
GP	Genetic Programming
GSA	Gravitational Search Algorithm
GWOCS	Grey wolf optimization mix with cuckoo search
HSA	Harmony search algorithm
HS	Harmony Search
HSI	Habitat suitability index
IAE	Individual absolute error
ICA	Imperialist Competitive Algorithm
IP	Interior point method
IRTG	Infrared thermography
KELM	Kernel extreme learning machine
KNN	K-Nearest Neighbour
KPCA	Kernel PCA
LCA	League Championship Algorithm
LGBM	Light Gradient Boosting Method
LM	Levenberg-Marquardt
LSTM	Long short-term memory
MJA	Modified JAYA algorithm
ML	Machine learning

MLR	Multilinear Response
MPP	Maximum power point
MSE	Mean Squared Error
NB	Naïve Bayes
NOCT	Nominal operating cell temperature
NR	Newton Raphson
OC	Open circuit
PCAs	Principal Component Analyses
PGJAYA	Performance-Guided JAYA
PSO	Particle Swarm Optimization
PV	Photovoltaic
RAMS	Reliability, Availability, Maintainability, and Safety of PV plants
RF	Random Forest
RKPCA	Reduced Kernel PCA
$RMSE$	Root mean square error
$RMSE_C$	Conventional Root mean square error
$RMSE_L$	Root mean square error based on Lambert W function
$RMSE_{max}$	The worst value of Root mean square error
$RMSE_{mean}$	The average value of Root mean square error
$RMSE_{min}$	The lowest value of Root mean square error
STD	The standard deviation of Root mean square error
SIV	Suitability index variable
SVM	Support Vector Machine
SSTDR	Spread spectrum time-domain reflectometry
STC	Standard test condition
TLABC	Teaching learning hybrid with artificial bee colony
TLBO	Learning-Teaching-based Optimization
UV	Ultraviolet
V-EL	Voting-Based Ensemble Model (ExT, AdB, LR, NB)
XGBoost	Extreme Gradient Boosting

Symbols:

μ	Emigration rates
λ	Immigration rate

Ω	Ohm
f_m	Error function
n_1	First diode ideality factor
n_2	The second diode ideality factor
n_3	The third diode ideality factor
q	Electron charge
x	Solution vector
A	Ampere
I	Current
I-V	Current-voltage
I_{d1}	First diode current
I_{d2}	Second diode current
I_{d3}	Third diode current
I_{mp}	Maximum power point current
I_{Ph}	Photo-generated-current
I_{s1}	First diode reserve saturation current
I_{s2}	Second diode saturation current
I_{s3}	Third diode saturation current
I_{sc}	Short-circuit current
K	Boltzmann constant
K_r	Predefined keeping rate
LL	Lower bound limit
NP	Total number of students
P-V	Power-voltage
P(CL_x=K)	The prior probability of a class being indexed as K
P_m	Measured power
P_{max}	Maximum power output
R_s	Series resistance
R_{sh}	Shunt resistance

T	Temperature
TF	Teaching factor
UL	Upper bound limit
V_{mp}	Maximum power point voltage
V_{oc}	Open circuit voltage
V_t	Junction thermal voltage
X_{mean}	The mean position of the current class
X_s	The individual being selected
X_i	Position of the i^{th} learner
$X_{(new,i)}$	New position
$X_{(old,i)}$	Old position
$X_{teacher}$	The position of the current teacher

Chapter I : General introduction

This thesis aims to provide an overview of the context of our research and the motivation behind it, followed by an exploration of the purposes and objectives of our work. We started first by focusing on the main contributions related to the application of meta-heuristic algorithms in the estimation of the parameters of different photovoltaic models.

And secondly, through the use of artificial intelligence based on machine learning approaches for detection and diagnosis of defects in the direct current side (DC) of photovoltaic systems (PV).

I.1 Background and project motivation

I.1.1 Background

Photovoltaic energy is a form of renewable energy that is derived from the direct conversion of sunlight. The term "photovoltaic" refers to the generation of electricity from light (photo) and the movement of electrons through a semiconductor material (voltaic).

Photovoltaic (PV) energy is positioned at the heart of global energy revolutions, illustrating a crucial transition towards sustainable and renewable sources. Over recent years, its adoption has expanded exponentially, redefining the energy landscape and significantly contributing to the diversification of electricity sources globally, thanks to its advantages [1]–[4][5], which are:

Solar irradiation is easily available without fuel costs.

- The production of photovoltaic energy produces no pollution or waste.
- Photovoltaic systems require less maintenance than other alternatives.
- They operate unattended with minimal periodic maintenance, thereby reducing labour costs.
- Although the initial cost is high, photovoltaic systems are profitable in the long run.
- Photovoltaic systems generate power locally, eliminating the need for long transmission lines.

Statistics for the first half of 2023 reveal a notable expansion of photovoltaic (PV) installations worldwide. In China, a remarkable growth of 153% compared to 2022 was recorded, while in Germany, a significant increase of 102% was shown. The United States followed suit with a notable increase of 153%, [6].

Conversely, India witnessed a significant contraction of 47% in its photovoltaic installations during the first three quarters of 2023 compared to the same period the previous year, primarily due to delays in project execution related to logistical obstacles and on-site issues [6].

According to the BLACKRIDGE Research and Consulting report on the global solar energy market, the worldwide photovoltaic solar energy industry is poised to achieve a new record with approximately 310 GW of annual solar deployments projected for 2023, marking a significant increase of 29.7% compared to the 239 GW installed in 2022. These data underscore the swift adoption of PV technology on a global scale, with several countries displaying remarkable penetration rates [6].

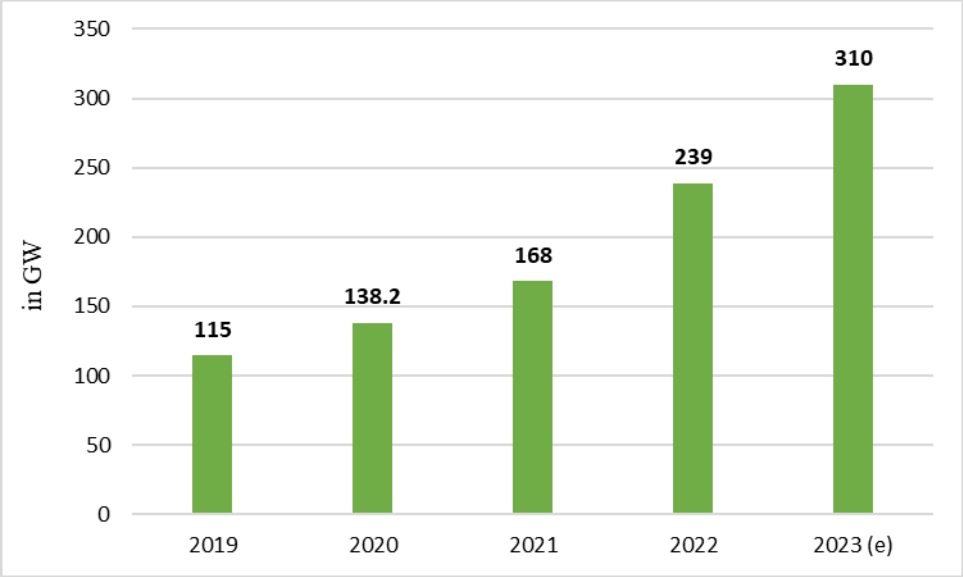


Figure I-1: Global annual solar PV capacity additions, 2019-2023[6].

By the end of 2023, key global nations in solar energy, including the United States, the European Union, and India, are projected to achieve a total annual solar capacity of 96.4 GW. Simultaneously, other countries such as Japan, Vietnam, South Korea, Brazil, and Australia are actively incorporating solar energy into their electrical grids. Considering the rate of solar energy additions in these nations, which collectively hold the majority of global solar energy capacity, worldwide solar additions are expected to surpass 310 GW. This notable growth in solar installations is primarily attributed to factors like the rapid expansion of solar capacity in China, driven by the increasing competitiveness of solar energy in both distributed and large-scale applications. Additionally, the growing awareness of the appeal of solar energy plays a crucial role in this trend[6].

I.1.2 Motivation

The durability of photovoltaic systems is consistently called into doubt for a variety of reasons, including fluctuating environmental conditions, the inherent complexity of the systems, and their vulnerability to failure. The aforementioned anomalies, which may manifest as external disturbances

or faults at the component level, possess the capacity to undermine the functionality and security of photovoltaic installations. In light of the escalating energy demand, the exponential expansion of photovoltaic systems underscores the criticality of ensuring the security of these systems and reducing energy losses via precise and prompt fault detection [7].

I.2 Problem statement

Fault detection in photovoltaic systems poses a significant challenge, garnering particular attention. These faults can manifest themselves at various levels of the system, ranging from the photovoltaic generator to the power conditioning units and the load side. Commonly observed faults include line-to-ground L-G faults, line-to-line (L-L) faults, open circuit faults, arcing faults, and mismatch faults. Therefore, establishing robust detection and localization mechanisms is crucial, aiming to reduce maintenance costs, and prevent energy loss, equipment damage, and safety risks, especially considering the continued growth in scale and complexity of these systems [7], [8].



Figure I-2: Different PV faults.

In this context, many researchers are involved in the development of various fault detection and diagnosis methods, categorized into three main groups: methods based on a manual process, methods based on a semi-automatic process, and methods based on an automatic process [7], [8].

I.2.1 Manual fault detection methods

These techniques are based on visual inspection, time-domain reflectometry, earth capacitance measurement, and spectral analysis. These methods are offline procedures that require disconnecting photovoltaic modules or strings from the system to assess defects, with various measurements and visual inspections contributing to fault identification. Manual inspection is well-suited for small-scale PV plants at the string level, enabling accurate defect identification through direct examination and precise measurements [9].

I.2.2 Semi-automatic fault detection methods

These techniques involve the use of tools such as thermal cameras and infrared or electroluminescence imaging. These methods are typically applied to larger-scale PV plants. Unlike manual methods, semi-automatic approaches are conducted online in real-time, providing a more dynamic and continuous assessment of the system. The use of imaging technologies allows for a more comprehensive understanding of the PV plant's condition, facilitating the identification and localization of faults, especially in larger installations where manual inspection may be impractical [9].

I.2.3 Automatic fault detection methods

These strategies leverage data loggers and monitoring systems for real-time data collection, employing algorithms to assess system performance. These methods, applicable across a spectrum of PV generator capacities, function online and fall into two key approaches. The first entails the development of algorithms grounded in mathematical analyses for fault detection, identification, and potential localization. The second, more sophisticated approach integrates artificial intelligence (AI) techniques within intricate scenarios to proactively address faults, enable self-healing, plan maintenance, and autonomously make decisions. These automated approaches make use of extensive datasets, encompassing measured currents, voltages, images, solar irradiance, and infrared or electro-luminance images, facilitating comprehensive fault recognition, categorization, and localization [9].

I.3 Aims and objectives

Artificial intelligence (AI) has become an indispensable tool in addressing complex challenges across diverse engineering domains. Particularly in the domain of optimization, metaheuristic techniques play a crucial role. These AI-inspired algorithms, mirroring natural processes, prove highly effective in solving complex engineering optimization problems. Moreover, the synergy of AI extends to machine learning techniques, which enable systems to learn from data and improve performance over time. In recent years, the advent of deep learning techniques has further propelled AI applications, allowing for the detection and classification of intricate patterns and faults. This integrated approach, combining artificial intelligence, metaheuristic optimization, machine learning, and deep learning techniques, represents a powerful paradigm in addressing multifaceted engineering challenges [10][11][12]. In the context of this work, metaheuristic optimization algorithms such as Shuffled Frog Leaping Algorithm (SFLA), Biogeography-based Optimization (BBO), and Learning-Teaching-based Optimization (TLBO) have been employed in modelling PV cell and module. Furthermore, machine learning approaches have been applied for the detection and classification of multiple PV faults.

The overall research objectives are as follows:

- To study metaheuristic optimization algorithms and apply them in PV parameter identification.
- To develop a new algorithm to identify the parameters of solar cells and modules.
- To develop a PV fault detection and classification algorithm based on machine learning and metaheuristic optimization algorithms.

I.4 Thesis contributions

Two contributions are developed in this thesis, they are descried bellow:

- The development of a new metaheuristic optimization algorithm named BB-TLBO. This algorithm is the result of the hybridization between the Biogeography-Based Optimization (BBO) and the Teaching Learning-Based Optimization (TLBO) algorithms. It is proposed to identify the parameters of solar cells and modules based on three mathematical models: the one-diode model (1DM), the two-diode model (2DM), and the three-diode model (3DM). BB-TLBO uses two different objective functions: the conventional root mean square error (RMSE) and the RMSE based on lambert W function. Two cases are studied to verify the effectiveness of the algorithm: the RTC France silicon solar cell and the STM6-40/36 monocrystalline commercial photovoltaic solar module. The new developed BB-TLBO algorithm has set a new benchmark for optimization methods in this field.

This contribution is discussed in details in chapter III and it is published in the following journal:

Rai, N., Abbadi, A., Hamidia, F., Douifi, N., Abdul Samad, B., & Yahya, K. (2023). Biogeography-based teaching learning-based optimization algorithm for identifying one-diode, two-diode and three-diode models of photovoltaic cell and module. *Mathematics*, *11*(8), 1861.

- Developing a new strategy for PV fault detection and classification based on ensemble learning with voting criteria. This new strategy combines four base learners such as extra tree classifier, naïve baise, logistic regression, and adaptative boost classifier. This voting-ensemble learning can detect and classify three type of faults such as: string fault, string to ground fault, and string to string fault. The voting-based ensemble learning approach has demonstrated outstanding performance compared to all the base learners across all performance indicators.

This contribution is detailed in chapter V.

I.5 Thesis outlines

This thesis is organised into six chapters. Following this chapter:

I.5.1 Chapter II

This chapter provides an overview of the photovoltaic effect and the construction of the photovoltaic cell. The modelling of the photovoltaic generator is explored in detail, with a focus on the various mathematical models of solar cells, including the one-diode model (1DM), the two-diode model (2DM), and the three-diode model. Then, different methods presented in literature to identify the unknown parameters of the PV models mentioned above are presented and discussed.

I.5.2 Chapter III

This chapter presents the BB-TLBO algorithm and its application to the identification of solar cell and module parameters based on three different models. This algorithm is evaluated on two cases: RTC France silicone solar cell and STM6-40/36 module at a certain irradiance level. BB-TLBO uses two different objective functions: the conventional root mean square error (RMSE) and the RMSE based on lambert W function.

I.5.3 Chapter IV

In this chapter, an overview of different factors that affect the efficiency of PV module is discussed, as well as the definition of diagnosis terminology used in this field to provide a nuanced understanding of the diagnostic terms crucial for decoding the language of PV systems. Then, the classification of different types of faults that can occur in PV systems with their methods of detection and diagnosis which play a pivotal role in ensuring the optimal performance and longevity of PV systems.

I.5.4 Chapter V

This chapter begins by presenting a PV array with three different fault scenarios and focuses, at the first stage, on utilizes the BB-TLBO algorithm for identifying model parameters in both healthy and faulty PV arrays. The resulting parameters help classify the scenarios as faulty or healthy.

Subsequently, the chapter discusses a published dataset used for fault detection through a voting- based ensemble learning approach. An overview of ensemble learning and the dataset description is provided, followed by details of the learning process and the proposed model's architecture. The results are evaluated based on five performance indicators: accuracy, precision, recall, F1 score, and confusion matrix.

I.5.5 Chapter VI

In this chapter a broad summary of the achievements and contributions made by this thesis are presented. Next, suggestions are made for additional development.

Chapter II : Photovoltaic array and modelling

II.1 Introduction

In recent times, there has been an incessant rise in global energy demand. Faced with this growing need, wealthy countries have increasingly turned to renewable energy sources as key components of their energy strategies. Amidst the array of alternatives, photovoltaic (PV) energy emerges as a beacon of hope. Unlike conventional energy sources like oil, coal, and nuclear power, which emit a plethora of pollutants into the atmosphere, PV energy champions cleanliness and environmental preservation. By harnessing the sun's abundant rays, PV systems offer a renewable, sustainable, and eco-friendly solution to our energy requirements.

Accurate modeling is crucial for the efficient operation of a photovoltaic system. In recent years, substantial progress has been made in understanding the behavior (characteristics) of photovoltaic systems through mathematical modeling.

These models aim to reproduce the actual performance of photovoltaic cells by adjusting their current-voltage (I-V) data measured under various operating conditions. Among the most used models are:

- The one-diode model (1DM) uses five (5) parameters to represent the solar cell operation.
- The two-diode model (2DM) contains seven (7) parameters to describe the operation of solar cells.
- Three-diode model (3DM): nine (9) parameters are used to represent the operation of solar cells.

These models require the estimation of their parameters. Accurate parameter identification from PV models holds significant importance, not only for assessing PV cell performance, but also for enhancing PV cell design, refining manufacturing processes, and ensuring quality control. Consequently, the development of effective parameter identification techniques has become increasingly imperative [13].

For this, researchers have explored various approaches, which can be broadly classified into two main methods:

- Conventional methods: these techniques include analytical and numerical methods;

- Advanced methods: These strategies include metaheuristic and hybrid methods.

II.2 Energy production

Photovoltaic production refers to the generation of electricity from solar energy using solar cells or it called also photovoltaic cells. These cells, often made from silicon for the reason that silicon is one of the most abundant elements on earth, directly convert sunlight into electricity through the photovoltaic effect.

Photovoltaic production can be used to power a wide range of applications, from small electronic devices to large electrical grids. It is widely regarded as a clean and renewable energy source, helping to reduce greenhouse gas emissions and mitigate the effects of climate change [14].

II.2.1 Photovoltaic effect

The photovoltaic effect, first discovered by French scientist Alexandre Edmond Becquerel in 1839, is the process of generating voltage or electric current within a photovoltaic cell when it is exposed to sunlight. This effect underlies the functionality of solar panels, as it is responsible for converting sunlight into usable electrical energy [14]–[17].

Over the years following its discovery, numerous scientists endeavoured to advance electricity generation technologies rooted in the photovoltaic effect. By 1870, investigations into the PV effect extended to selenium, revealing power efficiencies of only 1% to 2%, well below the threshold deemed acceptable for viable energy conversion. Subsequent developments in the 1950s led to the creation of high-purity crystalline silicon, culminating in the breakthrough development of silicon PV cells with an initial efficiency of 4% at Bell Labs in 1954. With further enhancements, efficiency rates soared to 11%. These advancements marked a pivotal moment, propelling the photovoltaic effect into a new era of solar power generation [14]–[17].

Photovoltaic cells are composed of two types of semiconductors, p-type and n-type, which are combined to form a p-n junction. This junction creates an electric field, directing electrons toward the positive p-side and holes toward the negative n-side. This field causes negatively charged particles to move in one direction and positively charged particles in the other direction [14]–[17].

Light consists of photons, which are small bundles of electromagnetic radiation or energy. Photons can be absorbed by photovoltaic cells, which make up solar panels. When light with the right wavelength hits these cells, the energy from the photons is transferred to an atom in the semiconducting material at the p-n junction. Specifically, the energy is passed to the electrons in the material, causing them to jump to a higher energy state known as the conduction band. This shift leaves behind a "hole"

in the valence band from which the electron jumped. This movement of the electron due to the absorbed energy creates two charge carriers: an electron and a hole [14]–[17].

In their normal state, electrons maintain the structure of the semiconducting material by forming bonds with adjacent atoms, making them unable to move freely. However, when electrons are excited to the conduction band, they can move through the material. The electric field created by the p-n junction causes electrons and holes to move in opposite directions. Instead of being attracted to the p-side, the free electron typically moves toward the n-side. This movement of the electron creates an electric current in the cell. As the electron moves, it leaves behind a hole that can also move, but in the opposite direction toward the p-side. This movement generates current within the cell [14]–[17].

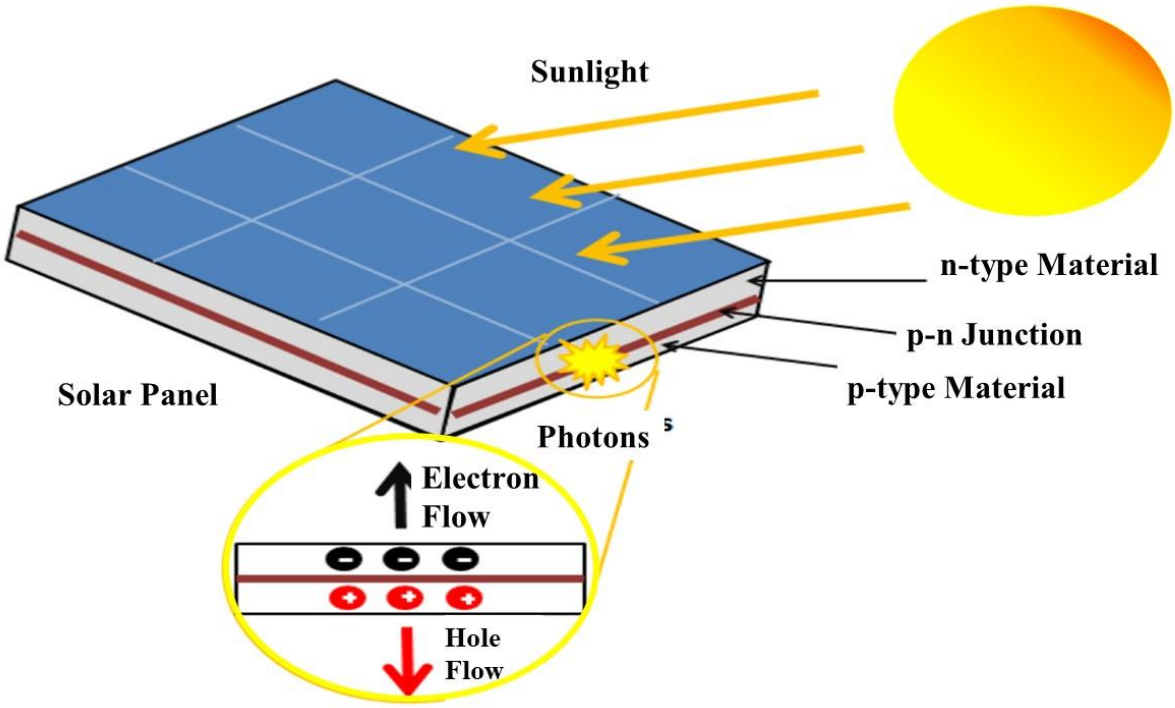


Figure II-1: A diagram showing the photovoltaic effect [16].

II.2.2 Photovoltaic cell

Photovoltaic cells, commonly known as solar cells were developed by Bell Labs in 1954, while the photovoltaic effect was discovered as early as 1839. The first generation of solar cells was fabricated employing silicon wafers in either monocrystalline or polycrystalline forms[18], [19].

Solar cells are electrical devices that convert solar energy directly into electrical energy through the photovoltaic effect. They are the fundamental components of solar panels, which utilize sunlight to generate electrical charges within the semiconductor materials of the cell, allowing the charges to move freely. The structure of a photovoltaic cell includes several layers of materials, the most important of which are semiconductors such as silicon. This design ensures there are no mechanical movements or

moving parts, allowing for continuous operation with minimal maintenance and an extended lifespan compared to other power generation technologies [18]–[21].

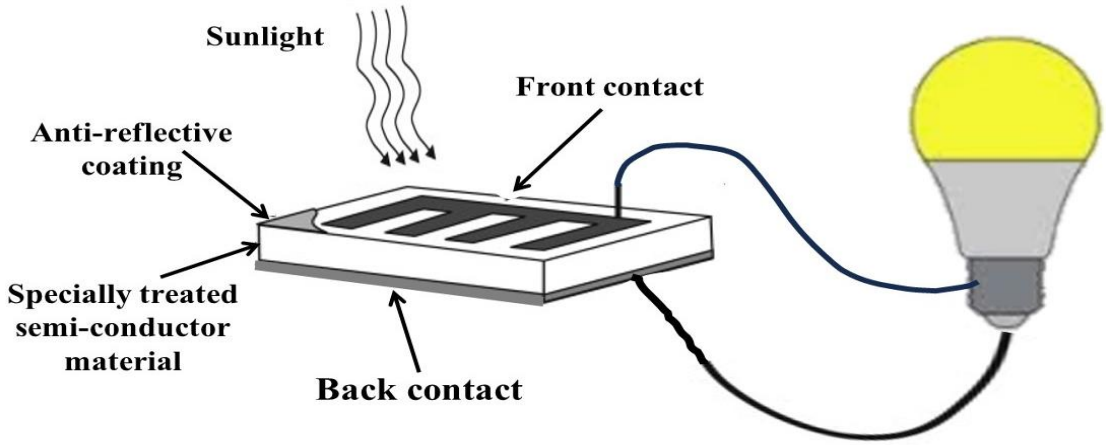


Figure II-2: The basis operation of PV cell [21].

Figure II-2 illustrates the structure of PV cell. Initially, a substrate is formed using a semiconductor of p-type, enriched with a small quantity of boron atoms. Subsequently, phosphorous atoms are introduced into the substrate to establish a p-n junction through high-temperature diffusion processing. Close to the junction, electrons from the n-side migrate towards the p-side, creating a layer of ions with a positive charge on the n-side. Likewise, holes in the p region disperse into the n region, resulting in a layer of ions with a negative charge on the p-side. This process leads to the formation of a potential barrier stemming from the reorganized positively and negatively charged ions [22].

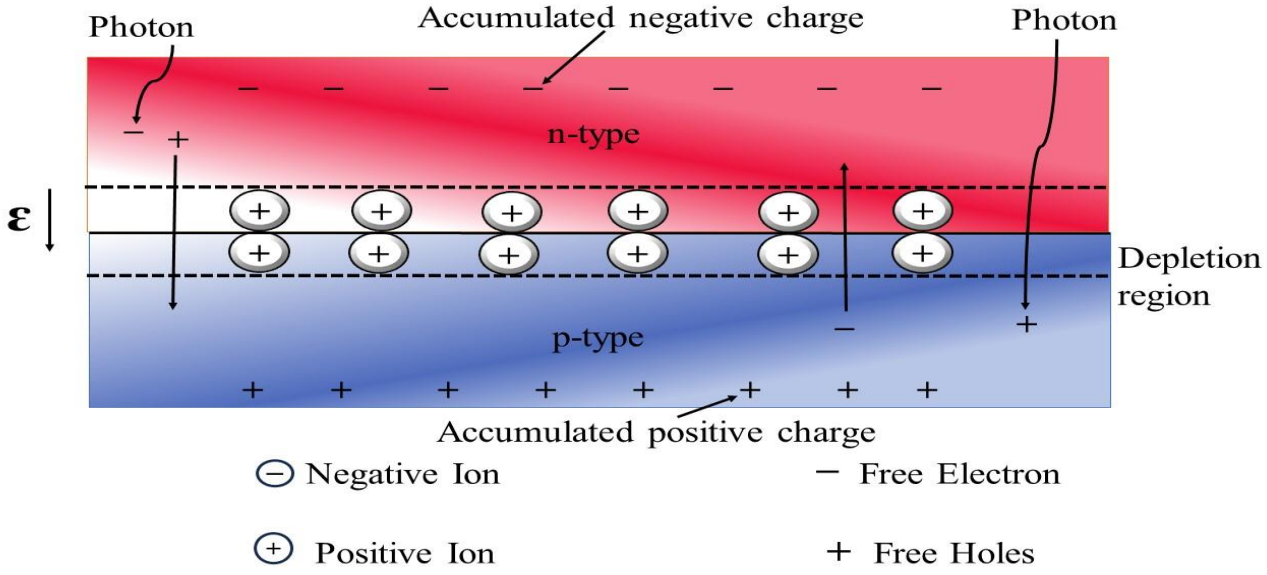


Figure II-3: Hole-electron pairs created by photons in p-n junction of PV cell [20].

The depletion region, referred to as a potential barrier, lacks mobile charge carriers, thus impeding the movement of electrical charges, as depicted in Figure II-3. When sunlight interacts with the p-n

structure, electrons become energized by photons, generating hole-electron pairs. These charges are segregated by the potential barrier at the p-n junction. Concurrently, the p-type semiconductor extracts holes, while the n-type semiconductor extracts electrons. Upon connection of an external circuit to the n-type and p-type semiconductors, illustrated in Figure II-4 a circuit is formed.

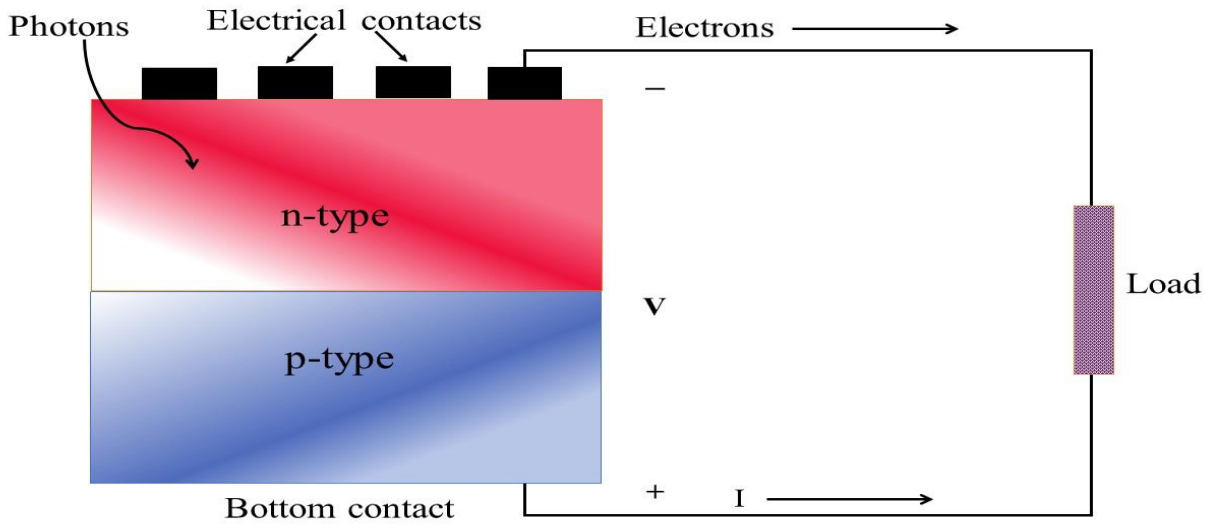


Figure II-4: Connecting solar with a load and conducting current [20].

Electrons flow through the external circuit from the n-type semiconductor to the p-type semiconductor, where they combine with the holes, thus generating current in the external circuit. Metallic contacts are affixed to both sides of the junction to capture the electrical current produced by photons on one side. Mono-crystalline and poly-crystalline silicon are the predominant materials utilized in PV cells. In practical applications, solar cells are interconnected in series and parallel configurations to amplify voltage, current, and power output. These cells are assembled into PV modules and subsequently encased in flat glass to shield them from environmental elements such as dust and water [23]–[25].

II.2.3 Photovoltaic cell characteristic

As a reminder, typically, the PV cell is engineered to operate as a generator, converting solar energy into electrical energy for a load. Conversely, when interconnected with other cells, it can act as a receiver, absorbing energy transmitted by neighbouring cells. The electrical behaviour of a PV cell, depicting its various operating states, is depicted in Figure II-5. Within this illustration, two crucial operating parameters stand out: the short circuit current (Isc) and the open circuit voltage (Voc) which subtly frame its three zones [26], [27]:

- The short-circuit current (Isc) is the maximum current achievable in a circuit, measured where the curve meets the vertical axis. It is typically expressed in amps and determined by connecting

the positive and negative terminals of the cell. When these terminals are shorted, the circuit's output voltage drops to zero. As its name suggests, the short-circuit current occurs when the cell is short-circuited. This current is directly related to the cell's surface area exposed to light and depends on the intensity of light absorbed by the cell for a given area [20][14].

- Voc, or open-circuit voltage, denotes the maximum achievable output voltage from the circuit, intersecting with the horizontal axis on the curve. This voltage is observed when PV cells are connected to a substantial resistance or under no-load conditions, resulting in zero current flow. Typically, around 0.6 V for crystalline cells, Voc represents the voltage across the cell without any current passing through. These critical points, along with Isc, mark the boundaries of the cell's operation, defining three distinct zones [14].
- Zone **I** in Figure II-5 depicts the cell's generator operation, where $0 \leq I \leq Isc$ and $0 \leq V \leq Voc$. This represents the typical functioning mode of the cell, where it supplies the electrical energy produced to the load [14].
- Zone **II** corresponds to the cell's receiver operation, characterized by $I > Isc$ and $V < 0$. In this mode, when the current surpasses the short circuit level due to external circuit constraints, the cell generates a negative voltage across its terminals. Excessive current can lead to cell damage if the voltage across its terminals reaches a threshold known as the breakdown voltage (Vbreakdown). Experimental investigations on various types of crystalline cells have revealed that the breakdown voltage typically ranges between 10 V and 30 V [28].
- Zone **IV** again corresponds to the cell's functioning as a receiver, characterized by $I < 0$ and $V > Voc$. In this state, when the voltage across the cell exceeds its open circuit voltage, a reverse current flow through the cell. Exceeding a certain limit of this reverse current can cause irreparable damage to the cell. Defining this limit is challenging as it depends on the technology and internal current flow phenomena within the faulty cell [14].
- Zone **III**: It should be noted that zone **III** in Figure II-5 corresponds to a prohibited operating zone for a semiconductor.

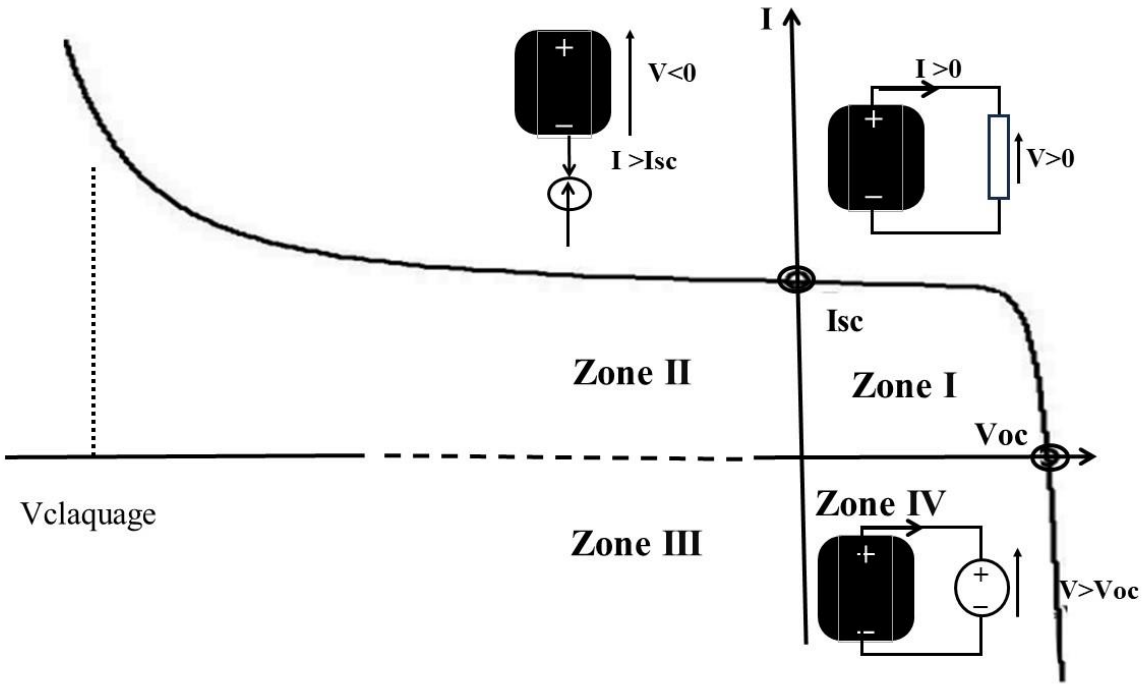


Figure II-5: Complete I-V characteristic of a PV cell [14].

II.3 Mathematical model

The assessment of PV module operation and the design of energy systems rely on analysing the electrical current-voltage characteristics of the modules across varying radiation levels and temperatures affecting the PV cell. Modelling PV cells and modules involves employing equations that offer varying levels of approximation to the real device. Numerous electrical models have been suggested in literature for simulating the performance of PV cells under diverse circumstances. The intricacy of these models varies based on the quantity of parameters that need to be determined. Each model essentially refines the ideal model, which includes a current source representing the incident solar power and a diode symbolizing the PN junction. Extra elements may be incorporated to better account for the comprehensive depiction of the PV cell's behaviour in specific operational regions [29].

Among the most common models, The Ideal single-diode (ISD) model offers the least computational complexity. The one-diode model (1DM) stands as the predominant model utilized for PV cells and modules due to its simplicity and accurate portrayal of power generation. Over time, advancements led to more refined models like the Bishop model, which delves into the behaviour of PV cells under reverse bias conditions. Furthermore, the two-diode model (2DM) enhances the accuracy of the single diode model by considering resistive losses and recombination mechanisms across different electrical components. Three-diode model (3DM) were developed to achieve higher accuracy compared to the 1DM and 2DM. Dynamic models, incorporating capacitance, have also emerged to capture the PV cell's dynamic behaviour. These models vary in complexity, with differing

numbers of parameters required for calculating the I-V characteristic. In this thesis, we mainly used the one-diode model, the two-diode model and the three-diode model which we will describe in more detail in next section [30]–[32].

II.3.1 One-diode model (1DM)

The one-diode model is the most commonly used model in the literature. It is used for PV cells and PV modules due to its simplicity and good accuracy in power generation zone II compared to the standard model. The one-diode electrical model of the PV cell is given in Figure II-6 [33]

The 1DM is composed of a photo-generated current represented as a current source in parallel with a diode and a so-called shunt resistance R_{sh} , all in series with a resistance R_s as shown in Figure II-6 [33]. The current output can be formulated as:

$$I_L = I_{ph} - I_{d1} - I_{sh} \tag{II-1}$$

Where: the photo-generated-current, the diode-current and the shunt-resistance-current are denoted by I_{ph}, I_{d1} and I_{sh} respectively. The I_{d1} can be calculated according to Shockley as follows:

$$I_{d1} = I_{s1}(exp\left(\frac{V_L+I_L R_s}{n_1 V_t}\right) - 1) \tag{II-2}$$

Where:

- I_{s1}: Diode reserve saturation current,
- V_L: Output voltage;
- n_1: Diode ideality constant;
- V_t represents junction thermal voltage, it is formulated as:

$$V_t = \frac{KT}{q} \tag{II-3}$$

Where:

- $K = 1.3806503 * 10^{-23}$ J/K stands to constant of the Boltzmann;
- $q = 1.60217646 * 10^{-19}$C stands to electron charge;
- T stands to junction temperature in Kelvin.

The parallel resistance current I_{sh} is calculated as:

$$I_{sh} = \left(\frac{(V_L+I_L R_s)}{R_{sh}}\right) \tag{II-4}$$

After replacing the equations (II-2) and (II-4) in equation (II-1), the output current of one-diode model is:

$$I_L = I_{ph} - I_{s1}\left(exp\left(\frac{V_L+I_L R_s}{n_1 V_t}\right) - 1\right) - \left(\frac{(V_L+I_L R_s)}{R_{sh}}\right) \quad \text{(II-5)}$$

As can be seen in Equation II-5, 1DM has five (5) unknown parameters which are:

- The photo-current I_{ph}
- The saturation current of the diode I_{s1}
- The series resistance R_s
- The shunt resistance R_{sh}
- The ideality factor n_1

While the One Diode Model (1DM) is frequently utilized for modelling PV systems due to its simplicity, it possesses certain inherent limitations, including:

- The One Diode Model (1DM) was initially devised under the premise that recombination loss in the depletion region is negligible. However, real-world conditions reveal this assumption to be flawed.
- The effectiveness of the 1DM diminishes when confronted with fluctuations in temperature.
- At lower irradiance levels, especially concerning the open-circuit voltage (Voc), the accuracy of the 1DM experiences a decline.

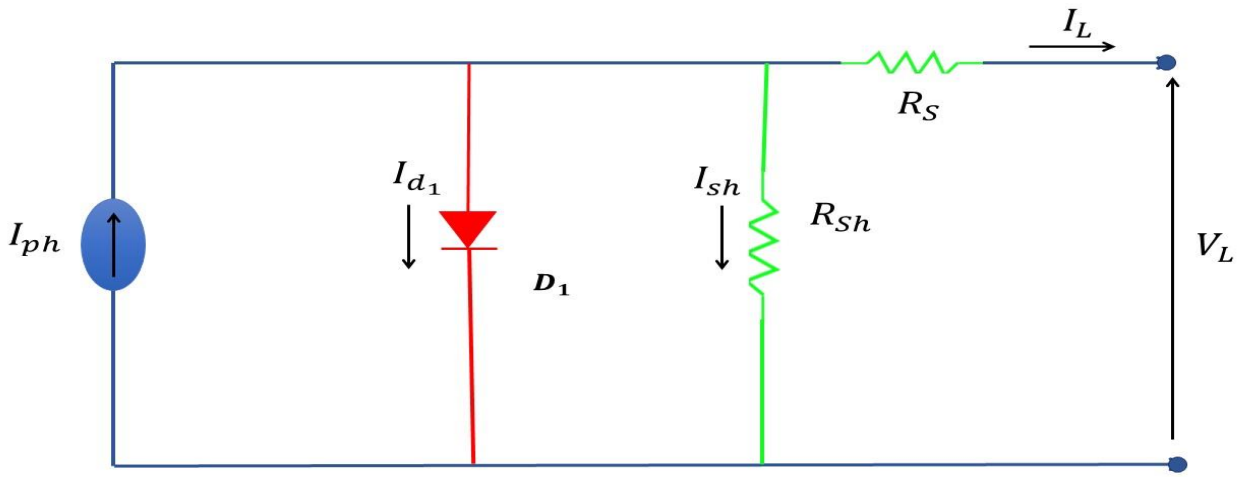


Figure II-6: Equivalent circuit of 1DM solar cell [34].

II.3.2 Two-Diode Model (2DM)

The two-diode model improves the one-diode model because it takes into account resistive losses and recombination mechanisms in the different electrical components of the circuit [31].

In this model, there is a diode to represent the polarization phenomena of the PN junction. The second diode represents the recombination of minority carriers, on the one hand on the surface of the material and on the other hand in the volume of the material [31].

As illustrated in Figure II-7, the 2DM consists of two diodes linked in parallel with the current source, and a shunt resistance R_{sh}. this combination is linked in series with a resistance R_s.

I-V relationship can be stated by using Kirchhoff's current law with Shockley diode equation as follows:

$$I_L = I_{ph} - I_{d1} - I_{d2} - I_{sh} \tag{II-6}$$

After following the same procedure used in the first model (1DM), the output current is formulated as follows:

$$I_L = I_{ph} - I_{s1}\left(exp\left(\frac{V_L+I_L R_s}{n_1 V_t}\right) - 1\right) - I_{s2}\left(exp\left(\frac{V_L+I_L R_s}{n_2 V_t}\right) - 1\right) - \left(\frac{(V_L+I_L R_s)}{R_{sh}}\right) \tag{II-7}$$

In contrast to the characteristic equation of the one Diode Model (1DM), the two Diode Model (2DM) incorporates two extra parameters associated with the second diode (I_{s2} and n_2), as outlined in Eq.II.7. Consequently, the 2DM involves consideration of seven unknown parameters:

- The photo-current I_{ph}
- The saturation current of the first diode I_{s1}
- The saturation current of the second diode I_{s2}
- The series resistance R_s
- The shunt resistance R_{sh}
- The ideality factor n_1
- The ideality factor n_2

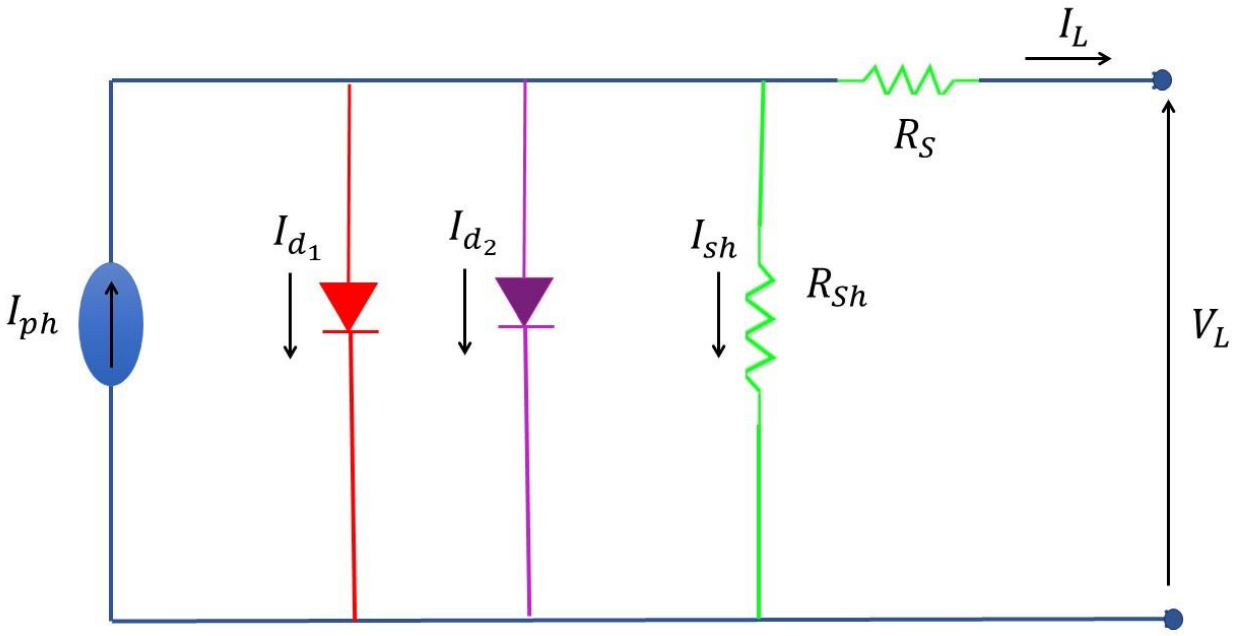


Figure II-7: Equivalent circuit of 2DM solar cell [34].

II.3.3 Three-Diode Model (3DM)

The three-diode model (3DM) proves more suitable for the industrial settings. Variable parameters n_1 and n_2 underscore the inadequacy of the double-diode model (2DM) in capturing the diverse current aspects of solar cell. The 3DM accounts for the effects of boundary gains and current losses due to leakage in the circuit. Consequently, the equivalent circuit of the 3D model, post the parallel integration of the third diode, is depicted in Figure II-8. Within this model, I_{d1} represents the current diode arising from recombination and diffusion processes in the near-neutral region, characterized by a volume of $n_1 = 1$. Additionally, the inclusion of a new diode, I_{d2}, accounts for spatial effects replicated with $n_2 = 2$. Introducing a third diode, I_{d3}, alongside the existing ones aims to address recombination within defect regions. Performance current calculations are anchored in Kirchhoff's Current Law (KCL) based on the aforementioned estimation as follows [35] [36]:

$$I_L = I_{ph} - I_{d1} - I_{d2} - I_{d3} - I_{sh} \quad \text{(II-8)}$$

After following the same procedure used in the first model (1DM), the output current is formulated as follows:

$$I_L = I_{ph} - I_{s1}\left(exp\left(\frac{(V_L+I_LR_s)}{n_1V_t}\right)-1\right) - I_{s2}\left(exp\left(\frac{(V_L+I_LR_s)}{n_2V_t}\right)-1\right) - I_{s3}\left(exp\left(\frac{(V_L+I_LR_s)}{n_3V_t}\right)-1\right) - \left(\frac{(V_L+I_LR_s)}{R_{sh}}\right) \quad \text{(II-9)}$$

In contrast to the characteristic equation of the two Diode Model (2DM), the three Diode Model (3DM) incorporates two extra parameters associated with the third diode (I_{s3} and n_3), as outlined in Equation II-9. Consequently, the 3DM involves consideration of nine unknown parameters:

- The photo-current I_{ph}
- The saturation current of the first diode I_{s1}
- The saturation current of the second diode I_{s2}
- The saturation current of the third diode I_{s3}
- The series resistance R_s
- The shunt resistance R_{sh}
- The ideality factor n_1
- The ideality factor n_2
- The ideality factor n_3

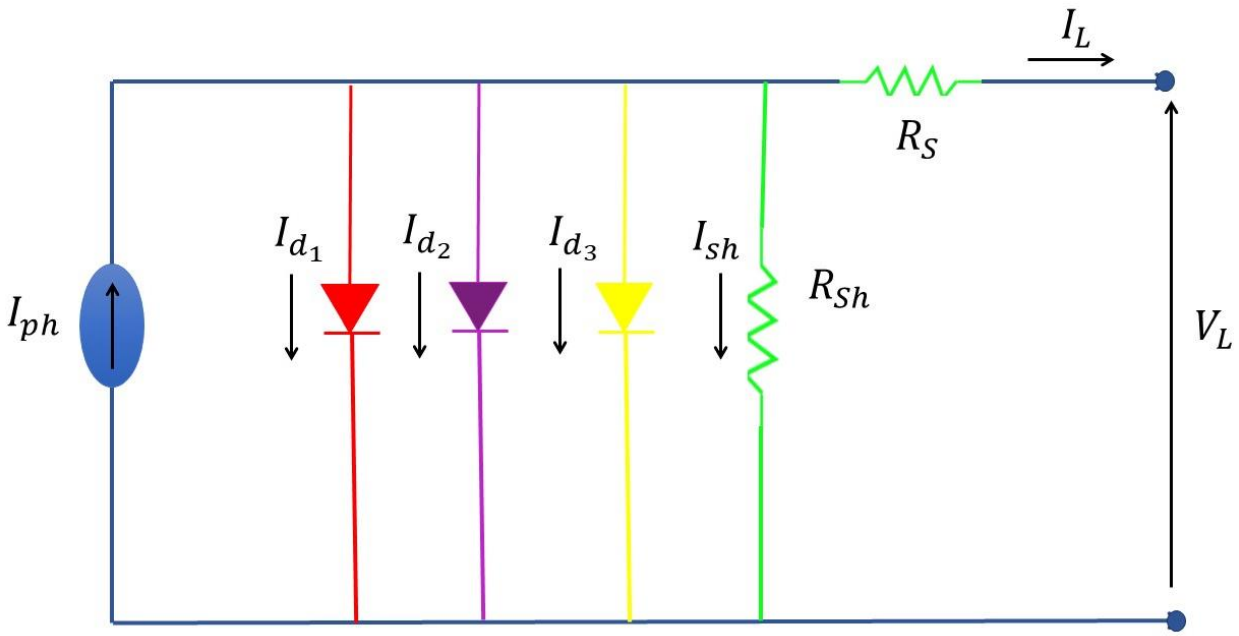


Figure II-8: Equivalent circuit of 3DM solar cell [34].

II.3.4 PV module Model

The equivalent circuit of the PV module is exposed in Figure II-9. Knowing that a PV module is made up of multiple solar cells connected together in series or in parallel and enclosed in a protective casing. The *I-V* characteristic of PV module can be expressed either in 1DM, 2DM or 3DM [34].

The output current of the PV module represented in 1DM is given as follows:

$$I_L = I_{ph}N_p - I_{s1}N_p\left[exp\left(\frac{\left((V_L/N_s)+N_sR_s(I_L/N_p)\right)}{n_1V_tN_s}\right) - 1\right] - \left(\frac{\left((V_L/N_s)+N_sR_s(I_L/N_p)\right)}{R_{sh}(N_s/N_p)}\right) \quad \text{(II-10)}$$

The output current of the PV module represented in 2DM is given as follows:

$$I_L = I_{ph}N_p - I_{s1}N_p\left[exp\left(\frac{\left((V_L/N_s)+N_sR_s(I_L/N_p)\right)}{n_1V_tN_s}\right)-1\right] - I_{s2}N_p\left[exp\left(\frac{\left((V_L/N_s)+N_sR_s(I_L/N_p)\right)}{n_2V_tN_s}\right)-1\right] - \left(\frac{\left((V_L/N_s)+N_sR_s(I_L/N_p)\right)}{R_{sh}(N_s/N_p)}\right) \quad \text{(II-11)}$$

The output current of the PV module represented in 3DM is given as follows:

$$I_L = I_{ph}N_p - I_{s1}N_p\left[exp\left(\frac{\left((V_L/N_s)+N_sR_s(I_L/N_p)\right)}{n_1V_tN_s}\right)-1\right] - I_{s2}N_p\left[exp\left(\frac{\left((V_L/N_s)+N_sR_s(I_L/N_p)\right)}{n_2V_tN_s}\right)-1\right] - I_{s3}N_p\left[exp\left(\frac{\left((V_L/N_s)+N_sR_s(I_L/N_p)\right)}{n_3V_tN_s}\right)-1\right] - \left(\frac{\left((V_L/N_s)+N_sR_s(I_L/N_p)\right)}{R_{sh}(N_s/N_p)}\right) \quad \text{(II-12)}$$

For each model of PV module there are the same unknown parameters of each model of PV cell.

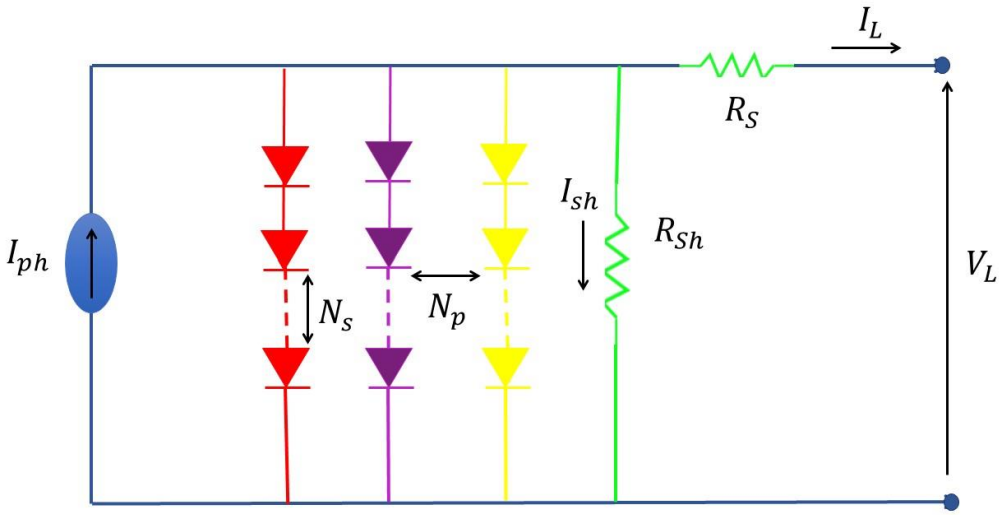


Figure II-9: Equivalent circuit of the PV module [34].

II.4 PV cell parameters identification methods

In the previous section, the discussion centred on the multitude of PV electrical models and their unknown parameters that must be identified. However, these models face a challenge in direct application due to the absence of suitable parameters specifically characterizing PV cells. This section elaborates on the process of determining their values. Despite efforts to ascertain them through datasheets or experimental measurements, this remains a challenge. The complexity arises from the response of their voltages and currents. Nevertheless, it is crucial to identify the most accurate parameters to enhance precision in simulations.

Addressing this complexity necessitates the utilization of methods capable of managing such challenges, garnering considerable interest from researchers. In recent years, several accurate parameter extraction methods have been proposed to address the nonlinear I-V characteristic of PV cells, modules, strings, or arrays. These methods can be classified into four primary approaches which are described below

II.4.1 Analytical approaches

Analytical approaches have historically been employed, relying on numerical techniques to estimate the unknown parameters of PV generator models. While these methods are relatively straightforward to implement, they often involve approximations and assumptions, leading to reduced accuracy. Parameters are typically determined at three reference points: open circuit (OC), short-circuit (SC), and the maximum power point (MPP), utilizing data sheets under standard test conditions (STC). However, the complexity escalates when dealing with more advanced PV generator models like 2DM and 3DM [37]–[39].

The analytical approach involves solving transcendental algebraic equations derived from expressions for significant points (such as open-circuit voltage (Voc), short-circuit current (Isc), and current (Imp) and voltage (Vmp) at the MPP) specified in manufacturer datasheets or from I-V curve measurements. Despite efforts to extract parameters using approximate equations and simplification methodologies, this method necessitates accurate values for Isc, Voc, Imp, and Vmp, resulting in a set of three equations for five unknown parameters. Furthermore, limitations arise from the number of samples required, especially when calculating the MPP and slopes from the entire I-V curve. While feasible with accurate key points, the presence of sampling noise, particularly at the MPP, compromises the accuracy of these analytical methods, making them unsuitable for high-precision calculations and time-consuming [40].

II.4.2 Numerical approaches

Numerical approaches have emerged as superior alternatives to analytical approaches, addressing the limitations of the latter in solving implicit nonlinear transcendental equations with higher accuracy. These methods, such as Gauss-Seidel, Newton-Raphson, and Levenberg-Marquardt (LM), rely on iterative techniques or optimization algorithms to minimize the discrepancy between measured and calculated I-V curves. By employing gradient-descent-based techniques, numerical methods facilitate the optimization of parameter values, ensuring rapid convergence, computational efficiency, and adequate precision. However, the effectiveness of numerical methods hinges on the initial values chosen, as inappropriate selections may lead to local search convergence rather than global optimization. Despite their advantages in handling nonlinearity and multivariable equations, numerical methods necessitate longer execution times compared to analytical methods and exhibit a high degree of dependency on initial solutions [41]–[44].

II.4.3 Metaheuristic approaches

Metaheuristic methods have emerged as a solution to address the limitations of numerical approaches, particularly in complex optimization tasks like extracting parameters from PV models.

Unlike numerical methods, metaheuristics offer the advantage of being easy to implement and are not constrained by specific problem characteristics. This has led to a surge in their popularity among researchers, with an increasing number of studies focusing on metaheuristic algorithms for solving PV cell parameter extraction problems [45]–[48].

Typically, the challenge of estimating parameters for PV cell models often translates into a numerical optimization problem, characterized by nonlinearity, constraints, and continuity. In response, metaheuristic optimization algorithms have emerged as common solutions for such optimization challenges. They do not require convexity, continuity or differentiability of objective functions. The objective of this optimization process is to minimize the disparity between measured and estimated I-V data by optimizing the objective function [49][33].

Metaheuristic approaches encompass both nature-inspired and non-nature-inspired approaches. Nature-inspired metaheuristic algorithms can be categorized into four main groups: evolutionary algorithms, physics-based algorithms, swarm-based algorithms, and human-based algorithms [49][33], as illustrated in Figure II-10.

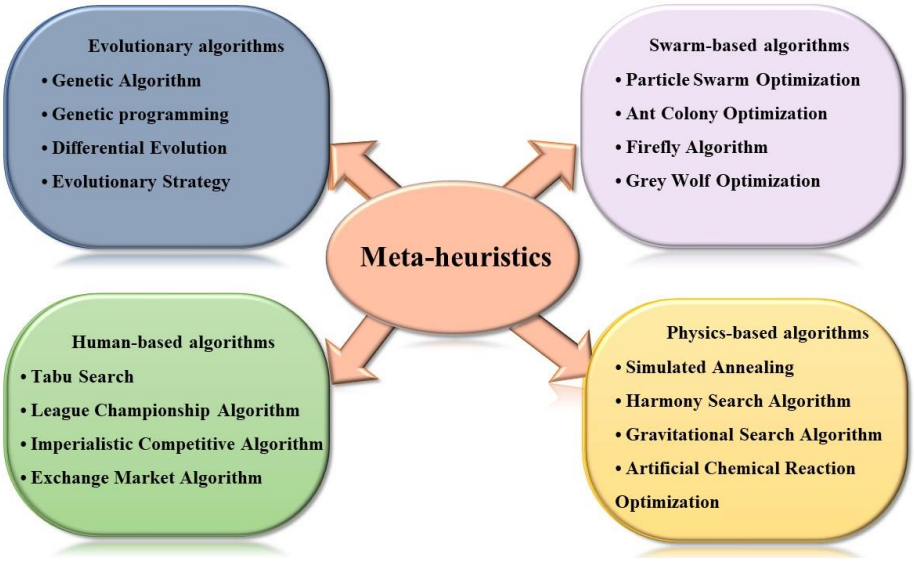


Figure II-10: Different categories of metaheuristic algorithms [49].

- Evolutionary algorithms: These methods replicate biological evolution, involving genetic recombination, mutation, and natural selection. Examples include Genetic Algorithm (GA), Evolutionary Strategy (ES), and Genetic Programming (GP) [49].
- Swarm-based algorithms: These approaches mimic the collective behaviour of social organisms. Individuals interact within a population, seeking improvement based on shared knowledge. Examples include Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Bat-inspired Algorithm (BA), and Firefly Algorithm (FA) [49].

- Physics-based algorithms: These techniques utilize principles of physics, such as gravity, inertia, and electromagnetic force. Examples include Simulated Annealing (SA), Gravitational Search Algorithm (GSA), Harmony Search (HS), Ray Optimization (RO), and Central Force Optimization (CFO) [49].
- Human-based algorithms: These methods draw inspiration from human behaviour. Examples include Imperialist Competitive Algorithm (ICA), League Championship Algorithm (LCA), and Exchange Market Algorithm (EMA) [49].

These methods exhibit unique strengths and weaknesses based on their sources of inspiration. In comparison to analytical and numerical iterative methods, metaheuristic techniques are distinguished by several key characteristics:

- They operate as population-based approaches, where each member represents a potential solution, unlike the singular solutions of other methods. This feature enhances the likelihood of discovering optimal solutions through the scale of the population.
- They employ stochastic processes, aligning well with real-world search strategies.
- They utilize randomized initialization, making them less reliant on specific initial conditions.
- They are characterized by exploration and exploitation capabilities. Exploration refers to their ability to cover a broad search space, while exploitation enables them to efficiently identify optimal solutions in neighbouring regions.

In the process of finding unknown PV parameters using metaheuristic algorithms, different objective functions are used such as: Root Mean Squared Error (RMSE), Mean Squared Error (MSE), Absolute Error (A.E), and Derivative at the maximum power point (MPP).

Metaheuristic methods demonstrate high accuracy and superior robustness compared to previous approaches, due to their stochastic nature and random initialization of their solutions, especially when tackling complex nonlinear problems like the PV 3DM model. However, they face a major drawback: high computational demand due to the complexity of large stochastic populations [13][50].

II.4.4 Hybrid approaches

Many metaheuristic methods suffer from various shortcomings, prompting researchers to continually develop new reliable methods to address these numerical issues and the high computational complexity of the problem. These methods are known as hybrid methods, which combine the strengths of two different approaches, where one method can effectively overcome the shortcomings of another, leading to improved results. The hybridization can take different forms, such as numerical-metaheuristic, and metaheuristic-metaheuristic, aiming to enhance the efficiency of parameter extraction [34].

For instance, researchers looked at combining two methods, one analytical and the other metaheuristic, to figure out the parameters for the single diode model of PV cells. They used the metaheuristic algorithms to optimize the values of series and shunt resistances (R_s, R_{sh}). Meanwhile, the analytical method helped estimate values for the ideality diode factor (n), diode saturation current (I_0), and photocurrent (I_{ph}).

Recently, there has been significant attention from researchers towards developing new hybrid methods, which offer increased accuracy, reliability, and faster convergence speeds. These methods are based on two or more metaheuristic algorithms to solve the parameter extraction problem of solar PV models. Among these hybrid algorithms: Performance-Guided JAYA (PGJAYA), Improved Teaching-Learning-Based Optimization (ITLBO), hybrid firefly and pattern search algorithm, hybrid PSO and GWO, Artificial Bee Swarm Optimization (ABSO) algorithm [51], enhanced form of chaotic grey wolf optimizer (ACGWO) [52], Enhanced Differential Evolutionary (IDEA) [53], adaptative differential evolution algorithm (DET) [54], modified Rao-1 optimization algorithm [55], modified JAYA algorithm (MJA) [56], Mutated Biogeography-based optimization algorithm (BBO-M) [57], optimization that is adaptive in both teaching and learning and differential evolution (ATLDE)[58], teaching learning hybrid with artificial bee colony (TLABC) [59], grey wolf optimization mix with cuckoo search (GWOCS) [60], Orthogonal learning-based Gray Wolf Optimizer [61],… etc.

In this thesis, particular attention will be paid to the aforementioned characteristics of metaheuristic approaches. Furthermore, this thesis makes a contribution to hybrid metaheuristic methods for determining the parameters of PV cells and modules that are unknown.

II.5 Conclusion

In conclusion, this chapter has provided a comprehensive overview of key aspects related to photovoltaic (PV) technology. Beginning with an exploration of the photovoltaic effect, we delved into the fundamental operation of PV cells, elucidating their structure, functioning, and characteristics. Furthermore, we examined various PV cell models, including the one diode, two diode, and three diode models, highlighting their significance in simulating PV cell behaviour under different operating conditions. Additionally, the chapter discussed PV parameter identification techniques, encompassing analytical, numerical, and metaheuristic methods, as well as hybrid approaches.

Through this thorough exploration, we have established a robust comprehension of the fundamental principles and vital tools within PV technology, providing a sturdy foundation for the next chapters. The following chapters will further explore PV parameter identification utilizing hybrid metaheuristic algorithms, as well as delve into the realm of fault detection and diagnosis methods within PV systems.

Chapter III : PV parameters identification using BB-TLBO algorithm

III.1 Introduction

Building on the discussion from the second chapter, where we explored different solar cell models such as one diode model, two-diode model and three-diode model, these models provide a framework to represent the electrical behaviour of a PV cell. However, these models rely on a set of unknown parameters that need to be identified for accurate performance prediction. This chapter focuses on a critical objective which is identifying the unknown parameters with high accuracy.

As highlighted previously, the hybrid method offers a significant advantage in this regard compared to traditional approaches. Therefore, this chapter delves into utilizing a hybrid method based on two metaheuristic algorithms to achieve this objective. This algorithm refers to as BB-TLBO and it is a hybridization between biogeography-based optimization (BBO) and teaching learning-based optimization (TLBO). BB-TLBO is applied on two cases; to identify the unknown parameters of RTC France silicon solar cell and to identify the unknown parameters of the monocrystalline STM6-40/36 module. This identification is based on the three models cited above.

This algorithm aims to combine the benefits of both TLBO and BBO to strike a balance between exploration and exploitation for a good search process. BB-TLBO not only overcomes respective limitations of BBO and TLBO but also demonstrates improved performance in determining the parameters of PV models, setting a new benchmark for optimization methods in the field.

III.2 Related work

Many research papers that use metaheuristic algorithms to identify PV parameters based on different models of solar cell and PV module such as: Genetic Algorithm (GA) is a population-based algorithm inspired by biology that mimics the "survival of the fittest" phenomenon. It has three main steps: selection, crossover and mutation. In the problem of PV identification parameters, GA is better utilization of search space due to mutation and crossover operations and it is good in exploration but it suffers of high computational burden and poor in exploitation [12]. To reduce the computational burden of GA, a hybrid strategy (GA+NR) has been proposed in [62]. In this strategy, GA was used to extract three parameters, R_s, R_{sh} and n whereas I_{ph} and I_s were extracted analytically using Newton Raphson (NR) method. According to the findings, the NR technique converges quickly when the number of unknowns is small, Low computational burden and accuracy depends on the parameters that GA has optimized. Another version of GA which is hybridization between GA and

interior point method is proposed in [63]. This algorithm GA +IP was used to identify the parameters of 1DM taking into account the standard and nominal operating cell temperature (NOCT) conditions of a PV. The results of GA+IP indicates that this algorithm is Accurate in real-time operating conditions, multi objective based optimization and slow convergence. In addition to the two previous version of GA a new variant is developed and presented in [64]. This algorithm is Genetic Algorithm with Convex Combination Crossover (GACCC), it is proposed to identify one diode model, two diode model and PV module. GACCC is accomplished through the inclusion of a new crossover operation to maintain a good balance between intensifying the best solutions and diversifying the search space. The results show a high accuracy and efficiency of GACCC in parameters identifying. Moreover, there are other variants and modifications of GA that have not been implemented, such as the Stud-Genetic Algorithm (SGA) [27], which might have a high impact on PV identification parameters.

Differential evolution (DE) is a heuristic, population-based algorithm originally proposed by Storn and Price in 1997. DE has mainly four stages: initialization, mutation, crossover, and selection. In the problem of identifying parameters, DE lacks the ability to attain a good optimization effect and does not accurately identify the model parameters, resulting in a bias in describing the internal behaviour of photovoltaic systems due to a weak global optimization capability, which will easily converge to the local optimum. Furthermore, it is overly dependent on the initial value of the mutation and crossover factor [12]. To overcome the above shortcoming several variants of DE have introduced. In [65] authors have introduced an improved DE (DVADE) to identify the parameters of 1DM and 2DM of solar cell and PV module. This algorithm is based on reusing previous individual vectors and an adaptive mutation strategy. In this algorithm DVADE, to enhance the effectiveness of differential evolution, the successful difference vectors from earlier generations are introduced to create the offspring in the following generations. The results demonstrate the accuracy, reliability and convergence speed of DVADE. Another variant of DE used to identify PV parameters has been proposed in [54], this proposed algorithm based on adaptative of DE technique (DET). The adaption is achieved through crossover and mutation factor. DVADE is compared to GA, chaos particle swarm optimization (CPSO), harmony search algorithm (HSA), and artificial bee swarm optimization (ABSO). The outcome shows that it is an optimal method which suits the parameter extraction of solar cells and modules. Recently authors propose an adaptative differential evolution (ADE) with the dynamic opposite learning strategy (DOL), called DOLADE. The opposite learning approach in DOLADE increases both the elite population and the population of underperformers, enhancing the particles' capacity for exploration. The outcomes show that DOLADE brings superior competition in terms of accuracy, dependability, and computational efficiency when it comes to extracting the best parameters for each PV cell model [66].

Biogeography-based optimization (BBO) introduced by Dan Simon in 2008 [67]. The BBO was applied to solve many optimization problems, and it proves its efficiency in finding optimal solutions. This is due to its good exploration feature for the current population. Nevertheless, BBO suffers from some drawbacks such as poorness in exploitation features, negligence of the best individuals over generations, and generation of infeasible solutions [57] [67]. To overcome the over mentioned disadvantages, various hybrid and variant of BBO have been proposed. In [68] biogeography-based heterogeneous cuckoo search (BHCS) algorithm has been proposed. This algorithm is a hybrid of two metaheuristic algorithms namely: BBO and Cuckoo Search (CS). BHCS is used to identify 1DM and 2DM of solar cell and two PV modules. The results show that BHCS is accurate and reliable compared to BBO and CS and other metaheuristic algorithm. In addition to BHCS, another version of BBO, BBO-M algorithm has been proposed in [57]. BBO-M is used to identify two types of cells: solar cell and fuel cell. In this algorithm, mutation step of DE and the chaos theory are integrated into the BBO structure for enhancing the global searching capability of the algorithm. The results show that BBO-M has fast convergence speed and it can produce solution of high quality in both types.

Teaching–learning-based optimization (TLBO) technique that was introduced by RV Rao et Al in 2012 [69]. TLBO algorithm mimics the school strategy in learning. TLBO comprises two fundamental stages, the teacher phase, and the learner phase. In the first phase, the best individual is selected to be a teacher, and in the second phase, learners receive knowledge from each other. In PV identification problem, TLBO is simple to understand, easy to implement and good at exploitation however TLBO has some of shortcomings like: convergence speed, it may require a large number of iterations to converge to the best solution, lack of diversity, it may suffer from a lack diversity in the population, which can limit its exploration capabilities and prevent it from escaping from local optima [12]. To overcome these limitations, various modification of TLBO such as improving and hybridization with other algorithms have been proposed. In [70] the authors develop a new optimization method called oppositional teaching learning based optimization (GOTLBO) to identify the parameters of both 1DM and 2DM. The initialization step and generation jumping of this algorithm are taken from generalized opposition-based learning (GOBL) and integrated in the original TLBO. This algorithm was compared with four evolutional algorithms: jDE, CLPSO, TLBO, and OTLBO. The results show that GOTLBO is verry competitive compared to other algorithms. Another variant of TLBO has been proposed in [58], the authors propose a hybrid adaptive teaching–learning-based optimization with differential evolution (ATLDE) reliably identify the unknown parameters of both 1DM and 2DM of solar cell and 1DM of two PV module. In order to verify the performance of ATLDE, it is compared with TLBO, DE and other state of the art metaheuristic algorithms. The results demonstrate that ATLDE can obtain more accurate or similar values, and consume fewer computing resources.

III.3 Problem formulation

The problem of identifying unknown parameters for all models of PV cells and modules can be solved by converting it into an optimization problem.

The discrepancy between calculated and measured data is minimized during the optimization phase. The objective function is designed to find the optimal values of the unknown parameters of different models in terms of root mean square error (RMSE) [36].

In this chapter, we have considered two different objective functions for optimal PV model parameter extraction.

The basic expression for the RMSE between the measured and calculated current values is:

$$RMSE(x) = \sqrt{\frac{1}{N}\sum_{1}^{N} f_M(V_L, I_L, x)} \quad \text{(III-1)}$$

Where:

N denotes collected experimental data;

V_L is measured voltage;

I_L is measured current;

x is the solution vector;

f_M is the error function.

III.3.1 Method 01: Conventional RMSE ($RMSE_C$)

The error function f_M based on conventional RMSE is defined for different PV models as follows:

III.3.1.1 1DM of PV cell

$$\begin{cases} f(V_L, I_L, x) = I_{ph} - I_{s1}\left[exp\left(\frac{q(V_L+R_S I_L)}{n_1 KT}\right) - 1\right] - \frac{(V_L+R_S I_L)}{R_{sh}} - I_L \\ x = \{I_{ph}, I_{s1}, R_s, R_{sh}, n_1\} \end{cases} \quad \textit{(III-2)}$$

There are five parameters in the equation expressed above that need to be identified: I_{ph}, I_{s1}, R_s, $R_{sh}$and n_1.

III.3.1.2 2DM of PV cell

$$\begin{cases} f(V_L, I_L, x) = I_{ph} - I_{s1}\left[exp\left(\frac{q(V_L+R_S I_L)}{n_1 KT}\right) - 1\right] - I_{s2}\left[exp\left(\frac{q(V_L+R_S I_L)}{n_2 KT}\right) - 1\right] - \frac{(V_L+R_S I_L)}{R_{sh}} - I_L \\ x = \{I_{ph}, I_{s1}, I_{s2}, R_s, R_{sh}, n_1, n_2\} \end{cases} \quad \textit{(III-3)}$$

There are seven parameters in the above equation which need to be identified: I_{ph}, I_{s1}, I_{s2}, R_s, R_{sh}, $n_1 and\ n_2$.

III.3.1.3 3DM of PV cell

$$\begin{cases} f(V_L, I_L, x) = I_{ph} - I_{s1}\left[exp\left(\frac{q(V_L+R_SI_L)}{n_1KT}\right)-1\right] - I_{s2}\left[exp\left(\frac{q(V_L+R_SI_L)}{n_2KT}\right)-1\right] - \\ I_{s3}\left[exp\left(\frac{q(V_L+R_SI_L)}{n_3KT}\right)-1\right] - \frac{(V_L+R_SI_L)}{R_{sh}} - I_L \\ x = \{I_{ph}, I_{s1}, I_{s2}, I_{s3}, R_s, R_{sh}, n_1, n_2, n_3\} \end{cases} \quad (III\text{-}4)$$

There are nine parameters in the last expression which need to be identified: I_{ph}, I_{s1}, I_{s2}, I_{s3}, R_s, R_{sh}, n_1, n_2 *and* n_3 .

III.3.1.4 1DM of PV module

$$\begin{cases} f(V_L, I_L, x) = N_pI_{ph} - N_pI_{s1}\left[exp\left(\frac{((V_L/N_s)+N_sR_S(I_L/N_p))}{n_1V_tN_s}\right)-1\right] - \left[\frac{((V_L/N_s)+N_sR_S(I_L/N_p))}{R_{sh}}\right] - I_L \\ x = \{I_{ph}, I_{s1}, R_s, R_{sh}, n_1\} \end{cases} \quad (III\text{-}5)$$

III.3.1.5 2DM of PV module

$$\begin{cases} f(V_L, I_L, x) = N_pI_{ph} - N_pI_{s1}\left[exp\left(\frac{((V_L/N_s)+N_sR_S(I_L/N_p))}{n_1V_tN_s}\right)-1\right] - N_pI_{s2}\left[exp\left(\frac{((V_L/N_s)+N_sR_S(I_L/N_p))}{n_2V_tN_s}\right)-1\right] - \left[\frac{((V_L/N_s)+N_sR_S(I_L/N_p))}{R_{sh}}\right] - I_L \\ x = \{I_{ph}, I_{s1}, I_{s2}, R_s, R_{sh}, n_1, n_2\} \end{cases} \quad (III\text{-}6)$$

III.3.1.6 3DM of PV module

$$\begin{cases} f(V_L, I_L, x) = N_pI_{ph} - N_pI_{s1}\left[exp\left(\frac{((V_L/N_s)+N_sR_S(I_L/N_p))}{n_1V_tN_s}\right)-1\right] - N_pI_{s2}\left[exp\left(\frac{((V_L/N_s)+N_sR_S(I_L/N_p))}{n_2V_tN_s}\right)-1\right] - \\ N_pI_{s3}\left[exp\left(\frac{((V_L/N_s)+N_sR_S(I_L/N_p))}{n_3V_tN_s}\right)-1\right] - \left[\frac{((V_L/N_s)+N_sR_S(I_L/N_p))}{R_{sh}}\right] - I_L \\ x = \{I_{ph}, I_{s1}, I_{s2}, I_{s3}, R_s, R_{sh}, n_1, n_2, n_3\} \end{cases} \quad (III\text{-}7)$$

There are five, seven and nine parameters for the three models of the PV module which must be identified, they are presented in vector x.

III.3.2 Method 02: RMSE based on Lambert W function ($RMSE_L$)

The complexity and nonlinearity of Equation III-1 become more significant when using the benchmarked equivalent circuit models 2DM and 3DM. Thus, to address this complexity when formulating the RMSE, a numerical method can be associated with this formulation. In this chapter, the Lambert W function is employed to precisely formulate and calculate the RMSE value [71]. So, if the output current I_L is represented as x, then the equation can be seen as follows:

$$f(x) = xe^x = \partial \quad (III\text{-}8)$$

The following inverse function is applied to find the value of *x*, as follows:

$$x = f^{-1}(xe^x) = W(\partial) \quad (III\text{-}9)$$

This equation (III-9), known as Lambert *W* function, is significant mathematical tool used to solve several nonlinear problems. Its graphical depiction in Figure III-1 consists of two segments: one for x ≥ -1 and the other for x < -1. Various methods like software solvers (e.g., MATLAB, Mathematica,

Maple), Taylor series, and Special Trans Function Theory are employed to solve the Lambert W function, with the latter known for its high precision based on several research papers [71].

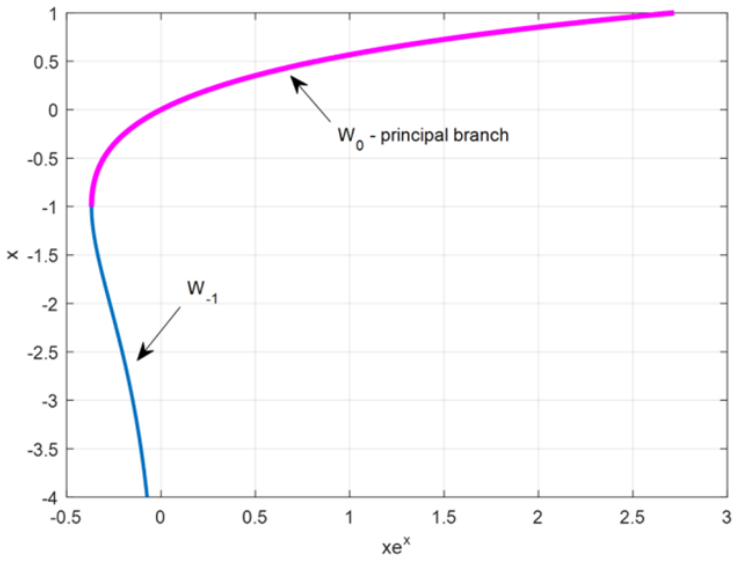


Figure III-1: Graphical representation of the Lambert W function, $x = W(xe^x)$.

The utilization of the Lambert W function in PV modelling within the relevant equivalent electrical models is carried out by organizing the corresponding equation in a specific manner. Starting from the one diode model (1DM), described as follows:

$$I_L = I_{ph} - I_{sd}.\left[exp\left(\frac{q.(V_L+R_s.I_L)}{n.k.T}\right) - 1\right] - \frac{V_L+R_s.I_L}{R_{sh}} \quad \text{(III-10)}$$

This equation can be arranged as follows [71].:

$$I_L = \frac{R_{sh}(I_{ph}-I_{sd})-V_L}{R_s+R_{sh}} - \frac{k.T}{q.R_s}.n.W(\partial) \quad \text{(III-11)}$$

$$\partial = \left[\frac{qR_{sh}R_sI_{sd}}{nkT(R_s+R_{sh})}\right].exp\left[\frac{R_{sh}q(R_sI_{ph}+R_sI_{sd})+V_L}{nkT(R_s+R_{sh})}\right] \quad \text{(III-12)}$$

As stated previously $W(\partial)$ is the solution of the Lambert equation.

From Equation (III-1), f_M which is the error function, is defined using the Lambert W function for different PV models as follows:

III.3.2.1 1DM of PV cell

$$f(V_L, I_L, x) = I_L - \left[\frac{R_{sh}(I_{ph}+I_{sd})-V_L}{R_s+R_{sh}} - \frac{k.T}{q.R_s}\right].[n.W(\partial)] \quad \text{(III-13)}$$

where:

$$\partial = \frac{qI_{sd}R_sR_{sh}}{n.k.T(R_{sh}+R_s)}exp\left(\frac{qR_{sh}(I_{ph}R_s+I_{sd}R_s+V_L)}{n.k.T(R_{sh}+R_s)}\right) \quad \text{(III-14)}$$

and the five unknown parameters are the component of the vector $x = (I_{ph}, I_s, R_s, R_{sh}, n)$

III.3.2.2 2DM of PV cell

$$f(V_L, I_L, x) = I_L - \left[\frac{R_{sh}(I_{ph}+I_{sd1}+I_{sd2})-V_L}{R_s+R_{sh}} - \frac{k.T}{q.R_s}\right].[n_1 W(\partial_1) + n_2 W(\partial_2)] \quad \text{(III-15)}$$

where:

$$\partial_1 = \frac{qI_{sd1}R_sR_{sh}}{n_1.k.T(R_{sh}+R_s)} exp\left(\frac{qR_{sh}(I_{ph}R_s+I_{sd1}R_s+V_L)}{n_1.k.T(R_{sh}+R_s)}\right) \quad \text{(III-16)}$$

$$\partial_2 = \frac{qI_{sd2}R_sR_{sh}}{n_2kT(R_{sh}+R_s)} exp\left(\frac{qR_{sh}(I_{ph}R_s+I_{sd2}R_s+V_L)}{n_2kT(R_{sh}+R_s)}\right) \quad \text{(III-17)}$$

and the seven unknown parameters are the component of the vector $x = (I_{ph}, I_{s1}, I_{s2}, R_s, R_{sh}, n_1, n_2)$

III.3.2.3 3DM of PV cell

$$f(V_L, I_L, x) = I_L - \left[\frac{R_{sh}(I_{ph} + I_{sd1} + I_{sd2} + I_{sd3}) - V_L}{R_s + R_{sh}} - \frac{k.T}{q.R_s}\right].$$
$$[n_1 W(\partial_1) + n_2 W(\partial_2) + n_3 W(\partial_3)] \quad \text{(III-18)}$$

where:

$$\partial_1 = \frac{qI_{sd1}R_sR_{sh}}{n_1.k.T(R_{sh}+R_s)} exp\left(\frac{qR_{sh}(I_{ph}R_s+I_{sd1}R_s+V_L)}{n_1.k.T(R_{sh}+R_s)}\right) \quad \text{(III-19)}$$

$$\partial_2 = \frac{qI_{sd2}R_sR_{sh}}{n_2kT(R_{sh}+R_s)} exp\left(\frac{qR_{sh}(I_{ph}R_s+I_{sd2}R_s+V_L)}{n_2kT(R_{sh}+R_s)}\right) \quad \text{(III-20)}$$

$$\partial_3 = \frac{qI_{sd3}R_sR_{sh}}{n_3kT(R_{sh}+R_s)} exp\left(\frac{qR_{sh}(I_{ph}R_s+I_{sdS}R_s+V_L)}{n_3kT(R_{sh}+R_s)}\right) \quad \text{(III-21)}$$

and the nine unknown parameters are the component of the vector $x = (I_{ph}, I_{s1}, I_{s2}, I_{s3}, R_s, R_{sh}, n_1, n_2, n_3)$.

III.3.2.4 1DM of PV module

$$f(V_L, I_L, x) = I_L - \left[\frac{R_{shm}(I_{phm}+I_{sdm})-V_L}{R_{sm}+R_{shm}} - \frac{k.T}{q.R_{sm}}\right].[n.W(\partial)] \quad \text{(III-22)}$$

where:

$$\partial = \frac{qI_{sdm}R_{sm}R_{shm}}{n_m.k.T(R_{shm}+R_{sm})} exp\left(\frac{qR_{shm}(I_{phm}R_{sm}+I_{sdm}R_{sm}+V_L)}{n_m.k.T(R_{shm}+R_{sm})}\right) \quad \text{(III-23)}$$

and the five unknown parameters are the component of the vector $x = (I_{phm}, I_{sdm}, R_{sm}, R_{shm}, n_s)$;

with: $(I_{phm} = N_p \cdot I_{ph})$; $(I_{sdm} = N_p \cdot I_{sd})$; $(R_{sm} = \frac{N_s}{N_p} \cdot R_s)$; $(R_{shm} = \frac{N_s}{N_p} \cdot R_{sh})$; $(n_m = N_s \cdot n)$

III.3.2.5 2DM of PV module

$$f(V_L, I_L, x) = I_L - \left[\frac{R_{shm}(I_{phm}+I_{sdm})-V_L}{R_{sm}+R_{shm}} - \frac{k.T}{q.R_{sm}}\right].[n_{s1} W(\partial_1) + n_{s2} W(\partial_2)] \quad \text{(III-24)}$$

where:

$$\partial_1 = \frac{qI_{sdm1}R_{sm}R_{shm}}{n_{m1}.k.T(R_{shm}+R_{sm})} exp\left(\frac{qR_{shm}(I_{phm}R_{sm}+I_{sdm1}R_{sm}+V_L)}{n_{m1}.k.T(R_{shm}+R_{sm})}\right) \quad \text{(III-25)}$$

$$\partial_2 = \frac{qI_{sdm2}R_{sm}R_{shm}}{n_{m2}kT(R_{shm}+R_{sm})} exp\left(\frac{qR_{shm}(I_{phm}R_{sm}+I_{sdm2}R_{sm}+V_L)}{n_{m2}kT(R_{shm}+R_{sm})}\right) \quad \text{(III-26)}$$

and the seven unknown parameters are the component of the vector $x = (I_{phm}, I_{sdm1}, I_{sdm2}, R_{sm}, R_{shm}, n_{m1}, n_{m2})$

with: $(I_{phm} = N_p \cdot I_{ph})$; $(I_{sdm1} = N_p \cdot I_{sd1})$; $(I_{sdm2} = N_p \cdot I_{sd2})$; $(R_{sm} = \frac{N_s}{N_p} \cdot R_s)$;

$(R_{shm} = \frac{N_s}{N_p} \cdot R_{sh})$; $(n_{m1} = N_s \cdot n_1)$; $(n_{m2} = N_s \cdot n_2)$;

III.3.2.6 3DM of PV module

$$f(V_L, I_L, x) = I_L - \left[\frac{R_{shm}(I_{phm} + I_{sdm}) - V_L}{R_{sm} + R_{shm}} - \frac{k.T}{q.R_{sm}}\right].$$
$$[n_{m1}W(\partial_1) + n_{m2}W(\partial_2) + n_{m3}W(\partial_3)] \quad \text{(III-27)}$$

where:

$$\partial_1 = \frac{qI_{sdm1}R_{sm}R_{shm}}{n_{m1}.k.T(R_{shm}+R_{sm})} exp\left(\frac{qR_{shm}(I_{phm}R_{sm}+I_{sdm1}R_{sm}+V_L)}{n_{m1}.k.T(R_{shm}+R_{sm})}\right) \quad \text{(III-28)}$$

$$\partial_2 = \frac{qI_{sdm2}R_{sm}R_{shm}}{n_{m2}kT(R_{shm}+R_{sm})} exp\left(\frac{qR_{shm}(I_{phm}R_{sm}+I_{sdm2}R_{sm}+V_L)}{n_{m2}kT(R_{shm}+R_{sm})}\right) \quad \text{(III-29)}$$

$$\partial_3 = \frac{qI_{sdm3}R_{sm}R_{shm}}{n_{m3}kT(R_{shm}+R_{sm})} exp\left(\frac{qR_{shm}(I_{phm}R_{sm}+I_{sdm2}R_{sm}+V_L)}{n_{m3}kT(R_{shm}+R_{sm})}\right) \quad \text{(III-30)}$$

and the nine unknown parameters are the component of the vector $x = (I_{phm}, I_{sdm1}, I_{sdm2}, I_{sdm3}, R_{sm}, R_{shm}, n_{m1}, n_{m2}, n_{m3})$

with: $(I_{phm} = N_p \cdot I_{ph})$; $(I_{sdm1} = N_p \cdot I_{sd1})$; $(I_{sdm2} = N_p \cdot I_{sd2})$; $(I_{sdm3} = N_p \cdot I_{sd3})$; $(R_{sm} = \frac{N_s}{N_p} \cdot R_s)$; $(R_{shm} = \frac{N_s}{N_p} \cdot R_{sh})$; $(n_{m1} = N_s \cdot n_1)$; $(n_{m2} = N_s \cdot n_2)$; $(n_{m3} = N_s \cdot n_3)$

III.4 The proposed hybrid BB-TLBO algorithm

In this section, the hybrid proposed algorithm is introduced, which combines Biogeography-Based Optimization (BBO) and Teaching-Learning Based Optimization (TLBO). A brief discussion of the BBO and TLBO algorithms. Following that, the innovative hybrid algorithm, BB-TLBO, is discussed in details, which leverages the strengths of both BBO and TLBO to enhance optimization performance.

III.4.1 Biogeography-based optimization

Biogeography-based optimization (BBO) is a metaheuristic algorithm announced by Dan Simon in year 2008 that is inspired by the theory of island biogeography. This theory employs mathematical models to describe the patterns of migration, speciation, and extinction of species across different islands. Biogeography-based optimization (BBO) considers each individual in the population as an island with a habitat suitability index (HSI), where each variable of the island is referred to as a suitability index variable (SIV). A desirable solution is indicated by an island with a high HSI that hosts a diverse range of species, while an undesirable solution is represented by an island with a low HSI, which has a small number of species. Consequently, individuals located on high HSI islands are more likely to migrate to islands with low HSI, and the low HSI islands receive immigrants from high HSI islands to promote genetic diversity and improve the overall quality of the population. BBO employs two operators namely migration and mutation to evolve the population [67][57][68].

III.4.1.1 Migration

Biogeography-based optimization involves assigning unique immigration λ and emigration rates μ to each island in the population, which are the functions of species in the habitat. The functions of immigration rate and emigration rate can be defined as:

$$\lambda_s = I\left(1 - \frac{s}{N_p}\right) \tag{III-31}$$

$$\mu_s = E\left(\frac{s}{N_p}\right) \tag{III-32}$$

The maximum possible immigration rate (denoted by I) is achieved when an island has no species present, whereas the maximum emigration rate (denoted by E) occurs when the island has the maximum number of species. s is the number of species of the s^{th} individual in the ordered population according to the fitness, and N_p is the number of candidate solutions in the population. With I and E typically being set to 1. Where the immigration rate (λ) determines whether a selected solution should modify its suitability index variable (SIV), while the emigration rate (μ) determines which solutions should migrate a random SIV to the selected solution.

III.4.1.2 Mutation

Nature is subject to cataclysmic events that can significantly alter the solution of an island, resulting in sudden changes to its habitat suitability index (HSI). In order to incorporate this type of random variation into the BBO algorithm, a mutation process is introduced, with mutation rates determined based on the species count probabilities. Specifically, the species count probability $P(s)$ represents the probability that an island contains s species, and its change rate can be calculated as:

$$\dot{P} = \begin{cases} -(\lambda_s + \mu_s)P_s + \mu_{s+1}P_{s+1}, & s = 0 \\ -(\lambda_s + \mu_s)P_s + \lambda_{s+1}P_{s-1} + \mu_{s+1}P_{s+1}, & 1 \le s \le s_{max} - 1 \\ -(\lambda_s + \mu_s)P_s + \lambda_{s+1}P_{s-1}, & s = s_{max} \end{cases} \tag{III-33}$$

where λ_s denotes the immigration rate and μ_s denotes emigration rate when there are S species in the island, and S_{max} denotes the maximal species in the island.

The mutation rate for each solution is determined based on its respective species count probability. Solutions with a lower probability (P_s) are more likely to undergo mutations to different solutions, while those with a higher probability are less likely to experience mutations that lead to different solutions. The mutation rates can be proportional to the species count probabilities as:

$$m(S) = m_{max}\left(\frac{1-P_s}{P_{max}}\right) \quad \text{(III-34)}$$

where m_{max} is the maximal mutation rate defined by user, and P_{max} is the maximal species count probability. The mutation operator is implemented to enhance the genetic diversity of the population, providing a means to improve solutions for islands with a low habitat suitability index (HSI), while allowing high HSI islands to potentially achieve even better solutions. In essence, the mutation operator helps to balance out the population and promote progress towards an optimal solution for the given problem.

III.4.2 Teaching–learning-based optimization

TLBO is population-based evolutionary algorithm that was proposed recently by Rao et al., and its main idea is to simulate a classical learning process that consists of a teacher phase and a learner phase. In the teacher phase, the best solution found in the entire population is considered as the teacher, and it shares its knowledge with the students to improve their outputs (i.e., grades or marks). In the learner phase, the students also learn knowledge from each other to improve their outputs. TLBO conducts this process through two basic operations, the teacher phase and the learner phase [58], [59], [69], [70], [72]–[74].

III.4.2.1 Teacher phase

During the teacher phase, the best learner discovered so far is selected as the teacher based on their fitness value. Additionally, the average position of all the learners must be computed to update the positions of every learner. Assuming that $X_i$represents the position of the i^{th} learner in an n-dimensional optimization problem, then X_i can be expressed as:

$$X_i = \{x_{i1}, x_{i2}, \dots, x_{in}\} \quad \text{(III-35)}$$

The teaching process is formulated as follows:

$$X_{new,i} = X_{old,i} + rand() * (X_{Teacher} - TF * X_{mean}) \quad \text{(III-36)}$$

The variables and computations utilized in the teacher phase include $X_{new,i}$ and $X_{old,i}$, which represent the new and old positions, respectively, of the i^{th} learner. $X_{Teacher}$ denotes the position of the current

teacher, while X_{mean} is the mean position of the current class, calculated as the sum of all learners' positions divided by the total number of students (N_p) (As it is shown in Equation III-37). The function $rand()$ generates a random number within the range $[0, 1]$, and TF denotes the teaching factor, which is typically set to either 1 or 2 using a heuristic approach.

It is necessary to re-evaluate all learners after each iteration of the teacher phase. If the new position $X_{new,i}$ is better than the old position $X_{old,i}$, it will be accepted and will flow into the learner phase. Otherwise, $X_{old,i}$ will not be updated.

$$X_{mean} = \frac{1}{N_p}\sum_{i=1}^{N_p} X_i \tag{III-37}$$

III.4.2.2 Learner phase

During the learner phase, the i^{th} learner X_i is randomly paired with another learner X_j from the class who is distinct from X_i. The learning process is described as follows:

$$X_{new,i} = \begin{cases} X_{old,i} + rand() * (X_{old,i} - X_{old,j}), \ if\ f(X_{old,i}) < f(X_{old,j}) \\ X_{old,i} + rand() * (X_{old,j} - X_{old,i}), \ otherwise \end{cases} \tag{III-38}$$

In this context, $X_{new,i}$ represents the new position of the i^{th} learner, while $X_{old,i}$ and $X_{old,j}$ correspond to the old positions of the i^{th} and j^{th} learners, respectively. Additionally, $rand()$ generates a random number between 0 and 1.

To update the learner positions, $X_{new,i}$ is accepted if it results in a better function value compared to the previous positions, $X_{old,i}$ and $X_{old,j}$.

III.4.3 Proposed BB-TLBO algorithm

The BBO was applied to solve many optimization problems, and it proves its efficiency in finding optimal solutions. This is due to its good exploration feature for the current population. Nevertheless, BBO suffers from some drawbacks such as poorness in exploitation features, negligence of the best individuals over generations, and generation of infeasible solutions.

On the other hand, the TLBO algorithm is known as a new and simple-to-understand method. Because of its design which mimics the school strategy in learning. However, TLBO suffers from deficiencies such as trapping in local optima and poor population diversity.

To overcome the over mentioned disadvantages, a new hybrid BB-TLBO algorithm is prosed in this paper. The BB-TLBO is described through the following phase:

III.4.3.1 Initialization

The population in the BB-TLBO are considered as learners in class $\{X_1, X_2, \dots, X_{np}\}$. The initialization of the BB-TLBO is inspired from the BBO technique where an assignment of positions is performed randomly to each individual along with the search space. Each individual position is represented by a vector X_i.

III.4.3.2 Migration process

In this phase, each individual (learners) changes his position inside the class to enhance his level of knowledge. This movement is performed using the immigration and an emigration rate represented by λ and μ respectively, using Equations III-31 and III-32.

III.4.3.3 Teaching phase

After performing the migration process, the next step (inspired from the TLBO) is to calculate the mean level of knowledge in the class using Equation III-37.

The population is then sorted in descendant order from the best to the worst according to their fitness's. $X_1$is selected as teacher, and the rest of the population is considered as learners. The teaching phase aims to enhance the mean level of knowledge of learners. Each learner X_i will update his knowledge level using Equation III-36.

After evaluation, if the fitness of $X_{new,i}$ is better than the fitness of $X_{old,i}$ then $X_{new,i}$ replaces $X_{old,i}$ and is accepted for pursuing the learning phase, otherwise $X_{old,i}$ stays in its current position.

III.4.3.4 Learning phase

The learning process is characterized by sharing knowledges between each learner X_i and other X_j that is randomly selected among the population. This phase is mathematically expressed as Equation III-38.

After evaluating the fitness of $X_{new,i}$, $X_{new,i}$ replaces $X_{old,i}$ if it has a better fitness, otherwise $X_{old,i}$ stays in its current position. At the end of this phase a sorting of the current population is performed.

III.4.3.5 Elitism phase

The final step in BB-TLBO is the elitism process, which is taken from the BBO. Here, a predefined keeping rate K_r is used to select the best individuals among the current population to be considered for the next iteration. The elitism process is expressed through the following equation:

$$X_s = \begin{cases} X_{N_p - s + i} & if \;\; s < round(K_r * N_p) \\ X_s & if \;\; s > round(K_r * N_p) \end{cases} \quad \text{(III-39)}$$

where X_s is the individual being selected to take part in the next generation. i denotes the current iteration.

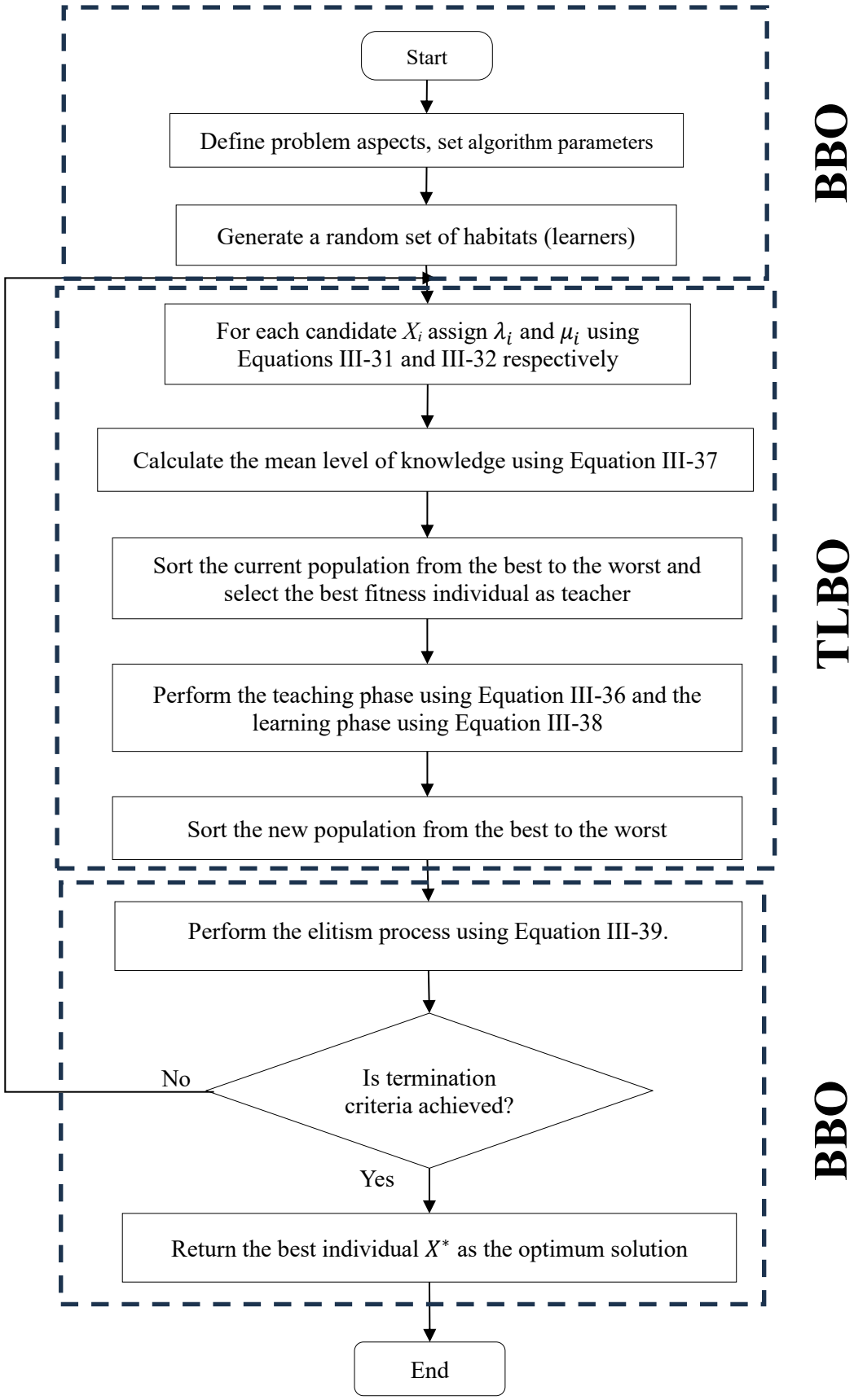


Figure III-2: Flowchart of BB-TLBO algorithm.

III.5 Results and discussion

This section is dedicated to the discussion of the experimental results obtained and the verification of the performance of the proposed hybrid algorithm.

In order to solve the parameter extraction problem, this algorithm is tested on two different cases, namely: the RTC FRANCE silicon solar cell and the commercial PV solar modules: the monocrystalline STM6-40/36 with their three different models (1DM, 2DM and 3DM).

To collect the experimental data I-V, 26 pairs of voltage and current samples were taken from a 57 mm diameter RTC France silicon solar cell at 33°C and 1000W/m^2 solar irradiance in order to extract the unknown parameters of three models (1DM, 2DM and 3DM) [60] as shown in annex. And likewise, 20 pairs voltage and current data set are collected from the STM6- 40/36 module under the cell temperature of 51°C and solar irradiation of 1000 W/m^2 [60] as shown in annex.

The lower and the upper limits (LL and UL) of the parameters of the three different models (1DM, 2DM and 3DM) are shown in Table III-1 [75].

To verify the efficiency of the proposed hybrid algorithm, it is compared to other benchmarked methods such as: DOLADE, LAPSO, ABC, BBO, TLBO, IQSODE, etc.

Each algorithm had been executed several times (in this example 30 executions) so that they can start their exploration from the same random point in the search space.

The simulation was performed in MATLAB 2019b, Intel(R) Core (TM) i3-6600U CPU @ 2.60GHz 2.81 GHz.

Table III-1: Upper and lower limits of the 1DM, 2DM and 3DM parameters

		RTC France solar cell		**STM6-40/36**	
	Parameters	**LL**	**UL**	**LL**	**UL**
1DM	$I_{ph}(A)$	0	1	0	2
	$I_{s1}(A)$	0	$1*10^{-06}$	0	$50*10^{-06}$
	$R_s(\Omega)$	0	0.5	0	0.36
	$R_{sh}(\Omega)$	0	100	0	1000
	n_1	1	2	1	2
2DM	$I_{ph}(A)$	0	1	0	2
	$I_{s1}, I_{s2}(A)$	0	$1*10^{-06}$	0	$50*10^{-06}$
	$R_s(\Omega)$	0	0.5	0	0.36
	$R_{sh}(\Omega)$	0	100	0	1000
	n_1, n_2	1	2	1	2
3DM	$I_{ph}(A)$	0	1	0	2
	$I_{s1}, I_{s2}, I_{s3}(A)$	0	$1*10^{-06}$	0	$50*10^{-06}$
	$R_s(\Omega)$	0	0.5	0	0.36
	$R_{sh}(\Omega)$	0	100	0	1000
	n_1, n_2	1	2	1	2
	n_3	2	5	1	2

III.5.1 Accuracy analysis

III.5.1.1 RTC France silicon solar cell identification

For the RTC France silicon solar cell, Table III-2 presents the results of the identification of five, seven and nine parameters as well as the as the $RMSE_C$ and $RMSE_L$ values of the 1DM, 2DM and 3DM obtained by the proposed algorithm BB-TLBO and other metaheuristic algorithms.

- For the 1DM identification and according to the lowest $RMSE_C$ value (9.860219E-04), BB-TLBO has the best results (mentioned in bold). BB-TLBO has also achieve the best results according to $RMSE_L$ with value of 7.730006E-04.
- For the 2DM, BB-TLBO is able to reach the minimum $RMSE_C$ value of 9.824848E-04 and also the minimum $RMSE_L$ value of 6.745134E-04.
- For the 3DM, the best parameters identified were obtained by BB-TLBO ($RMSE_C$ value of 9.80767E-04 and $RMSE_L$ value of 6.285394E-04.

As can be seen from this table, based on all results, the $RMSE_L$ is more precise than $RMSE_C$.

III.5.1.2 Monocrystalline solar module STM6-40/36 identification

Table III-3 reports the results of the identification of 5, 7 and 9 parameters as well as the $RMSE_C$ and $RMSE_L$ values for the 1DM, 2DM and 3DM respectively of the STM6-40/36 solar module obtained by BB-TLBO and other algorithms.

- As shown in Table III-3, for the 1DM identification, BB-TLBO reaches the lowest value for both $RMSE_C$ and $RMSE_L$ ($RMSE_C$ =1.729814E-03 and $RMSE_L$ =1.721922E-03) followed by LNMHGS according to $RMSE_C$ with RMSE value of 1.781552E-03. Both SDO and MCSWOA algorithms followed the BB-TLBO algorithm with $RMSE_L$ =1.729800E-03.
- BB-TLBO obtained the lowest value of $RMSE_C$ (best $RMSE_C$) which is equal to 1.693885E-03 for 2DM and 1.689064E-03 for 3DM compared to other algorithms mentioned in the same table. According to $RMSE_L$, BB-TLBO achieves the best value of 1.682879E-03 for 2DM and with vale of 1.682659E-03 for 3DM.

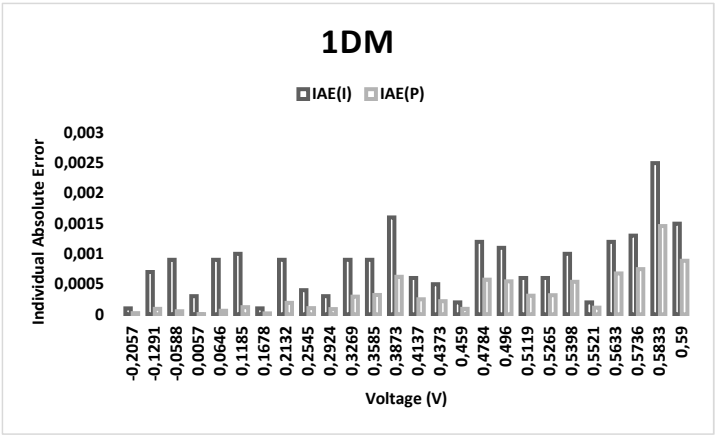
It is well noted that the RMSE based on Lambert W function is more precise than the conventional one.

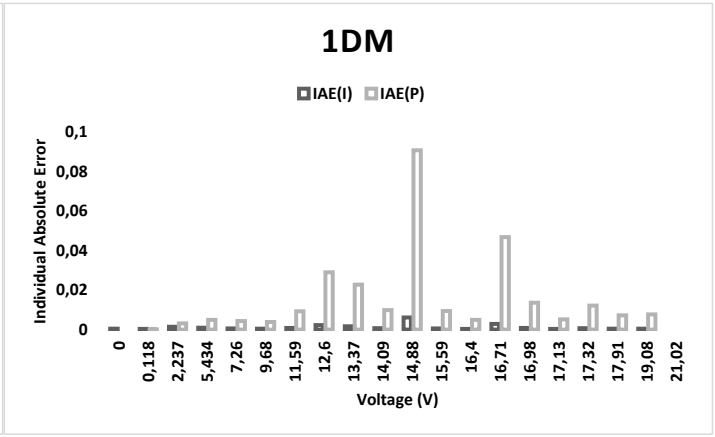
Table III-2: Best values of the identified parameters for the three adopted models of the RTC France silicon solar cell

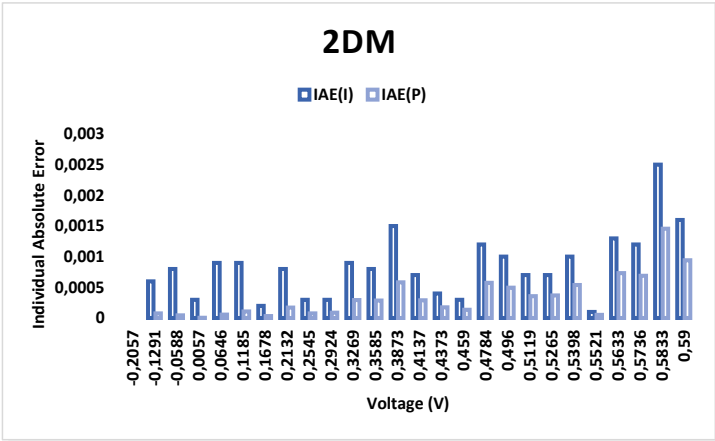
ref	Alg	$I_{ph}(A)$	$I_{s1}(A)$	$I_{s2}(A)$	$I_{s3}(A)$	$R_s(\Omega)$	$R_{sh}(\Omega)$	n_1	n_2	n_3	$RMSE_C$	$RMSE_L$
1DM												
[60]	GWOCS	0.76077	3.22E-07	/	/	0.03639	53.632	1.4808	/	/	9.860700E-04	/
[75]	PSO	0.7682	3.30E-07	/	/	0.03624	53.59878	1.48334	/	/	9.861450E-04	/
[71]	CPMPSO	0.760776	0.323021	/	/	0.036377	53.71852	1.481184	/	/	0.00098602	7.753900E-04
[71]	HISA	0.760788	0.3106846	/	/	0.036547	52.889794	1.4772678	/	/	0.00098911	7.730100E-04
[76]	BHCS	0.76078	0.32302	/	/	0.03638	53.71852	1.48118	/	/	0.00098608	7.754200E-04
[77]	LNMHGS	0.760758	3.35E-07	/	/	0.0362279	54.766955	1.4849735	/	/	9.886170E-04	/
[78]	HSOA	0.760763	3.32E-07	/	/	0.0362666	54.4885	1.4839901	/	/	9.874471E-04	/
[80]	RUN	0.760611	3.20-07	/	/	0.0364161	53.670706	1.4802504	/	/	9.862400E-04	/
	BBO	0.76073	3.67E-07	/	/	0.0356482	54.81141	1.49407	/	/	1.060180E-03	9.043310E-04
	TLBO	0.7601	4.23E-07	/	/	0.0351926	63.40126	1.50879	/	/	9.868560E-04	8.062100E-04
proposed	**BB-TLBO**	**0.760776**	**3.23E-07**	**/**	**/**	**3.64E-02**	**53.7186**	**1.4811808**	**/**	**/**	**9.860219E-04**	**7.730006E-04**
2DM												
[75]	GWOCS	0.76076	5.38E-07	2.49E-07	/	0.03666	54.7331	2	1.4588	/	9.833400E-04	/
[76]	AOE	0.76081	1.4111	1.8662	/	0.03723	55.5501	0.89974	0.13523		/	7.606800E-04
[75]	OLBGWO	0.760781	2.26E-07	6.43E-07	/	0.036722	55.307755	1.451328	1.96175	/	9.825560E-04	/
[75]	CWOA	0.76077	2.42E-07	6.00E-07	/	0.03666	55.2016	1.45651	1.9899	/	9.827200E-04	/
[80]	RUN	0.7608025	2.60E-07	5.58E-07	/	0.0364458	55.383219	1.4634784	1.9996951	/	9.871700E-04	/
[76]	TVACPSO	0.7608	1.32716	1.735315	/	0.03797	56.5496	4.04E-03	0.92746	/	/	7.436500E-04
[76]	MFO	0.760693	2.48E-13	3.60E-08	/	62.55984	1.461875	0.000001	2	/	1.053209E-03	/
	BBO	0.76083	1.00E-06	1.00E-06	/	0.0287533	100	1.62607	1.85384	/	1.130320E-03	8.672120E-04
	TLBO	0.76065	4.64E-07	8.37E-10	/	0.0348009	64.1189	1.51873	1.62049	/	1.004280E-03	7.631200E-04
proposed	**BB-TLBO**	**0.760781**	**7.48E-07**	**2.26E-07**	**/**	**0.0367398**	**55.482685**	**2**	**1.4510611**	**/**	**9.824848E-04**	**6.745134E-04**
3DM												
[75]	RAO	0.760795	2.62E-13	2.63E-13	9.78E-13	0.03674	55.35801	1.771502	1.451415	2.41101	9.845690E-04	/
[76]	ABC-DE	0.761148	1.263258	1.712923	2.733502	0.060322	70.669069	0.0144989	20.762276	0.055782	/	5.728946E-03
[76]	FC-EPSO1	0.76081	1.3324	1.7035	1.7906	0.037959	56.588	4.29E-02	0.38552	0.58921	/	7.4300E- 04
[75]	TLO	0.760789	2.54E-13	4.56E-14	1.48E-13	0.03671	55.3144	1.460287	1.740863	2.25143	9.861250E-04	/
[80]	RUN	0.7608367	3.30E-12	2.65E-07	8.42E-08	0.3631346	53.612584	1.071707	1.47338	1.572965	9.891330E-04	/
[75]	CS	0.760776	1.40E-07	1.90E-07	3.10E-08	0.0363	53.7218	1.4872	1.4771	4.4663	9.878570E-04	/
	BBO	0.76172	4.89E-07	6.62E-07	1.91E-07	0.03082	46.520172	1.88106	1.56784	2	1.224560E-03	8.212570E-04
	TLBO	0.76071	3.52E-07	1.55E-09	7.50E-07	0.0359163	55.36847	1.489839	1.967693	3.50035	9.857690E-04	7.514890E-04
proposed	**BB-TLBO**	**0.760781**	**2.65E-07**	**8.17E-27**	**1.00E-06**	**0.0366345**	**55.212701**	**1.4636056**	**2**	**2.2370724**	**9.807670E-04**	**6.285394E-04**

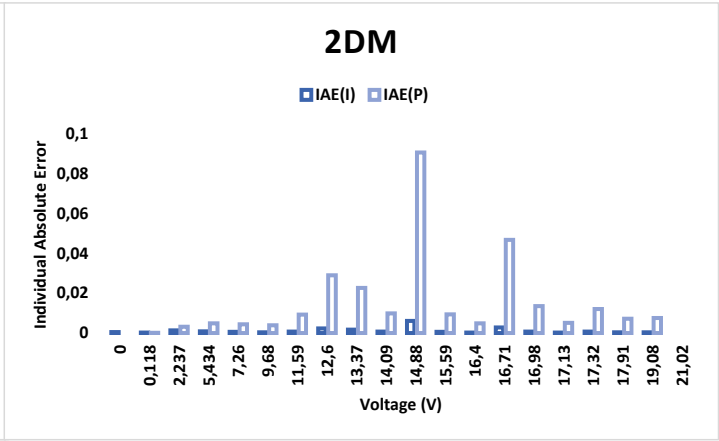
Table III-3: Best values of the identified parameters for the three adopted models of the solar module STM6-40/36

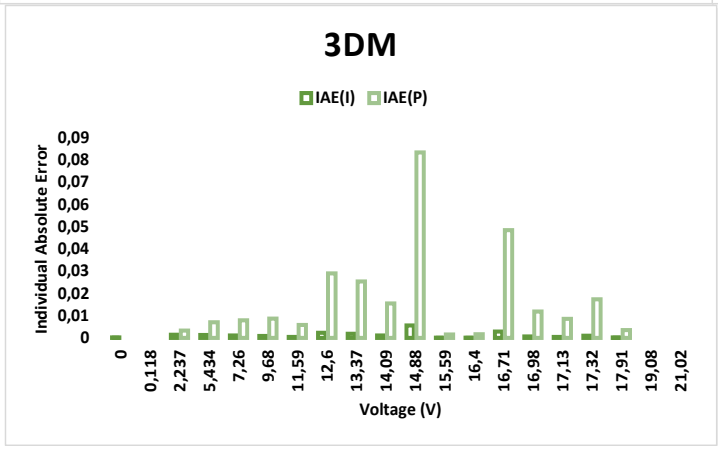
ref	Alg	$I_{ph}(A)$	$I_{s1}(A)$	$I_{s2}(A)$	$I_{s3}(A)$	$R_s(\Omega)$	$R_{sh}(\Omega)$	n_1	n_2	n_3	$RMSE_C$	$RMSE_L$
1DM												
[59]	TLBO-ABC	1.66317	2.14E-06	/	/	3.63E-03	17.25952	1.54354	/	/	1.806100E-03	/
[60]	GWO	1.656206	7.34E-06	/	/	1.48E-03	930.331	1.69641	/	/	7.141200E-03	/
[68]	CS	1.66172	3.73E-06	/	/	1.73E-03	21.74472	1.60905	/	/	2.515900E-03	/
[76]	STBLO	1.7	1.5	/	/	0.005	15.4	1.4127	/	/	/	1.900000E-03
[76]	SDO	1.6639	1.5203	/	/	0.0043	15.9283	1.7387	/	/	/	1.729800E-03
[76]	MCSWOA	1.6639	1.5203	/	/	0.0043	15.9294	1.739	/	/	/	1.729800E-03
[76]	MFO	1.862398	0.00E+00	/	/	0	32.86344	27.04247	/	/	3.107574E-01	/
[76]	LNMHGS	1.663453	2.08E-06	/	/	0.1327753	608.1023	56.5036	/	/	1.781552E-03	/
[76]	HSOA	1.662983	2.51E-06	/	/	0.11062102	649.5839	57.28595	/	/	1.935188E-03	/
	BBO	1.66096	5.41E-06	/	/	0.00035139	25.84392	1.656222	/	/	2.301350E-03	2.223850E-03
	TLBO	1.663687	1.78E-06	/	/	0.0041756	16.09665	1.523005	/	/	2.141840E-03	2.001470E-03
proposed	**BB-TLBO**	**1.663905**	**1.74E-06**	**/**	**/**	**4.27E-03**	**15.9283**	**1.5203**	**/**	**/**	**1.729814E-03**	**1.721922E-03**
2DM												
[82]	ELPSO	1.664843	6.21E-06	1.67E-09	/	0.5	606.8883	41.99348	67.344	/	1.830700E-03	/
[76]	EPSO	1.6644	1.7577	1.4527	/	0.26591	560.55	7.4011	1.4338	/	/	2.057300E-03
[76]	FC-EPSO1	1.6634	1.5818	1.5445	/	0.01101	597.29	1.8519	0.097148	/	/	1.772100E-03
[83]	MCSWOA	1.6639	6.10E-07	1.18E-05	/	0.0054	16.9519	1.4224	2.1992	/	1.706100E-03	/
[84]	SDO	1.6639	1.74E-06	5.00E-05	/	0.0043	15.9372	1.5203	54.5816	/	1.729800E-03	/
[85]	EPSO	1.6644	7.40E-06	1.43E-06	/	0.26591	560.55	1.7577	1.4527	/	2.057300E-03	/
	BBO	1.673873	5.70E-06	2.36E-05	/	4.75E-07	384.0411	1.931814	1.908805	/	3.292930E-03	3.002831E-03
	TLBO	1.66387	1.46E-06	5.28E-07	/	0.00425743	16.13067	1.508304	1.708518	/	1.891110E-03	1.734210E-03
proposed	**BB-TLBO**	**1.663744**	**5.78E-08**	**5.94E-06**	/	**0.00643011**	**17.39714**	**1.255163**	**1.80345**	/	**1.693885E-03**	**1.682879E-03**
3DM												
[36]	BSDE	1.661089	1.49E-06	3.90E-06	1.37E-06	0.00201029	25.78758	1.531215	1.868278	1.882878	2.894519E-03	/
[36]	SDO	1.663744	6.36E-06	8.34E-07	4.72E-07	0.00537714	17.00997	1.995203	1.969208	1.404689	1.701880E-03	/
[36]	MRFO	1.662318	2.60E-06	6.33E-08	2.77E-07	0.0028541	19.10016	1.566776	1.955282	1.974909	2.033192E-03	/
[36]	BSA	1.657576	3.51E-07	2.19E-07	1.14E-05	0.00466652	43.38224	1.392706	1.823212	1.969855	3.654669E-03	/
[36]	CSO	1.660801	5.35E-06	4.54E-08	3.56E-07	0.00011186	25.49507	1.657277	1.999604	1.85013	3.320772E-03	/
	BBO	1.655004	6.94E-09	6.57E-09	6.89E-09	0.0155553	13.11045	1.141827	1.334208	1.114979	3.366020E-03	2.907204E-03
	TLBO	1.663129	1.68E-07	1.97E-06	1.21E-10	0.00381094	17.01134	1.997693	1.534797	1.555108	1.805210E-03	1.724010E-03
proposed	**BB-TLBO**	**1.663767**	**7.76E-07**	**4.22E-06**	**1.76E-08**	**0.0068122**	**17.39**	**1.671175**	**1.761521**	**1.184057**	**1.689064E-03**	**1.682659E-03**











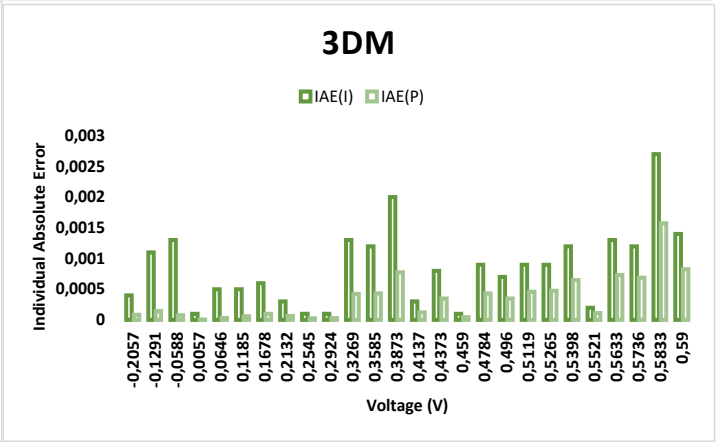


Figure III-3: IAE of the current and the power of the three adopted models of the RTC France silicon solar cell.

Figure III-4: IAE of the current and the power of the three adopted models of the solar module STM6-40/36

After identifying the parameters of all models for RTC France silicone solar cell and STM6-40/36 solar module, the output current and power relative to the measured voltage can be determined. Figures III-3 and III-4 show the IAE absolute error of current and power calculated using Equations III-40 and III-41 respectively.

$$IAE(I) = |I_c - I_m| \quad \text{(III-40)}$$

$$IAE(P) = |P_c - P_m| \quad \text{(III-41)}$$

Where: I_C and P_c are the current and the power calculated using BB-TLBO, respectively. I_m and P_m are the measured current and power, respectively.

As can be seen in Figures III-3 and III-4, the current IAE value for:

- RTC France solar cells based on 1DM, 2DM and 3DM are less than 0.0025(A), 0.002896(A) and 0.00310(A) respectively.
- STM6-40/36 solar module are less than 0.00609(A), 0.005581(A) and 0.0028(A) for 1DM, 2DM and 3DM respectively.

As can be shown in Figures III-3 and III-4, The value of IAE of the power for:

- RTC France solar cell based on 1DM, 2DM and 3DM are less than 0.0014583(w), 0.001413(w) and 0.0813554(w).
- STM6-40/36STM6-40/36 solar module are less than 0.09062(w), 0.083042(w) and 0.0015(w) for 1DM, 2DM and 3DM respectively.

Moreover, the characteristic curves (I-V) and (P-V) for all the models and in the different cases show that the measured data and the calculated data obtained by BB-TLBO are similar (Figures III-5 and III-6).

It can be concluded from Tables III-2 and III-3 that the proposed hybrid metaheuristic algorithm BB-TLBO can achieve better RMSE value and exhibits very good accuracy in parameter estimation compared to other algorithms for 1DM, 2DM and 3DM whether for the RTC France solar cell or the STM6-40/36 solar module.

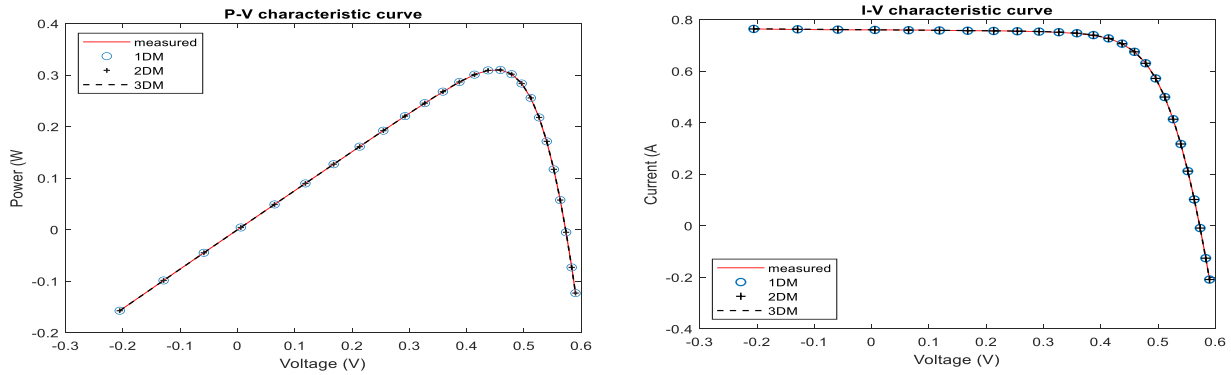


Figure III-3: P-V and I-V characteristic curves for the three RTC France silicon solar cell models.

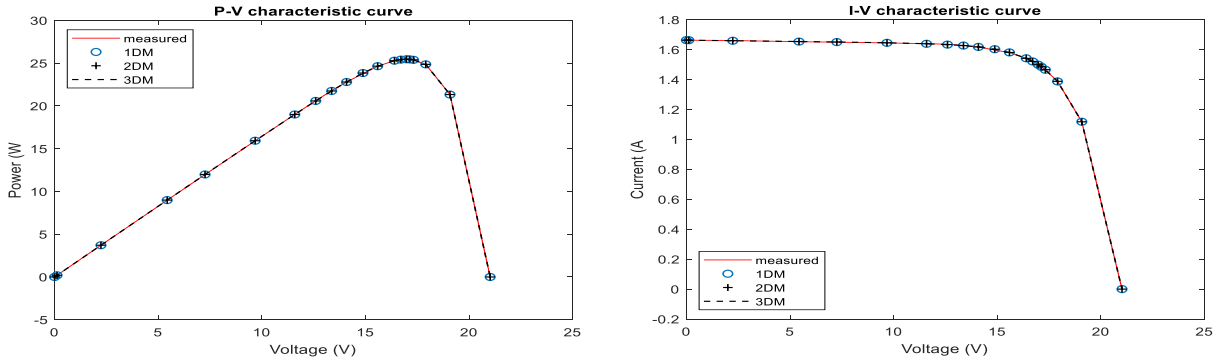


Figure III-4: P-V and I-V characteristic curves for the three models of the STM6-40/36 solar modules.

III.5.2 Statistics analysis

In order to evaluate the robustness and reliability of the proposed BB-TLBO algorithm, the statistical results obtained by BB-TLBO and other algorithms such as: BBO, TLBO, DOLADE, LAPSO, … are presented in Tables III-4 and III-5. The best statistical results for 1DM, 2DM and 3DM of the RTC France silicon solar cell and the STM6-40/36 are shown in bold.

These results include: $RMSE_{min}$ which represents the lowest value of RMSE, $RMSE_{mean}$ corresponds to the average value of RMSE, $RMSE_{max}$ is the worst value, *It*(s) is the number of iterations to obtain the best results, CPU(s) is the running time and STD is the standard deviation of the RMSE values, it indicates the reliability of the algorithm.

III.5.2.1 RTC France silicon solar cell identification

The statistical results for the 1DM, 2DM and 3DM of RTC France silicon solar cell are shown in Table III-4.

- For 1DM identified on the basis of BB-TLBO, it is found that the lowest RMSE value is reached after 50 iterations and in 1.14 seconds with an STD of 8.821895E-19.
- For the France RTC cell based on 2DM, the best results are obtained by BB-TLBO which records an RMSE=9.824848E-04, STD=2.686117E-08, *It*(s)=117 and CPU=10.75355s.
- For the identification of the 3DM of RTC France, BB-TLBO records an RMSE=9.807670E-04, STD=3.034764E-08, *It*(s)=143 and CPU=11.1490s.

III.5.2.2 Monocrystalline solar module STM6-40/36 identification

For the identification of the STM6-40/36 solar module based on 1DM, 2DM and 3DM, the statistical results obtained by BB-TLBO and other algorithms are shown in Table III-5.

- It can be seen that BB-TLBO achieved the best results after 134 it(s) and 117 it(s) for 1DM and 2DM respectively. It features the lowest CPU time (0.169081s (1DM) and 5.415939s (2DM)) with the lowest STD value (STD=1.64E-15 (1DM) and STD=9.00E- 06 (2DM)) and the smallest RMSE value (1.73E-03 (1DM) and 1.69E-03(2MD)).

- From the results shown in the same table, both of BB-TLBO and IQSODE give the same RMSE values in terms of (min, mean and max). However, in term of STD the BB-TLBO algorithm obtained the second-best result with value of 6.616421E-19 following the IQSODE algorithm that obtained a value of 4.336809E-19.
- For 3DM, the best STD (STD= 1.294738E-05) is obtained by BB-TLBO which reaches the minimum RMSE value of 1.689064E-03 after only 110 iterations in 6.066541s.

Thus, the proposed BB-TLBO algorithm is stable and reliable to identify the parameters of RTC France silicon solar cell and STM6-40/36 solar module.

Table III-4: Statistical results obtained during the identification of 1DM, 2DM and 3DM of RTC France

Ref	alg	$RMSE_{min}$	$RMSE_{mean}$	$RMSE_{max}$	STD	$It(s)$	$CPU(s)$
1DM							
[77]	ECM-JADE	9.860219E-04	9.860219E-04	9.860219E-04	4.609693E-17	/	/
[66]	DOLADE	9.860219E-04	9.860219E-04	9.860219E-04	3.277377E-17	/	/
[78]	IMFOL	9.860200E-04	9.870500E-04	9.897700E-04	1.020700E-06	/	/
[79]	RLDE	9.860220E-04	9.860220E-04	9.860220E-04	5.451690E-17	/	/
[79]	LaPSO	9.860220E-04	9.860220E-04	9.860220E-04	3.668200E-13	/	/
[78]	MFO	1.11160E-03	2.20190E-03	2.44800E-03	3.81790E-04	/	/
	BBO	1.06018E-03	2.11037E-03	2.46065E-03	4.74228E-04	> 2000	> 20.3542
	TLBO	9.86856E-04	1.05013E-03	1.23504E-03	6.75216E-05	> 2000	> 10.4671
proposed	**BB-TLBO**	**9.860219E-04**	**9.860219E-04**	**9.860219E-04**	**8.821895E-19**	**50**	**1.137621**
2DM							
[77]	IQSODE	9.824849E-04	9.836710E-04	9.860260E-04	1.345204E-06	/	/
[66]	DOLADE	9.824849E-04	9.826028E-04	9.860219E-04	6.457690E-07	/	/
[78]	IMFOL	9.825300E-04	9.867500E-04	9.959600E-04	3.226800E-06	/	/
[78]	MFO	1.053200E-03	2.147700E-03	2.994000E-03	4.346400E-04	/	/
[79]	FLIDE	9.824850E-04	9.884790E-04	1.021620E-03	1.055370E-05	/	/
[79]	RLDE	9.824850E-04	9.858400E-04	1.007730E-03	4.400270E-06	/	/
[79]	LaPSO	9.824850E-04	9.842650E-04	9.862120E-04	1.680880E-06	/	/
	BBO	1.13032E-03	2.45570E-03	3.72398E-03	7.68496E-04	> 2000	>20.2376
	TLBO	1.00428E-03	1.12907E-03	1.38165E-03	1.11267E-04	> 2000	>22.34672
proposed	**BB-TLBO**	**9.824848E-04**	**9.824899E-04**	**9.826319E-04**	**2.683117E-08**	**117**	**10.75355**
3DM							
[80]	CS	9.87857E-04	3.51602E-03	8.49868E-01	2.28842E-03	/	/
[81]	TLO	9.86125E-04	2.92145E-03	4.41563E-01	2.19458E-03	/	/
[81]	RAO	9.84569E-04	2.01125E-03	2.12589E-01	9.62140E-04	/	/
[81]	ABC	9.84522E-04	3.51369E-03	8.52365E-01	2.28412E-03	/	/
	BBO	1.22456E-03	2.36534E-03	3.68666E-03	6.53851E-04	> 2000	>8.7465
	TLBO	9.85769E-04	1.04152E-03	1.16712E-03	5.44385E-05	> 2000	> 9.2232857
proposed	**BB-TLBO**	**9.807670E-04**	**9.807670E-04**	**9.807670E-04**	**3.034764E-08**	**143**	**11.1490**

Table III-5: Statistical results obtained during the identification of 1DM, 2DM and 3DM of the STM6-40/36 solar module

ref	alg	$RMSE_{min}$	$RMSE_{mean}$	$RMSE_{max}$	STD	$It(s)$	$CPU(s)$
1DM							
[82]	MLBSA	1.743415E-03	4.297333E-03	3.329609E-02	5.677849E-03	/	/
[60]	GWOCS	1.733700E-03	1.745700E-03	1.752800E-03	1.044700E-05	/	/
[83]	ELPSO	2.180300E-03	2.250300E-03	3.716000E-03	2.921100E-04	/	/
[84]	SDO	1.729800E-03	1.770300E-03	1.950000E-03	4.510800E-05	/	/
[78]	IMFOL	1.887000E-03	3.033500E-03	4.674900E-03	8.545400E-04	/	/
[78]	MFO	3.107600E-01	3.107600E-01	3.107600E-01	1.693800E-16	/	/
[77]	**IQSODE**	**1.729814E-03**	**1.729814E-03**	**1.729814E-03**	**4.336809E-19**	/	/
	BBO	2.301350E-03	6.243230E-03	1.071160E-02	1.543310E-03	> 2000	>5.85746
	TLBO	2.141840E-03	2.610550E-03	3.329680E-03	3.125230E-04	> 2000	>5.836608
proposed	**BB-TLBO**	**1.729814E-03**	**1.729814E-03**	**1.729814E-03**	**6.616421E-19**	**134**	**0.169081**
2DM							
[84]	SDO	1.729800E-03	1.811800E-03	2.028800E-03	7.242100E-05	/	/
[83]	ELPSO	1.830700E-03	2.035100E-03	2.117800E-03	8.427100E-05	/	/
	BBO	3.292930E-03	5.573130E-03	9.093230E-03	1.479270E-03	> 2000	>8.0379308
	TLBO	1.891110E-03	3.490130E-03	5.263350E-03	8.689140E-04	> 2000	>6.3345426
proposed	**BB-TLBO**	**1.693885E-03**	**1.694230E-03**	**1.697676E-03**	**8.587814E-07**	**117**	**5.415939**
3DM							
	BBO	3.366020E-03	4.878470E-03	5.759890E-03	5.820820E-04	> 2000	>8.4645975
	TLBO	1.805210E-03	3.299990E-03	5.131280E-03	1.097240E-03	> 2000	>6.3790644
proposed	**BB-TLBO**	**1.689064E-03**	**1.689064E-03**	**1.707124E-03**	**1.294738E-05**	**110**	**6.066541**

III.5.3 Convergence analysis

The convergence curves of BB-TLBO and other algorithms such as: BBO, TLBO, SCE, SFLA, DE and ABC (when identifying RTC France silicon solar cell and STM6-40/36 solar module for the three models (1DM, 2DM and 3DM)) are shown in Figures III-6 and III-7.

According to these figures, the BB-TLBO algorithm has the best convergence speed compared to the other algorithms in all cases.

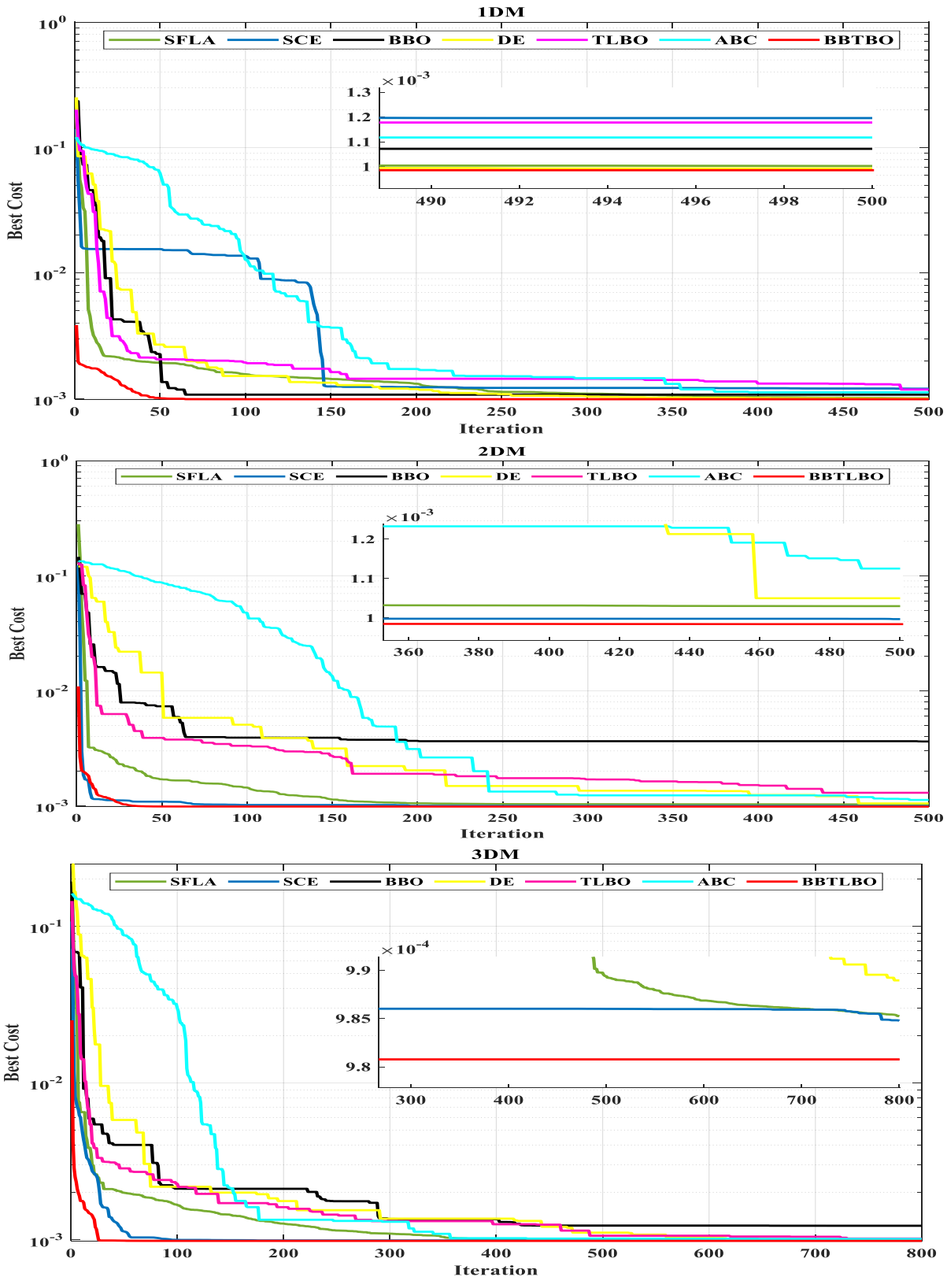


Figure III-7: Convergence curves for three models of RTC France silicon solar cell.

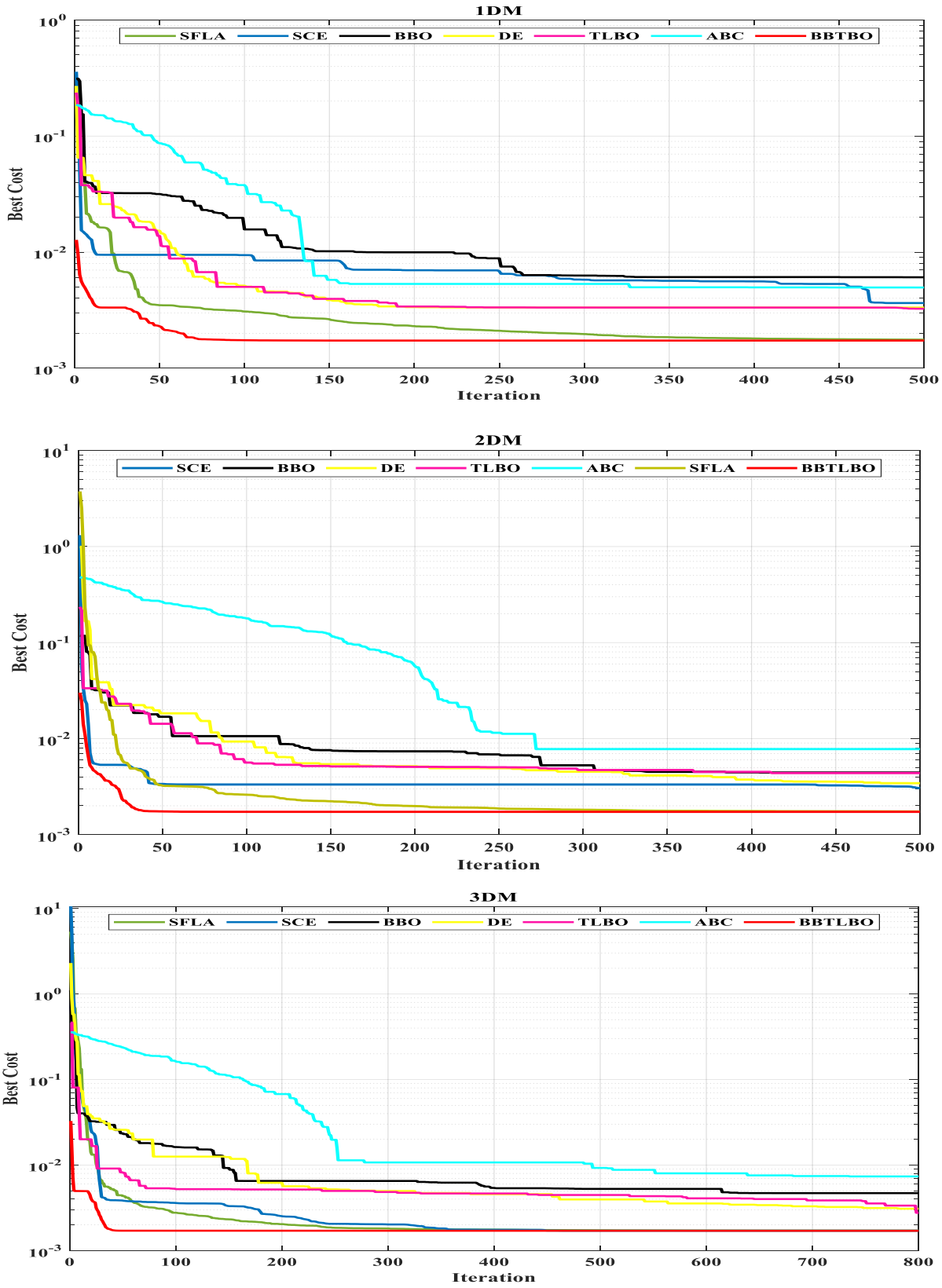


Figure III-8: Convergence curves for three models of STM6-40/36 module.

III.6 Conclusion

This chapter deals with the problem of identifying unknown parameters related to several models of solar cells and modules. Biogeography-based teaching learning-based optimization (BB-TLBO) is suggested as a new optimization algorithm to solve this problem. This algorithm combines two metaheuristic algorithms called: biogeography-based optimization (BBO) and teaching learning-based optimization (TLBO). The root mean square error (RMSE) is used as an objective function in its both forms conventional and Lambert W function. To check the efficiency of BB-TLBO, it is applied on one diode model (1DM), two diode model (2DM) and three diode model (3DM) of silicon solar cell RTC France and commercial PV solar module, the monocrystalline STM6-40/36.

The following points can be drawn from the results discussed above:

(i) BB-TLBO performs better in terms of robustness and reliability according to comparison tables 4-5. It revealed the smallest RMSE value in all models of France RTC cell and STM6-40/36 module

(ii) BB-TLBO has a very high accuracy in terms of identification of parameters according to IAE, I-V and P-V characteristics.

(iii) According to the statistical results (RMSE, STD, CPU and It(s) values), BB-TLBO outperforms its own single algorithms (BBO and TLBO) as well as other newly techniques such as: DOLADE, IQSODE, IMFOL, … etc. For example, BB-TLBO records the best statistical results (RMSE=1.729814E-03, STD=6.616421E-19, CPU=0.17s, It(s)=134) when identifying parameters of STM6-40/36 presented on 1DM, and it records the smallest RMSE in less than 6 seconds, RMSE= 1.693885E-03 for 2DM and RMSE=1.689064E-03 for 3DM representations.

(iv) Convergence curves demonstrated that BB-TLBO has a very fast convergence speed.

(v) From the results mentioned above, it is clear that the three-diode model (3DM) is the more accurate model to model the RTC France silicon solar cell and the STM6-40/36 solar module

Overall, The BB-TLBO algorithm is proven to be superior to other recently introduced parameter extraction strategies in terms of precision, stability, and speed through experiments. Consequently, the proposed BB-TLBO can be applied as an effective alternative solution to the issue of PV model parameter extraction.

Chapter IV : Literature review on PV faults

IV.1 Introduction

In contrast to traditional energy sources that release harmful pollutants into the air, photovoltaic (PV) energy emerges as a renewable, sustainable, and eco-friendly alternative. However, like any complex system, PV installations are susceptible to various faults and failures that can hinder their performance and efficiency over time. these faults include short circuit faults, open circuit faults, line-to-line faults, line-to-ground faults, and more.

To address this critical issue, numerous research endeavours have concentrated on developing methods for detecting and diagnosing faults in PV systems, with the objective of mitigating the potential for system failures. These methods are employed to address concerns regarding the reliability, availability, maintainability, and safety of PV plants, collectively referred to as RAMS.

Fault detection and diagnosis in PV systems can be categorized into three primary methods:

- Manual methods involve techniques such as visual inspection, time domain reflectometry, earth capacitance measurement, and spread spectrum analysis. These methods are particularly suitable for smaller-scale PV plants, typically at the string level.
- Semi-automatic methods employ tools like thermal cameras, infrared, or electroluminescence imaging to analyse and pinpoint faults. This approach is commonly utilized in larger-scale PV plants.
- Automatic methods utilize data loggers and monitoring systems to gather data and implement fault detection algorithms for performance analysis. These methods are commonly deployed in PV generators with rated power ranging from a few kilowatts to hundreds of megawatts.

In general, diagnosis process involves three key steps:

- Detection: This initial phase focuses on comparing measured data with reference values to identify anomalies. Accurate threshold values are essential to avoid false alarms. Data quality and the accuracy of reference models can impact this stage.
- Identification: In this step, algorithms are used to determine the nature and cause of faults. This can be challenging, especially when multiple faults occur simultaneously or share similar characteristics. Advanced AI techniques such as machine learning and deep learning are often employed.

- Localization and Isolation: The final step involves pinpointing the fault's location within the system and isolating it swiftly. This phase requires extensive information and expert knowledge for effective resolution.

This chapter will cover all the essential aspects related to faults that can occur in photovoltaic systems, including their identification, their impact on the overall operation, and the different methods used for their detection and diagnosis.

IV.2 Factors that affect the efficiency of the PV module

There is growing interest in the use of photovoltaic modules as a future energy source. Many countries with abundant sunshine view these modules as a way to efficiently harness free energy. Indeed, sunlight offers clean and inexhaustible energy, thus attracting the attention of many players in the electricity sector. To enhance the electricity production potential of photovoltaic modules, it is essential to improve their efficiency. Module efficiency refers to the portion of solar energy received by the module that is converted into electricity [85]–[87].

Recent research indicates that the performance of solar cells is influenced by three primary factors:

- Semiconductor materials are used to absorb photons and convert them into electron-hole pairs.
- The formation of the semiconductor which facilitates the separation of holes and electrons.
- Contacts positioned at the front and back of the solar cell, which serve as paths for current flow to the external circuit.

To ensure the effective functioning of the PV module, quality standards must be met. According to[88], the requirements for a high-quality photovoltaic energy system include:

- Proper sizing and orientation to ensure adequate electrical power and energy supply.
- Integration of an efficient control circuit to mitigate issues like overcurrent protection and electrical losses.
- Inclusion of an effective charge controller, especially when batteries are part of the system.

The efficiency of the PV module is influenced by various factors, categorized into three groups: cell materials, PV system devices, and environmental factors, as depicted in Figure IV-1.



Figure IV-1: Factors, which affects the efficiency of photovoltaic module performance [88].

IV.2.1 Cell materials

The solar cell consists of semiconductor materials, with the most common ones being monocrystalline silicon, amorphous silicon, polycrystalline silicon, and copper-indium selenide (CIS). These materials, utilized in the production of photovoltaic modules, are influenced by various factors that hinder their capacity to produce electricity. These factors include:

IV.2.1.1 Effect of corrosion on PV module efficiency

Corrosion occurs when a material undergoes degradation due to chemical reactions with its surrounding environment. When liquid or gas contacts the exposed parts of a metallic surface, corrosion can occur, often initiated by the presence of salts, acids, and elevated temperatures. This phenomenon impacts a wide range of metals. According to recent studies, the conductive layer of PV modules is susceptible to corrosion when moisture infiltrates the module edges, particularly affecting thin film modules that lack frames. Additionally, corrosion typically leads to an increase in the series resistance of the module, resulting in significant performance loss and reduced efficiency [89]–[91].

IV.2.1.2 Effect of degradation on PV efficiency

Degradation stands out as a significant contributor to the decline in the lifetime of PV modules, influenced by various climatic and physical factors. This process leads to a reduction in the module's electricity generation capacity, reflected in key characteristics such as short-circuit current, open-circuit voltage, fill factor (FF), maximum power output (Pmax), maximum output voltage (Voc), and

P-V and I-V curves. As module degradation increases, the efficiency of the PV module declines [92], [93]. A previous study has identified different types of module degradation, illustrated in Figure IV-2.

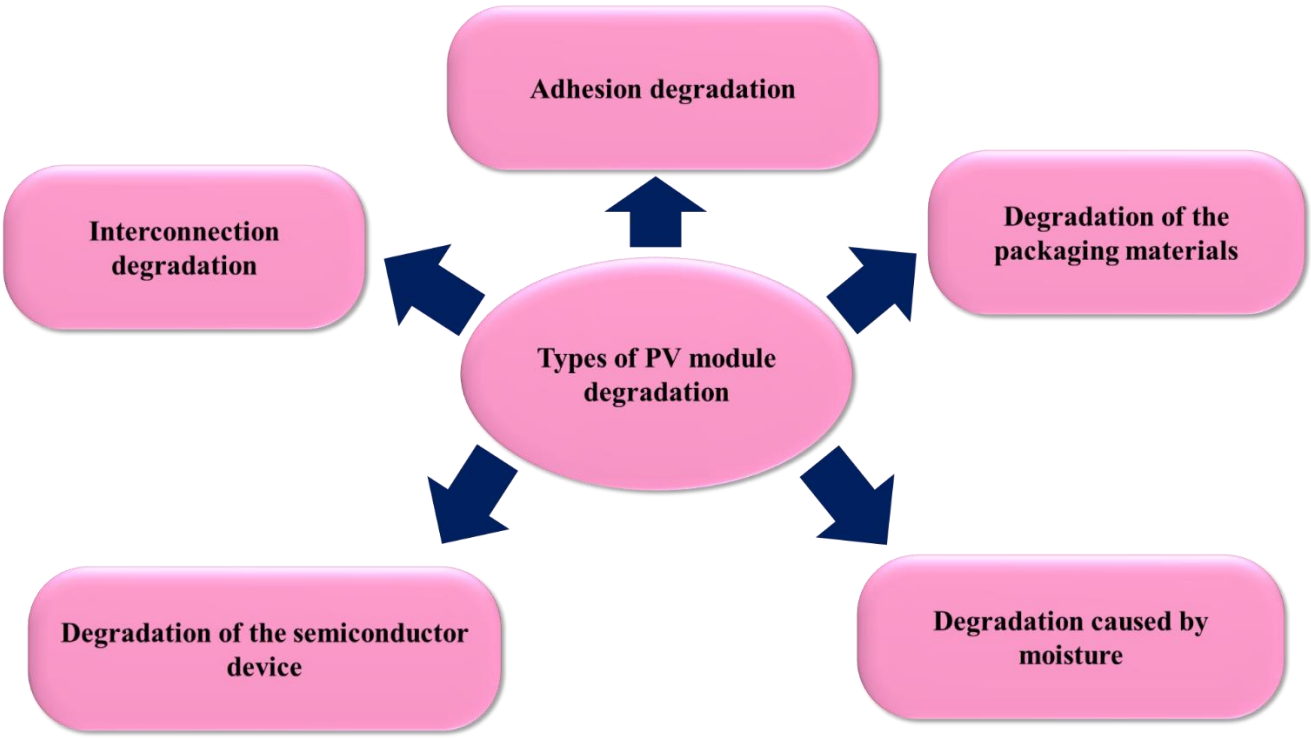


Figure IV-2: Types of PV module degradation [94].

IV.2.1.3 The effect of cracked cells on PV module efficiency

The occurrence of cracks in PV module cells poses an inevitable challenge commonly encountered in PV technology, impacting both the module's longevity and efficiency. Cracks within a solar cell manifest varying effects on the module's power output. A single crack causing the disconnection of a crucial cell element can result in a significant loss in output power, while cracks that do not disrupt vital elements have a lighter impact on module performance. External factors such as hailstorms, strong winds, and snow loads have been identified in recent research as contributors to surface cracks on modules[95], [96]. These cracks can lead to the disconnection of essential cell parts, diminishing the module's ability to generate maximum power and subsequently reducing efficiency. The research also identifies different types of cracks affecting PV modules, as depicted in Figure IV-3.

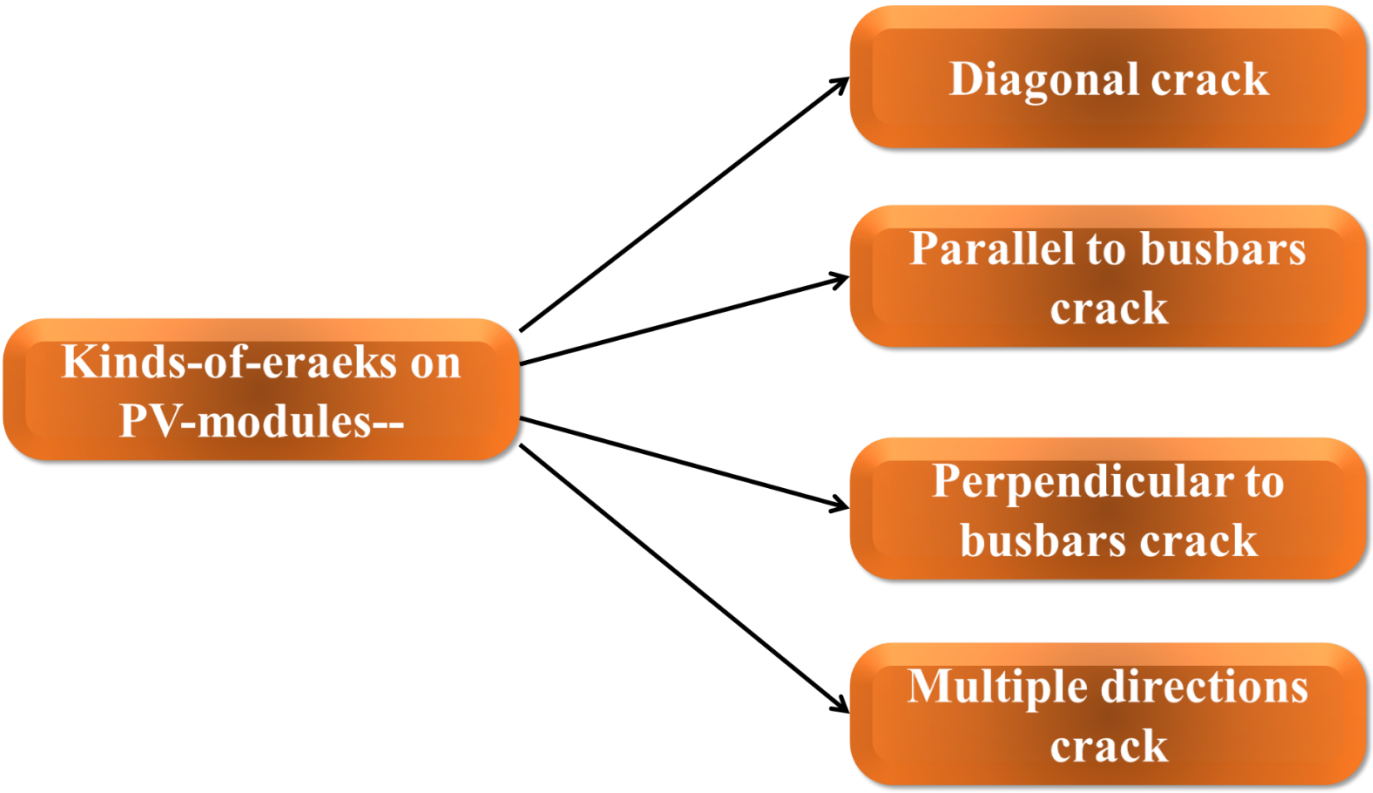


Figure IV-3: Kinds of cracks on PV modules [94].

IV.2.1.4 The effect of contact stability on PV module efficiency

The contacts situated at the front and rear of the solar cell, serving as pathways for current to flow to the external circuit, play a crucial role in solar cell performance. Once the holes and electrons within the semiconductor module are separated, they must be collected to facilitate their flow through an electrical circuit, necessitating electrical contacts. The contact on the surface of the panel, where light penetrates, must effectively gather carriers without shading more than 10% of the cell's surface. Conversely, the rear surface contact of the panel typically consists of a metal sheet. Establishing a strong ohmic contact between the semiconductor and the metal interface is essential to minimize losses, thereby improving module performance as losses decrease [87]. A research work [88] suggests that the mechanical stability of photovoltaic cells plays a crucial role in both the reliability of the module and potential power degradation. It also indicates that if the thickness of the cells in a PV module decreases, their mechanical stability also decreases.

IV.2.1.5 Effect of the selected semiconductor on solar panel performance

Besides production costs, module efficiency stands out as a key metric for evaluating Photovoltaic modules across various applications. In terms of efficiency, crystalline silicon remains the top performer at the module level, followed by CIGS thin-film [97]. The material composition of the module and the energy gap of the semiconductor materials are crucial factors. Energy gaps of solar cells made of semiconductors depends on type of solar cell material, the most efficient solar cell

materials typically fall within 1.3 and 1.6 eV [98]. Utilizing materials with high energy conversion efficiency leads to an increase in module efficiency and power output.

IV.2.2 PV system devices

The equipment employed in solar energy systems for electricity generation serves a valuable purpose, yet they are not without their limitations. When not utilized effectively or when their performance is hindered by equipment failure or other factors, the efficiency of the modules decreases. These factors are outlined below:

IV.2.2.1 Size of inverters, batteries, and charge controller

The efficiency of the PV module relies heavily on the inverter's performance, which converts direct current (DC) into alternating current (AC). Meanwhile, the charge controller plays a crucial role in preventing battery overcharging and ensuring that batteries do not discharge back to the charging source while storing energy. If the inverter is undersized, it may clip a portion of the module's power output to prevent exceeding its rated capacity. Conversely, oversizing the inverter involves installing one with a capacity greater than the PV module's nominal capacity, often done for potential future expansions but not typically recommended. Incorrectly sized inverters, batteries, or charge controllers can lead to system underperformance or failure, thus diminishing system efficiency [99].

IV.2.2.2 Effect of Inverter power conversion efficiency on PV performance

The efficiency of the inverter in a PV system, responsible for converting DC to AC power, holds significant importance. It dictates how much DC power generated by the panel is successfully converted into AC power, with some energy being lost as heat and power consumed while the inverter is active. In certain Photovoltaic systems, inverters with lower power ratings than the solar panel are utilized, as prolonged operation at the panel's rated efficiency may not be sustainable. This setup ensures alignment with the initial input voltage and maximum point voltage range requirements. Utilizing inverters with high power conversion efficiency enhances PV system performance. However, the power conversion efficiency of the inverter can diminish due to factors such as direct exposure to sunlight, inadequate ventilation, or installation near combustible surfaces, as noted in recent studies [100]–[102].

IV.2.2.3 Effect of incorrect size/ type of conductor

The conductors employed in PV installations, as indicated in recent studies, are typically engineered to meet the following criteria [103]:

- Endure exposure to UV radiation originating from sunlight;

- Possess water-resistant properties;
- Tolerate extreme temperatures.

Solar energy systems are typically designed to utilize standard color-coded solar cables and interconnect sizes. However, during the installation process, the cables or conductors may undergo significant twisting, abrasions, and flexing as they are routed through conduits, potentially impacting the insulation of the cable. The utilization of incorrectly sized or inappropriate types of cable conductors has the potential to compromise the performance of the installation or lead to system failure [103].

IV.2.2.4 Effect of extent of voltage drop on PV performance

Voltage drop refers to the inevitable loss of electrical potential that occurs when current flows through a conductor from the source to the load. The extent of voltage drop is determined by the total length of the conductor traversed by the current. It is essential to ensure that the voltage drop remains within acceptable limits for cable installations. For example, the voltage drops along the cable connecting the PV panel to the inverter should ideally be around 3%, with 2% on the DC side and 1% on the AC side. Therefore, the cabling work in a PV system should be carried out in a manner that minimizes resistive losses to less than 3% for optimal efficiency. Minimizing losses in a PV energy system is crucial as it prevents energy wastage and preserves the overall efficiency of the system [104], [105].

IV.2.2.5 Odd wiring technique

Incorrect wiring techniques contribute significantly to the decreased performance of solar panel installations. When PV modules from different manufacturers, or with varying wattage or voltage, are interconnected, the challenge does not stem from the vendors themselves, but rather from the differences in performance degradation and characteristics among the combined PV modules. It is crucial to connect only PV modules with similar or identical current ratings. For instance, if a 5A PV module is connected to another PV module with a current rating of 5.8A, the overall current will be limited to 5A, resulting in a reduction in total current output and consequently, a decrease in system performance. Similarly, connecting a 20V PV module to a 30V PV module will lead to a total voltage of 20V, resulting in reduced output power and a subsequent decline in system performance [106].

IV.2.2.6 Insufficient use of disconnect

The PV energy system necessitates a sufficient number of disconnects to safeguard against system failure. When a heavy load is connected to the solar system, it requires a separate disconnect to isolate it from the system when it should not receive power. As recommended in previous research [107], the

design must protect the system components from short circuits, reverse polarity connections, and electrical storms. Fuses or circuit breakers should be positioned on the positive line near the battery, while lightning surge protection and other relevant devices must be installed with proper means of disconnection. Equipping solar systems with the requisite number of disconnects helps minimize failure. Conversely, installing fewer disconnects than required significantly heightens the risk of failure, leading to inevitable poor performance. Ensuring the appropriate number of disconnects and over-current protection is crucial for the safety of the system's electrical design.

IV.2.2.7 Effect of Incorrect grounding on PV performance.

The performance of a solar energy installation is susceptible to failure or poor performance if grounding is not properly implemented. As outlined in [102], challenges may arise due to various factors, including:

- Absence of lightning protection;
- Grounding conductor contact with module frames and aluminium rails;
- Installation of PV energy systems in exposed locations. Incorrect grounding adversely affects the entire solar energy installation, resulting in poor performance and eventual failure.

IV.2.2.8 Underrated component usage

To ensure the PV installation begins and remains operational at its best, it's crucial to employ the correct size and type of cables and components. Utilizing inappropriate components, such as inferior or faulty ones, diminishes the overall performance of the installation and may lead to system failure [94].

IV.2.3 Environmental factors

Numerous environmental elements contribute to the decline in PV module performance, but their impact can be mitigated to enhance the module's functionality. Below are some key environmental factors that hinder PV module efficiency:

IV.2.3.1 Effect of dust on PV module efficiency

In recent research [108], it was discovered that dust particles significantly impact the efficiency of PV modules. An experiment conducted revealed that a PV module left uncleaned for over six months could experience up to a 50% reduction in power output. Moreover, observations from [87] suggest that dust accumulation peaks during dry seasons. The buildup of dust on PV module surfaces diminishes their power output and overall efficiency. To maintain optimal performance, it is recommended to regularly clean the surfaces of PV modules to remove dust particles.

IV.2.3.2 The effect of solar radiation on PV module efficiency

The power output of PV modules is influenced by irradiance, along with other factors such as the angle of incidence and module tilt angle. Although higher solar irradiation typically implies increased power output, the reality is more nuanced. Approximately 15% of solar irradiation is converted into usable electricity, while the remaining 85% is converted into heat. This heat elevates module temperature, leading to decreased PV module efficiency [90]. Recent research [109] has highlighted the variations in solar radiation incidents on the surfaces of photovoltaic modules depending on the position of the panels and the time of day.

Additionally, the amount of solar radiation reaching the surface of a module depends on the type of photovoltaic system. For fixed photovoltaic systems, the radiation depends on the orientation and inclination of the panel, while for tracking systems it depends on the type of tracker. Concentrating photovoltaic systems use normal radiation directly [86]. Studies conducted in [110] indicate that higher irradiance results in increased output current, leading to greater electricity production, provided that all relevant parameters remain constant.

IV.2.3.3 The effect of humidity on PV modules

The atmospheric relative humidity represents the moisture content of the air, typically indicated by the ratio of water vapor present to the saturated vapor pressure at the same temperature. Research [111] indicates that when air with consistent water vapor content undergoes a temperature increase, relative humidity decreases. Despite being invisible, water vapor affects solar cells, primarily causing corrosion. This corrosion adversely affects the titanium-silver contacts of solar cells, leading to decreased efficiency of the PV module.

IV.2.3.4 The effect of ambient temperature on PV module efficiency

The surrounding air temperature, known as ambient temperature, significantly impacts the efficiency and power output of installed PV modules. According to recent research [112], factors like wind speed, cloud cover, and ambient temperature itself influence the temperature of the PV module, while its rate of temperature change is determined by the module's frame position and material. At lower temperatures, the module generates higher voltages, while higher temperatures lead to a decrease in voltage. This is because rising temperatures cause the semiconductor bandgap to shrink, resulting in a reduction in open circuit voltage, as the p-n junction voltage is temperature-dependent [113]. Consequently, as module temperature rises, PV power output decreases, primarily due to the conversion of incident radiation into heat energy. Optimal module performance occurs at lower temperatures, underscoring the importance of heat dissipation from the PV surface during sunny periods. It's worth noting that while sunlight generates electricity, it also produces heat concurrently,

leading to increased module surface temperatures. Various mechanisms contribute to heat generation alongside electricity production in PV modules [114], as depicted in Figure IV-4.

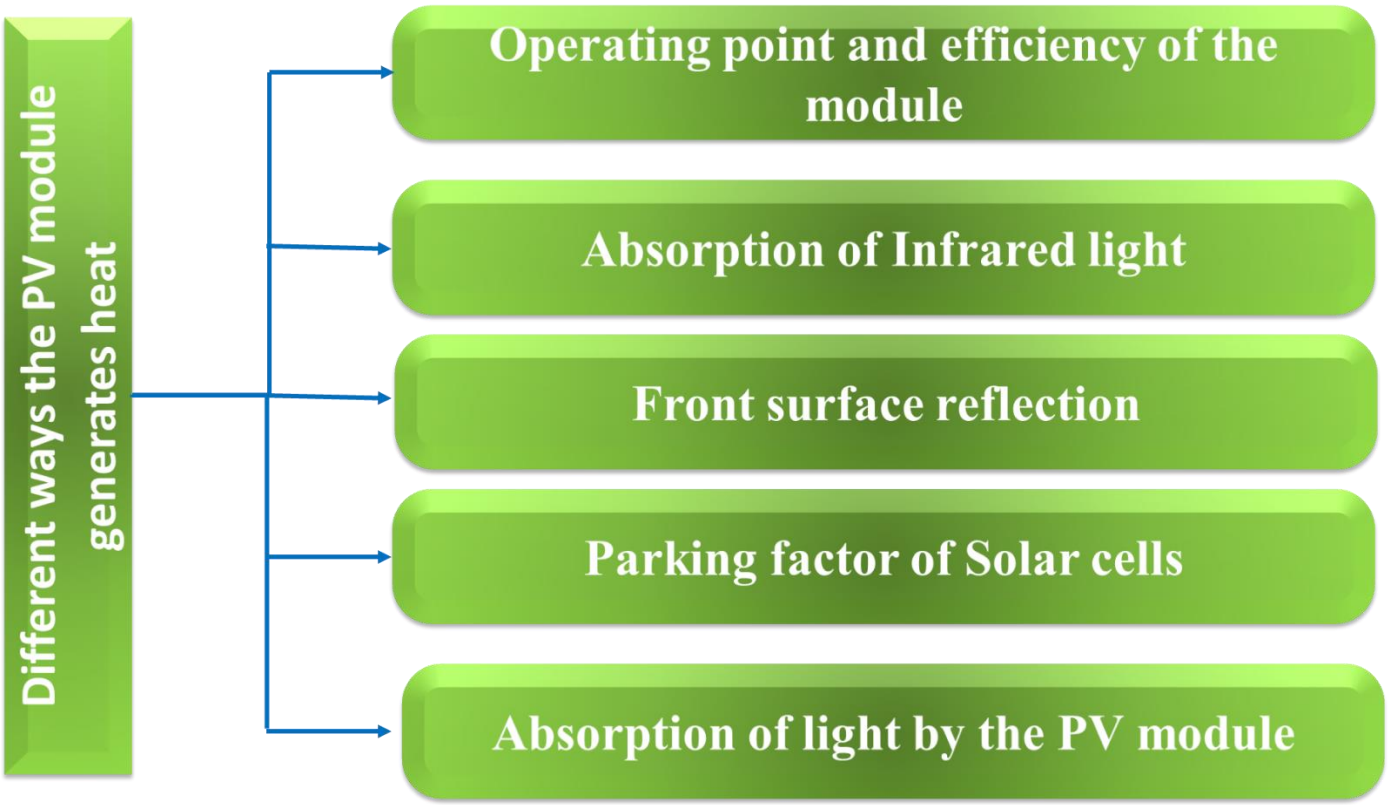


Figure IV-4: Several ways PV module generates heat [94].

a) The operating point and efficiency of the PV module

The quantity of sunlight absorbed by the PV module, which is converted into usable electricity, hinges on both the solar cell's efficiency and its operating point. When operating the PV module to obtain either the open-circuit voltage or short-circuit current, no electricity is generated from the PV module. Instead, the solar radiation absorbed by the PV module under these conditions merely generates heat [94].

b) Absorption of infrared light

The band gap of a semiconductor, as explained in [115], denotes the minimum energy required by an electron to become free. Light hitting the PV module with energy lower than this band gap doesn't directly contribute to electricity generation but heat up the PV solar panel. Typically, infrared light is absorbed by the aluminum at the rear of the panel. However, in PV modules lacking aluminum coverage at the rear, the infrared light passes through the solar cell and exits.

c) Front surface reflection

When sunlight hits a PV module, some of the light is reflected away from the module surface, termed as electrical loss. These reflected rays neither contribute to electricity generation nor add to the module's heating. The maximum temperature rise of the module is determined by the product of the reflection and the incident power [114].

d) Parking factor of the solar cells

Solar cells are designed to efficiently absorb solar radiation. However, during this absorption process, the PV module cells produce a significant amount of heat that exceeds the specified requirement. This generated heat per unit area increases as the parking factor of the solar cells increases [94].

e) Absorption of light by the PV module

The amount of light absorbed by components other than the solar cells in the PV module contributes to the module's heating. This absorption is influenced by factors such as the material composition and colour of the rear surface [94].

IV.2.3.5 Effect of shading on the PV module efficiency

Shading is a significant factor that impacts the power output of PV modules. Clouds, nearby structures like electricity poles, trees, and buildings, as well as other obstructions, can cast shadows on the PV module surface, leading to reduced electricity generation. To optimize performance, PV modules are typically installed facing the sun and positioned to minimize shading, as any obstruction on the module surface can diminish power output and overall efficiency, as noted in [116][87].

IV.2.3.6 Effect of wind direction and wind Speed

Wind speed and direction significantly impact PV module performance. Previous research [112]indicates that while wind direction has a minor effect on PV module temperature, wind speed has a substantial impact, particularly on temperature rise. Additionally, the temperature difference between ambient temperature and PV cell temperature decreases with higher wind speeds. Another study [117] suggests that increasing wind speed enhances PV module efficiency, especially at higher elevations. Furthermore, optimal efficiency is achieved when wind direction is from the north to the rear of the PV module, facilitating improved heat extraction.

IV.2.3.7 Effect of Tilt angle on PV module efficiency

The PV module receives solar radiation from various sources, including direct sunlight, sky radiation, and reflected sunlight from surrounding objects. Optimizing the tilt angle and orientation of

the PV module enhances its efficiency. According to [118], the tilt angle of the PV module is crucial for maximizing solar radiation reception, influenced by factors such as sun position, daily, monthly, and yearly variations. To harness maximum solar power, it's essential to mount the module at the ideal tilt angle for the location. Negative tilt angles position the module's plane towards the pole, while positive angles face the equator. Recent research [110] suggests adjusting the tilt angle based on seasonal changes. Lower tilt angles during summer optimize electricity generation, while higher angles during winter compensate for reduced irradiance. Typically, tilting below the location's latitude by 15° is recommended for summer and tilting 15° above for winter. Given the intermittency of renewable energy sources, wind energy complements solar energy in hybrid systems to stabilize the energy supply.

IV.3 Diagnosis terminology definitions

In the literature on diagnosing faults in photovoltaic (PV) systems, there is inconsistency in the terminology used. Therefore, it is important to offer clear definitions that precisely outline the process of diagnosing faults in PV systems. This aligns with existing references in fault diagnosis, which falls under the domain of automatic control or automation engineering. Before exploring the various faults and diagnostic methods in PV systems, it is crucial to establish a strong grasp of the key terms used in the diagnostic field. This understanding will not only enhance comprehension but also enable effective communication throughout the study.

IV.3.1 Faults

Fault is any deviation between the characteristic observed on the device and the reference characteristic [119], [120].

IV.3.2 Failure

Permanent interruption of the system's ability to perform the requested function with the performance defined in the technical specifications [121].

IV.3.3 Symptom

A distinctive character of an amoral functional state [122].

IV.3.4 Residual

A signal is potentially indicative of faults. it reflects the consistency of the data concerning the behavioural model of the system [123].

IV.3.5 Fault detection

Fault detection involves identifying a deviation in a component or system's operation before it evolves into a failure. This initial phase is paramount in fault diagnosis since the accuracy of all subsequent steps relies on it [124]–[126].

IV.3.6 Fault isolation

Following the fault detection procedure, fault isolation involves identifying the nature, position, and timing of fault detection. Its objective, for maintainability reasons, is to pinpoint the fault location at the most basic level of the replaceable component [125], [126].

IV.3.7 Fault identification (classification)

Following the fault detection procedure, fault identification or classification involves attributing the fault to its root cause and assessing its magnitude. The objective of fault identification is to pinpoint the primary failure mode [125], [126].

IV.3.8 Fault diagnosis

Fault diagnosis involves determining the type, location, and magnitude of fault detection. It encompasses both fault detection and fault identification. While fault diagnosis methods may vary across different engineering disciplines, their phases generally remain consistent [124], [126].

IV.4 Faults in PV systems

PV systems are often susceptible to various failures, which can substantially impact their performance by reducing efficiency. The flowchart depicted in Figure IV-5 illustrates the different types of faults typically encountered in PV systems. As indicated in this figure, faults in PV systems primarily fall into two categories: faults on the direct current (DC) side and faults on the alternating current (AC) side. Subsequently, a comprehensive description of the most frequently occurring faults in PV systems is provided [127].

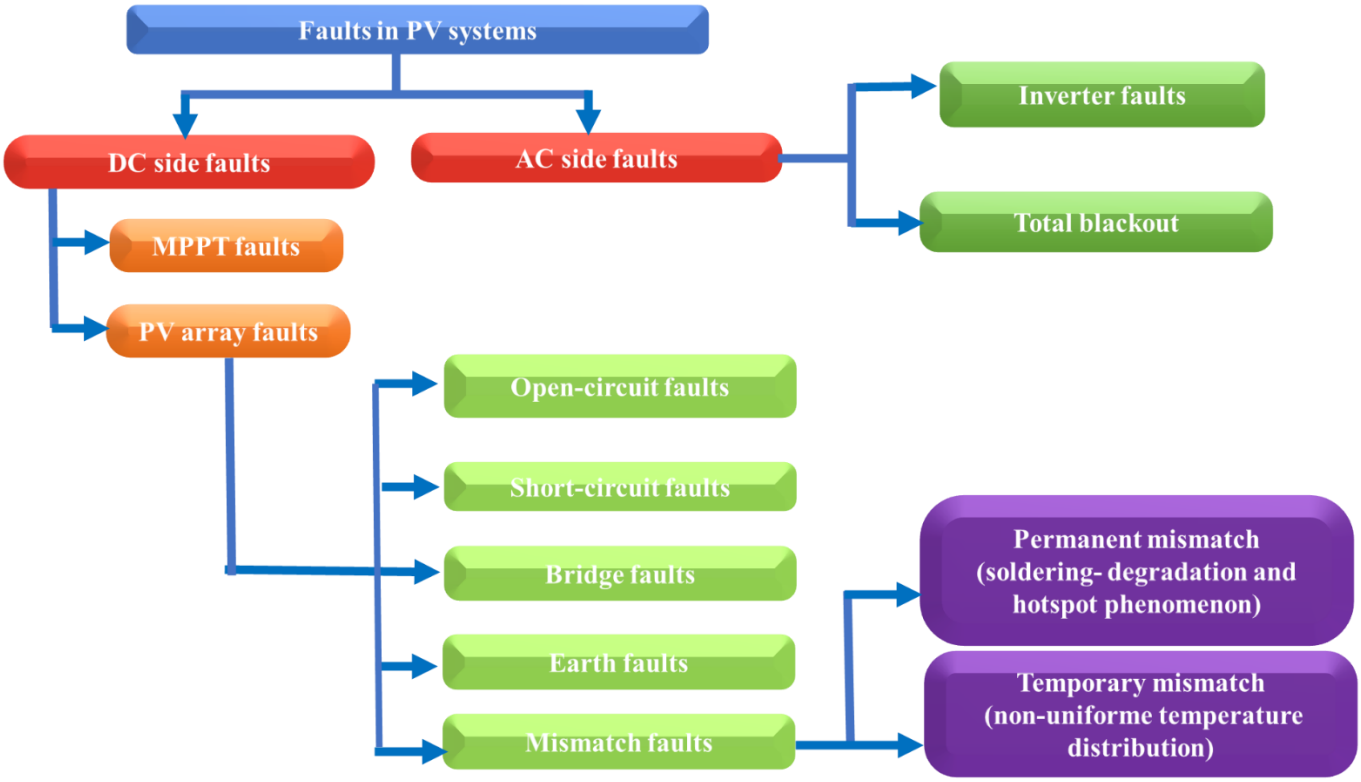


Figure IV-5: Types of faults usually occurred in PV systems [127].

IV.4.1 DC side faults

Typically, the primary DC side faults in PV systems are attributed to faults in PV arrays and MPPT controllers. Detecting faults in this initial category is frequently essential to prevent energy loss, system shutdowns, and potentially catastrophic fires. Subsequent subsections provide detailed explanations of DC side faults, including their primary causes and associated impacts [127].

IV.4.1.1 PV array faults

Throughout its operation, a PV array may encounter five significant types of faults: mismatch faults, open-circuit faults, earth faults, bridge faults (Line-Line faults), and short-circuit faults [127].

a) Mismatch faults

A mismatch fault occurs when multiple solar cells or PV modules are connected, each with varying electrical parameters or exposed to different operating conditions, such as temperature and irradiance [127].

This type of fault poses a significant challenge due to its difficulty in detection and its considerable impact on power reduction. Mismatch faults are typically divided into two main categories: temporary and permanent mismatch faults [127].

1) Temporary mismatch faults

These faults stem from the uneven temperature distribution caused by snow cover. Specifically, during the winter season, snow accumulation on the surface of PV modules can lead to uneven temperature distribution, significantly reducing the energy production of the PV system [127].

2) Permanent mismatch faults

This subset of mismatch faults primarily includes three types of faults: degradation fault, soldering fault, and hotspot phenomenon [127].

- **Degradation fault**

The degradation of PV arrays primarily results from the aging of PV cells. This occurs when the series resistance between PV cells increases over time, often due to diminished contact adherence or corrosion induced by water vapor [127].

- **Soldering fault**

This fault arises when the solder connection between solar cells and their connecting ribbons is unintentionally severed [127].

- **Hotspot phenomenon**

The hotspot phenomenon arises when one solar cell within a series connection receives less incident light (Iph) compared to the others, often due to partial shading. In this scenario, the affected cell becomes reverse-biased, leading to power dissipation rather than generation. This imbalance can cause a notable temperature increase in the cell, potentially resulting in damage [127].

b) Open-circuit faults

This fault category arises when the series connection between the current-carrying path and the load is inadvertently severed, opened, or compromised [127].
A comparable scenario can occur from inadequate connections between solar cells or the rupture of cables between PV modules [127].

c) Earth faults

This category of fault is often regarded as one of the most common issues encountered in PV arrays. An earth fault can occur when an unintended connection to the ground is inadvertently established [127].

d) Bridge faults (Line-Line faults)

Bridge fault, also referred to as Line-Line fault, arises when a low-resistance connection is established between two points of different potential within distinct PV strings. Corrosion, mechanical damage, and insulation failure in cables are the main sources of this type of fault [127].

e) Short-circuit faults

This fault primarily impacts solar cells, bypass diodes, or PV modules, typically resulting from water infiltration into the modules or poor wiring between modules and inverters. Such a fault may occur due to an inadvertent connection between two points of varying voltages within a single string or PV module [127].

IV.4.1.2 MPPT faults

The objective of the MPPT controller (DC/DC converters) is to capture the Maximum Power Point (MPP) of the PV array. However, in practical PV system operations, ineffective MPP tracking may occur, particularly when encountering partial shading conditions [127].

IV.4.2 AC side faults

In the AC side of PV systems, two primary categories of faults may occur: inverter faults and complete blackout [127].

IV.4.2.1 Inverter faults

The failure of each part of the inverter (transformer, capacitors, IGBTs, and drive circuitry) leads to the inverter faults [127].

IV.4.2.2 Total blackout

This fault category entails a sustained loss of power, typically stemming from accidents or natural disasters. In such scenarios, the inverter promptly ceases the supply of power from the PV generator to the distributed grid [127].

IV.5 Fault detection and diagnosis

The occurrence of any fault within PV systems can lead to a notable reduction in system performance and security. To tackle these challenges, a multitude of fault detection and diagnosis (FDD) methods have been introduced in recent literature. These methods typically rely on various data types, such as electrical measurements, environmental data, or images of photovoltaic panels. Based on the type of data utilized, fault detection and categorization techniques in photovoltaic systems can

be broadly classified into two categories: non-electrical methods, encompassing visual and thermal methods (VTMs), and electrical methods [128].

IV.5.1 Visual and thermal methods for detection of faults and diagnosis

These visual and thermal techniques have the capability to detect and pinpoint various types of faults, regardless of whether they are permanent or temporary. These faults include discoloration, browning, surface soiling, hot spots, breaking, and delamination. However, employing these methods requires sophisticated equipment, such as thermal cameras and drones. The visual and thermal methods encompass visual inspection, ultraviolet (UV) fluorescence (FL) imaging, electroluminescence imaging, and thermography [128][9].

IV.5.1.1 Ultraviolet (UV) fluorescence (FL) methods

The UV fluorescence technique, also known as FL imaging, is applied to examine the discoloration of photovoltaic (PV) modules due to ethylene vinyl acetate (EVA), as demonstrated in Figure IV-6. It can identify the quantity and positioning of cell cracks within PV modules even in low-light outdoor conditions, although it may not detect cracks along the cell's perimeter [129][130].

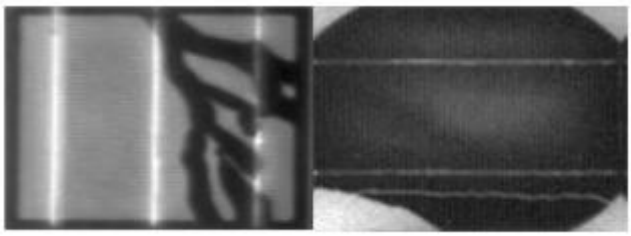


Figure IV-6: UV FL failure images [128].

IV.5.1.2 Electroluminescence (EL) methods

The electroluminescence (EL) technique stands out as a prominent method among Visual and Thermal Monitoring (VTM) strategies, allowing for the examination of PV modules or cells and failure detection through EL imaging, as illustrated in Figure IV-7. By providing PV modules with a DC current to stimulate radiative recombination within the solar cells, electroluminescence emission is captured using a charged silicon camera (CCD), which is readily available commercially [131][132][130].

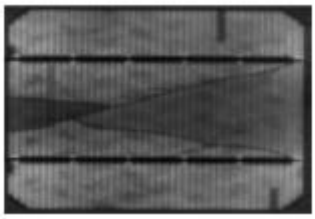


Figure IV-7: EL failure image example [128].

IV.5.1.3 Infrared thermography

Infrared (IR) thermal imaging stands out as a crucial non-destructive and contactless method for detecting failures. Essentially, this technique relies on the emission of infrared waves from the surface of PV systems or their electrical components, which occurs when these surfaces release energy in the form of electromagnetic waves. These infrared waves are produced by the movement of atoms within any object with a temperature above 0 Kelvin or when the object receives external energy. Thermography, enabled by infrared imaging, proves invaluable for localizing and categorizing failures in PV modules, as well as in various system components like cabling, diodes, DC box combiners, junction boxes, and connectors. Infrared thermography (IRTG) is widely favoured due to its ability to provide swift, reliable, precise, cost-effective, and 2D visualizations of key features within PV modules. Figure IV-8 showcases two distinct thermography techniques employed for PV module failure detection: active IRTG and passive IRTG [133].

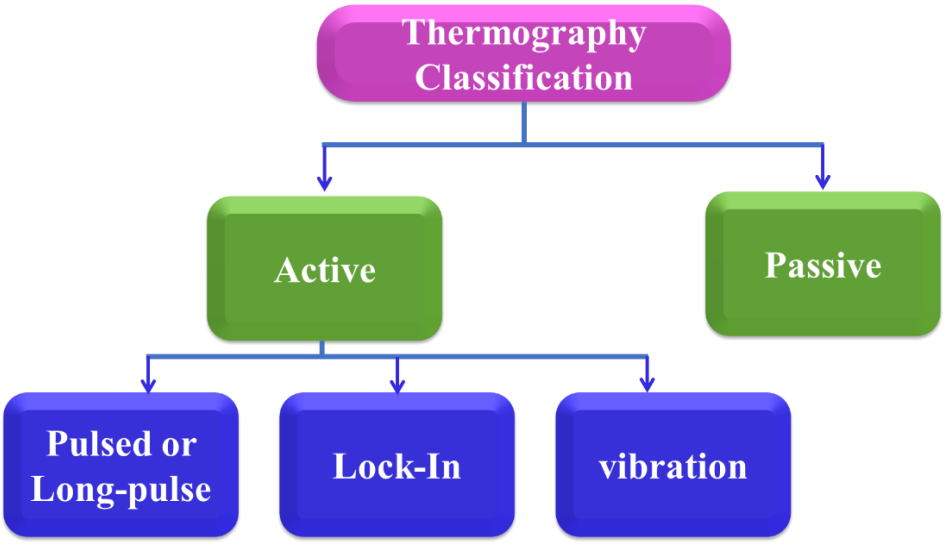


Figure IV-8: Thermography techniques [128].

a) Active IRTG

Active thermal imaging involves inducing an internal heat flow within an object using an external heat source, thereby elevating its temperature. Pulsed thermography, a subtype of active infrared

thermography (IRTG), represents a rapid and straightforward approach to this process. Typically, it involves applying a heat pulse to the object using a heat source like a lamp or a heating gun [133].

In long-pulse thermography (TG), a continuous low-power heating source is administered to PV modules, primarily focusing on cooling processes. On the other hand, lock-in thermography (TG) utilizes an oscillating temperature domain to heat the object, enabling the detection of internal failures by monitoring changes in thermal wave patterns [134].

Vibro-thermography employs mechanical vibrations to generate thermal energy within PV modules, thereby highlighting hot spots indicative of defects such as cracks and delamination [135][136].

b) Passive IRTG

The passive infrared thermography (IRTG) method, also known as "thermography under steady-state conditions," operates without the need for external heat sources; instead, it captures infrared radiation emitted by PV modules. This passive thermography approach is widely adopted due to its simplicity and cost-effectiveness, as it solely relies on an infrared camera [133]. It facilitates fault detection without physical contact with the object or the need for additional hardware or human intervention.

Offering real-time imaging, passive IRTG enables simultaneous observation and recording of images, minimizing data inaccuracies. Researchers in [137] propose a solution for PV fault detection utilizing a deep learning approach and a dataset of thermal images to detect cell instances, rendering the algorithm suitable for automated inspection tasks.

However, the application of IRTG has its limitations. It necessitates high-quality cameras, often at a considerable expense, to ensure precise measurements and prevent errors. Moreover, skilled operators with adequate training are essential for the correct utilization of these cameras. Conducting a preliminary study to determine the optimal altitude is also crucial to enhance image resolution [133].

IV.5.2 Electrical methods for detection of fault and diagnosis

Electrical methods are instrumental in the detection and diagnosis of various faults in photovoltaic (PV) systems. These methods can identify issues in PV modules, strings and arrays, inverter switches, batteries, arc faults, grounding faults, diode faults, and hotspots. The primary electrical-based fault detection and diagnosis (FDD) techniques depend on specific models of PV plants to recognize different types of faults. Broadly, electrical methods for fault detection can be categorized into three main groups [9]:

IV.5.2.1 Statistical and signal processing approaches

Statistical and signal processing methodologies primarily rely on analysing the waveform of signals (refer to Figure IV-10). Offline techniques such as time domain reflectometry (TDR), spread spectrum time-domain reflectometry (SSTDR), and earth capacitance measurement (ECM) are commonly employed. While these methods are effective in detecting and localizing faults, they necessitate an external generator function and require the system to be offline during diagnosis (as depicted in Figure IV-9). Typically used in small-scale PV plants, these techniques are not automated and are unable to identify the type or source of faults. However, in modern monitoring and supervision practices, automatic and online fault diagnosis methods are highly recommended [9].

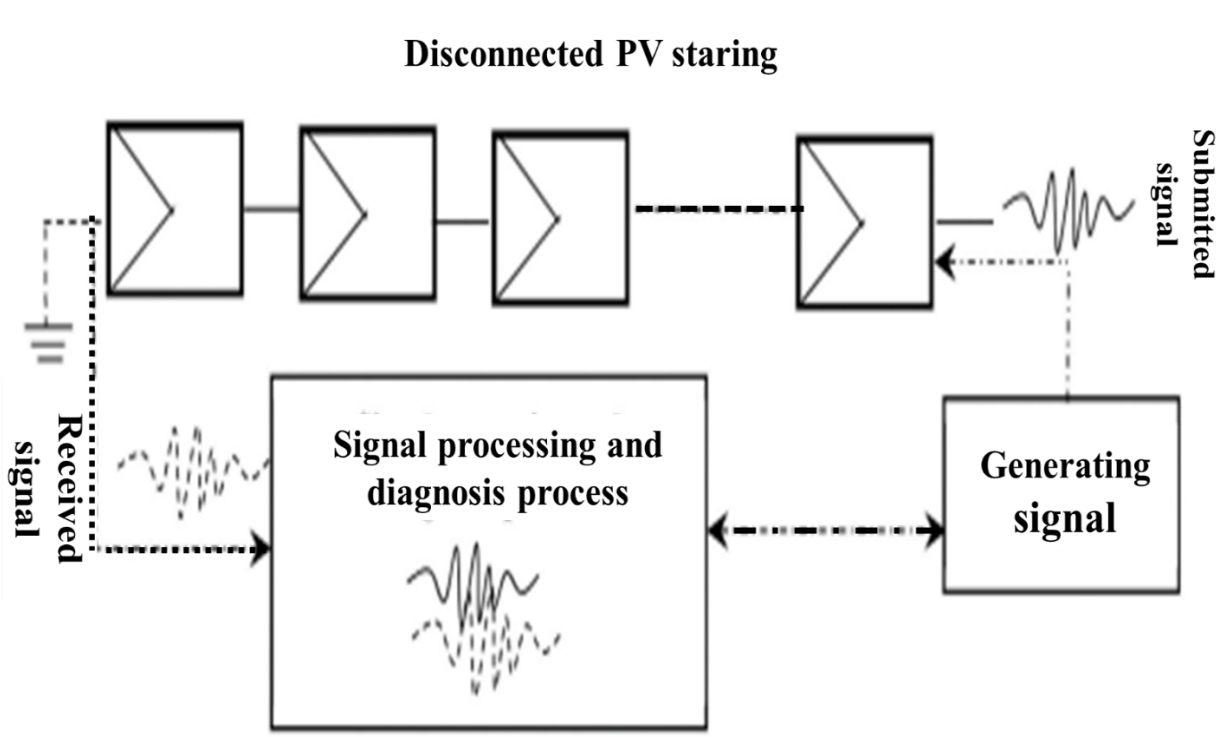


Figure IV-9: TDR-based method for FD in a disconnected PV string [9].

IV.5.2.2 Methods-based I-V characteristics analysis

I–V curve analysis serves as a conventional method for Fault Detection and Diagnosis (FDD), leveraging the electrical characteristics of a module to reveal crucial parameters like short-circuit current and open-circuit voltage that can indicate system failures. This analysis involves monitoring and measuring the current-voltage curve as the voltage or current across the module changes under the influence of an external electronic load or power source. Typically, the response characteristics of cells or modules identical to those under test are used as references for comparison [130][138].

Under normal operation, the I–V characteristics adhere to a specific curve, as depicted in Figure IV-10, but deviations from this curve occur in the event of a fault, with the extent of deviation influenced

by the type and severity of the fault. While the I–V curve of a module can be instrumental in detecting various faults, it falls short in pinpointing the precise location of these failures. Consequently, complementary techniques are often required to locate faults, a process that can be resource-intensive in terms of time and cost [133][139].

To streamline the monitoring and maintenance of systems and reduce operational costs and time, visual and automatic anomaly classification methods offer a promising solution. These methods enable simpler system monitoring and maintenance while also saving time and resources [133][140].

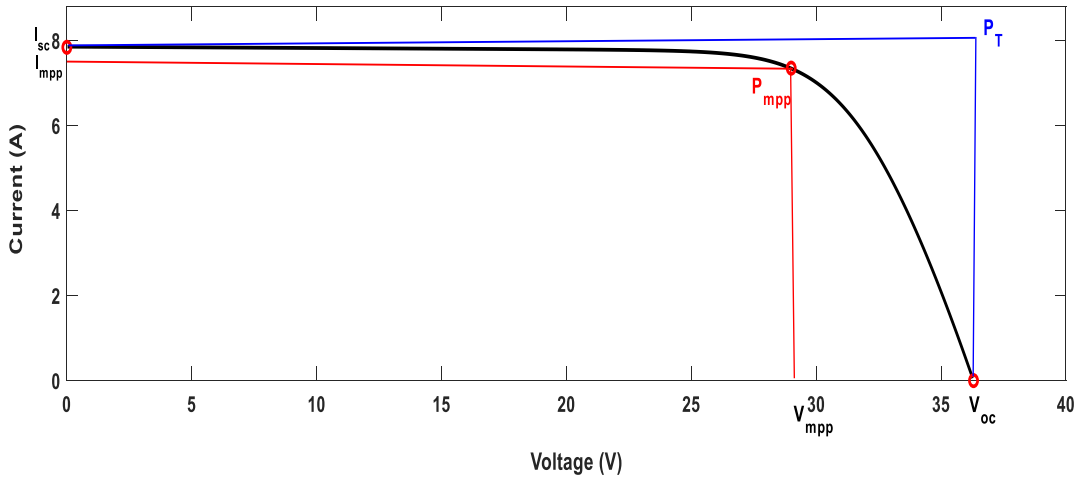


Figure IV-10: I-V curve parameters. Isc, Voc, Pmpp [128].

IV.5.2.3 Methods based on artificial intelligence techniques

These techniques encompass the utilization of machine learning and deep learning. Machine learning primarily finds application in methods reliant on the analysis of measured I–V curves, currents, voltages, and other parameters, whereas deep learning is primarily employed in fault diagnosis based on infrared and electroluminescent images. To effectively deploy these approaches, a substantial database is essential for training and evaluating classification and recognition models. Figure IV-11 illustrates the operational principle of these methodologies. AI-based methods are predominantly employed to detect, classify, and ascertain the nature of faults, effectively addressing issues arising from faults exhibiting similar signatures or symptoms [9].

Approaches utilizing infrared (IR) thermography images can also incorporate machine learning techniques; however, deep learning proves to be the most optimal given the availability of a substantial database. Moreover, the accuracy of classification relies on factors such as the volume and quality of data, alongside the algorithms employed. Raw data, lacking pre-processing such as noise removal, can significantly compromise the efficacy of fault detection and diagnosis methods [9].

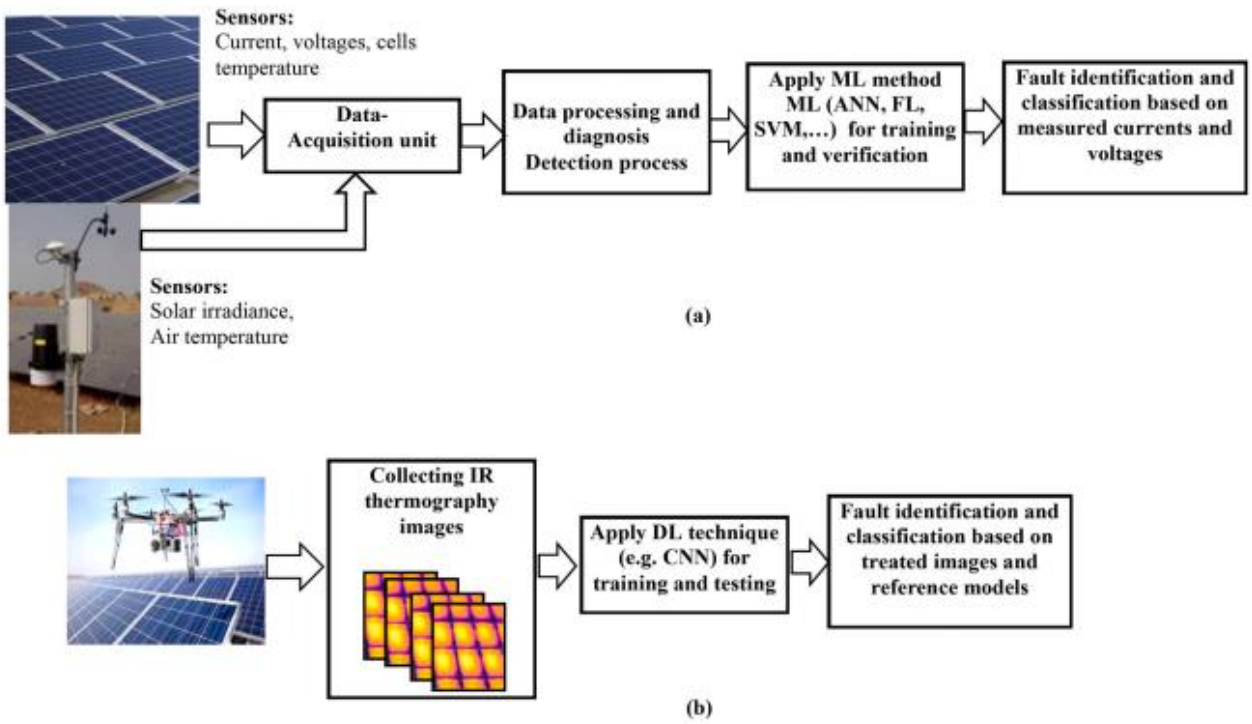


Figure IV-11: Simplified framework of fault identification and classification based on AI-techniques: a) I–V based machine learning approach and b) thermography based [9].

IV.6 Conclusion

In conclusion, this chapter has provided a comprehensive overview of various factors affecting the efficiency of PV modules, definitions of diagnostic terminology, common faults in PV systems, and methods of fault detection and diagnosis. Understanding these aspects is essential to guarantee the optimal reliability of photovoltaic systems. By addressing factors such as shading, wind direction, tilt angle, and module degradation, practitioners can maximize the efficiency and longevity of these systems. The presentation of fault detection and diagnosis methods highlights the importance of employing appropriate techniques to identify and classify faults that may occur in photovoltaic systems.

Chapter V : PV fault detection and classification using Voting-Ensemble Learning

V.1 Introduction

In the realm of renewable energy, photovoltaic (PV) systems stand as beacons of hope, converting sunlight into electricity. However, these systems are not immune to faults, which can arise due to environmental factors, component deterioration, or installation errors. Timely and accurate detection and classification of these faults are crucial for maintaining optimal performance, ensuring safety, and minimizing maintenance costs.

In the preceding chapter, we explored various kind of PV faults and the traditional methods employed for their detection, including electrical and non-electrical approaches. Among the non-electrical techniques, artificial intelligence (AI) has emerged as a promising tool, offering the potential to overcome the limitations of traditional methods. AI encompasses a broad spectrum of algorithms, including machine learning (ML) and deep learning (DL) techniques. Machine learning algorithms, such as decision trees, support vector machines, and k-nearest neighbors, have demonstrated effectiveness in learning patterns from historical data to make predictions. Deep learning has garnered significant interest lately owing to its ability to handle intricate datasets and advancements in computing power. Various artificial neural network (ANN) architectures, ranging from basic models to sophisticated ones like auto-encoders, self-organizing maps, and long short term memory (LSTM) networks, have proven effective in predicting the output of photovoltaic (PV) power plants [129].

In this chapter, a brief application and discussion about the use of parameter identification for fault detection is presented before focusing on an extensive application of AI techniques for detecting and diagnosing faults in PV arrays. Special attention will be dedicated to the ensemble learning, a powerful AI technique. This is because ensemble learning combines the strengths of multiple AI models, enhancing accuracy and robustness in complex fault scenarios. By aggregating the predictions of various ML and DL models, ensemble methods can mitigate individual model weaknesses and deliver superior performance, making them well-suited for the intricate task of PV fault diagnosis.

V.2 PV array fault diagnosis using BB-TLBO algorithm

The previous chapters discussed the importance of accurately identifying parameters when creating models for PV arrays. These parameters are essential for simulating a reliable PV array model, which can help identify and diagnose deviations from expected behavior, indicating potential faults. This method is ideal for model-based PV fault detection and diagnosis techniques. However, the choice between model-based or model-free methods depends on various factors, such as the availability of accurate datasets.

After applying the BB-TLBO algorithm to identify the parameters of RTC France silicon solar cell and the STM6-40/36 module and the results show its efficiency and its robustness. In this context, BB-TLBO is used to detect faults that may occur in the PV array. In this case, parameter identification utilizes the I-V data representing the corresponding healthy or faulty scenario. The resulting parameters, as well as the RMSE values, will be used to classify Faults.

V.3 Tested PV array

As shown in Figure V-1, the PV array setup configuration is created in the Matlab/Simulink environment by connecting four (04) PV modules in series. The modules used are SunPower SPR-X20-250-BLK model.

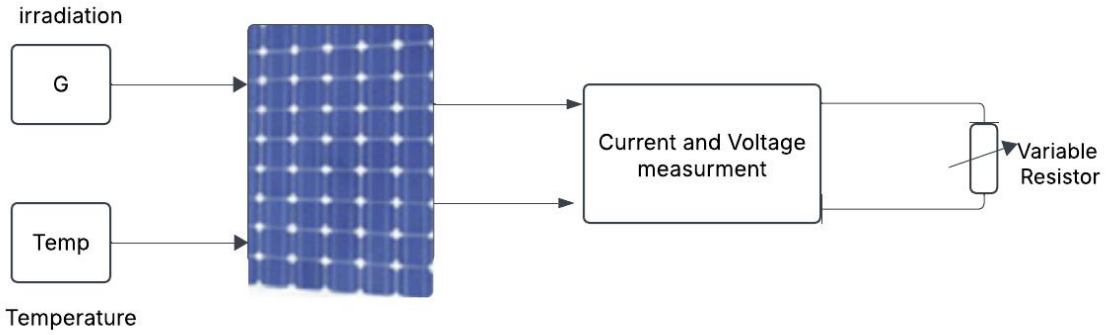


Figure V-1: Schematic diagram or PV characterization.

V.4 Examined faulty scenarios

V.4.1 Fault scenario n° 01 (Partial shading)

As described in Figure V-2, fault scenario n° 01 is selected when the photovoltaic array is exposed to an irradiance level of G=1000W/m² with a temperature of T=25°C. This specific fault condition can be outlined as follows:

- Modules 01 and 02 are operating under normal conditions (fully functional module).
- Modules 03 and 04 are partially shaded (shade covers 50% of each module's surface).

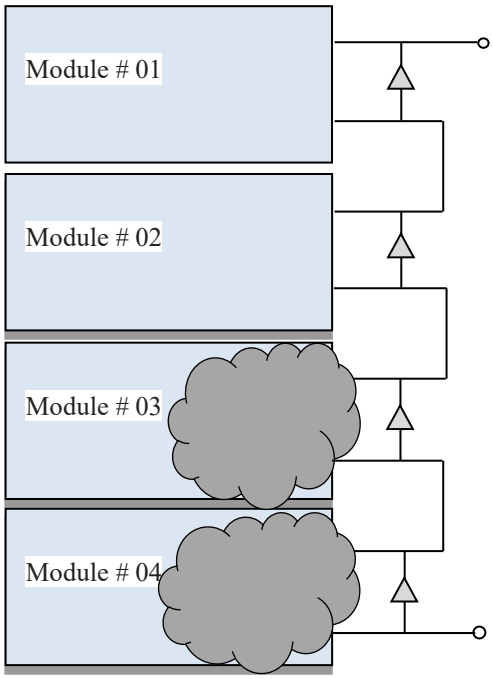


Figure V-2: Fault scenario n° 01: Partial shading.

V.4.2 Fault scenario n° 02 (Short-circuit fault of module 02)

The short-circuit fault as chosen in scenario n°02 is described in Figure V-3. This scenario is when the photovoltaic array is exposed to an irradiance level of G=1000W/m² with a temperature of T=45°C. This faulty situation is defined as follows:

- Modules 01, 03 and 04 are operating under normal conditions (fully functional modules).
- Module 02 is short-circuited using a parallel resistance R01=$10^{-6}\Omega$ (faulty module).

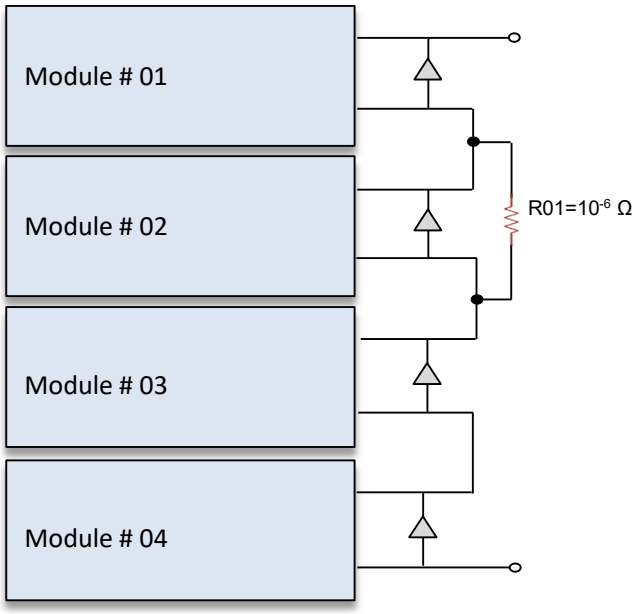


Figure V-3: Fault scenario n° 02: Short-circuit fault on module # 02.

V.4.3 Fault scenario n° 03 (Short-circuit fault of module 02 and 03)

As shown in Figure V-4, in this scenario both module 02 and module 03 are exposed to a short circuit. This occurs when the photovoltaic array is exposed to conditions of irradiance G=1000W/m² and temperature T=45°C. This faulty situation is defined as follows:

- Modules 01 and 04 are operating normally.
- Module 02 and module 03 are short-circuited using a parallel resistance of R01=10^{-06} Ω.

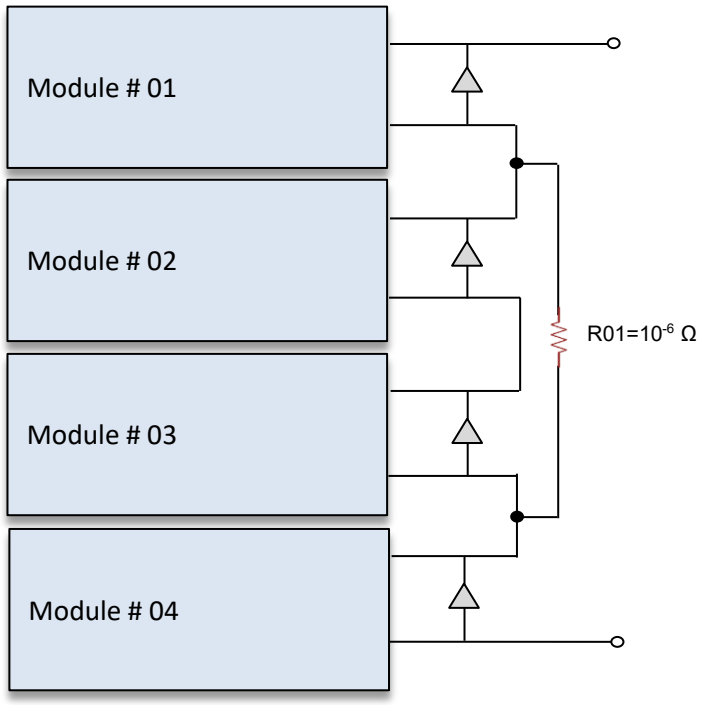


Figure V-4: Fault scenario n° 03: Short-circuit fault on module # 02 and module #03.

In this section, the proposed PV array fault detection and diagnosis is performed based on the analysis of the data resulting from the parameter identification using the BB-TLBO algorithm of each scenario including the healthy one. Table V-1 presents the lower and upper bounds of the search space for each parameter. In order to evaluate the performance of the PV array, it is crucial to gather I-V data for various scenarios, including both healthy and faulty ones.

Table V-1: Upper and lower limits of the unknown parameters

Parameters	*Lower Bound*	*Upper Bound*
$I_{ph}(A)$	0	10
$I_{sd}(A)$	0	1E-04
$R_s(\Omega)$	0	5
$R_{sh}(\Omega)$	0	5000
n	0	100

V.5 Parameter identification and fault detection and diagnosis:

The utilization of the BB-TLBO algorithm yields the parameters of the PV array model for the four (04) cases, as detailed in Table V-2.

Table V-2: Presents the outcomes of the PV array 1DM parameter identification process

Parameters / Scenarios	I_{ph}	I_{s1}	R_s	R_{sh}	n	*RMSE*
Healthy case	6,294112462	4,34042874E-09	1,203654215	2483,2242	1,136507789	**9,64985341260E-05**
One module short-circuited	6,294116628	4,33595694E-09	0,902655661	1862,0203	0,852347741	**1,11417074179E-04**
Partial shaded String	6,735239366	1,80998258E-23	1,124991551	41,3943	0,221291688	**2,41655225481E-01**
02 module short-circuited	6,294126010	4,32558989E-09	0,601677965	1240,7615	0,568178676	**1,37364573363E-04**

After analyzing the obtained parameters and the RMSE value for each scenario, it appears that I_{ph} and I_{sd} do not provide significant information for scenario classification. In contrast, R_s, R_{sh}, and n, along with RMSE, when combined, provide useful information for fault classification. Therefore, the four (04) scenarios can be classified as reported in Table V-3.

Table V-3: Classification of fault scenarios according to their obtained parameters

Parameters	Scenario
n = 1.13 and *RMSE* = 9.64E-05	Healthy PV array
n = 0,85 and *RMSE* = 1,11E-04	One module short-circuited
R_{sh} = 41,39 and *RMSE* = 2,41E-01	Partial shaded String
R_s = 0.60 and *RMSE* = 1,37E-04	02 module short-circuited

It is important to recognize that this method has limitations. More data would be needed for each scenario under study to obtain reliable results. In other words, multiple tests in different conditions are needed for each scenario, including the healthy one. The tests should include different conditions of irradiance and temperature to confirm the fault signature in each scenario. This ensures that identified parameters correspond to the given scenario to merit generalization.

Furthermore, visual screening of the resulting data corresponding to the patterns that emerge for each scenario may not be readily apparent through simple observation. Therefore, it is essential to use advanced tools such as artificial intelligence (AI) algorithms (machine learning, deep learning...),

given the large volume of data being analyzed. Additionally, the application of AI algorithms can enhance the identification of subtle trends and anomalies that might be overlooked by a human analyst. These advanced techniques can process complex datasets more efficiently and uncover multilayered relationships within the data that are crucial for interpreting the outcomes of the various tests. By employing techniques such as clustering and classification, AI can help discern underlying patterns that may indicate specific fault signatures or operational inefficiencies.

Utilizing AI enables iterative learning, allowing models to be continually refined as new data becomes available. This corresponds thoroughly with the situation of the scenarios being tested in different weather conditions as stated above, which can improve the predictive capabilities associated with each assessed scenario. Such adaptability is especially important in dynamic environments where conditions can change significantly over time. The ability to optimize models in real time can lead to quicker and more accurate diagnostics.

It is also essential to consider additional variables that may affect outcomes, such as the system's age, environmental factors, and historical performance data. Understanding these multidimensional aspects contributes to a more comprehensive view of the scenarios under evaluation. Therefore, establishing a feedback loop where insights gained from AI applications inform future experimental designs is crucial. This cyclical approach not only enhances the reliability of the findings but also advances methodologies in the field.

V.6 Artificial intelligence related work

Many research papers on PV fault detection that use artificial intelligence, have been conducted using different methodologies, such as: In a study referenced as [141], a probabilistic neural network (PNN) algorithm is utilized to analyze PV arrays with blocking diodes, focusing on assessing line-to-line and line-to-ground faults. Nevertheless, this method overlooks faults characterized by high impedance, which are common in real PV arrays. The applicability of this approach to practical PV systems remains unverified, and its implementation is intricate and costly.

In [142], an approach for fault detection is suggested, utilizing an SVM classifier to detect short-circuit faults. Nevertheless, the necessity for multiple filters renders this method expensive.

Chen et al. [143] devised an optimized kernel extreme learning machine (KELM) method for diagnosing various faults including open circuit, short circuit, partial shading, and degradation. While demonstrating high performance in both experimental and simulated settings. This approach is characterized by longer computational times compared to alternative techniques.

In reference [144], researchers have developed multilayer perceptron neural networks capable of detecting PV faults with success. Nonetheless, this investigation warrants significant attention due to

the complexities involved in feature extraction from wavelets and the identification of hidden layers within neural network architectures.

In their work referenced as [145], the authors crafted a deep residual network tailored for PV fault detection. Demonstrating commendable classification accuracy across both simulated and experimental scenarios, the model showcases promise. Yet, it's worth noting that this model solely identifies line-to-line and open circuit faults, potentially limiting its applicability to broader fault detection contexts.

An artificial neural network (ANN) has been integrated into a monitoring system for photovoltaic (PV) panels to forecast output power and detect degradation. However, the ANN model's effectiveness is compromised in the presence of noisy data [146].

In [147], an intricate approach involving CNN, AlexNet, and SVM is explored. Regrettably, the outcomes yielded relatively low accuracy, ranging between 69.39 and 73.53 percent.

Justin et al. [148] investigated the use of bagging and stacking-based ensemble learning (EL) methods to detect anomalies in energy production caused by photovoltaic (PV) faults. Their findings revealed an accuracy of 79.50% for the bagging-based EL and 94.00% for the stack-based EL.

Eskandari et al. [149] introduced the EL approach for detecting line-line faults, employing a combination of three algorithms: Support Vector Machine (SVM), Naïve Bayes (NB), and k-Nearest Neighbor (KNN). Across various scenarios, these classifiers demonstrated strong performance, achieving accuracies of 99.00% and 99.50%.

Dhibi et al. [150] introduced the EL method, incorporating SVM, KNN, and DT. They also devised two types of Principal Component Analyses (PCAs): kernel PCA (KPCA) and reduced kernel PCA (RKPCA). The objective behind PCA development was to enhance performance and reduce computational burden. Despite the study's remarkable performance, executing EL remained time-consuming. Specifically, the EL method based on PCAs required 110.36 seconds for RKPCA and 221.85 seconds for KPCA.

Kapucu et al. [151] explored the effectiveness of the QDA-ETent-DT classifier, consisting of quadratic discriminant analysis, extra trees with entropy, and decision tree components, which demonstrated superior performance compared to similar classifiers. When utilized to detect two photovoltaic (PV) faults—partial shading and short circuit—the classifier achieved an accuracy rate of 97.46% before optimization and 97.67% after optimization.

In reference [152], an AdaBoost ensemble approach is suggested for diagnosing PV array faults. Impressively, it achieved a detection accuracy of approximately 97.58%. However, it's important to note that this method solely identifies open circuits, short circuits, and degradation faults, potentially constraining its utility in identifying a broader range of faults.

The study referenced as [153] implemented ensemble learning (EL) with a voting mechanism, incorporating linear regression (LR), support vector machine (SVM), and decision tree (DT) models to detect and isolate faults in a solar energy system. Their findings indicate that the EL approach they proposed outperforms alternative methods in terms of specific performance metrics such as precision and F1-score.

In [154], researchers devised an ensemble learning (EL) strategy for detecting and classifying faults in photovoltaic (PV) arrays. They explored a principal component analysis (PCA) algorithm for feature selection and employed an LSTM algorithm to forecast the system's power. They also utilized a category boosting (CatBoost) algorithm for fault detection and classification. Their approach yielded accuracy rates exceeding 98% for both fault detection and classification.

Adhya et al. [155] demonstrated strong performance, achieving an accuracy of approximately 99.00% in diagnosing PV faults by employing Ensemble Learning (EL) with Light Gradient Boosting Method (LGBM), Categorical Boosting (CatBoost), and Extreme Gradient Boosting (XGBoost). Nevertheless, the study lacked comprehensive discussion on the feature selection process and did not specify the type of ensemble learning used.

This research [11] introduces a novel DSEL (Deep Stack-based Ensemble Learning) approach for diagnosing faults in PV arrays. It combines deep neural network, LSTM, and bidirectional LSTM models with logistic regression for fault diagnosis, covering various PV faults and incorporating the MPPT algorithm. The DSEL algorithm offers robust fault diagnostics for both noiseless and noisy data, outperforming prior methods with 98.62% accuracy for noiseless data and 94.87% for noisy data. It enhances fault detection while improving prediction accuracy, providing more efficient and reliable fault categorization in PV arrays.

V.7 Ensemble learning

The ensemble learning (EL) technique combines multiple base learner algorithms to construct an optimal prediction model that significantly surpasses the performance of individual base learners. The purpose of using ensemble models is to minimize generalization error. This technique effectively reduces prediction error, particularly when the base models are diverse and independent. Among the primary categories of ensemble learning algorithms are bagging, boosting, stacking, and voting [159][160]. These are described below:

V.7.1 Voting

The Voting Classifier ensemble approach combines predictions from multiple independent models, or base estimators, to make a final prediction by leveraging the "wisdom of the crowd" concept. This method uses two types of voting: hard voting, where each model makes a prediction and the most

common prediction is chosen, and soft voting, where each model predicts the probability or confidence for each class, and the class with the highest average probability is selected. Weighted voting further refines this approach by allowing different models to have varying influences on the final prediction based on their performance. This strategy enhances accuracy, robustness, and model diversity while reducing bias and variance, making it a versatile and effective tool for improving predictive performance in machine learning [156]–[159] [160].

V.7.2 Bagging

Bagging, which stands for Bootstrap Aggregating, is an ensemble learning method designed to enhance model performance by minimizing variance and reducing overfitting. This technique involves combining predictions from multiple decision trees, each trained on different samples drawn from the same dataset. These samples are created using bootstrapping, a method that resamples the training data with replacement to form several subsets of the same size as the original dataset. By training decision trees on these varied subsets and merging their predictions, bagging produces a more robust model that is less affected by fluctuations in the training data. Advanced decision tree variants, like Random Forests and Extra Trees, utilize bagging to improve accuracy by considering subsets of both training samples and features. While this method may increase computational complexity, it is highly parallelizable, allowing for significantly reduced training times with the right hardware. However, bagging is generally not applied to deep learning models due to their longer training durations [161]–[164].

IV.7.1 Stacking

Stacking is an ensemble learning method that combines multiple base-level classifiers with a metalearner classifier to improve prediction accuracy. In stacking, the original data is fed into several distinct models, and the predictions from these models are then used as input for the metaclassifier. This metaclassifier estimates the final output by considering both the inputs and the outputs of each base model, assigning weights to optimize performance. Unlike simple averaging or majority voting, stacking leverages the strengths of various learning methods by integrating their predictions in a sophisticated manner. Advanced techniques in stacking include Multilinear Response (MLR) and Probability Distribution (PD) stacking. The process involves dividing the dataset into n sections for n-fold cross-validation, using these subsets to train and test various models. The predictions from these models serve as metadata for the metamodel, which produces the final forecast, often referred to as the winner-takes-all approach. Stacking, also known as mixing or model blending when involving a linear model for the final decision, effectively uses a metalearning model to combine the outputs of base models, enhancing overall predictive performance [165]–[169].

IV.7.2 Boosting

Boosting is an ensemble learning technique that transforms weak learners into strong ones by using weighted averages. Unlike bagging, boosting involves training a sequence of models where each subsequent model focuses on the errors made by its predecessor. The original dataset is partitioned into several subgroups, with each subset used to train a weak classifier. New subsets are created from the misclassified elements of the prior model, leading to a series of modestly performing models. The process continues iteratively, with the new models addressing the hardest-to-classify instances, thus reducing bias and improving overall performance. Boosting can be applied to both regression and classification problems, using methods like AdaBoost and Gradient Boosting. These methods iteratively update the models by minimizing loss functions, resulting in superior predictions compared to single weak learners [164], [170], [171].

V.8 Description of the simulation setup PV array

The setup adopted includs a Grid-connected 255 kW PV power plant modeled in the MATLAB/Simulink environment. This setup provides datasets for training and testing, covering normal operating conditions and abnormal plant scenarios [172]. Figure V-5 displays the entire system. The PV system includes a three-level IGBT bridge inverter controlled by a maximum power point tracker tool, as well as a 0.25/255 kW three-phase transformer for grid connection.

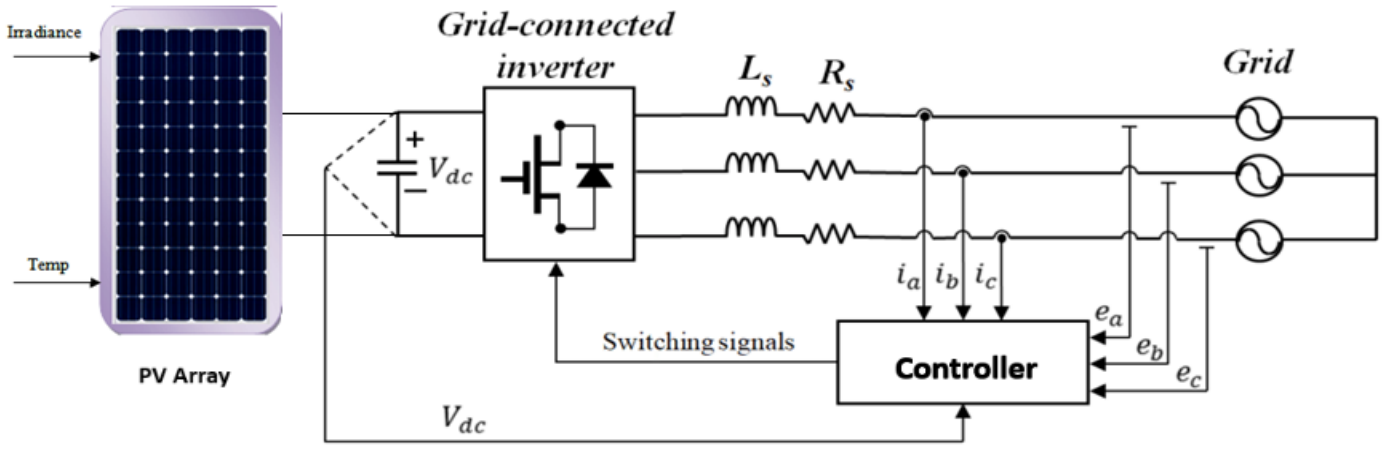


Figure V-5: Schematic diagram of a Grid-connected PV system.

Figure V-6 illustrates the PV array internal components, which consists of 88 parallel strings, with each string containing seven series-connected modules. Table V-4 details the specifications of the modules used.

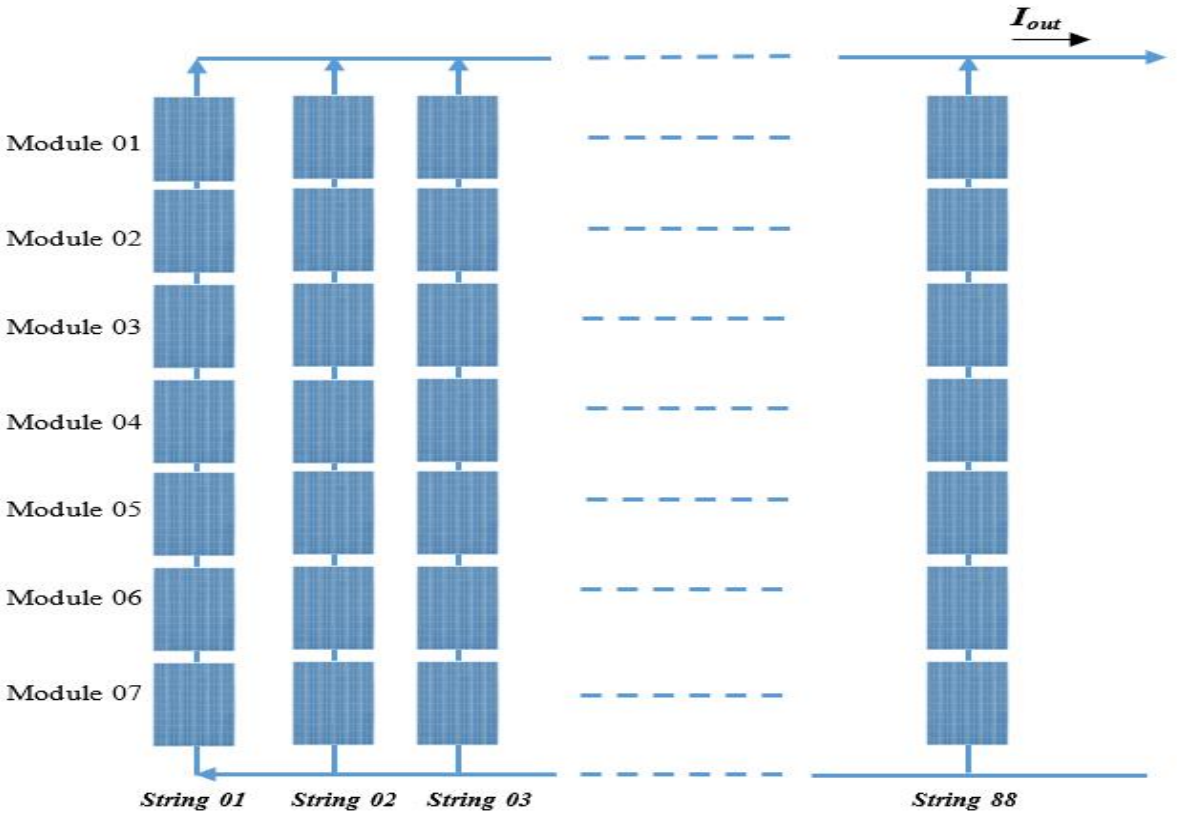


Figure V-6: Internal components of the PV array.

Table V-4: Module characteristics used in the studied PV array

Module characteristics	Values
Short-circuit current, I_{sc} (A)	6.09
Maximum Power Point, MPP (W)	414.801
Voltage at MPP, V_{MPP} (V)	72.9
Current at MPP, I_{MPP} (A)	5.69
Open-circuit Voltage, V_{oc} (V)	85.3
Number of cells	128

V.9 Dataset description

This study utilizes a publicly available dataset [173]. The dataset was obtained through multiple testing operations. It comprised 698 instances, each with 30 features and one (01) column for classes. Within the dataset, there were 124 healthy instances (17.76% of the total data), 178 instances (25.50%) with string faults, 174 instances (24.93%) with string-to-ground faults, and 222 instances (31.81%) with string-to-string faults. Table V-5 summarizes these results.

Table V-5: Dataset of the studied PV array

Scenarios	Number of instances	Ratio	Class label
Healthy PV array	124	17.76%	0
String fault	178	25.50%	1
String-to-ground fault	174	24.93%	2
String-to-string fault	222	31.81%	3
Total	**698**	**100%**	

The measurements cover various attributes (features) including temperature, radiation, and fault resistance, with the following ranges:

- Temperature: 10°C < T < 35°C.

- Radiation: 100 W/m² < G < 1000 W/m^2.

- Fault resistance: 1 Ω < R < 2000 Ω.

Among the 30 features, statistics such as average, maximum, minimum, and variance values of current were collected from three strings. Each string had two ammeters to measure current at the top and bottom, flowing in the same direction.

Additional measurements included total average DC power, total current, and total average DC voltage. The algorithm focused on fault detection and estimating faulty strings, emphasizing specific range values that significantly enhanced accuracy.

These specific rangers, reported in Table V-6, were directly taken from the PV plant to enhance the fault detection algorithm accuracy. These rangers are identified as follows:

$$R_1 = I_{1Amax} - I_{1Amin} \quad \text{(V-1)}$$

$$R_2 = I_{2Amax} - I_{2Amin} \quad \text{(V-2)}$$

$$R_3 = I_{1A} - I_{1B} \quad \text{(V-3)}$$

$$R_4 = I_{2A} - I_{2B} \quad \text{(V-4)}$$

Table V-6: Additional features directly measured from the PV array

Features	Type	Description
I_{1A}	Nominal	Top average current on string n°01
I_{2A}	Nominal	Top average current on string n°02
I_{3A}	Nominal	Top average current on string n°03
I_{1B}	Nominal	Bottom average current on string n°01
I_{2B}	Nominal	Bottom average current on string n°02
I_{3B}	Nominal	Bottom average current on string n°03

I_{TOTAL}	Nominal	Total average current
V_{DC}	Nominal	Total average DC voltage
P_{DC}	Nominal	Total average DC power
G	Nominal	Solar irradiance (100 W/m^2 < G < 1000 W/m²)
T	Nominal	PV array temperature (10°C < T < 35°C)
Class	Categorical	Includs 04 classes: {0, 1, 2, 3}

V.10 Considered faulty scenarios

It is crucial to understand how photovoltaic (PV) arrays behave under various conditions in order to maximize their performance and guarantee reliability. This study seeks to examine and analyze four specific scenarios chosen for their significance and possible impact on PV array performance. The goal is to gain valuable insights from these scenarios to develop more reliable methods for detecting and diagnosing various faults.

In addition to normal situations, this study also considers three faulty scenarios described as bellow:

V.10.1 Fault scenario n° 01 (string fault)

As illustrated in Figure V-7, a string fault occurs when a single string of photovoltaic panels experiences a fault, and in this case, the fault was tested on string 1. This type of fault can occur due to damage to one of the units within the string, or as a result of a loose or faulty connection between the units. As a result of this fault, the energy output from string 1 decreases. Although the solar system can continue to operate, its overall efficiency is lower due to the reduced output from this faulty string.

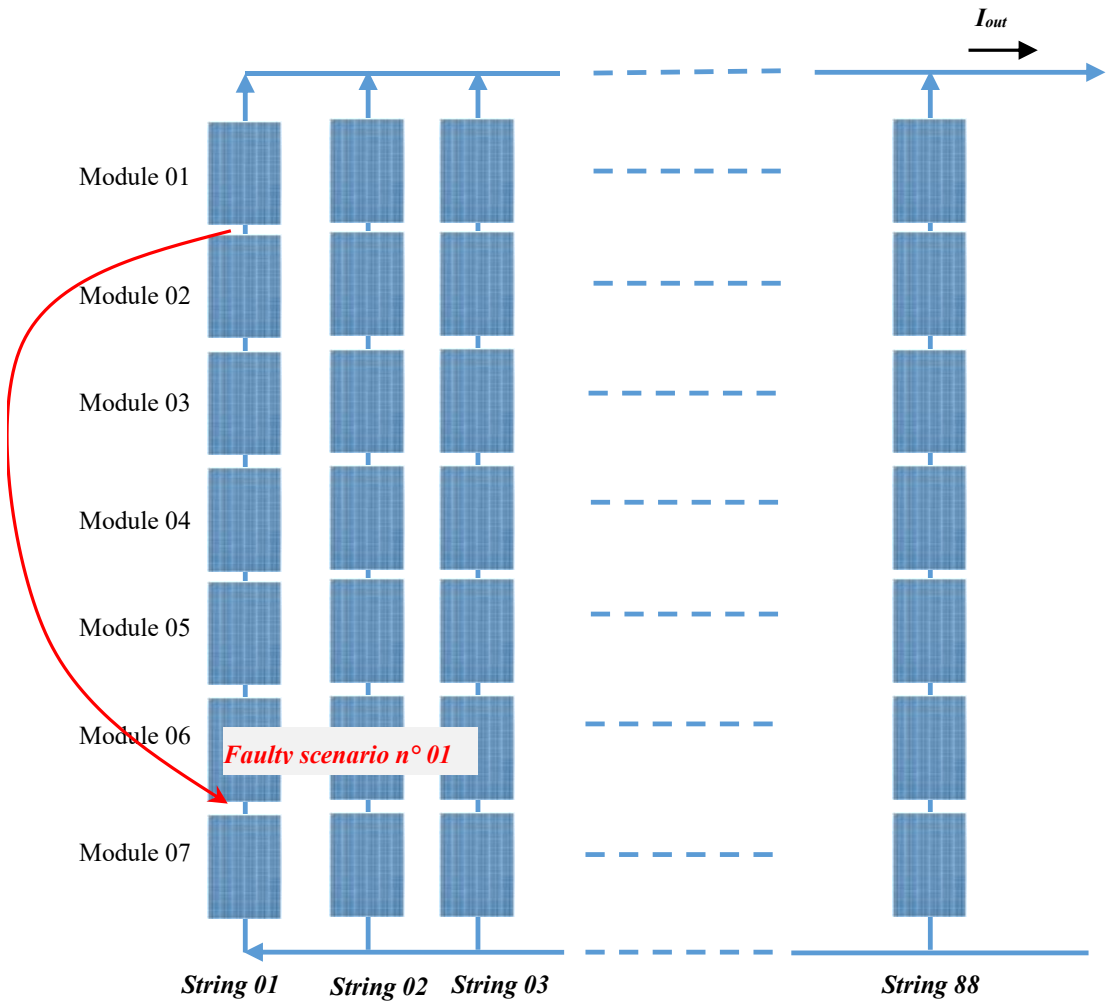


Figure V-7: Faulty scenario n°01 (string fault).

V.10.2 Fault scenario n° 02 (string-to-ground fault)

As shown in Figure V-8, a string-to-ground fault occurs when string 01 connects to the ground, causing electric current to leak from the string to the ground. This type of fault can result from deteriorated cable insulation or physical damage to the cables or photovoltaic units that exposes conductive parts. When current leaks to the ground, it creates a safety hazard by potentially causing electric shocks if the metallic parts of the string are not properly connected to the earth. Furthermore, this fault can reduce the efficiency of the solar system and increase operational risks.

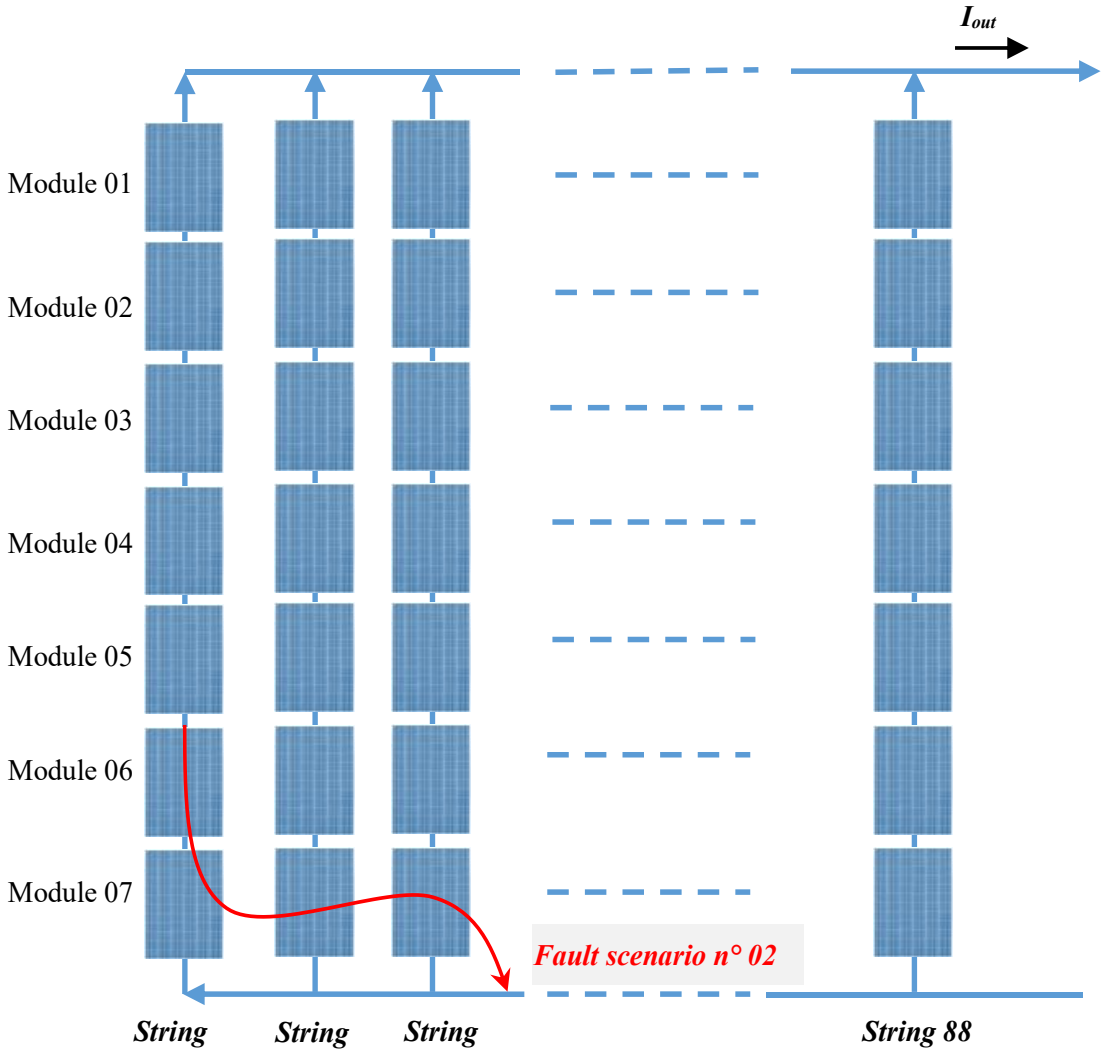


Figure V-8: Fault scenario n° 02 (string-to-ground fault).

V.10.3 Fault scenario n° 03 (string-to-string fault)

As exposed in Figure V-9, a string-to-string fault occurs when there is a short circuit between two strings of photovoltaic panels, and in this case, the fault was tested between string 1 and string 2. This type of fault can occur due to unintended connection between the two strings, leading to a short circuit. The short circuit between the strings results in an imbalance of current between them and can lead to energy loss and damage to the components of the solar system. It can also cause fluctuations in voltage and current, making the system less efficient and increasing operational risks if not detected and corrected promptly.

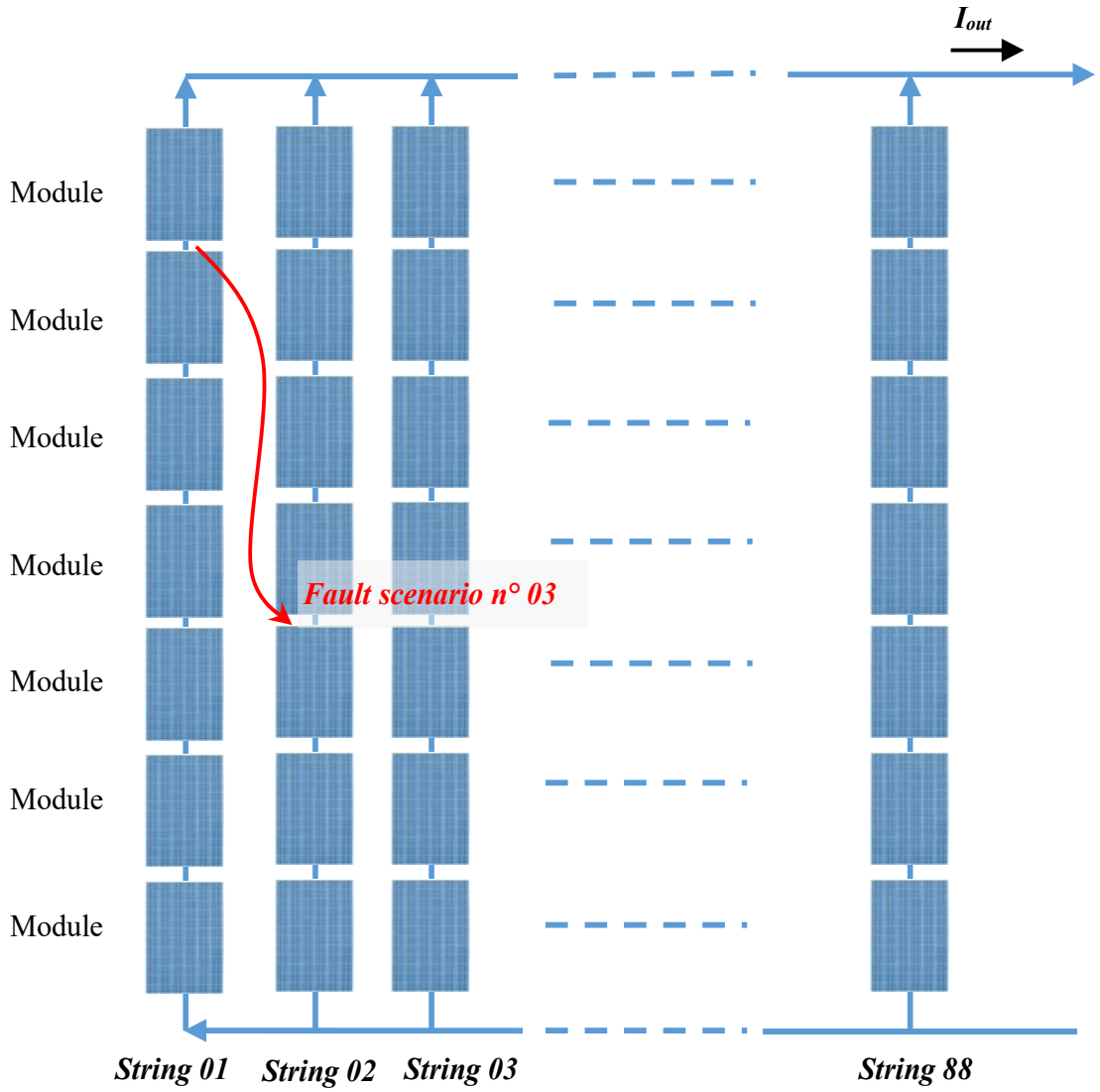


Figure V-9: Fault scenario n° 03 (string-to-string fault).

V.11 The proposed approach based on ensemble learning

This section introduces a precise method for detecting and classifying different PV faults, including string faults, string-to-ground faults, and string-to-string faults which are describe above, using a voting-based ensemble learning $\boldsymbol{V-EL_{(ExT,AdB,LR,NB)}}$ approach. The proposed method integrates four different machine learning algorithms: Extra Trees Classifier, AdaBoost Classifier, Logistic Regression Classifier, and Naive Bayes. The voting ensemble approach is employed because it leverages the strengths of each model while compensating for their weaknesses, resulting in more accurate and reliable predictions. Figure V-10 illustrates the flowchart of the proposed method and the learning algorithm process.

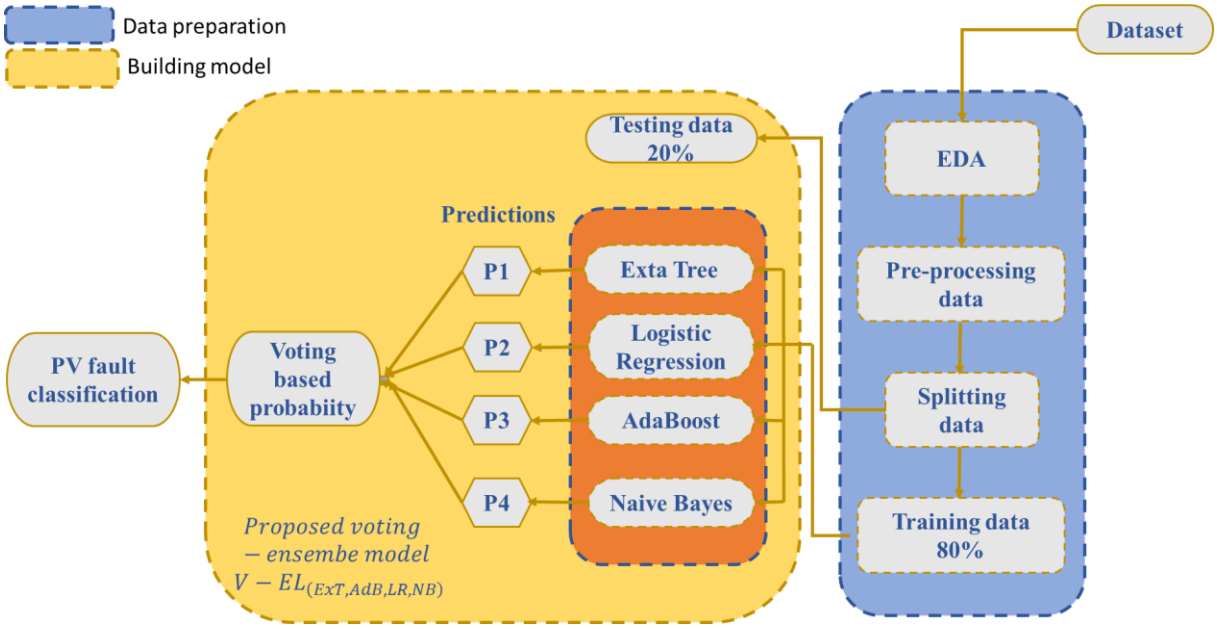


Figure V-10: Flowchart of $V - EL_{(ExT,AdB,LR,NB)}$.

This flowchart outlines the process of fault classification in photovoltaic (PV) systems using a proposed voting-based ensemble model. The algorithm consists of four main stages, which are discussed step-by-step for each part of the flowchart as follows:

V.11.1 Dataset

The starting point is the dataset containing information on PV faults. 30 features including temperature, radiation, and fault resistance, and other different features mentioned above in section of dataset description.

V.11.2 Dataset preparation

V.11.2.1 Exploratory Data Analysis (EDA)

Exploring the dataset is crucial for gaining a deeper understanding of it. This involves several exploratory processes, such as reviewing data information, checking for missing values, describing the data, and identifying unique values. These steps ensure accurate dataset preparation. For instance, if missing values are detected, an appropriate handling process must be implemented.

V.11.2.2 Pre-Processing Data

The dataset is now fully prepared. The next steps involve various pre-processing tasks, including label class partition, label encoding, and normalization. Label class partition is used to categorize each condition in the dataset (free-of-fault state, String fault, String-to-ground fault, and String-to-string fault). Label encoding converts categorical data into numerical form, which is essential since machine learning algorithms cannot process categorical data. For normalization, the MinMax scaler and standard scaler are applied. These scalers, represented mathematically as (V-5) and (V-6) respectively,

ensure that the data is more structured, making it easier for the algorithm to distinguish between healthy and faulty conditions.

$$z = \frac{(x - x_{min})}{(x_{max} - x_{min})} \quad \text{(V-5)}$$

$$z = \frac{(x - u)}{s} \quad \text{(V-6)}$$

Where:

z: the normalized data;

x: the data intended to be normalized;

x_{min}: the minimum value of x;

x_{max}: the maximum value of x;

u : the mean of training data;

s: the standard deviation.

V.11.2.3 Splitting Data

The processed data is randomly shuffled and split into two distinct sets for training and testing. The division is as follows:

- Training Set: 80% of the data is designated for model training.
- Testing Set: 20% of the data is set aside as an independent test set to assess the final performance of the trained model.

V.11.3 Building model

In our study, we implemented a comprehensive machine learning pipeline for model training and evaluation. Initially, we standardized the feature set using the StandardScaler from the sklearn.preprocessing module to ensure the data follows a normal distribution with a mean of 0 and a standard deviation of 1. This was achieved through the fit_transform method on the training set and the transform method on the validation set.

Following this preprocessing step, we trained multiple base classifiers on the scaled training data. Specifically, we utilized the ExtraTreeClassifier, the GaussianNB, the LogisticRegression, and the AdaBoostClassifier. For the ExtraTreeClassifier, we set parameters to control overfitting, including random_state=42, min_samples_split=10, and max_leaf_nodes=15. The LogisticRegression model employed L2 regularization with a penalty term C=1.0 and was configured to run up to 1000 iterations to ensure convergence. The AdaBoostClassifier was instantiated with 100 estimators, a learning rate

of 0.1. Additionally, the 'SAMME' algorithm was used with a fixed random_state of 42 to ensure reproducibility.

To leverage the strengths of individual models, we constructed a hard voting classifier using the VotingClassifier from the sklearn.ensemble module. This ensemble method combines the predictions of the ExtraTreeClassifier, GaussianNB, LogisticRegression, and AdaBoostClassifier to produce a final prediction based on majority voting.

V.11.3.1 Base classifier

a) Extra Trees Classifier

The Extra Trees Classifier (ExT) is an ensemble method similar to Random Forest (RF) in that it uses multiple decision trees to make the final prediction. However, ExT differs in its approach to building the forest of trees. Unlike RF, which uses bootstrapped samples to construct each tree, ExT uses the entire dataset. Additionally, when splitting nodes, ExT randomly selects cut points for each of the k randomly chosen features. The Gini index or Gini impurity is used to determine which feature to split at a node, aiming to make the resulting leaf nodes as pure as possible. The feature with the lowest Gini index value is selected as the root node [174][175][176] [177].

b) Naive Bayes

Naïve Bayes (NB) is a widely recognized classification algorithm in data mining. It calculates the probability that a new example belongs to a particular class, assuming that all attributes are independent given the class. For a given input vector x_0 that needs classification, Bayes' theorem is employed to determine the probability of each class based on the observed values of the predictive attributes [178],

$$P(CL_x = K|x_0) = \frac{P(CL_x=K)\prod_{j=1}^{2} P\left(x_j \middle| CL_x = k\right)}{\sum_{K=1}^{3} P(CL_x=K)\prod_{j=1}^{2} P\left(x_j \middle| CL_x = K\right)} \quad \text{(V-7)}$$

Where:

$x_0 = [x_1, x_2]$ represents the vector of an observation's predictive attributes;

$P(CL_x = K)$ represents the prior probability of a class being indexed as K;

$P\left(x_j \middle| CL_x = k\right)$ represents the probability of the observed attribute x_j given that the class is $CL_x = k$.

Given the assumption of normal distribution for attribute values within each class, the probability of observing a specific attribute can be determined using its mean and standard deviation, as expressed by the following equation:

$$P(x_j|CL_x = K) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_j-\mu)^2}{2\sigma^2}} \quad \text{(V-8)}$$

Here, μ and σ represent the mean value and standard deviation of the corresponding attribute, respectively, obtained from the training dataset for the given class $CL_x = K$.

c) Logistic Regression

In linear regression models, the dependent variable Y is typically continuous, predicted based on one or more independent variables X. However, there are situations where the dependent variable is qualitative and can be categorized into two or more distinct classes. In such cases, the least squares method used in linear regression does not yield plausible estimators. Logistic regression offers a suitable alternative, allowing the use of a regression model to calculate or predict the probability of a specific event $(\pi(x))$ occurring [179] [180].

This section presents the binary logistic regression model, a specific type of generalized linear model known as a logit model. To analyse $\pi(x)$, independent observations $x_1, x_2, \dots, x_n$ are collected [180].

In this context, it is appropriate to assume as a starting point that π(x) is a monotonic function with values between 0 and 1, indicating that $\pi(x)$ is a probability distribution function. Since $\pi(x)$ spans from zero to one, a straightforward linear representation for π across all potential values of x is inadequate, given that linear values extend from $]-\infty; +\infty[$ [180].

In this scenario, a transformation is necessary to ensure that any value of x corresponds to a value within the range $[0,1]$. Using the logistic transformation, also known as the logit, the equation becomes

$$logit = ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p \quad \text{(V-9)}$$

The ratio $\left(\frac{\pi(x)}{1-\pi(x)}\right)$, referred to as the chance or odds, spans from $[0; +\infty[$. Therefore, the natural logarithm of the odds, $log_e(odds)$, extends from $]-\infty; +\infty[$.

As a consequence, based on Equation IV.9, we observe that:

$$e^{logit} = e^{\beta_0+\beta_1 x_1+\cdots+\beta_p x_p} \quad \text{(V-10)}$$

$$\left(\frac{\pi(x)}{1-\pi(x)}\right) = e^{\beta_0+\beta_1 x_1+\cdots+\beta_p x_p} \quad \text{(V-11)}$$

The logistic function, defined as the inverse of the logit function (Equation IV.9), is represented by:

$$\pi(x) = \frac{exp(\beta_0+\beta_1 x_1+\cdots+\beta_p x_p)}{1+exp(\beta_0+\beta_1 x_1+\cdots+\beta_p x_p)} \quad \text{(V-12)}$$

where $\pi(x)$ ranges from 0 to 1.

When incorporating an explanatory variable x_1 into the model, if $\beta_1 > 0$, π increases, while if $\beta_1 < 0$, π decreases. When $\beta_0 = 0$ and $\beta_1 = 0$, $\pi(x)$ equals 0.5. The parameter estimation vector β is derived using the maximum likelihood method [181] [180].

d) Adaptative Boost classifiers

The AdaBoost algorithm, developed by Yoav Freund and Robert Schapire in 1996, is a machine learning technique that enhances weak classifiers, such as one-split decision trees, to form a robust classifier. It achieves this by adjusting the weights of observations in the dataset based on previous prediction errors. Specifically, the algorithm increases the weights of misclassified observations in earlier iterations, making them more likely to be correctly classified in subsequent rounds. AdaBoost was the first to demonstrate that boosting concepts could be implemented both efficiently and simply. It continues to be widely used across various fields of application [182][183][184][172][11] [185].

The implementation of the AdaBoost model consists of four essential steps: collecting data, creating a strong classifier from weak classifiers using training, testing, or validation data, and applying the classifier to engineering problems. Weak classifiers are generated using decision trees, also known as classification and regression trees (CART). These weak classifiers are then combined to produce a strong classifier through majority voting for classification tasks or by taking the arithmetic mean for regression tasks [182][183][184][172][11] [185].

V.11.3.2 Voting-Based Probability

Predictions (P1, P2, P3, P4): Each base classifier makes its predictions on the testing data. These predictions are then aggregated using a voting mechanism. The voting mechanism calculates the probability of each class by considering the votes from all base classifiers.

V.11.3.3 Ensemble model

Proposed Voting-Based Ensemble Model $V - EL_{(ExT,AdB,LR,NB)}$: This ensemble model combines the predictions from the four base classifiers (Extra Tree, AdaBoost, Logistic Regression, and Naive Bayes) using a voting mechanism to improve classification accuracy and reliability.

This ensemble method outperforms using a single classifier by addressing the limitations of traditional machine learning algorithms. It combines each prediction result, and the final performance is determined through a majority vote. The hyperparameters for the proposed $\boldsymbol{V - EL_{(ExT,AdB,LR,NB)}}$ algorithm are displayed in Table V-7.

Table V-7: The hyperparameters for the proposed $V - EL_{(ExT,AdB,LR,NB)}$ algorithm

Model	Items	Parameters
ExT	random stat	42
	min_samples_spit	10
	max_leaf_nodes	15
NB	/	/
LR	penalty	'I2'
	C	1
	max it	1000
adaBoost	n_estimators	100
	learning rate	0.1
	algorithm	'SAMME'
	random-state	42

V.11.3.4 Testing data evaluation

Testing Data (20%): The remaining portion of the dataset is used to test the performance of the ensemble model. The ensemble model's predictions are compared with the actual labels in the testing data to evaluate its accuracy and effectiveness in classifying PV faults.

V.11.4 PV fault classification

The output of the flowchart is PV fault classification which will be:

0 if it is a healthy case;
1 if it is fault 1 (string fault);
2 if it is fault 2 (string to ground fault);
3 if it is fault 3 (string to string).

V.12 Results and discussion

This section presents the results and discussion of the proposed $V - EL_{(ExT,AdB,LR,NB)}$ algorithm for detecting three different PV faults (string fault, string to ground fault, and string to string fault). To thoroughly evaluate the algorithm's effectiveness in PV fault detection, several assessment metrics are utilized, including accuracy, precision, recall, F1 score. Additionally, a confusion matrix visualization is provided to offer a clearer, more intuitive understanding of the $V - EL_{(ExT,AdB,LR,NB)}$ algorithm's performance.

We implemented the proposed approach using Python 3.9 on a 64-bit Windows 10 system equipped with an Intel Core i5 vPro processor and 16.00 GB of memory.

V.12.1 Accuracy performance

A commonly used performance indicator in ML applications for detecting PV faults is the accuracy score. This value indicates the proportion of correct predictions made by the proposed algorithm. Mathematically, it can be expressed as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \quad \text{(V-13)}$$

Where:

TP: True Positive

TN: True Negative

FN: False Positive

FP: False Negative

The accuracy table (Table V-8) shows that the voting-based ensemble model $V - EL_{(ExT,AdB,LR,NB)}$ achieves the highest accuracy at 0.992857. This indicates that the ensemble model correctly identifies the vast majority of instances. Among the individual models, Extra Trees (ExT) and Logistic Regression (LR) both have an accuracy of 0.864286, indicating good performance but still notably lower than the ensemble. AdaBoost (AdB) has a slightly lower accuracy (AdB=0.828571), and Naive Bayes (NB) has the lowest accuracy (NB=0.75,) suggesting it is the least effective of the individual models.

Table V-8: The accuracy of the proposed ensemble learning and its base models

Model	$\boldsymbol{V - EL_{(ExT,AdB,LR,NB)}}$	**ExT**	**adaBoost**	**LR**	**NB**
Accuracy	**0.992857**	0.864286	0.828571	0.864286	0.75

V.12.2 Precision performance

The precision indicator represents the positive predictive value of the proposed algorithm for a given dataset. Mathematically, precision can be expressed as follows:

$$precision = \frac{TP}{TP+FP} \quad \text{(V-14)}$$

The precision table (Table V-9) highlights that the voting-based ensemble model $V - EL_{(ExT,AdB,LR,NB)}$ achieves the highest precision equals to 0.9928. This high precision indicates that the ensemble model has a very low rate of false positives. Logistic Regression (LR) and Extra Trees (ExT) also show good precision value of 0.8640 and 0.8638, respectively, meaning they are relatively effective in predicting positive instances correctly. AdaBoost (AdB) has a slightly lower precision of 0.8283, while Naive Bayes (NB) lags behind with a precision of 0.74, indicating a higher rate of false positives compared to other models.

Table V-9: The precision of the proposed ensemble learning and its base models

Model	$\boldsymbol{V - EL_{(ExT,AdB,LR,NB)}}$	**ExT**	**adaBoost**	**LR**	**NB**
Precision	**0.9928**	0.8638	0.8283	0.8640	0.74

V.12.3 Recall performance

To assess the performance of the proposed algorithm, a recall indicator is employed to examine the true positive rate. The recall value demonstrates the ability of the proposed algorithm to correctly identify all positive samples within a given dataset. Mathematically, this can be represented as:

$$Recall = \frac{TP}{TP+FN} \tag{V-15}$$

In the recall table (Table V-10), the voting-based ensemble model $V - EL_{(ExT,AdB,LR,NB)}$ again leads with the highest recall at 0.9927. This high recall signifies that the ensemble model has a very low rate of false negatives, capturing almost all actual positive instances. Logistic Regression (LR) and Extra Trees (ExT) show good recall values at 0.8639 and 0.8635, respectively. AdaBoost (AdB) follows closely at 0.8281, while Naive Bayes (NB) has the lowest recall at 0.7395, indicating it misses more actual positive instances compared to the other models.

Table V-10: The recall of the proposed ensemble learning and its base models

Model	$\boldsymbol{V - EL_{(ExT,AdB,LR,NB)}}$	**ExT**	**adaBoost**	**LR**	**NB**
Recall	**0.9927**	0.8635	0.8281	0.8639	0.7395

V.12.4 F1 score performance

The F1 score evaluates the harmonic mean of precision and recall. A high F1 score, nearing 1, indicates effective performance of the proposed algorithm in detecting PV faults. Conversely, a low F1 score, nearing 0, suggests poor algorithm performance. The mathematical expression for the F1 score is as follows:

$$F1\ score = \frac{2TP}{2TP+FP+FN} \quad \text{(V-16)}$$

The F1 score table (Table V-11) reflects a balanced measure of precision and recall. The voting-based ensemble model $V - EL_{(ExT,AdB,LR,NB)}$ achieves the highest F1 score equal to 0.9928, indicating overall superior performance. Logistic Regression (LR) and Extra Trees (ExT) have similar F1 scores of 0.8639 and 0.8635, respectively, showing they are relatively balanced in terms of precision and recall. AdaBoost (AdB) has a slightly lower F1 score of 0.8281, while Naive Bayes (NB) has the lowest F1 score of 0.7395, reflecting its weaker overall performance compared to the other models.

Table V-11: The F1 score of the proposed ensemble learning and its base models

Model	$V - EL_{(ExT,AdB,LR,NB)}$	**ExT**	**adaBoost**	**LR**	**NB**
F1 score	**0.9928**	0.8635	0.8281	0.8639	0.7395

V.12.5 Confusion matrix performance

In classification tasks, a confusion matrix (CM) is frequently employed to assess the performance of the proposed algorithm. Figure V-11 presents a confusion matrix illustrating the detection of PV faults. Effective algorithms are characterized by high true positive and true negative counts, whereas high false positive and false negative counts suggest unreliable performance.

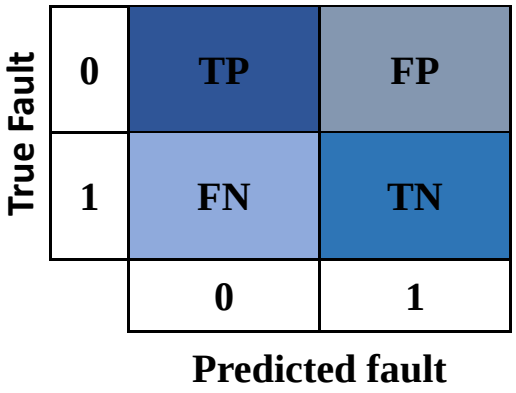


Figure V-11: Confusion matrix

Figure V-12 shows the confusion matrix of the proposed voting-ensemble learning and its base learners. The analysis of the same figure reveals that the Voting Classifier, which combines the strengths of the individual classifiers, provides a balanced and robust performance across different classes, effectively leveraging the strengths of the individual classifiers.

For Class 0, it achieved 24 correct predictions out of 30, with 4 instances misclassified as Class 1 and 2 as Class 2. Class 1 was correctly predicted 25 times, with 1 instance misclassified as Class 0 and 4 as Class 2. Class 2 had high accuracy with 43 correct predictions out of 50, and 7 misclassifications as Class 1. Finally, Class 3 was predicted with high accuracy, achieving 47 correct predictions and 3 misclassifications as Class 2.

While the Extra Tree, AdaBoost, Logistic Regression, and Naive Bayes classifiers each have varying levels of accuracy and specific weaknesses, the ensemble approach of the Voting Classifier offers more consistent and reliable performance. This highlights the advantage of using ensemble methods, as they can mitigate the weaknesses of individual classifiers and provide a more comprehensive and accurate prediction.

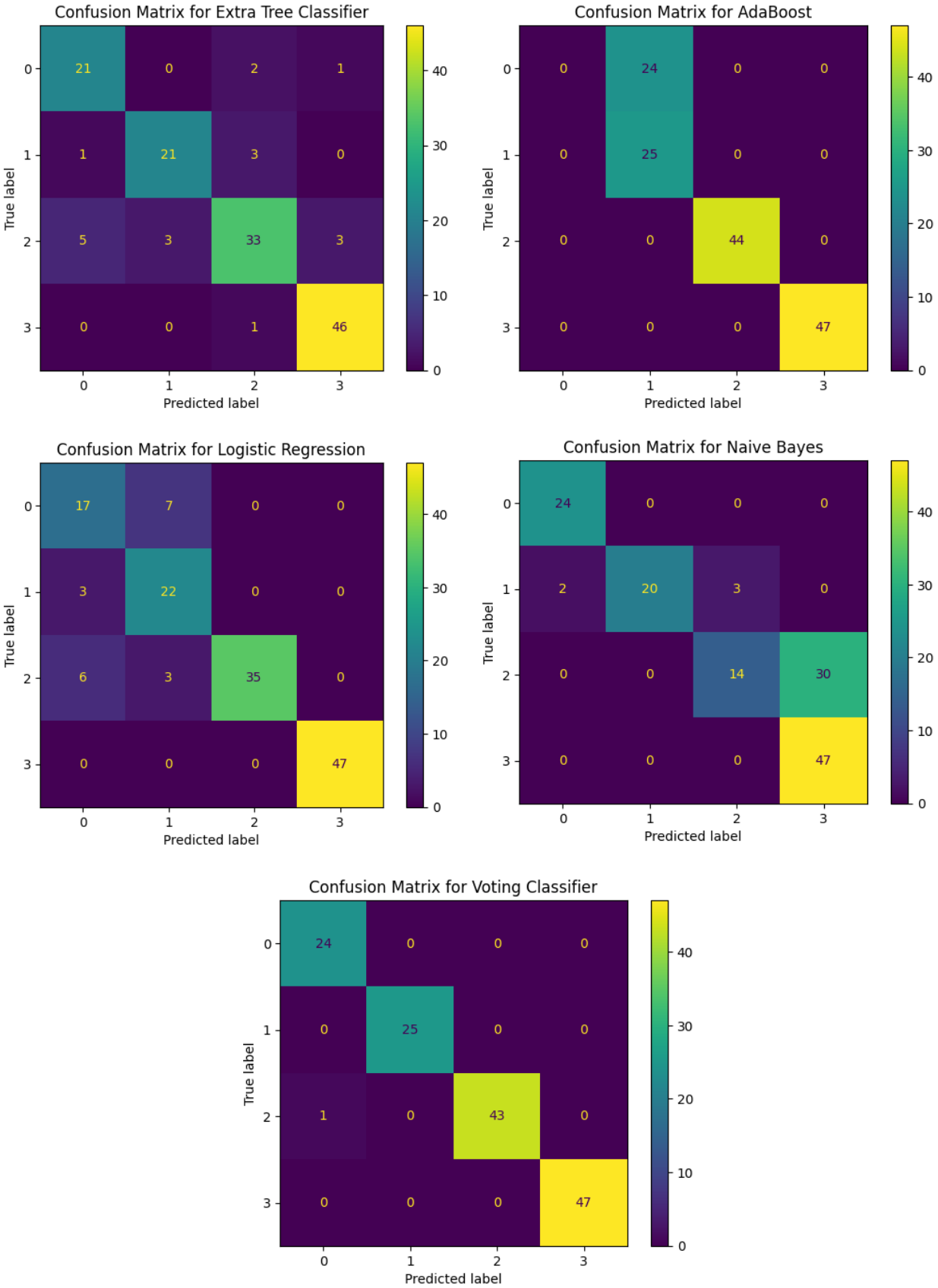


Figure V-12: The confusion matrix of the propose voting-ensemble model and its base models.

V.13 Conclusion

In this chapter, the focus is on accurately detecting and classifying faults in photovoltaic (PV) DC-side systems. The chapter begins by explaining the use of the BB-TLBO algorithm for identifying model parameters in both healthy and faulty PV arrays. The resulting parameters are then analyzed to classify the scenarios as faulty or healthy.

Even though using metaheuristic techniques for parameter identification and fault detection is an important first step, providing fundamental insights into the operational characteristics of PV systems, the true potential of artificial intelligence (AI) lies in its ability to process vast amounts of data, recognize intricate patterns, and adapt to changing conditions.

As a result, $V - EL_{(ExT,AdB,LR,NB)}$, a novel voting-based ensemble learning method has been developed to address three common fault types: string faults, string-to-ground faults, and string-to-string faults. $V - EL_{(ExT,AdB,LR,NB)}$ combines the strengths of four base learners - Extra Trees, Naive Bayes, Logistic Regression, and AdaBoost - to achieve superior performance compared to individual models. Evaluation using accuracy, precision, F1 score, recall, and confusion matrices confirmed $V - EL_{(ExT,AdB,LR,NB)}$'s effectiveness.

This proactive method will be essential in maximizing the lifespan and efficiency of PV systems, ensuring that they remain a reliable source of clean energy.

Chapter VI : General conclusion and future works

VI.1 General conclusion

The research work presented in this thesis focuses on developing strategies to identify photovoltaic (PV) DC-side system parameters, as well as to detect and classify faults that could occur on these systems. The main goal of the thesis was to enhance the efficiency and performance of PV systems and shield them from potential failures by creating robust fault detection and diagnosis strategies. Specifically, the use of advanced techniques such as metaheuristic algorithms and artificial intelligence tools, was pivotal to address this issue.

However, utilizing these tools requires a high-quality dataset that accurately describes the relationship between PV generator parameters and fault types, while faithfully representing the system's behaviour under normal and faulty conditions. Obtaining such a dataset can be challenging in practice, especially when certain faults may cause severe damage to the system. To overcome this challenge, a precise simulation model was developed to simulate the behaviour of the PV system under normal and faulty conditions.

The primary contribution of this thesis is the development of the BB-TLBO hybrid algorithm, which integrates Biogeography-Based Optimization (BBO) with Teaching-Learning-Based Optimization (TLBO). This novel algorithm is highly effective in accurately identifying the parameters of photovoltaic (PV) cells and modules, specifically for one-diode, two-diode, and three-diode models. The BB-TLBO algorithm sets a new standard for optimization techniques in PV systems by delivering substantial improvements in accuracy, convergence speed, and robustness. Notably, the algorithm also plays a pivotal role in fault detection and classification, as it leverages the same parameter identification process for both healthy and faulty PV arrays.

The adoption of this metaheuristic technique marks an important step in fault diagnosis, particularly because of its compatibility with smaller datasets. In contrast, artificial intelligence (AI) methods are better suited for handling larger datasets. Accordingly, this thesis introduces a novel approach to fault detection and classification in PV arrays, utilizing ensemble learning (EL) techniques.

By combining multiple classification algorithms within a voting model, this approach successfully diagnosed several common fault types, including string faults, string-to-ground faults, and string-to-string faults. The results demonstrated significant performance improvements over traditional detection methods, particularly in terms of accuracy, recall, and F1 score.

The thesis is structured as follows:

- Chapter 1: General introduction providing background and motivation for the research project, including a detailed statement of the problem regarding fault detection and diagnosis of the PV systems.
- Chapter 2: Presentation of the PV array modelling process and methods, starting with an explanation of the core principles of the photovoltaic effect and solar PV cell operations. It emphasizes the importance of understanding various electrical and mathematical models used to represent the behavior of solar cells and modules, such as single diode, double diode, and three diode models. It also highlights the crucial task of identifying unknown parameters using different methods.
- Chapter 3: Details the proposed BB-TLBO technique for identifying the unknown parameters of PV generators' models, demonstrating its superiority over competing techniques using both the classic RMSE cost function and the Lambert RMSE cost function.
- Chapter 4: Reviews the state-of-the-art methods used in fault detection and diagnosis of the PV system, identifying their strengths and weaknesses in this field of engineering.
- Chapter 5: Presents the contribution, including two methods for fault detection and diagnosis of the PV array. The first method relies on using the BB-TLBO technique for model parameter identification and classification of results according the corresponding scenarios. The second method uses Ensemble Learning (EL) technique as part of Artificial Intelligence (AI) techniques. The results show promising performances due to the precision reported in this chapter.
- Chapter 6: General conclusion summarizing the thesis, related results and findings, and suggesting future research directions.

This research work achieved notable findings that contribute significantly to the field of photovoltaic (PV) systems. These key achievements including:

- Strengthening the understanding of solar PV energy in both theoretical and practical frameworks. The comprehensive study of PV systems not only advanced theoretical knowledge but also provided practical insights into how PV systems operate under normal and faulty conditions.
- Deepening the modelling of PV generators by employing various electrical circuit-based models through metaheuristic techniques. The hybridization of Biogeography-Based Optimization (BBO) and Teaching-Learning-Based Optimization (TLBO) approaches has

significantly enhanced the accuracy, speed, and robustness of PV system parameter identification. This hybrid algorithm proved effective across multiple PV models, offering a versatile solution for real-world applications.

- Successful development and validation of a metaheuristic method that precisely models PV cells, modules, and arrays with optimal computational efficiency. This method not only improved the modeling process but also provided a robust platform for identifying PV parameters across different system configurations.
- Application of the BB-TLBO algorithm in fault detection and classification, offering a reliable and streamlined approach to diagnosing faults in PV arrays. The dual-purpose nature of the algorithm (handling both parameter identification and fault detection) ensures that PV systems can be more effectively monitored and maintained, enhancing overall system reliability.
- Development and experimental simulation of an effective PV array that accurately replicates real-world behaviour. This simulation model enabled safe and controlled experimentation, including the testing of risky fault scenarios that might otherwise compromise system integrity. This simulation tool will prove valuable for future research and development in PV fault detection.
- Detection and classification of various types of faults using Ensemble Learning (EL) techniques, applied to the dataset used. The introduction of EL-based fault classification added a new dimension to fault diagnosis in PV systems, offering improved accuracy and reliability over traditional detection methods. By combining multiple classification algorithms in a voting model, the system was able to handle complex fault scenarios that might evade detection through single-method approaches.

These contributions significantly advance the state of knowledge in PV system performance optimization and fault detection, laying a solid foundation for future innovations in renewable energy systems.

VI.2 Future Works

This thesis has led to the development of innovative tools and methods for parameter identification and fault detection in photovoltaic systems. However, several avenues for future research remain open. Therefore, the research work suggests several possible directions for future exploration, including:

- Scaling the BB-TLBO algorithm and fault detection strategies to large-scale PV installations, such as utility-scale solar farms. Expanding the application of these methodologies at larger scales could yield valuable insights into optimizing energy production, maintenance, and fault

management across extensive PV networks. Such advancements could contribute to enhancing energy efficiency on a global scale.

- Developing new, more efficient algorithms that address the challenges of PV modeling with greater accuracy and reduced computational time. Further improvements in optimization techniques can enable faster identification of PV system parameters and more precise fault detection, which is crucial for real-time applications.
- Integrating advanced artificial intelligence (AI) techniques, such as deep learning (DL), convolutional neural networks (CNNs), and reinforcement learning, into PV fault detection and classification systems. These AI methods have the potential to significantly improve the ability of PV systems to autonomously detect, classify, and predict faults. By leveraging the power of AI, future systems can achieve greater reliability and self-learning capabilities, leading to smarter, more resilient energy infrastructures.
- Implementing real-time applications of the developed algorithms to enable immediate detection and response to PV system anomalies. Real-time monitoring is essential for minimizing system downtime and ensuring that faults are addressed promptly, preventing long-term damage and energy losses.
- Integrating the proposed methodologies into IoT-enabled smart grids to facilitate predictive maintenance, remote fault detection, and automated energy management. By incorporating these techniques into smart grid technologies, PV systems can benefit from enhanced operational longevity, optimized energy production, and reduced maintenance costs. IoT integration would allow for seamless communication between PV systems and grid operators, enabling proactive management of faults and performance issues.
- Developing more advanced simulation models that accurately represent the behaviour of PV generators under a wider variety of fault conditions. As PV systems become more complex, these models will need to reflect a broader range of operational scenarios, including rare or extreme fault conditions. These models can serve as critical testing environments for new algorithms and fault detection techniques, reducing the risk of costly real-world failures.
- Exploring the potential of AI-driven predictive models for energy output forecasting and system optimization. AI tools, particularly those based on machine learning, can be harnessed to predict energy generation trends, optimize panel orientation, and even anticipate maintenance needs, further contributing to the operational efficiency and sustainability of PV systems.

This research lays the foundation for a future where photovoltaic systems can be managed more efficiently and intelligently, playing a key role in the global energy transition. Addressing these research directions can further improve the reliability, safety, and efficiency of PV systems, pushing the boundaries of renewable energy technology.

References

[1] H. Ravaee, S. Farahat, and F. Sarhaddi, “Artificial neural network based model of photovoltaic thermal (PV/T) collector,” J. Math. Comput. Sci., vol. 4, no. 3, pp. 411–417, 2012.

[2] R. Karki and R. Billinton, “Reliability/cost implications of PV and wind energy utilization in small isolated power systems,” IEEE Trans. Energy Convers., vol. 16, no. 4, pp. 368–373, 2001.

[3] B. Parida, S. Iniyan, and R. Goic, “A review of solar photovoltaic technologies,” Renew. Sustain. Energy Rev., vol. 15, no. 3, pp. 1625–1636, 2011.

[4] Y. Hamakawa, “Recent advances in solar photovoltaic technology and its new role for environmental issue,” Renew. Energy, vol. 5, no. 1–4, pp. 34–43, 1994.

[5] O. Hachana, “Etude d’un système photovoltaïque en vue du diagnostic,” PhD thesis in automatic, University of Ferhat Abbes – Setif1, Algeria, 2015.

[6] Blackridgeresearch, “Global solar capacity growth”, available on: https://www.blackridgeresearch.com/blog/global-solar-photovoltaic-pv-power-market-industry-annual-review-roundup-2023#Global Solar Capacity Growth. [Accessed: Feb. 4, 2024].

[7] A. Mehmood, H. A. Sher, A. F. Murtaza, et al., “Fault detection, classification and localization algorithm for photovoltaic array,” IEEE Trans. Energy Convers., vol. 36, no. 4, pp. 2945–2955, 2021.

[8] G.Elyes, “Faults detection and diagnosis of photovoltaic systems using artificial intelligence tools, ” PhD thesis in Electrical Engineering, university of Blida, Algeria, 2018.

[9] A. Mellit and S. Kalogirou, “Artificial intelligence and internet of things to improve efficacy of diagnosis and remote sensing of solar photovoltaic systems: Challenges, recommendations and future directions,” Renew. Sustain. Energy Rev., vol. 143, p. 110889, 2021.

[10] M. A. Hariri-Ardebili, G. Mahdavi, L. K. Nuss, et al., “The role of artificial intelligence and digital technologies in dam engineering: Narrative review and outlook,” Eng. Appl. Artif. Intell., vol. 126, p. 106813, 2023.

[11] E. Lodhi, F. Yuewang, G. Xiong, et al., “A novel deep stack-based ensemble learning approach for fault detection and classification in photovoltaic arrays,” Remote Sens., vol. 15, p. 1277, 2023.

[12] D. S. Pillai and N. Rajasekar, “Metaheuristic algorithms for PV parameter identification: A comprehensive review with an application to threshold setting for fault detection in PV systems,” Renew. Sustain. Energy Rev., vol. 82, pp. 3503–3525, 2018.

[13] S. Li, W. Gong, and Q. Gu, “A comprehensive survey on meta-heuristic algorithms for parameter extraction of photovoltaic models,” Renew. Sustain. Energy Rev., vol. 141, p.

110828, 2021.

[14] W. O. Compaore, “Aide à la décision pour le diagnostic des défauts pour une maintenance proactive dans un générateur photovoltaïque,” PhD thesis in automatic, university of Joseph Ki-Zerbo, Dec. 19, 2023.

[15] "Photovoltaic Effect," Mr. Solar, Available: https://www.mrsolar.com/photovoltaic-effect/. [Accessed: Feb. 4, 2024].

[16] P. Rappaport, “The photovoltaic effect and its utilization,” Solar Energy, vol. 3, no. 4, pp. 8–18, 1959.

[17] K. Kaygusuz, “Renewable energy: Power for a sustainable future,” Energy Explor. Exploit., vol. 19, no. 6, pp. 603–626, 2001.

[18] K.Riad, “Détection et isolation de défauts combinant des méthodes à base de données appliquées aux systèmes électro-énergétiques,” PhD thesis in automatic, university of Ferhat Abbes - Setif 1, Algeria, March 2015.

[19] R. Pradhan, “Development of New Parameter Extraction Schemes and Maximum Power Point Controllers for Photovoltaic Power Systems,” PhD thesis in electrical engineering, natioanal institut of technology, Rourrkela, Inde, 2014.

[20] S. Gupta, O. Singh, and M. A. Ansari, “Maximum power point tracking techniques for photovoltaic system: A review,” in Advances in Signal Processing and Communication: Select Proceedings of ICSC 2018, vol. 526, Lecture Notes in Electrical Engineering. Singapore: Springer, 2019, pp. 455–465.

[21] "Operation of a basic photovoltaic cell," Wikimedia Commons. Available: https://commons.wikimedia.org/wiki/File:Operation_of_a_basic_photovoltaic_cell.gif. [Accessed: Feb. 4, 2024].

[22] S. T. Buddha, “Topology Reconfiguration To Improve The Photovoltaic (PV) Array Performance,” Master Thesis, Arizona State University, 2011.

[23] B. Tull, Photovoltaic Cells: Science and Materials. 2004. [Online]. Available: http://www.molchem.science.ru.nl/rowan/Coll/caput-college/PhotovoltaicCells.pdf. [Accessed: May 5, 2023].

[24] G. M. Masters, Renewable and Efficient Electric Power Systems, 2nd ed. Hoboken, NJ, USA: John Wiley & Sons, 2013, ch. 8, pp. 448–458.

[25] L. Castañer et S. Silvestre, Modelling Photovoltaic Systems Using PSpice. Chichester, Royaume-Uni: John Wiley & Sons, 2002, ch. 3, pp. 103–132.

[26] T.Faris, O.Hammou, “Dimensionnement d’un système de pompage d’eau photovoltaïque destiné à l’irrigation (Cas de la région de M’sila) ,” Master thesis, University of Mohamed oudiaf - M'sila, 2022.

[27] R. G. Vieira, F. M. U. de Araújo, M. Dhimish, et al., "A comprehensive review on bypass diode application on photovoltaic modules," Energies, vol. 13, no. 10, p. 2472, 2020.

[28] O. Z. Sharaf and M. F. Orhan, "Concentrated photovoltaic thermal (CPVT) solar collector systems: Part I–Fundamentals, design considerations and current technologies," Renewable and Sustainable Energy Reviews, vol. 50, pp. 1500–1565, 2015.

[29] C. Kumar and D. M. Mary, "A novel chaotic-driven Tuna Swarm Optimizer with Newton-Raphson method for parameter identification of three-diode equivalent circuit model of solar photovoltaic cells/modules," Optik, vol. 264, p. 169379, 2022.

[30] M. A. Navarro, D. Oliva, A. Ramos-Michel, et al., "An analysis on the performance of metaheuristic algorithms for the estimation of parameters in solar cell models," Energy Conversion and Management, vol. 276, p. 116523, 2023.

[31] F. Ali, A. Sarwar, F. I. Bakhsh, et al., "Parameter extraction of photovoltaic models using atomic orbital search algorithm on a decent basis for novel accurate RMSE calculation," Energy Conversion and Management, vol. 277, p. 116613, 2023.

[32] M. A. El-Dabah, R. A. El-Sehiemy, H. M. Hasanien, et al., "Photovoltaic model parameters identification using Northern Goshawk Optimization algorithm," Energy, vol. 262, p. 125522, 2023.

[33] M. Eslami, E. Akbari, S. T. Seyed Sadr, et al., "A novel hybrid algorithm based on rat swarm optimization and pattern search for parameter extraction of solar photovoltaic models," Energy Science & Engineering, vol. 10, no. 8, pp. 2689–2713, 2022.

[34] N. Rai, A. Abbadi, F. Hamidia, et al., "Biogeography-based teaching learning-based optimization algorithm for identifying one-diode, two-diode and three-diode models of photovoltaic cell and module," Mathematics, vol. 11, no. 8, p. 1861, 2023.

[35] A. A. Z. Diab, H. M. Sultan, T. D. Do, et al., "Coyote optimization algorithm for parameters estimation of various models of solar cells and PV modules," IEEE Access, vol. 8, pp. 111102–111140, 2020.

[36] A. R. Ginidi, A. M. Shaheen, R. A. El-Sehiemy, et al., "Supply demand optimization algorithm for parameter extraction of various solar cell models," Energy Rep., vol. 7, pp. 5772–5794, 2021.

[37] J. Bai, S. Liu, Y. Hao, et al., "Development of a new compound method to extract the five parameters of PV modules," Energy Convers. Manage., vol. 79, pp. 294–303, 2014.

[38] D. S. H. Chan and J. C. H. Phang, "Analytical methods for the extraction of solar-cell single- and double-diode model parameters from IV characteristics," IEEE Trans. Electron Devices, vol. 34, no. 2, pp. 286–293, 1987.

[39] F. Khan, S.-H. Baek, Y. Park, et al., "Extraction of diode parameters of silicon solar cells under

high illumination conditions," Energy Convers. Manage., vol. 76, pp. 421–429, 2013.

[40] V. Kongphet, “Contribution to fault detection of PV modules using I-V curves,” PhD thesis in electric power., university of Paris Saclay, 2022.

[41] A. K. Tossa, Y. M. Soro, Y. Azoumah, et al., "A new approach to estimate the performance and energy productivity of photovoltaic modules in real operating conditions," Sol. Energy, vol. 110, pp. 543–560, 2014.

[42] K. Et-Torabi, I. Nassar-Eddine, A. Obbadi, et al., "Parameters estimation of the single and double diode photovoltaic models using a Gauss–Seidel algorithm and analytical method: A comparative study," Energy Convers. Manage., vol. 148, pp. 1041–1054, 2017.

[43] A. A. Elbaset, H. Ali, and M. Abd-El Sattar, "Novel seven-parameter model for photovoltaic modules," Sol. Energy Mater. Sol. Cells, vol. 130, pp. 442–455, 2014.

[44] V. J. Chin and Z. Salam, "A new three-point-based approach for the parameter extraction of photovoltaic cells," Appl. Energy, vol. 237, pp. 519–533, 2019.

[45] P. J. Gnetchejo, S. N. Essiane, P. Ele, et al., "Important notes on parameter estimation of solar photovoltaic cell," Energy Convers. Manage., vol. 197, p. 111870, 2019.

[46] K. Ishaque and Z. Salam, "An improved modeling method to determine the model parameters of photovoltaic (PV) modules using differential evolution (DE)," Sol. Energy, vol. 85, no. 9, pp. 2349–2359, 2011.

[47] M. Zagrouba, A. N. I. S. Sellami, M. Bouaïcha, et al., "Identification of PV solar cells and modules parameters using the genetic algorithms: Application to maximum power extraction," Sol. Energy, vol. 84, no. 5, pp. 860–866, 2010.

[48] L. Wu, Z. Chen, C. Long, et al., "Parameter extraction of photovoltaic models from measured IV characteristics curves using a hybrid trust-region reflective algorithm," Appl. Energy, vol. 232, pp. 36–53, 2018.

[49] T. Bhattacharyya, B. Chatterjee, P. K. Singh, et al., "Mayfly in harmony: A new hybrid meta-heuristic feature selection algorithm," IEEE Access, vol. 8, pp. 195929–195945, 2020.

[50] B. Aoufi, “Contribution to the development of a fault diagnosis technique for photovoltaic generators,” PhD thesis in automatic, university of Ouargla, Algeria, 2024.

[51] A. Askarzadeh and A. Rezazadeh, "Artificial bee swarm optimization algorithm for parameters identification of solar cell models," Appl. Energy, vol. 102, pp. 943–949, 2013.

[52] J. Pan, Y. Gao, Q. Qian, et al., "Parameters identification of photovoltaic cells using improved version of the chaotic grey wolf optimizer," Optik, vol. 242, p. 167150, 2021.

[53] N. Shankar, N. Saravanakumar, et al., "Solar photovoltaic module parameter estimation with an enhanced differential evolutionary algorithm using the manufacturer’s datasheet information," Optik, vol. 224, p. 165700, 2020.

[54] C. Chellaswamy and R. Ramesh, "Parameter extraction of solar cell models based on adaptive differential evolution algorithm," Renew. Energy, vol. 97, pp. 823–837, 2016.

[55] X. Jian and Y. Zhu, "Parameters identification of photovoltaic models using modified Rao-1 optimization algorithm," Optik, vol. 231, p. 166439, 2021.

[56] T. V. Luu and N. N. Son, "Parameters extraction of solar cells using modified JAYA algorithm," Optik, vol. 203, p. 164034, 2020.

[57] Q. Niu, L. Zhang, and K. Li, "A biogeography-based optimization algorithm with mutation strategies for model parameter estimation of solar and fuel cells," Energy Convers. Manag., vol. 86, pp. 1173–1185, 2014.

[58] S. Li, W. Gong, L. Wang, et al., "A hybrid adaptive teaching–learning-based optimization and differential evolution for parameter identification of photovoltaic models," Energy Convers. Manag., vol. 225, p. 113474, 2020.

[59] X. Chen, B. Xu, C. Mei, et al., "Teaching–learning–based artificial bee colony for solar photovoltaic parameter estimation," Appl. Energy, vol. 212, pp. 1578-1588, 2018.

[60] W. Long, S. Cai, J. Jiao, et al., "A new hybrid algorithm based on grey wolf optimizer and cuckoo search for parameter extraction of solar photovoltaic models," Energy Convers. Manag., vol. 203, p. 112243, 2020.

[61] F. J. Xavier, A. Pradeep, M. Premkumar, et al., "Orthogonal learning-based Gray Wolf Optimizer for identifying the uncertain parameters of various photovoltaic models," Optik, vol. 247, p. 167973, 2021.

[62] M. S. Ismail, M. Moghavvemi, and T. M. I. Mahlia, "Characterization of PV panel and global optimization of its model parameters using genetic algorithm," Energy Convers. Manag., vol. 73, pp. 10-25, 2013.

[63] A. M. Dizqah, A. Maheri, and K. Busawon, "An accurate method for the PV model identification based on a genetic algorithm and the interior-point method," Renew. Energy, vol. 72, pp. 212-222, 2014.

[64] N. Hamid, R. Abounacer, M. Idali Oumhand, et al., "Parameters identification of photovoltaic solar cells and module using the genetic algorithm with convex combination crossover," Int. J. Ambient Energy, vol. 40, no. 5, pp. 517-524, 2019.

[65] Z. Liao, Q. Gu, S. Li, et al., "An improved differential evolution to extract photovoltaic cell parameters," IEEE Access, vol. 8, pp. 177838–177850, 2020.

[66] J. Zhou, Y. Zhang, Y. Zhang, et al., "Parameters identification of photovoltaic models using a differential evolution algorithm based on elite and obsolete dynamic learning," Appl. Energy, vol. 314, p. 118877, 2022.

[67] D. Simon, "Biogeography-based optimization," IEEE Trans. Evol. Comput., vol. 12, no. 6, pp.

702-713, 2008.

[68] X. Chen and K. Yu, "Hybridizing cuckoo search algorithm with biogeography-based optimization for estimating photovoltaic model parameters," Sol. Energy, vol. 180, pp. 192-206, 2019.

[69] S. Li, W. Gong, L. Wang, et al., "A hybrid adaptive teaching–learning-based optimization and differential evolution for parameter identification of photovoltaic models," Energy Convers. Manag., vol. 225, p. 113474, 2020.

[70] X. Chen, K. Yu, W. Du, et al., "Parameters identification of solar cell models using generalized oppositional teaching learning based optimization," Energy, vol. 99, pp. 170-180, 2016.

[71] M. Ćalasan, S. H. E. A. Aleem, and A. F. Zobaa, "On the root mean square error (RMSE) calculation for parameter estimation of photovoltaic models: A novel exact analytical solution based on Lambert W function," Energy Convers. Manag., vol. 210, p. 112716, 2020.

[72] K. Yu, X. Chen, X. Wang, et al., "Parameters identification of photovoltaic models using self-adaptive teaching-learning-based optimization," Energy Convers. Manag., vol. 145, pp. 233-246, 2017.

[73] R. V. Rao, V. J. Savsani, et D. P. Vakharia, "Teaching–learning-based optimization: an optimization method for continuous non-linear large scale problems," Inf. Sci., vol. 183, no. 1, pp. 1-15, 2012.

[74] S. Li, W. Gong, X. Yan, et al., "Parameter extraction of photovoltaic models using an improved teaching-learning-based optimization," Energy Convers. Manag., vol. 186, pp. 293-305, 2019.

[75] M. Naeijian, A. Rahimnejad, S. M. Ebrahimi, et al., "Parameter estimation of PV solar cells and modules using Whippy Harris Hawks Optimization Algorithm," Energy Rep., vol. 7, pp. 4047-4063, 2021.

[76] B. Aoufi, “Contribution to the development of a fault diagnosis technique for photovoltaic generators,” PhD thesis in automatic, university of Ouargla, Algeria, 2024.

[77] A. A. Abd El-Mageed, A. A. Abohany, H. M. H. Saad, et al., "Parameter extraction of solar photovoltaic models using queuing search optimization and differential evolution," Appl. Soft Comput., vol. 134, p. 110032, 2023.

[78] M. Qaraad, S. Amjad, N. K. Hussein, et al., "Photovoltaic parameter estimation using improved moth flame algorithms with local escape operators," Comput. Electr. Eng., vol. 106, p. 108603, 2023.

[79] Y. Li, K. Yu, J. Liang, et al., "A landscape-aware particle swarm optimization for parameter identification of photovoltaic models," Appl. Soft Comput., vol. 131, p. 109793, 2022.

[80] M. Premkumar, T. S. Babu, S. Umashankar, et al., "A new metaphor-less algorithm for the photovoltaic cell parameter estimation," Optik, vol. 208, p. 164559, 2020.

[81] C. Kumar, T. D. Raj, M. Premkumar, et al., "A new stochastic slime mould optimization algorithm for the estimation of solar photovoltaic cell parameters," Optik, vol. 223, p. 165277, 2020.

[82] K. Yu, J. J. Liang, B. Y. Qu, et al., "Multiple learning backtracking search algorithm for estimating parameters of photovoltaic models," Applied Energy, vol. 226, pp. 408-422, 2018.

[83] A. R. Jordehi, "Enhanced leader particle swarm optimisation (ELPSO): An efficient algorithm for parameter estimation of photovoltaic (PV) cells and modules," Solar Energy, vol. 159, pp. 78-87, 2018.

[84] G. Xiong, J. Zhang, D. Shi, et al., "Application of supply-demand-based optimization for parameter extraction of solar photovoltaic models," Complexity, vol. 2019, no. 1, p. 3923691, 2019.

[85] S. Mekhilef, R. Saidur, and M. Kamalisarvestani, "Effect of dust, humidity and air velocity on efficiency of photovoltaic cells," Renew. Sustain. Energy Rev., vol. 16, no. 5, pp. 2920-2925, 2012.

[86] A. R. Amelia, Y. M. Irwan, W. Z. Leow, et al., "Investigation of the effect temperature on photovoltaic (PV) panel output performance," Int. J. Adv. Sci. Eng. Inf. Technol., vol. 6, no. 5, pp. 682-688, 2016.

[87] P. Mohanty and A. Tyagi, “Introduction to solar photovoltaic technology,” in Food, Energy, and Water, M. R. Opp, Ed. Amsterdam, Netherlands: Elsevier, 2015, pp. 309–348.

[88] L. Chaar and M. H. Rashid, “Photovoltaic system conversion,” in Power Electronics Handbook, M. H. Rashid, Ed. Oxford, U.K.: Elsevier, 2011, pp. 711–721.

[89] "Corrosion 101: What is Corrosion?" EonCoat, [Online]. Available: https://eoncoat.com-/corrosion-101-what-is-corrosion/, [Accessed: 17-Jun-2022].

[90] M. T. Chaichan and H. A. Kazem, Generating Electricity Using Photovoltaic Solar Plants in Iraq. Cham, Switzerland: Springer International Publishing, 2018.

[91] J.-H. Kim, J. Park, D. Kim, et al., “Study on mitigation method of solder corrosion for crystalline silicon photovoltaic modules,” Int. J. Photoenergy, vol. 2014, no. 1, p. 809075, 2014.

[92] C. C. C. Okorieimoh, B. Norton, and M. Conlon, "The Effects of the Transient and Performance Loss Rates on PV Output Performance," in Proc. IEECP – SCI-INDEX, Technological University Dublin, Dublin, Ireland, Jun. 2021.

[93] J. Larrivee, "An analysis of degradation rates of PV power plants at the system level," M.Sc thesis in Energy Science, Utrecht University, Utrecht, The Netherlands, Apr. 2013.

[94] L. O. Idoko, "Enhanced PV module efficiency to boost hybrid microgrid power and microgrid ancillary services market creation," Ph.D. thesis in Electronic and Electrical Engineering., Univ. of Strathclyde, Glasgow, UK, Oct. 2018.

[95] M. Köntges, S. Kajari-Schröder, I. Kunze, et al., “Crack statistic of crystalline silicon photovoltaic modules,” in Proc. 26th Eur. Photovolt. Solar Energy Conf. Exhib., Hamburg, Germany, 2011, pp. 3290-3294.

[96] M. Dhimish, V. Holmes, B. Mehrdadi, et al., “The impact of cracks on photovoltaic power performance,” J. Sci.: Adv. Mater. Devices, vol. 2, no. 2, pp. 199-209, 2017.

[97] N. Adamovic, A. Zimmermann, N. Severino, et al., “SolarDesign Vision: On-the-fly alterable thin-film solar modules for design driven applications,” in 10th International Conference on Nanosciences@ Nanotechnologies (NN13)-Abstract Book, 2013.

[98] R. C. Neville, Solar Energy Conversion: The Solar Cell, 1st ed. Amsterdam, Netherlands: Elsevier, 1995.

[99] J. Sykes, "Sizing inverters to optimise solar panel system efficiency" [Online]. Available: https://www.solarchoice.net.au/blog/optimizing-solar-panel-systemefficien cy-through-inverter-sizing/. [Accessed: 18-March-2024].

[100] M. Fedkin and J. A, "Efficiency of Inverters | Utility Solar Power and Concentration." [Online]. Available: https://www.e-education.psu.edu/eme812/node/738. [Accessed: 19-Jul-2020].

[101] J. Good and J. X. Johnson, "Impact of inverter loading ratio on solar photovoltaic system performance," Appl. Energy, vol. 177, pp. 475–486, 2016.

[102] W. Leader, T. Tsoutsos, Z. Gkouskos, et al., "Catalogue of common failures and improper practices on PV installations and maintenance," Eur. Photovolt. Ind. Assoc., pp. 1–18, 2011.

[103] Keysight Technologies, "Installation and Maintenance of Solar Photovoltaic Systems Using handheld and benchtop solutions for solar energy applications", available: https://keysight.zinfi.net/concierge/OEMs/keysight/wwwcontent/Attachments/pdf /5992-1004.pdf, [Accessed: 14-Feb-2019].

[104] I. Bodnár, B. Csehi, B. Sukály, et al., "Examination of power loss and voltage drop of a solar panel as a function of environmental factors," in Proc. 20th Int. Carpathian Control Conf. (ICCC), Kraków-Wieliczka, Poland, IEEE, 2019, pp. 1-5.

[105] S. Ekici and M. A. Kopru, "Investigation of PV system cable losses," Int. J. Renew. Energy Res. (IJRER), vol. 7, no. 2, pp. 807-815, 2017.

[106] A. Massi Pavan, A. Mellit, D. De Pieri, et al., "A study on the mismatch effect due to the use of different photovoltaic modules classes in large-scale solar parks," Prog. Photovolt. Res. Appl., vol. 22, no. 3, pp. 332-345, 2014.

[107] L. O. Idoko, "Enhanced PV module efficiency to boost hybrid microgrid power and microgrid ancillary services market creation," Ph.D. thesis in Electronic and Electrical Engineering., Univ. of Strathclyde, Glasgow, UK, Oct. 2018.

[108] M. J. Adinoyi and S. A. M. Said, "Effect of dust accumulation on the power outputs of solar

photovoltaic modules," Renew. Energy, vol. 60, pp. 633-636, 2013.

[109] A. Karafil, H. Ozbay, and M. Kesler, "Temperature and solar radiation effects on photovoltaic panel power," J. New Results Sci., vol. 5, pp. 48-58, 2016.

[110] M. J. B. Buni, A. K. Al-Walie, and K. A. Al-Asadi, "Effect of solar radiation on photovoltaic cell," Int. Res. J. Adv. Eng. Sci., vol. 3, no. 3, pp. 47-51, 2018.

[111] H. K. Elminir, V. Benda, and J. Tousek, "Effects of solar irradiation conditions and other factors on the outdoor performance of photovoltaic modules," J. Electr. Eng.-Bratislava, vol. 52, no. 5/6, pp. 125-133, 2001.

[112] J. K. Kaldellis, M. Kapsali, and K. A. Kavadias, "Temperature and wind speed impact on the efficiency of PV installations: Experience obtained from outdoor measurements in Greece," Renew. Energy, vol. 66, pp. 612-624, 2014.

[113] B. V. Chikate, Y. Sadawarte, and B. D. Sewagram, "The factors affecting the performance of solar cell," Int. J. Comput. Appl., vol. 1, no. 1, pp. 15-20, 2015.

[114] L. Xu, W. Liu, H. Liu, et al., "Heat generation and mitigation in silicon solar cells and modules," Joule, vol. 5, no. 3, pp. 631-645, 2021.

[115] Y. Baghzouz, "'EE 446/646 Photovoltaic Devices I.' [Online]. Available: http://www.egr.unlv.edu/~eebag/Photovoltaic Devices I.pdf. [Accessed: 12-Jun-2022].

[116] N. Gupta and M. Khatri, "An experimental analysis of shading on solar photovoltaic system," Int. J. Emerg. Technol., vol. 8, no. 1, pp. 226-228, 2017.

[117] A. M. Homadi, "Effect of elevation and wind direction on silicon solar panel efficiency," Int. J. Energy Power Eng., vol. 10, no. 9, pp. 1205-1212, 2016.

[118] M. A. A. Mamun, M. Hasanuazzaman, and J. Selvaraj, "Impact of tilt angle on the performance of photovoltaic modules in Malaysia: A review," in Proc. 4th IET Clean Energy and Technology Conf. (CEAT 2016), Kuala Lumpur, Malaysia: IET, 2016, pp. 1-8.

[119] A. Drews, A. C. de Keizer, H. G. Beyer, et al., "Monitoring and remote failure detection of grid-connected PV systems based on satellite observations," Sol. Energy, vol. 81, no. 4, pp. 548-564, 2007.

[120] A. C. Cohen, "Reliability, Availability and Maintainability (RAM) Dictionary," J. Qual. Technol., vol. 21, no. 2, pp. 105–110, Apr. 1989.

[121] S. Stettler, P. Toggweiler, E. Wiemken, et al., "Failure detection routine for grid-connected PV systems as part of the PVSAT-2 project," in Proc. 20th Eur. Photovolt. Sol. Energy Conf. & Exhib., Barcelona, Spain, 2005, pp. 2490–2493.

[122] S. K. Firth, K. J. Lomas, and S. J. Rees, "A simple model of PV system performance and its use in fault detection," Sol. Energy, vol. 84, no. 4, pp. 624–635, 2010.

[123] A. Chouder and S. Silvestre, "Automatic supervision and fault detection of PV systems based

on power losses analysis," Energy Convers. Manag., vol. 51, no. 10, pp. 1929–1937, 2010.

[124] T. Mishina, H. Kawamura, S. Yamanaka, et al., "A study of the automatic analysis for the IV curves of a photovoltaic subarray," in Conf. Rec. 29th IEEE Photovolt. Spec. Conf., New Orleans, LA, USA, 2002, pp. 1630–1633.

[125] A. Mellit and S. Kalogirou, Handbook of Artificial Intelligence Techniques in Photovoltaic Systems: Modeling, Control, Optimization, Forecasting and Fault Diagnosis. Cambridge, MA, USA: Academic Press, 2022.

[126] V. Puig and S. Simani, Eds., "Diagnosis and Fault-tolerant Control 1: Data-driven and Model-based Fault Diagnosis Techniques," Hoboken, NJ, USA: John Wiley & Sons, 2021.

[127] R. Mahalakshmi, M. Karuppasamypandiyan, A. Bhuvanesh, et al., "Classification and detection of faults in grid connected photovoltaic system," Int. J. Sci. Eng. Res., vol. 7, pp. 149-154, 2016.

[128] G. M. El-Banby, N. M. Moawad, B. A. Abouzalm, et al., "Photovoltaic system fault detection techniques: a review," Neural Comput. Appl., vol. 35, no. 35, pp. 24829-24842, 2023.

[129] M. Köntges, S. Kajari-Schröder, and I. Kunze, "Cell cracks measured by UV fluorescence in the field," in Proc. 27th EU PVSEC, Frankfurt, Germany, 2012, pp. 3033-3040.

[130] G. M. El-Banby, N. M. Moawad, B. A. Abouzalm, et al., "Photovoltaic system fault detection techniques: a review," Neural Comput. Appl., vol. 35, no. 35, pp. 24829-24842, 2023.

[131] M. W. Akram, G. Li, Y. Jin, et al., "Automatic detection of photovoltaic module defects in infrared images with isolated and develop-model transfer deep learning," Sol. Energy, vol. 198, pp. 175-186, 2020.

[132] M. A. El-Rashidy, "An efficient and portable solar cell defect detection system," Neural Comput. Appl., vol. 34, no. 21, pp. 18497-18509, 2022.

[133] A. W. Kandeal, M. R. Elkadeem, A. K. Thakur, et al., "Infrared thermography-based condition monitoring of solar photovoltaic systems: A mini review of recent advances," Sol. Energy, vol. 223, pp. 33-43, 2021.

[134] X. Maldague, F. Galmiche, and A. Ziadi, "Advances in pulsed phase thermography," Infrared Phys. Technol., vol. 43, no. 3-5, pp. 175-181, 2002.

[135] A. Sinha, O. S. Sastry, and R. Gupta, "Detection and characterisation of delamination in PV modules by active infrared thermography," Nondestruct. Test. Eval., vol. 31, no. 1, pp. 1-16, 2016.

[136] J. Rantala, D. Wu, and G. Busse, "Amplitude-modulated lock-in vibrothermography for NDE of polymers and composites," Res. Nondestruct. Eval., vol. 7, pp. 215-228, 1996.

[137] R. Pierdicca, M. Paolanti, A. Felicetti, et al., "Automatic faults detection of photovoltaic farms: solAIr, a deep learning-based system for thermal images," Energies, vol. 13, no. 24, p. 6496, 2020.

[138] S. Fadhel, C. Delpha, D. Diallo, et al., "PV shading fault detection and classification based on IV curve using principal component analysis: Application to isolated PV system," Solar Energy, vol. 179, pp. 1-10, 2019.

[139] D. Korkmaz and H. Acikgoz, "An efficient fault classification method in solar photovoltaic modules using transfer learning and multi-scale convolutional neural network," Eng. Appl. Artif. Intell., vol. 113, p. 104959, 2022.

[140] D. Manno, G. Cipriani, G. Ciulla, et al., "Deep learning strategies for automatic fault diagnosis in photovoltaic systems by thermographic images," Energy Convers. Manag., vol. 241, p. 114315, 2021.

[141] M. N. Akram and S. Lotfifard, "Modeling and health monitoring of DC side of photovoltaic array," IEEE Trans. Sustain. Energy, vol. 6, no. 4, pp. 1245-1253, 2015.

[142] Z. Yi and A. H. Etemadi, "Line-to-line fault detection for photovoltaic arrays based on multiresolution signal decomposition and two-stage support vector machine," IEEE Trans. Ind. Electron., vol. 64, no. 11, pp. 8546-8556, 2017.

[143] Z. Chen, L. Wu, S. Cheng, et al., "Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and IV characteristics," Appl. Energy, vol. 204, pp. 912-931, 2017.

[144] S. Ahmad, N. Hasan, V. S. B. Kurukuru, et al., "Fault classification for single phase photovoltaic systems using machine learning techniques," in Proc. 8th IEEE India Int. Conf. Power Electron. (IICPE), Jaipur, India, 2018, pp. 1–6.

[145] Z. Chen, Y. Chen, L. Wu, et al., "Deep residual network based fault detection and diagnosis of photovoltaic arrays using current-voltage curves and ambient conditions," Energy Convers. Manag., vol. 198, p. 111793, 2019.

[146] S. Samara and E. Natsheh, "Intelligent real-time photovoltaic panel monitoring system using artificial neural networks," IEEE Access, vol. 7, pp. 50287–50299, 2019.

[147] F. Aziz, A. U. Haq, S. Ahmad, et al., "A novel convolutional neural network-based approach for fault classification in photovoltaic arrays," IEEE Access, vol. 8, pp. 41889–41904, 2020.

[148] J. D. De Guia, R. S. Concepcion, H. A. Calinao, et al., "Application of ensemble learning with mean shift clustering for output profile classification and anomaly detection in energy production of grid-tied photovoltaic system," in Proc. 2020 12th Int. Conf. Inf. Technol. Electr. Eng. (ICITEE),Yogyakarta, Indonesia, 2020, pp. 286–291.

[149] A. Eskandari, J. Milimonfared, and M. Aghaei, "Line-line fault detection and classification for photovoltaic systems using ensemble learning model based on IV characteristics," Sol. Energy, vol. 211, pp. 354–365, 2020.

[150] K. Dhibi, M. Mansouri, K. Bouzrara, et al., "An enhanced ensemble learning-based fault

detection and diagnosis for grid-connected PV systems," IEEE Access, vol. 9, pp. 155622–155633, 2021.

[151] C. Kapucu and M. Cubukcu, "A supervised ensemble learning method for fault diagnosis in photovoltaic strings," Energy, vol. 227, p. 120463, 2021.

[152] E. Lodhi, F.-Y. Wang, G. Xiong, et al., "An AdaBoost ensemble model for fault detection and classification in photovoltaic arrays," IEEE J. Radio Freq. Identif., vol. 6, pp. 794–800, 2022.

[153] N.-C. Yang and H. Ismail, "Voting-based ensemble learning algorithm for fault detection in photovoltaic systems under different weather conditions," Mathematics, vol. 10, no. 2, p. 285, 2022.

[154] A. Mellit and S. Boubaker, "An effective ensemble learning method for fault diagnosis of photovoltaic arrays," in Proc. Int. Conf. Electronic Engineering and Renewable Energy Systems, Singapore: Springer Nature Singapore, 2022, pp. 687–695.

[155] D. Adhya, S. Chatterjee, and A. K. Chakraborty, "Performance assessment of selective machine learning techniques for improved PV array fault diagnosis," Sustain. Energy Grids Netw., vol. 29, p. 100582, 2022.

[156] A. Sarwar, M. Ali, J. Manhas, et al., "Diagnosis of diabetes type-II using hybrid machine learning based ensemble model," Int. J. Inf. Technol., vol. 12, pp. 419–428, 2020.

[157] P. Branco, L. Torgo, and R. P. Ribeiro, "A survey of predictive modeling on imbalanced domains," ACM Comput. Surv. (CSUR), vol. 49, no. 2, pp. 1–50, 2016.

[158] S. E. A. Ashri, M. M. El-Gayar, and E. M. El-Daydamony, "HDPF: Heart disease prediction framework based on hybrid classifiers and genetic algorithm," IEEE Access, vol. 9, pp. 146797–146809, 2021.

[159] R. Jani, M. S. I. Shanto, M. M. Kabir, et al., "Heart disease prediction and analysis using ensemble architecture," in Proc. 2022 Int. Conf. Decision Aid Sci. Appl. (DASA), Dhaka, Bangladesh, 2022, pp. 1386–1390.

[160] P. Mahajan, S. Uddin, F. Hajati, et al., "Ensemble learning for disease prediction: A review," Healthcare, vol. 11, no. 12, p. 1808, 2023.

[161] M. D. Basar and A. Akan, "Detection of chronic kidney disease by using ensemble classifiers," in Proc. 2017 10th Int. Conf. Electr. Electron. Eng. (ELECO), Bursa, Turkey, 2017, pp. 544–547.

[162] O. A. Jongbo, A. O. Adetunmbi, R. B. Ogunrinde, et al., "Development of an ensemble approach to chronic kidney disease diagnosis," Sci. Afr., vol. 8, p. e00456, 2020.

[163] D. S. Sisodia and A. Verma, "Prediction performance of individual and ensemble learners for chronic kidney disease," in Proc. 2017 Int. Conf. Inventive Comput. Inform. (ICICI), Coimbatore, India, 2017, pp. 1027–1031.

[164] D. Ramesh and Y. S. Katheria, "Ensemble method based predictive model for analyzing disease datasets: a predictive analysis approach," Health Technol., vol. 9, pp. 533–545, 2019.

[165] M. Pal and B. R. Roy, "Evaluating and enhancing the performance of skin disease classification based on ensemble methods," in Proc. 2020 2nd Int. Conf. Adv. Inf. Commun. Technol. (ICAICT), Dhaka, Bangladesh, 2020, pp. 439–443.

[166] E. C. Igodan, A. F.-B. Thompson, O. Obe, et al., "Erythemato Squamous Disease prediction using ensemble multi-feature selection approach," Int. J. Comput. Sci. Inf. Secur. (IJCSIS), vol. 20, pp. 95–106, 2022.

[167] N. Nahar, F. Ara, M. A. I. Neloy, et al., "A comparative analysis of the ensemble method for liver disease prediction," in Proc. 2019 2nd Int. Conf. Innov. Eng. Technol. (ICIET), Dhaka, Bangladesh, 2019, pp. 1-6.

[168] N. Singh and P. Singh, "A stacked generalization approach for diagnosis and prediction of type 2 diabetes mellitus," in Computational Intelligence in Data Mining: Proceedings of the International Conference on ICCIDM 2018, Singapore., Springer, 2020, pp. 559-570.

[169] V. Shorewala, "Early detection of coronary heart disease using ensemble techniques," Informatics Med. Unlock., vol. 26, p. 100655, 2021.

[170] I. D. Mienye and Y. Sun, "A survey of ensemble learning: Concepts, algorithms, applications, and prospects," IEEE Access, vol. 10, pp. 99129–99149, 2022.

[171] L. Muflikhah, N. Widodo, W. F. Mahmudy, et al., "Prediction of Liver Cancer Based on DNA Sequence Using Ensemble Method," in Proc. 2020 3rd Int. Seminar Res. Inf. Technol. Intell. Syst. (ISRITI), Yogyakarta, Indonesia, Dec. 2020, pp. 37–41.

[172] S. S. Ghoneim, A. E. Rashed, and N. I. Elkalashy, "Fault detection algorithms for achieving service continuity in photovoltaic farms," Intell. Autom. Soft Comput., vol. 30, no. 2, pp. 467–479, 2021.

[173] A. R. Rashed, “Fault detection dataset in photovoltaic farms,” Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/amrezzeldinrashed/fault-detection-dataset-in-photovoltaic-farms. [Accessed: Apr. 4, 2024].

[174] K. Kaur and S. K. Mittal, "WITHDRAWN: Classification of mammography image with CNN-RNN based semantic features and extra tree classifier approach using LSTM," Materials Today: Proceedings, vol. 27, pp. 1234-1241, 2020.

[175] T. M. Hossain, J. Wataada, M. Hermana, et al., "Supervised machine learning in electrofacies classification: A rough set theory approach," J. Phys.: Conf. Ser., vol. 1521, p. 052048, 2020.

[176] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," Machine Learning, vol. 63, pp. 3–42, 2006.

[177] A. Sarwar, M. F. Mushtaq, U. Akram, et al., "IoT networks attacks detection using multi-novel

features and extra tree random-voting ensemble classifier (ER-VEC)," J. Ambient Intell. Humanized Comput., vol. 14, no. 12, pp. 16637–16651, 2023.

[178] Y. León-Ruiz, M. González-García, R. Álvarez-Salas, et al., "Fault diagnosis based on machine learning for the high frequency link of a grid-tied photovoltaic converter for a wide range of irradiance conditions," IEEE Access, vol. 9, pp. 151209–151220, 2021.

[179] A. C. Atkinson, "Regression diagnostics, transformations and constructed variables," J. Roy. Stat. Soc. Ser. B (Methodol.), vol. 44, no. 1, pp. 1–22, 1982.

[180] F. S. de Menezes, G. R. Liska, M. A. Cirillo, et al., "Data classification with binary response through the Boosting algorithm and logistic regression," Expert Syst. Appl., vol. 69, pp. 62–73, 2017.

[181] D. W. Hosmer Jr., S. Lemeshow, and R. X. Sturdivant, Applied Logistic Regression, 3rd ed. Hoboken, NJ, USA: John Wiley & Sons, 2013.

[182] N. Ghaffarzadeh and A. Azadian, "A comprehensive review and performance evaluation in solar (PV) systems fault classification and fault detection techniques," J. Solar Energy Res., vol. 4, no. 4, 2018.

[183] D.-C. Feng, Z.-T. Liu, X.-D. Wang, et al., "Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach," Constr. Build. Mater., vol. 230, p. 117000, 2020.

[184] R. E. Schapire, "Explaining AdaBoost," in Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 37-52.

[185] G. M. Toche Tchio, J. Kenfack, D. Kassegne, et al., "A comprehensive review of supervised learning algorithms for the diagnosis of photovoltaic systems, proposing a new approach using an ensemble learning algorithm," Appl. Sci., vol. 14, no. 5, p. 2072, 2024.

List of publication and conference paper

[P1] **N. Rai**, A. Abbadi, F. Hamidia, N. Douifi, B. Abdul Samad, and K. Yahya, "Biogeography-Based Teaching Learning-Based Optimization Algorithm for Identifying One Diode, Two-Diode and Three-Diode Models of Photovoltaic Cell and Module", Mathematics, 11(8), 1861, 2023.

[P2] N. Douifi, A. Abbadi, F. Hamidia, K. Yahya, M. Mohamed, and **N. Rai, "**A Novel MPPT Based Reptile Search Algorithm for Photovoltaic System under Various Conditions", Applied Sciences, 13(8), 4866, 2023.

[C1] N. Douifi, A. Abbadi, F. Hamidia, **N. Rai,** and B. KANOUNI, "Enhancing PV System Efficiency under Partial Shading with the Dragonfly Optimization Algorithm", Journée d'Etude sur l'Exploitation des Energies Renouvelables pour la Sécurité Energétique de l'Algérie, J3ERSEA'24, Médéa-Algérie, 10 Juin 2024"

[C2] N. Douifi, A. Abbadi, F. Hamidia, **N. Rai,** "Hybrid GWO-PSO based MPPT algorithm for photovoltaic system under uniform and non-uniform shading Conditions", 3rd Conference on Electrical Engineering (CEE'2023), Military school Polytechnique, Algiers May 09-10, 2023.

[C3] A. KIFOUCHE, T. Naoual, D. Lalmi, O. Bayasli, **N. Rai**, A. Guessoum," Tracking Solar Radiation Using Recurrent Neural Network", the National Conference On Process Engineering & Renewable Energy - NCPERE'2023, at the Faculty of Sciences and Technology, University of Djilali Bounaâma Khemis Miliana - Algeria on May 3rd, 2023.

[C4] **N. Rai**, A. Abbadi, F. Hamidia, B. Kanouni, and A. Kahlessenane, "A New Modified Incremental Conductance Algorithm Used for PV System", In Artificial Intelligence and Heuristics for Smart Energy Efficiency in Smart Cities: Case Study: Tipasa, Algeria pp. 256-265, 2022.

[C5] N. Douifi, A. Abbadi, F. Hamidia, **N. Rai,** "PSO MPPT controller for grid connected PV system under partial shading conditions», Algerian Symposium on Renewable Energy and Materials (ASREM2022), Medea, Algeria, March 16-17, 2022

[C6] **N. Rai**, A. Abbadi, F. Hamidia, "Parameters extraction of three-diode photovoltaic model using differential evolution algorithm", IC-SEAM'21 April 21-22, 2021, Ouargla, ALGERIA.

Annex

Table 01: The measurement voltage and current of RTC France silicon solar cell were taken from a 57 mm diameter RTC France silicon solar cell at 33°C and 1000W/m^2 solar irradiance.

Item	V_m	I_m
1	-0,2057	0,764
2	-0,1291	0,762
3	-0,0588	0,7605
4	0,0057	0,7605
5	0,0646	0,76
6	0,1185	0,759
7	0,1678	0,757
8	0,2132	0,757
9	0,2545	0,7555
10	0,2924	0,754
11	0,3269	0,7505
12	0,3585	0,7465
13	0,3873	0,7385
14	0,4137	0,728
15	0,4373	0,7065
16	0,459	0,6755
17	0,4784	0,632
18	0,496	0,573
19	0,5119	0,499
20	0,5265	0,413
21	0,5398	0,3165
22	0,5521	0,212
23	0,5633	0,1035
24	0,5736	-0,01
25	0,5833	-0,123
26	0,59	-0,21

Table 02: The measurement voltage and current of the solar module STM6-40/36 were taken under the cell temperature of 51°C and solar irradiation of 1000 W/m^2.

Item	V_m	I_m
1	1,663	0
2	1,663	0,118
3	1,661	2,237
4	1,653	5,434
5	1,65	7,26
6	1,645	9,68
7	1,64	11,59
8	1,636	12,6
9	1,629	13,37
10	1,619	14,09
11	1,597	14,88
12	1,581	15,59
13	1,542	16,4
14	1,524	16,71
15	1,5	16,98
16	1,485	17,13
17	1,465	17,32
18	1,388	17,91
19	1,118	19,08
20	0	21,02