الجمهورية الجزائرية الديموقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المدرسة الوطنية العليا للتكنولوجيا المتقدمة
École Nationale Supérieure des Technologies Avancées
قسم الهندسة الكهربائية و الإعلام الآلي الصناعي

Department of Electrical Engineering and Industrial Computer Science

Final Year Project to Obtain the Diploma of Engineering

Field : **Electronics** Speciality :**Embedded Systems**

Design and Implementation of a Real-Time Monitoring and Supervision System for a Grid-Connected Photovoltaic Power Plant

Realised by :

Chahine KHALED

Abderrahim ADDALA

Full Name	University	Grade	Quality
Ms. Karima BOULAAM	ENSTA	MCB	President
Mr. M.Lassaad BENZAOUI	ENSTA	MAA	Examinator
Ms. Lamia HABANI	ENSTA	MAA	Examinator
Mr. Reda DERMOUCHE	ENSTA	MCA	Supervisor
Mr. Aissa CHOUDER	U.M'sila	Professor	Supervisor
Ms. Imane DROUICHE	ENSTA	MCB	Supervisor

Academic year : 2024 - 2025

I dedicate this work:

To Allah
, the Most Gracious, the Most Merciful,
whose guidance and blessings illuminate every path.

To my father
, who is part of this in every way.
Without him, none of it would have been.

To my family
, for their good words and endless support.

To my dear friend
, whose positivity brings a smile even in the most difficult moments.

To **Addala Abderrahim**,
for his unwavering commitment, sharp insight, and dedication throughout every stage of this project.

And to the ones I adore and whose company brings me joy,
Thank you for adding light to this path.

Khaled Chahine

I dedicate this work:

To Allah,
the Source of all knowledge and wisdom,
whose grace has illuminated our path and sustained our efforts,
from the beginning to the end of this work.

To myself,
for the hard work and determination that brought me here.
I am proud and grateful.

To my beloved father,
for your constant support and encouragement.
This work reflects how much I appreciate and love you.

To my loving mother,
for your endless support throughout my studies.
This work is a sign of my deepest gratitude.

To my dear siblings,
your love, support, and encouragement have been my strength.

To my friends,
thank you for your unwavering presence and belief in me.

To **Khaled Chahine**,
for his perseverance, technical skill, and constant support that helped bring this work to completion.

To all those I hold dear,
I am deeply grateful for your presence and support.
This work is for each of you, with heartfelt thanks.

Addala Abderrahim

Acknowledgement

First and foremost, we would like to express our heartfelt gratitude to Professor Aissa Chouder, our supervisor, for his unwavering support, insightful guidance, and continuous encouragement throughout this project. His expertise and academic rigor greatly enriched our work and motivated us to maintain a high standard of quality.

We are also sincerely thankful to Professor Houcine Oudira and to the other professors of the University of Msila for their generous support and thoughtful involvement throughout our project. We are also grateful to Mr. Ahmed and Mr. Moadh, dedicated staff members at the photovoltaic station, for their assistance and professionalism during our field visits. Their practical support and availability played a vital role in the completion of this work.

Our sincere appreciation goes to Mr. Dermouche from ENSTA, His dedication to students and constant support made him a truly exceptional mentor, leaving a lasting impact on our experience.We also thank our teachers for their support and encouragement throughout this journey. Their guidance helped shape our thinking and deepen our understanding.

To all who contributed to this work, directly or indirectly, we are truly grateful.

Khaled Chahine & Addala Abderrahim

Abstract

This End-of-Study Project presents the design and implementation of a real-time monitoring and supervision system tailored for large-scale grid-connected photovoltaic systems. As the integration of renewable energy into power systems grows, maintaining consistent performance and ensuring operational reliability becomes essential. The work addresses the need for a robust and centralized platform—based on SCADA architecture—that enables continuous oversight of photovoltaic systems, enabling early detection of irregularities, efficient decision-making, and enhanced energy yield.

The proposed system facilitates real-time data acquisition and visualization, giving operators access to crucial information such as energy production levels, system efficiency, and the status of system components. It further supports predictive analysis, and fault detection and diagnosis, enabling timely intervention and minimizing system downtime. By aligning the system's architecture with established performance metrics and supervisory principles, the study ensures that operators can both monitor and evaluate the plant's behaviour in real-time.

Furthermore, the work emphasizes the significance of modular architecture, accessibility, and clarity in monitoring tools, ensuring they serve as decision-support systems rather than passive displays. This work ultimately contributes to improving the management of photovoltaic power plants by offering a scalable and user-focused solution that enhances operational awareness, responsiveness, and long-term energy performance.

Keywords: Grid-Connected Photovoltaic Systems, Real-Time Monitoring, SCADA, Visualization, Predictive Analysis, Fault Detection and Diagnosis, Operational Reliability.

ملخص

يعرض هذا المشروع تصميم وتنفيذ نظام مراقبة وإشراف فوري مخصص لمحطات الطاقة الشمسية الكهروضوئية الكبرى المرتبطة بالشبكة. مع التوسع في إدماج مصادر الطاقة المتجددة ضمن الشبكات الكهربائية، تبرز الحاجة إلى الحفاظ على أداء مستقر وضمان موثوقية التشغيل. يستجيب هذا العمل إلى هذا التحدي من خلال تطوير منصة مركزية وموثوقة مبنية على معمارية SCADA، تمكّن من المراقبة المستمرة للنظام الكهروضوئي، مما يسمح بالكشف المبكر عن الاضطرابات، وتحسين اتخاذ القرار، وزيادة المردودية الطاقوية.

يُسهل النظام المقترح عملية اقتناء البيانات وعرضها في الوقت الحقيقي، مما يوفر للمشغلين معلومات حيوية مثل مستويات إنتاج الطاقة، كفاءة النظام، وحالة مكونات المحطة. كما يدعم التحليل التنبؤي، واكتشاف الأعطال وتشخيصها، مما يُمكّن من التدخل السريع وتقليل التوقفات. ومن خلال التوافق مع مؤشرات الأداء ومعايير الإشراف المعتمدة، يضمن هذا النظام مراقبة دقيقة وتقييمًا فعالًا لسلوك المحطة في الزمن الحقيقي.

كما يسلط هذا العمل الضوء على أهمية وضوح أدوات المراقبة وسهولة استخدامها وبنيتها النمطية، حتى تكون بمثابة أدوات دعم اتخاذ القرار، وليس مجرد واجهات عرض سلبية. يسهم هذا العمل في تحسين إدارة محطات الطاقة الشمسية الكهروضوئية من خلال تقديم حل قابل للتوسع يتمحور حول المستخدم، ويعزز الوعي التشغيلي، وسرعة الاستجابة، والأداء الطاقوي على المدى الطويل.

الكلمات المفتاحية: الأنظمة الكهروضوئية المرتبطة بالشبكة، المراقبة في الزمن الحقيقي، SCADA، التمثيل البياني، التحليل التنبؤي، اكتشاف وتشخيص الأعطال، الموثوقية التشغيلية.

Résumé

Ce projet de fin d'études présente la conception et la mise en œuvre d'un système de surveillance et de supervision en temps réel, adapté aux centrales photovoltaïques de grande taille raccordées au réseau. Avec l'intégration croissante des énergies renouvelables dans les réseaux électriques, il devient essentiel de maintenir une performance constante et d'assurer la fiabilité opérationnelle. Ce travail répond au besoin d'une plateforme centralisée et robuste — basée sur une architecture SCADA — permettant une supervision continue des systèmes photovoltaïques, en assurant la détection précoce des irrégularités, une prise de décision efficace et un meilleur rendement énergétique.

Le système proposé facilite l'acquisition de données en temps réel et leur visualisation, offrant aux opérateurs un accès à des informations essentielles telles que les niveaux de production d'énergie, l'efficacité du système et l'état des composants. Il prend également en charge l'analyse prédictive ainsi que la détection et le diagnostic des défauts, ce qui permet une intervention rapide et réduit les interruptions. En alignant l'architecture du système sur des indicateurs de performance reconnus et des principes de supervision établis, l'étude garantit que les opérateurs peuvent surveiller et évaluer le comportement de l'installation en temps réel.

En outre, ce travail met l'accent sur l'importance de l'accessibilité, de la clarté et de l'architecture modulaire des outils de supervision, afin qu'ils servent de véritables systèmes d'aide à la décision plutôt que de simples interfaces passives. Ce travail contribue ainsi à améliorer la gestion des centrales photovoltaïques en offrant une solution évolutive, centrée sur l'utilisateur, qui renforce la réactivité, la visibilité opérationnelle et la performance énergétique à long terme.

Mots-clés : Systèmes photovoltaïques raccordés au réseau, Surveillance en temps réel, SCADA, Visualisation, Analyse prédictive, Détection et diagnostic des défauts, Fiabilité opérationnelle.

Contents

General Introduction **1**

1 Photovoltaic system & Technology: State of the art **3**
1.1 **Introduction** . . . 4
1.2 Photovoltaic Systems . . . 4
1.2.1 Photovoltaic Cells and Modules . . . 4
1.2.2 Photovoltaic Module Technology . . . 6
1.2.3 Electrical Characteristics of PV Modules . . . 6
1.2.4 Strings and Arrays . . . 8
1.2.5 PV System Types . . . 9
1.2.5.1 Classification Based on Grid Connection . . . 9
1.2.5.2 Classification Based on System Scale . . . 11
1.3 Performance Analysis Method for Photovoltaic Systems . . . 13
1.3.1 System Performance Index under IEC 61724 Method . . . 13
1.3.2 Normalised Losses . . . 15
1.3.3 Performance Ratio (PR) . . . 15
1.3.4 Corrected Performance Ratio (CPR) . . . 16
1.4 SCADA Systems in Photovoltaic Power Plants . . . 17
1.5 **Conclusion** . . . 19

2 Operational Worksite — Ain El Melh Solar Power Plant **20**
2.1 **Introduction** . . . 21
2.2 Site description . . . 21
2.2.1 Installed Photovoltaic Power Plant Characteristics . . . 21
2.2.2 Main Zones of the PV Power Plant . . . 22
2.2.2.1 Stages of Energy Conversion and Flow . . . 23
2.2.2.2 Data Acquisition and Transmission . . . 24
2.3 Existing SCADA Interface . . . 27
2.3.1 Interface Pages . . . 27
2.3.2 Requirements and Assessment . . . 30
2.4 **Conclusion** . . . 31

3 SCADA Interface Development **32**
3.1 **Introduction** . . . 33
3.2 Communication Architecture . . . 33
3.2.1 Stage 1 Modbus RTU protocol . . . 34
3.2.2 Stage 2 Modbus RTU-to-TCP Conversion (Serial Device Manager) . . . 35
3.2.3 Stage 3 Request-Response Communication Mechanisms . . . 38
3.3 Design and Implementation of a Real-Time SCADA Server . . . 38
3.3.1 Back-End Data Processing . . . 39
3.3.1.1 Python-Based Data Acquisition Module . . . 41
3.3.1.2 Parameter Calculation Module . . . 43
3.3.1.3 Predictive Module . . . 44
3.3.2 Data Storage with MongoDB . . . 46
3.3.3 Application Server (Flask) . . . 47
3.3.4 Front-End and Visualization . . . 48
3.4 Data Server Emulation . . . 49
3.5 **Conclusion** . . . 52

4 Novel Deep Learning Prediction Model, Fault Detection, and Diagnosis in Grid-Connected Photovoltaic Systems **54**
4.1 **Introduction** . . . 55
4.2 Methodology . . . 55
4.3 Proposed Hybrid Deep Learning Model (CNN-LSTM) . . . 55
4.3.1 The Dataset . . . 55
4.3.2 Data Processing . . . 56
4.3.3 CNN-LSTM Hybrid Model . . . 56
4.3.4 Performance Evaluation Metrics . . . 58
4.3.5 Prediction Results . . . 58
4.4 Proposed Fault Detection and Diagnosis Methodology . . . 61
4.4.1 Fault Detection Technique . . . 61
4.4.2 Diagnosis Approach . . . 65
4.4.3 Fault Diagnosis Scenarios . . . 67
4.5 Standalone Python GUI for Prediction and Fault Diagnosis . . . 69
4.5.1 Page 1 - Prediction and Fault Detection Module . . . 70
4.5.2 Page 2 - Fault Diagnosis Module . . . 70
4.6 **Conclusion** . . . 71

5 Field Experimental Validation **72**
5.1 **Introduction** . . . 73
5.2 SCADA Interface Presentation . . . 73
5.3 Monitoring Interface . . . 74

5.3.1 General Monitoring Page 74
5.3.2 Inverter Monitoring Page 76
5.3.3 String Monitoring Page 79
5.3.4 Historical Logging Page 82
5.4 Analysis Interface 85
5.4.1 Home Page 86
5.4.2 Prediction and Fault Detection page 86
5.4.3 Diagnosis page 88
5.5 System Evaluation 89
5.6 **Conclusion** 89

General Conclusion **91**

List of Figures

Figure 1.1 PV cell 4
Figure 1.2 PV module 5
Figure 1.3 PV cell interconnection 5
Figure 1.4 Diagram of photovoltaic cell classified by invention technology 6
Figure 1.5 I-V curve of a PV module 7
Figure 1.6 Solar cell equivalent circuit 7
Figure 1.7 Schematic of PV string 8
Figure 1.8 Schematic of PV array configuration 9
Figure 1.9 Grid tied photovoltaic system 10
Figure 1.10 Off grid tied photovoltaic system 10
Figure 1.11 Hybrid PV system 11
Figure 1.12 Residential PV system 11
Figure 1.13 Commercial PV systems 12
Figure 1.14 Large scale PV system 12
Figure 1.15 Energy flow in a grid-connected photovoltaic system with parameters, efficiencies and losses 13
Figure 1.16 Weather corrected and uncorrected annual trend for a 24 MW facility . . 16
Figure 1.17 SCADA Architecture for Grid-Connected PV Plants 17

Figure 2.1 Site view of the Ain El Melh solar power plant 21
Figure 2.2 PV Power Plant connection to notional grid 22
Figure 2.3 General layout of the PV Power Plant 23
Figure 2.4 Energy flow: from field generation to grid injection via shelter and sub-station. 24
Figure 2.5 SCADA network 26
Figure 2.6 Site technical equipment 27
Figure 2.7 SCADA menu 28
Figure 2.8 SCADA mainMenu 29
Figure 2.9 SCADA inverter surveillance 29
Figure 2.10 SCADA weather monitoring 30

Figure 3.1 Schematic Architecture of Data Transmission in a Large-Scale PV Grid System 33

Figure 3.2 MODBUS message RTU framing 34
Figure 3.3 Modbus RTU-to-TCP Conversion 36
Figure 3.4 Modbus RTU Payload Encapsulation into a TCP/IP–Ethernet Frame . . 37
Figure 3.5 Sending Modbus TCP Requests 38
Figure 3.6 Receive Modbus TCP Response 38
Figure 3.7 SCADA Server Architecture and Network Connectivity 39
Figure 3.8 Config file 40
Figure 3.9 Data Normalization Process 40
Figure 3.10 Architecture of Polling Modules 42
Figure 3.11 Data Export and Back-End Processing System 43
Figure 3.12 Flowchart with Parameter Module 44
Figure 3.13 Flowchart with Analytic Module 45
Figure 3.14 Database Architecture 46
Figure 3.15 Flask as a Middleware: Managing HTTP Requests and Data Flow 47
Figure 3.16 Front-End HMI Architecture and Data Flow 48
Figure 3.17 Full-Stack Web Application Architecture 49
Figure 3.18 Raspberry pi4 server 50
Figure 3.19 Raspberry pi4 data sending emulation 51

Figure 4.1 The proposed CNN-LSTM Architecture 57
Figure 4.2 Predicted and measured power (a) clear sky day (b) cloudy day 59
Figure 4.3 Predicted and measured current (a) clear sky day (b) cloudy day 59
Figure 4.4 Predicted and measured voltage (a) clear sky day (b) cloudy day 60
Figure 4.5 Detection of PV Inverter Faults Using the Hampel Identifier 62
Figure 4.6 Fault Detection Technique Based on CPR Metric and Hampel Identifier . 63
Figure 4.7 Standard Deviation Spread in a Normal Distribution 66
Figure 4.8 Fault detection and diagnosis flowchart of the proposed method 67
Figure 4.9 Partial Shading or Soiling Effects on PV Module 68
Figure 4.10 Short-Circuit Fault in PV Modules 68
Figure 4.11 String-Level Fault in PV System 69
Figure 4.12 Standalone Offline Application for PV System prediction, Fault Detection and Diagnosis 70

Figure 5.1 Inverter Gauges 74
Figure 5.2 Parameters curve representation 75
Figure 5.3 Curve full size 75
Figure 5.4 Inverters surface level monitoring 76
Figure 5.5 Inverter monitoring 77
Figure 5.6 Inverter alarms 77
Figure 5.7 Performance parameters 78

Figure 5.8 Prediction module 79
Figure 5.9 Inverter's junction boxes 80
Figure 5.10 Inverter's strings 80
Figure 5.11 Isolated junction box level 1 81
Figure 5.12 Alarms on junction box and string level 82
Figure 5.13 Data saving parameters 83
Figure 5.14 Historical report 83
Figure 5.15 meteo Excel file example 84
Figure 5.16 Junction box inverter 10 side A Excel file example 85
Figure 5.17 inverter 10 side A Excel file example 85
Figure 5.18 Home page PV System Diagnosis and Fault Detection 86
Figure 5.19 Prediction Page 87
Figure 5.20 Faults Detection Page 88
Figure 5.21 Diagnosis Analysis Page 88

List of Tables

1.1 Efficiency Comparison by Generation 6
1.2 Description of the variables of Equation N° 1.1 8

3.1 Common Modbus Function Codes 35
3.2 Comparative Analysis of MongoDB (NoSQL) and Relational SQL Databases . 47

4.1 PV plant monitoring data 56
4.2 Layer configuration of the proposed model 58
4.3 Performance metrics under different weather conditions 59
4.4 Performance metrics of current prediction under different weather conditions . 59
4.5 Performance metrics of voltage prediction under different weather conditions (Alternative Model) 60
4.6 Comparative Evaluation of Proposed CNN-LSTM and Benchmark Models . . 60

5.1 comparison summary of old and new SCADA interface 89

Abbreviations

AI: Artificial Intelligence
API: Application Programming Interface
CPR: Corrected Performance Ratio (dimensionless)
GUI: Graphical User Interface
MB TCP: Modbus Transmission Control Protocol
ML: Machine Learning
MPPT: Maximum Power Point Tracking
NoSQL: "Not Only SQL" (non-relational databases)
PLC: Programmable Logic Controller
PR: Performance Ratio
PV: Photovoltaic
REST: Representational State Transfer
RTU: Remote Terminal Unit
SCADA: Supervisory Control and Data Acquisition
SQL: Structured Query Language
O&M: Operations & Maintenance

List of symboles

$\mathbf{P}_{\mathrm{nom}}$: Nominal installed photovoltaic (PV) power
$\mathbf{H}_{\mathbf{T}}$: Total in-plane solar irradiation
$\mathbf{G}_{\mathrm{STC}}$: Standard irradiance under Standard Test Conditions, with a value of $\mathbf{1}$ kW/m$^{\mathbf{2}}$
T: Duration of time
$\mathbf{Y}_{\mathbf{A}}$: Array yield
$\mathbf{Y}_{\mathbf{F}}$: Final yield
$\mathbf{Y}_{\mathbf{R}}$: Reference yield
η_{a}: PV array efficiency
η_{inv}: Inverter efficiency
$\mathbf{r}_{\mathrm{sys}}$: System performance ratio
γ: Temperature coefficient

General Introduction

The global transition toward renewable energy has placed photovoltaic (PV) systems at the forefront of sustainable power generation. As these systems scale in size and complexity, their operational reliability, efficiency, and availability become critically dependent on continuous monitoring and intelligent supervision. Effective monitoring not only ensures energy production stability but also contributes to the long-term performance and safety of the installation. In this context, advanced supervisory technologies are essential to manage the dynamic conditions and distributed architecture of modern PV power plants.

Supervisory Control and Data Acquisition (SCADA) systems have emerged as a cornerstone of industrial automation, offering real-time monitoring, control, and data acquisition across diverse domains, including renewable energy. In photovoltaic systems, SCADA platforms serve as the primary interface between human operators and plant components, such as PV modules, inverters, transformers, and environmental sensors. A well-designed SCADA interface enhances situational awareness through intuitive data visualization, facilitates rapid fault response, and supports decision-making through actionable information. Its role is not limited to visualization. It encompasses system-wide coordination, historical analysis, and operational control.

The current SCADA system deployed at the target photovoltaic (PV) station provides basic monitoring capabilities, primarily focusing on energy output and general plant status. It allows operators to view some electrical parameters and environmental data, but lacks access to detailed component-level information. The architecture consists of a local control panel with limited visualization and manual data retrieval functions. However, the existing SCADA infrastructure at the target PV station suffers from several limitations that hinder efficient supervision and maintenance. The current system is restricted in scope, offering limited access to component-level data and control functions. Its visualization capabilities are outdated, lacking detailed and dynamic graphical representations. Furthermore, the system provides no integrated analytics for performance evaluation or predictive diagnostics, and it completely lacks fault detection and forecasting capabilities. These shortcomings reduce system transparency, increase the burden of manual supervision, and delay maintenance actions—ultimately affecting the plant's reliability and energy yield.

To overcome these limitations, this project proposes the design and development of a next-generation SCADA interface tailored for grid-connected photovoltaic plants. The solution aims to enhance interactivity, improve access to real-time and historical data, and introduce intelligent monitoring features through modern web-based technologies and data analytics. The proposed system integrates advanced visualization tools, Python-based performance analysis, AI-driven fault detection, and predictive modeling. Built upon a scalable and modular open-source architecture, the new SCADA interface is designed to be both future-proof and operator-friendly, bridging the gap between real-time supervision and data-driven decision support.

The first chapter provides a comprehensive overview of photovoltaic systems and highlights the critical role of SCADA in the supervision of renewable energy infrastructures. Chapter two examines the limitations of the existing SCADA platform, detailing its current functionalities and identifying key areas for enhancement. Chapter three focuses on the design and implementation of the upgraded SCADA interface, emphasizing improved data visualization, storage architecture, and user interaction features. Chapter four introduces advanced fault detection and diagnosis mechanisms based on artificial intelligence, including the integration of a CNN-LSTM model for predictive analysis. Finally, Chapter Five presents the deployment process of the developed application, validates its performance through experimental scenarios, and provides a comprehensive evaluation of its effectiveness.

This structured approach enables the development of a robust, intelligent, and user-centric SCADA solution that meets the evolving demands of modern photovoltaic power systems.

Motivation

The increasing integration of photovoltaic energy into national grids calls for advanced supervision systems that go beyond traditional control architectures. While traditional SCADA platforms form the foundation of monitoring infrastructures, they increasingly lack the adaptability, analytical intelligence, and responsiveness required by today's energy systems. These limitations become particularly critical in large-scale PV installations, where timely interventions and predictive maintenance have a direct impact on energy output and operational costs.

This project stems from the ambition to redefine supervisory systems in the context of renewable energy. It proposes a smart, modular, and data-centric SCADA interface designed to operate as an integral part of a grid-connected photovoltaic power plant. More than just a monitoring tool, the system is envisioned as a strategic component in the energy value chain, supporting both real-time plant supervision and future interoperability with national grid dispatching systems. By aligning with the evolution toward smart grids, the platform positions itself for scalable deployment and integration into future energy management frameworks.

Chapter 1

Photovoltaic system & Technology: State of the art

1.1 Introduction

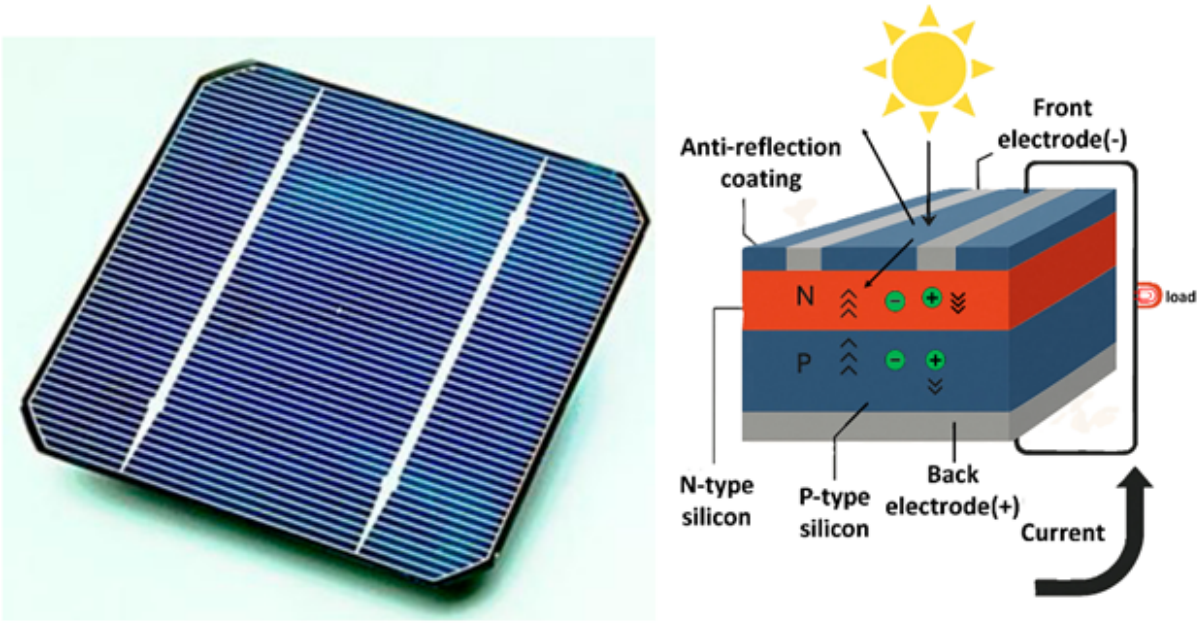
Solar energy is a highly accessible and renewable energy resource with significant global potential. It is the radiated energy generated by the sun that, despite the attenuation caused by the atmosphere, releases a large quantity of electromagnetic radiation to the earth's surface. The conversion of this energy into electricity using photovoltaic (PV) technology has emerged over the last few decades as a key technology for meeting the growing demand for electrical energy, reducing dependence on fossil fuels, and limiting their harmful impact on the environment [1]. This chapter reviews the state of the art of photovoltaic systems, including the physical principle of the photovoltaic effect, types of photovoltaic systems and architectures, up to monitoring and supervision.

1.2 Photovoltaic Systems

1.2.1 Photovoltaic Cells and Modules

- **PV Cells**

The photovoltaic (PV) cell, the fundamental component of a PV system, transforms sunlight into electrical energy. Its efficiency varies based on material purity, design, and manufacturing techniques, such as the use of passivation layers and anti-reflective coatings[2]. The figure 1.1 depicts a real PV cell with a cross-sectional view.



(a) real Solar cell (b) Side section view of PV cell

Figure 1.1: PV cell

- **PV Module**

A photovoltaic (PV) module is a set of interconnected PV cells that are encapsulated and framed to create a robust, weather-resistant unit. The arrangement of the cells is carefully designed to maximise sunlight absorption and electrical output. The figure 1.2 depicts the cell interconnections and the layered structure of a typical PV module.

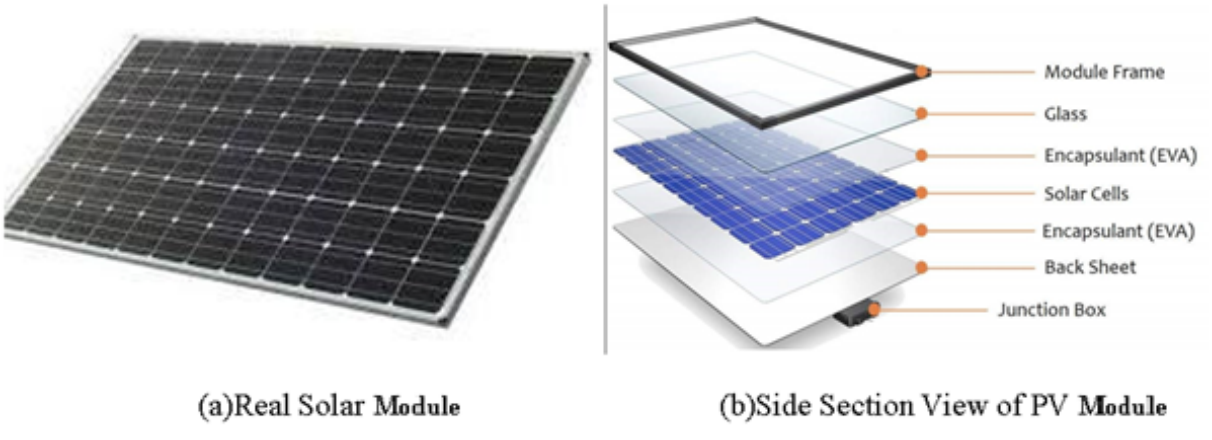


Figure 1.2: PV module

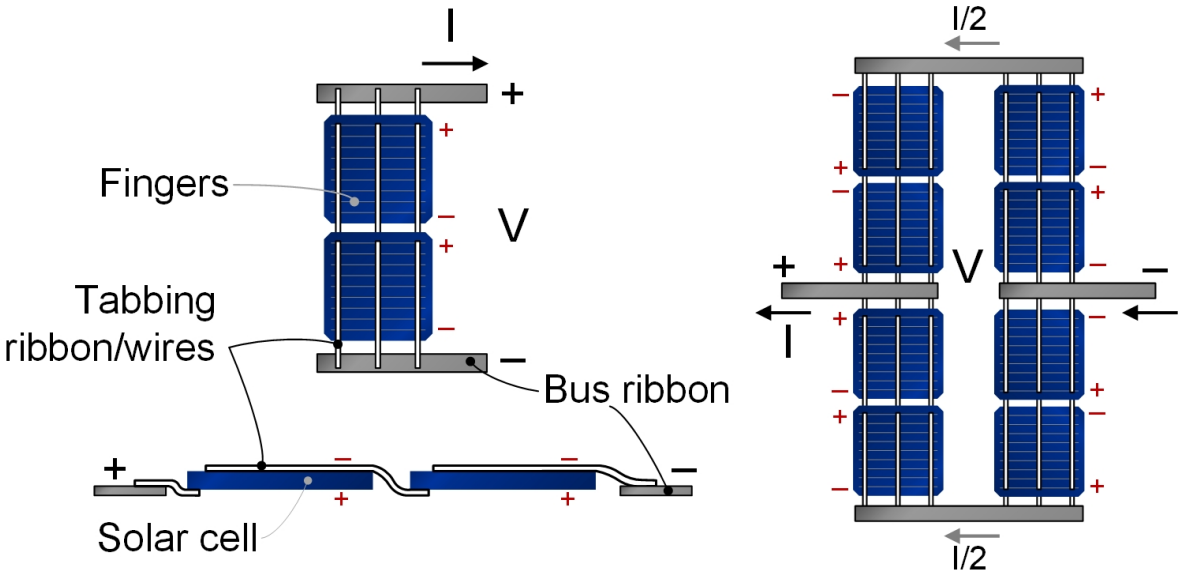


Figure 1.3: PV cell interconnection

1.2.2 Photovoltaic Module Technology

PV modules are classified into different generations based on their manufacturing technology and materials[2][3], as illustrated in figure 1.4.

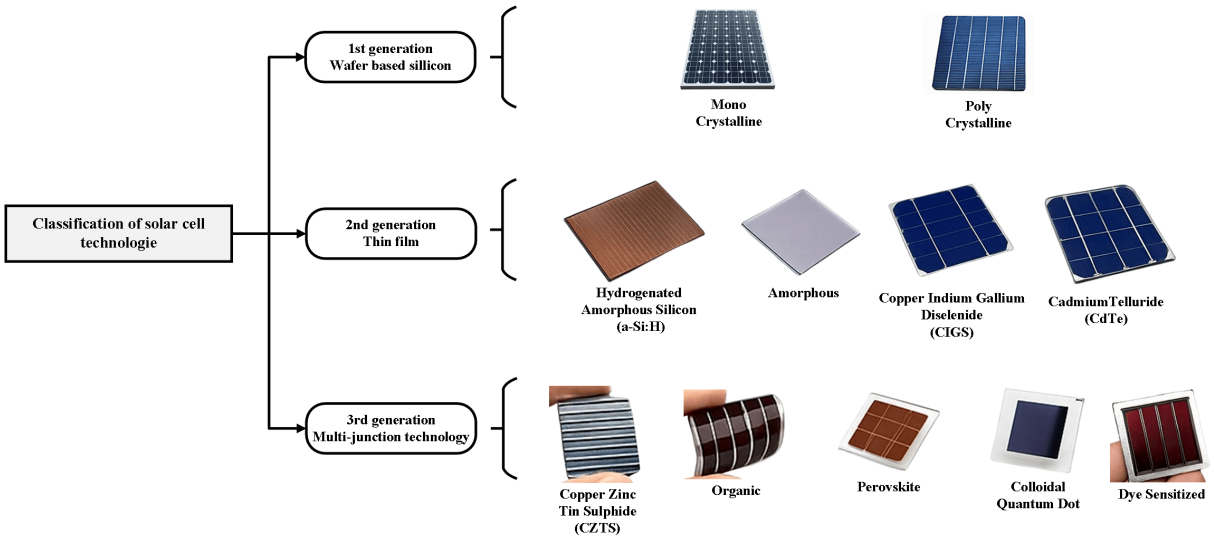


Figure 1.4: Diagram of photovoltaic cell classified by invention technology

Where each generation offers its proper efficiency on the laboratory level and commercial level, the table 1.1 shows that.

Table 1.1: Efficiency Comparison by Generation

Efficiency Type	**Generation 1**	**Generation 2**	**Generation 3**
Laboratory efficiency	26%	16%	45%
Commercial efficiency	16%	10%	45%

1.2.3 Electrical Characteristics of PV Modules

Photovoltaic (PV) modules have several key electrical characteristics that define their performance. These characteristics are typically represented in current-voltage (I-V) and power-voltage (P-V) curves, which are essential for understanding how a solar panel will perform under different conditions[4]. The given figure 1.5 depicts the behaviour of a PV module under a given temperature and irradiance.

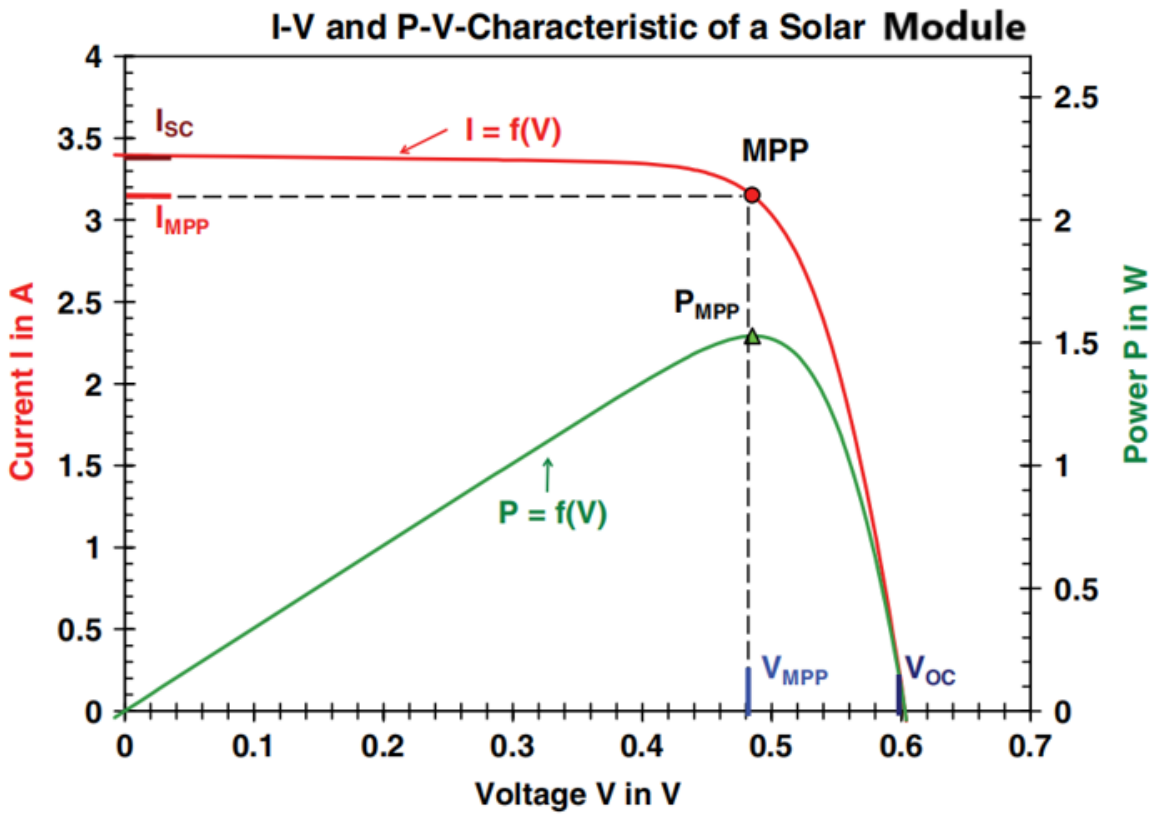


Figure 1.5: I-V curve of a PV module

In the literature, the physical model of a PV cell/module is described by the One Diode Model (ODM)[5], given in the figure 1.6.

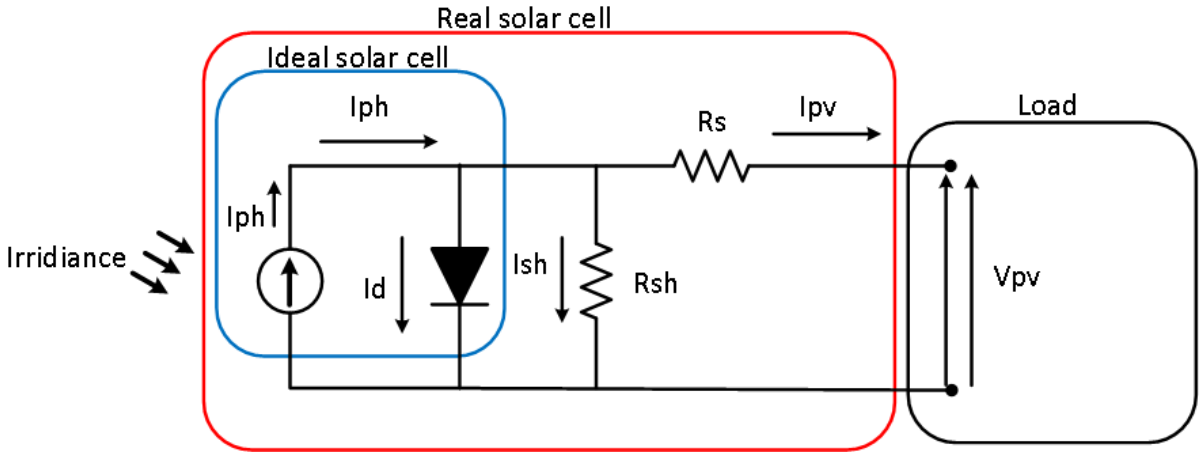


Figure 1.6: Solar cell equivalent circuit

The output current (Ipv) of the PV cell is given by:

$$I_{\text{PV}} = I_{\text{ph}} - I_S \left(e^{\frac{e(V_{PV}+R_S \times I)}{nkT}} - 1 \right) - \frac{(V_{PV} + R_S \times I)}{R_P} \tag{1.1}$$

. where each term of this equation is explained in the table 1.2.

Table 1.2: Description of the variables of Equation N° 1.1

Term	Description
I_{pv}	Output current (A)
I_{ph}	Photocurrent (light-generated current, A)
I_{s}	Diode reverse saturation current (A)
V_{pv}	Output voltage (V)
R_{s}	Series resistance (Ω)
R_{p}	Shunt resistance (Ω)
n	Diode ideality factor (1 for ideal diode, 1–2 for real cells)
k	Boltzmann constant (1.380649×10^{-23} J K^{-1})
T	Temperature (K)

1.2.4 Strings and Arrays

PV panels are interconnected in series (strings) and parallel (arrays) to achieve optimal voltage and current levels[6]:

- **Strings:** A string in solar modules refers to an array of photovoltaic modules connected in series. This configuration means that the electrical current generated by each panel flows successively from one to the other, adding up their individual voltages. This arrangement is common in installations where a specific voltage is sought to match the requirements of the solar inverter.

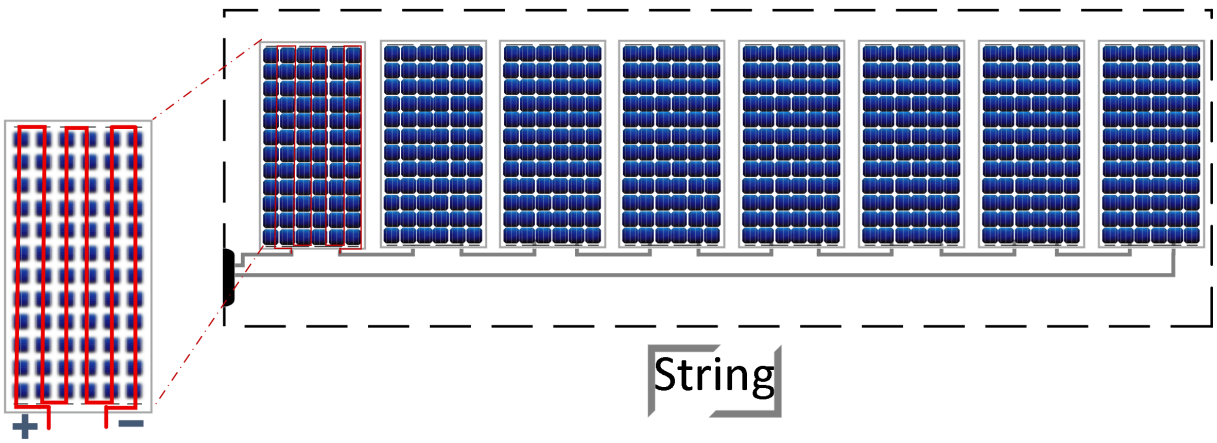


Figure 1.7: Schematic of PV string

- **Arrays:** The term array refers to a complete arrangement of solar panels in a photovoltaic installation. An array can be made of multiple strings connected together, either in series, in parallel, or in a combination of both configurations, depending on the voltage and current requirements of the system. In essence, an array is the total set of solar panels working together to generate power.

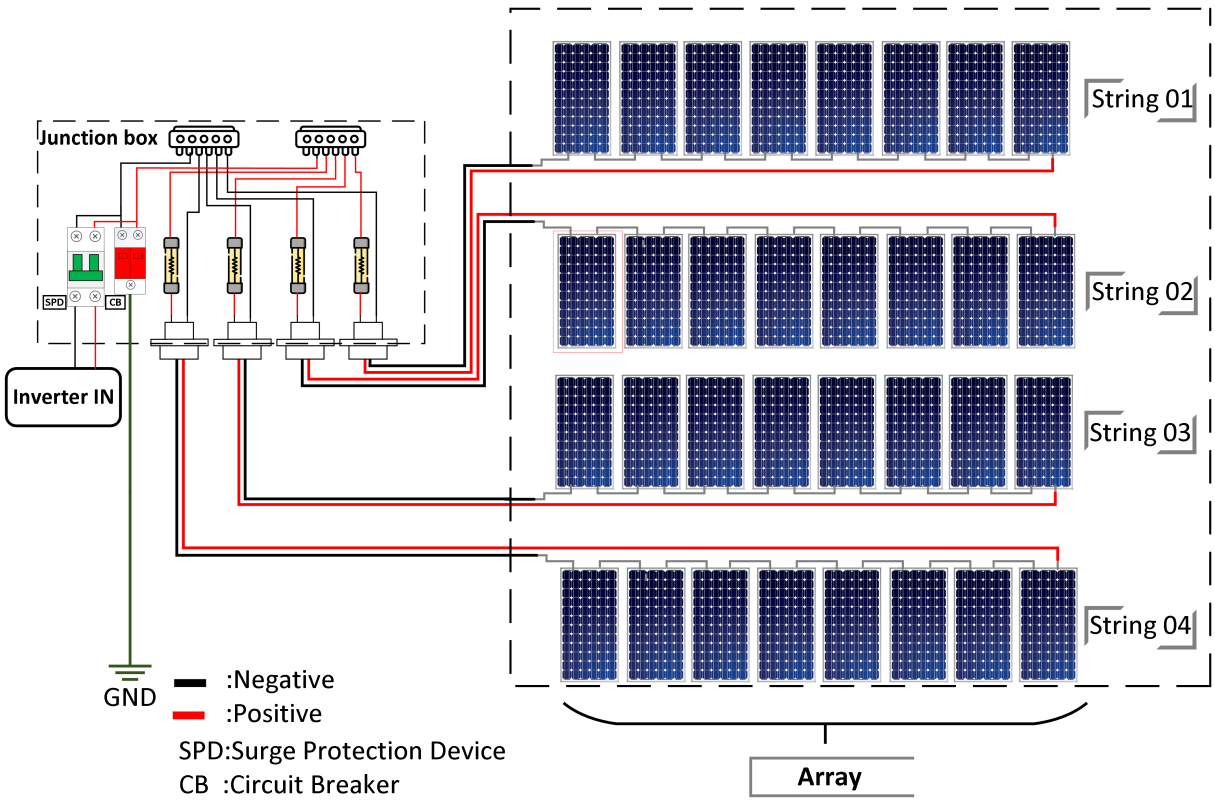


Figure 1.8: Schematic of PV array configuration

1.2.5 PV System Types

The overview below provides a systematic and structured overview of the different types of photovoltaic (PV) systems, their configurations, and the typical applications associated with them[7][8]:

1.2.5.1 Classification Based on Grid Connection

- **Grid-Tied PV Systems:** A grid-connected photovoltaic system is a system that is connected to the electricity grid. Such systems do not require batteries or other equipment for energy storage. As a result, the overall cost of equipment and installation is significantly lower than for any other type of solar energy system[9]. The main components of a grid-connected photovoltaic system are: the photovoltaic array, the grid-connected inverter, and the net meter, as shown in the figure 1.9.

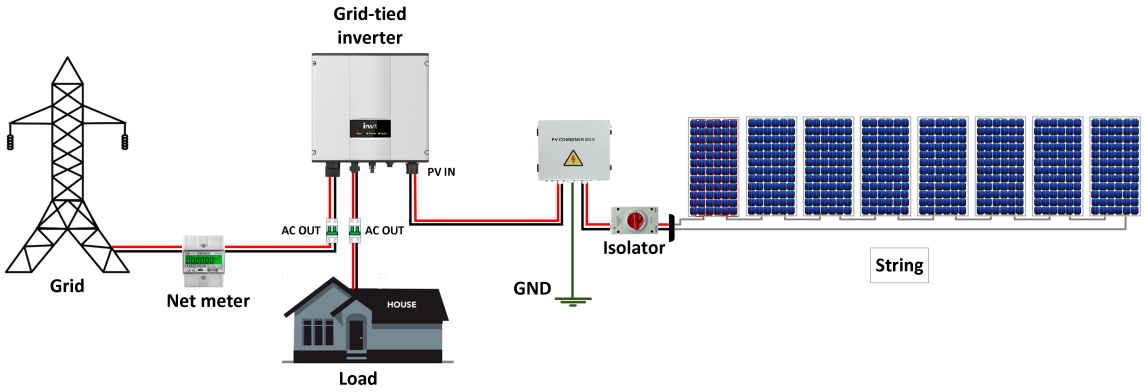


Figure 1.9: Grid tied photovoltaic system

- **Off-Grid PV Systems:** An off-grid photovoltaic system, also known as a stand-alone photovoltaic system, is a power generation installation that uses solar radiation to produce electricity in places where there is no access to the mains utility grid [10]. The components of a typical off-grid photovoltaic system are photovoltaic modules, which absorb sunlight, an inverter, a charge controller, and batteries to store the excess electricity produced, as shown in Figure 1.10.

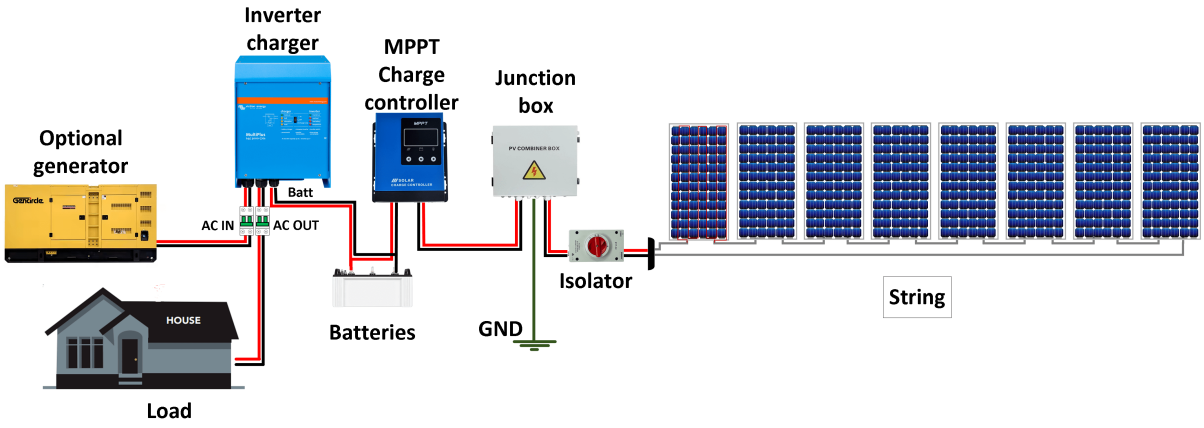


Figure 1.10: Off grid tied photovoltaic system

- **Hybrid PV Systems:** A hybrid photovoltaic (PV) system combines the advantages of both grid-connected and stand-alone solar systems. It uses photovoltaic modules to generate electricity, a hybrid inverter to manage the power flow, and a battery to store excess energy for later use or in the event of a power outage. This system provides continuous power, grid connectivity, and energy independence [11]. The main components of a hybrid system are the PV array and the hybrid inverter, as shown in Figure 1.11.

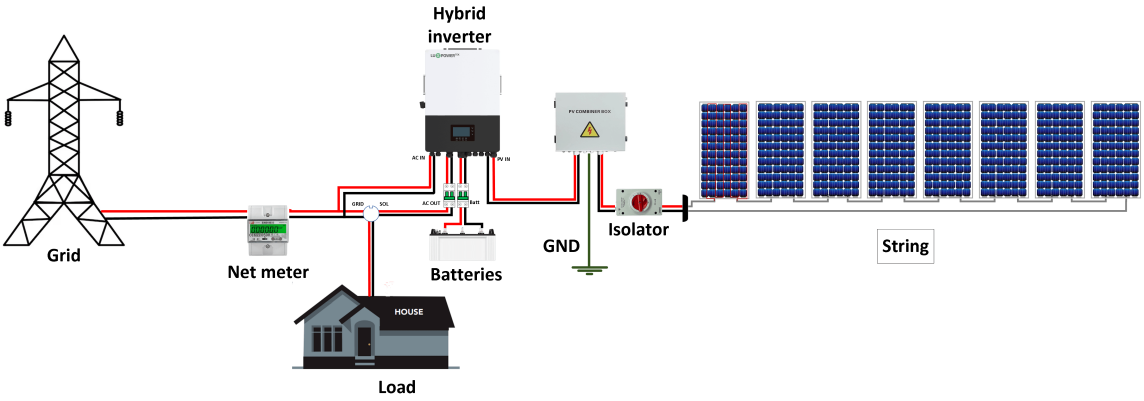


Figure 1.11: Hybrid PV system

1.2.5.2 Classification Based on System Scale

- **Residential PV Systems:** A residential photovoltaic system is an installation of rooftop or ground-mounted photovoltaic modules and related components that convert sunlight into electricity[12]. This energy can be used directly within the household or stored for later use (see figure 1.12).



Figure 1.12: Residential PV system

- **Commercial PV Systems :** Commercial photovoltaic (PV) systems, also known as solar energy systems, provide electricity for businesses and industries[12]. These systems deliver key advantages, including lower energy expenses, a smaller environmental footprint, and the potential to earn revenue by selling surplus power (see figure 1.13).



Figure 1.13: Commercial PV systems

- **Utility-Scale PV Systems:** Large-scale photovoltaic (PV) systems are large-scale photovoltaic power plants designed to produce electricity for the utility grid, typically with a capacity higher than one megawatt (MW)[12]. These systems often include ground-mounted solar panels, and the electricity produced is fed into the grid and distributed to consumers (See figure 1.14).



Figure 1.14: Large scale PV system

1.3 Performance Analysis Method for Photovoltaic Systems

There are several methodologies for performance analysis. The following section elaborates on the two most widely used approaches. The first approach is derived from the standardised norms outlined in IEC 61724 [13]. The second, known as the PVUSA method, involves systematically collecting solar power output, meteorological data, and system performance metrics and correlating them with variables such as irradiance, wind speed, and ambient temperature [14].

1.3.1 System Performance Index under IEC 61724 Method

In a photovoltaic installation, the monitoring system must enable energy flows to be tracked and provide an analysis of the key parameters influencing the energy conversion chain. The main parameters, as well as the most significant energy values in terms of efficiency and losses, are illustrated in Figure 1.15.

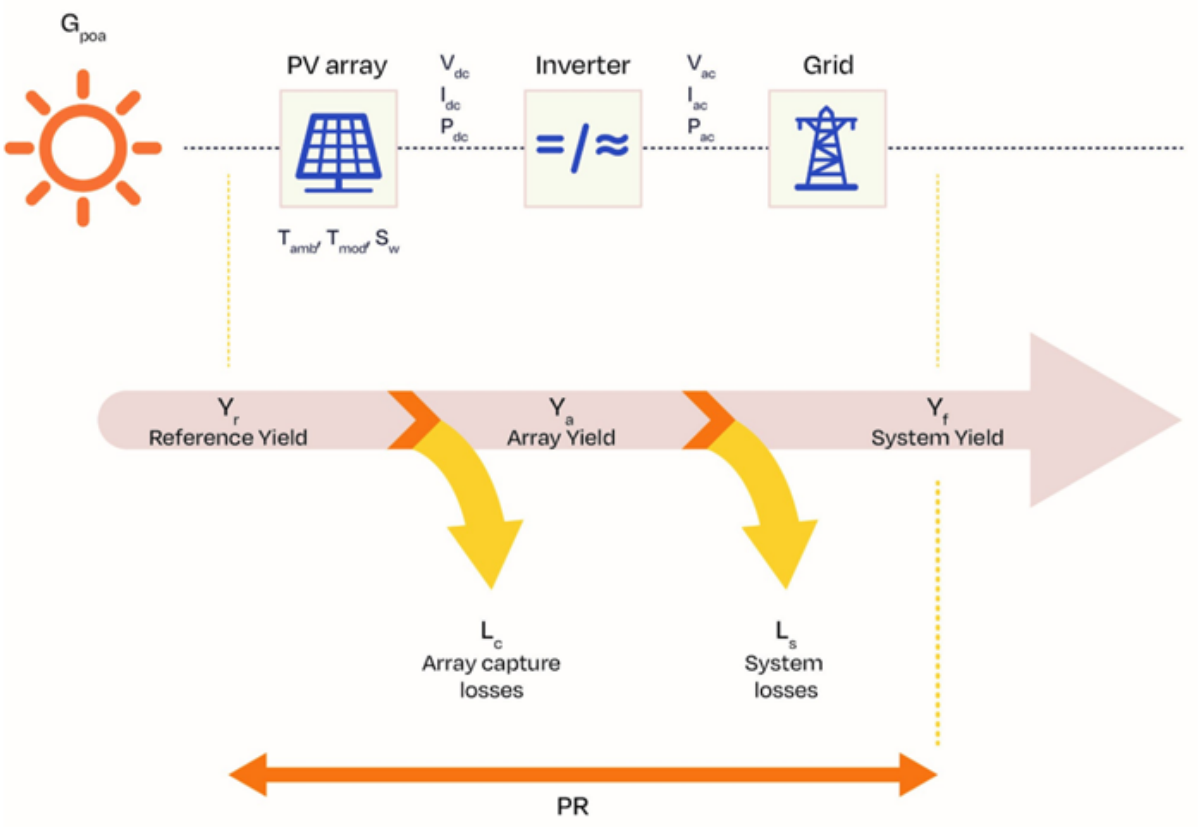


Figure 1.15: Energy flow in a grid-connected photovoltaic system with parameters, efficiencies and losses

Yields are quantities of energy normalised to the nominal power of the photovoltaic plant. Losses are the differences between yields [15].

- **Reference Yield (Yr)**

The Reference Yield is the ratio of the total in-plane solar irradiation (H_t in Wh/m^2) to the reference irradiance ($G_{ref} = 1000$ W/m^2). It represents the equivalent number of peak sun hours (h/day), where peak sun hours denote the duration at which solar irradiance equals G_{ref}. Its expression is given by:

$$Y_r\,[\mathrm{h}] = \frac{H_t\,[\mathrm{Wh/m^2}]}{G_{\mathrm{ref}}\,[\mathrm{W/m^2}]} \tag{1.2}$$

- **Array Yield (Ya)**

The Array Yield is defined as the ratio of the energy output of the photovoltaic array (E_{DC}) to its nominal power (P_0) under Standard Test Conditions (STC). Mathematically, it is expressed as:

$$Y_a\,[\mathrm{h}] = \frac{E_{\mathrm{DC}}\,[\mathrm{Wh}]}{P_0\,[\mathrm{W}]} \tag{1.3}$$

where:

- E_{DC} = Total DC energy produced by the array (Wh)
- P_0 = Installed DC rated power under STC (W)

This metric represents the equivalent number of hours the system would need to operate at its nominal power (P_0) to produce the same energy output (E_{DC}). The Array Yield serves as a performance indicator, assessing the actual energy generation efficiency of the PV modules relative to their rated capacity under ideal conditions.

- **Final Yield (Yf)**

The Final Yield represents the daily energy output of the photovoltaic system (in kWh) normalized by the installed DC rated power (P_0 in kW). It corresponds to the equivalent number of hours the system would operate at its rated power to produce the same energy output. Mathematically, it is defined as:

$$Y_f\,[\mathrm{h}] = \frac{E_{\mathrm{AC}}\,[\mathrm{kWh}]}{P_0\,[\mathrm{kW}]} \tag{1.4}$$

where:

- E_{AC} = Total AC energy delivered by the system (kWh)
- P_0 = Installed DC rated power under STC (kW)

1.3.2 Normalised Losses

These losses represent the difference between the above yields and indicate how many hours the system would need to operate at its rated power in order to compensate for them.

- **Capture Losses (L_C)**

  The normalized energy difference between the theoretically available solar resource (Reference Yield, Yr) and the DC energy generated by the PV array (Array Yield, Ya), expressed in equivalent peak sun hours. These losses primarily result from:

  - Module temperature deviations from STC
  - Spectral effects
  - Mismatch losses
  - Soiling and degradation

These losses are given by the following expression:

$$L_C\,[\mathrm{h}] = Y_r\,[\mathrm{h}] \;-\; Y_a\,[\mathrm{h}] \tag{1.5}$$

- **System Losses (L_S)**

  The normalized energy loss occurring between the PV array's DC output and the final AC delivery to the grid, accounting for:

  - Inverter conversion inefficiencies
  - Wiring and transformer losses

The expression giving these losses is:

$$L_S\,[\mathrm{h}] = Y_a\,[\mathrm{h}] \;-\; Y_f\,[\mathrm{h}] \tag{1.6}$$

1.3.3 Performance Ratio (PR)

The **Performance Ratio (PR)** quantifies a PV system's efficiency by comparing its actual AC energy output to its theoretical maximum under ideal conditions (STC). Expressed as a percentage (typically 70% to 85% for well-functioning systems), PR isolates system losses (e.g., thermal, inverter, soiling) while normalizing for location and weather variability. A high PR indicates minimal energy losses and optimal operation, while a low PR signals potential issues like shading, degradation, or equipment faults.

$$PR = \left(\frac{Y_f}{Y_r}\right) \cdot 100\% \tag{1.7}$$

1.3.4 Corrected Performance Ratio (CPR)

The corrected PR improves performance analysis by systematically filtering out influences that are not related to the system, such as transient weather anomalies, grid limitations, or data irregularities. This adjustment enables reliable and fair comparisons of the performance of photovoltaic systems in different locations, climatic conditions, and operating periods. By isolating controllable losses (e.g. equipment failure, soiling or inverter inefficiencies) from external variables, the corrected PR is a powerful diagnostic tool for stakeholders: investors use it to validate return on investment against energy yield guarantees, engineers use it to compare system designs under standardised conditions, and operators rely on it to diagnose specific maintenance needs, distinguishing real performance issues from temporary environmental effects [16]. The corrected performance ratio is illustrated in Figure 1.16, adapted from [16], which depicts the annual trend comparing uncorrected and corrected values. The calculation of the CPR is based on the following expression.The calculation of the CPR is based on the following expression:

$$CPR = \frac{PR}{1 \; - \; \delta \times \left(T_{\text{cell,Type,avg}} - T_{cell}\right)} \tag{1.8}$$

Where:

$$T_{cell} = T_m + \frac{G_{\text{POA}}}{G_{\text{STC}}} \times T_{\text{cnd}} \tag{1.9}$$

- T_m: Measured module surface temperature
- G_{POA}: Plane-of-array irradiance
- G_{STC}: Standard test condition irradiance
- T_{cnd}: Conduction-temperature drop
- δ: Temperature coefficient

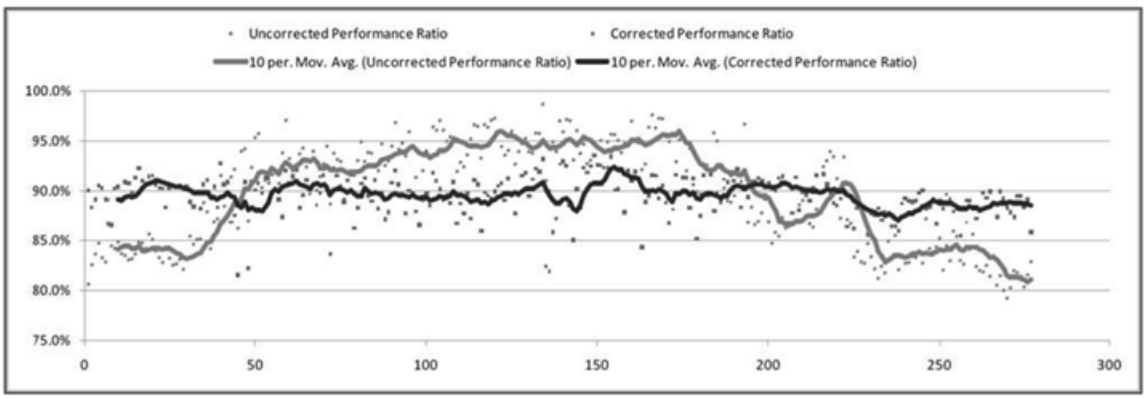


Figure 1.16: Weather corrected and uncorrected annual trend for a 24 MW facility

1.4 SCADA Systems in Photovoltaic Power Plants

Supervisory Control and Data Acquisition (SCADA) systems are essential for the reliable operation and efficient management of large-scale photovoltaic (PV) installations. In the context of renewable energy, SCADA systems enable remote monitoring, real-time data acquisition, automated control, and diagnostics functions that are critical to ensuring continuous power generation and improving system resilience [17][18].

By providing centralized visibility into system performance, SCADA enhances operational decision-making, supports rapid fault detection, and facilitates preventive maintenance. This leads to improved energy yield, reduced downtime, and better long-term efficiency across utility-scale or hybrid grid-connected PV systems [19][20].

The following section explains the architecture of SCADA systems in photovoltaic applications, detailing their core components and communication infrastructure.

- **SCADA system architecture in photovoltaic**

To ensure effective supervision and control of photovoltaic (PV) installations, SCADA systems adopt a modular and layered architecture. This structure typically includes field-level devices such as sensors and power electronics, communication networks for data transmission, and a centralized supervisory system for data processing and visualization[21]. The architecture enables real-time monitoring, fault detection, and system optimization, forming the backbone of reliable and scalable PV plant management (see Figure 1.17)

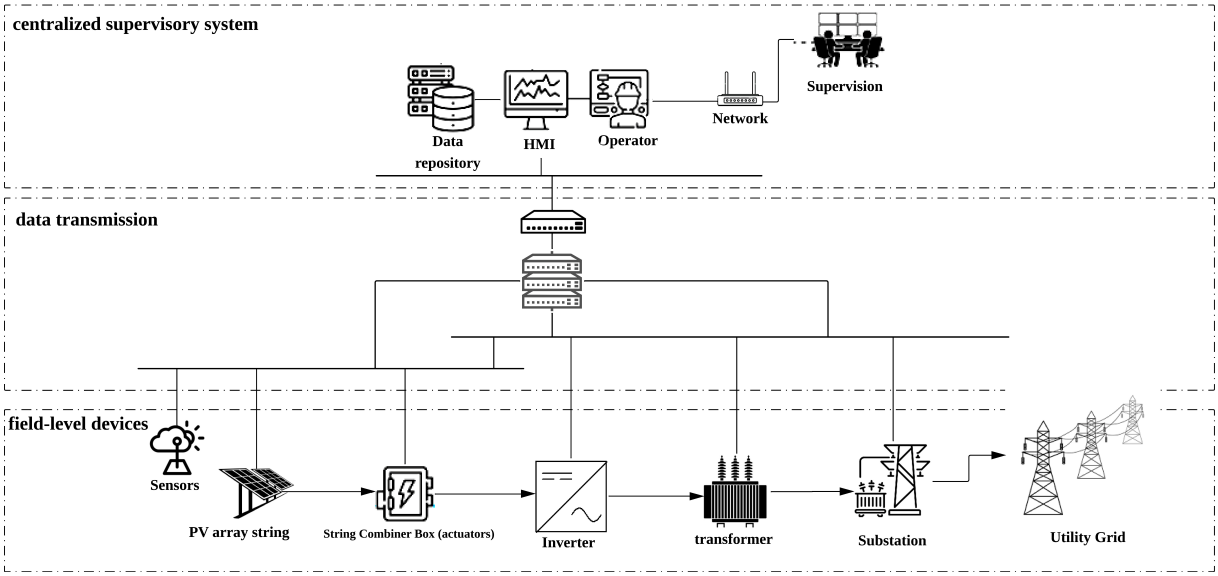


Figure 1.17: SCADA Architecture for Grid-Connected PV Plants

This layered configuration illustrates the integration of SCADA systems and reflects their operational logic, encompassing ground-level data acquisition, centralized control, and informed decision-making. As will be explained in the following sections, each layer plays a crucial role in ensuring the system's overall effectiveness.

- **Centralized Supervisory Layer**
  At the top of the architecture lies the centralized control room, where the SCADA master server, data repository, Human-Machine Interface (HMI) workstations, and operational staff converge. The data repository archives both real-time and historical measurements such as solar irradiance, module and ambient temperatures, string-level currents, inverter status, energy yield, and system alarms. The HMI provides graphical dashboards, trending tools, and alarm displays, enabling operators to monitor system health, adjust parameters, and initiate corrective actions. A robust network backbone interconnects these components, supporting both local and remote access. This dedicated supervisory environment enables collaborative decision-making and ensures rapid response to operational events[22].

- **Data Transmission Layer**
  The intermediate layer comprises the communication infrastructure responsible for linking field-level devices to the supervisory system. This includes servers, switches, routers, and potentially wireless access points. While not explicitly detailed, the architecture supports industry-standard communication protocols such as Modbus, DNP3, and OPC UA[23]. The converging lines in the figure represent redundant or segregated communication paths, enhancing system reliability and cybersecurity. This layer is critical for ensuring that sensor data, control commands, and status information are transmitted with minimal latency and high integrity, essential for real-time monitoring and automated control functions.

- **Field Device Layer (PV Power Generation System)**
  The lowest tier of the model signifies the physical PV installation. Environmental sensors (e.g., irradiance and temperature sensors) provide analog and digital measurements to RTUs or PLCs[24][25]. PV module strings deliver direct current (DC) power into string combiner boxes. In these boxes, individual string outputs are fused, protected, and, if equipped, interfaced to string-level disconnect actuators[26]. The power is then inverted by grid-tied inverters and stepped up via an on-site transformer[27]. The substation's primary functions include the collection of higher voltage output, provision of grid synchronization, protection relays, circuit breakers, and revenue metering prior to the energization of the utility grid. This configuration underscores the SCADA system's role in tracing every kilowatt from photon to grid, safeguarding assets, and optimizing performance.

1.5 Conclusion

Photovoltaic systems consist of core components—solar panels, inverters, meters, and protective devices—that convert solar energy into electricity. Their performance depends not only on proper system design but also on continuous evaluation of operating conditions and energy yield. In large installations, SCADA systems are essential for real-time monitoring and control, enhancing both efficiency and reliability. However, conventional SCADA solutions face limitations in scalability and fault detection, motivating the development of a more advanced, intelligent SCADA tailored to modern photovoltaic plants.

Chapter 2

Operational Worksite — Ain El Melh Solar Power Plant

2.1 Introduction

A rigorous understanding of the physical layout and functional organization of the Ain El Melh photovoltaic (PV) power plant is essential for any technical intervention concerning its SCADA system. This chapter delineates the plant's principal operational domains-including energy generation, power conversion and dispatch, and supervisory control-to establish a concrete contextual basis for the subsequent analysis of communication infrastructure, data acquisition nodes, and control channels. Such foundational insight ensures that any effort related to the design, optimization, or diagnostic evaluation of the SCADA system remains firmly aligned with the plant's actual operational topology.

2.2 Site description

2.2.1 Installed Photovoltaic Power Plant Characteristics

The 20 MW grid-connected photovoltaic (PV) power plant, integrated into the medium-voltage distribution network as shown in Figure 2.2, forms a critical element of a broader national renewable energy strategy led by SONELGAZ[28]. Situated in Ain El Melh (Bousaada), on the Algerian highlands at the threshold of the Great Algerian Desert, the facility spans approximately 40 hectares and delivers an installed capacity of 20 MWp[28]. Its strategic location and scale make it a key contributor to regional energy autonomy and sustainability. An aerial and a ground-level view of the site are shown in Figure 2.1, providing a direct visual reference to the installation described.



(a) PV Power Upper-view (b) Front-view of PV strings

Figure 2.1: Site view of the Ain El Melh solar power plant

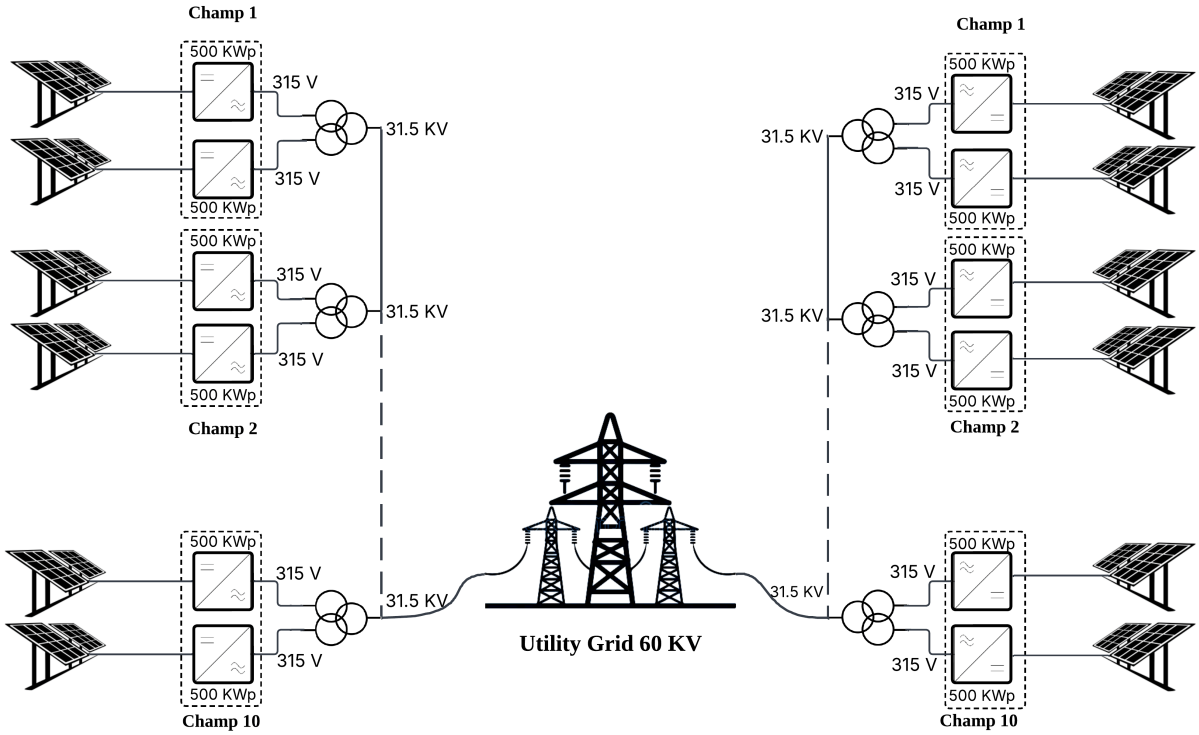


Figure 2.2: PV Power Plant connection to notional grid

2.2.2 Main Zones of the PV Power Plant

The photovoltaic (PV) power station is functionally structured into four core zones, each contributing to the generation, adaptation, delivery, and monitoring of electrical energy. The system architecture(illustrated in Figure 2.3)depicts the coordinated flow of both energy and data across the plant's functionally distinct yet operationally interlinked subsystems[28].

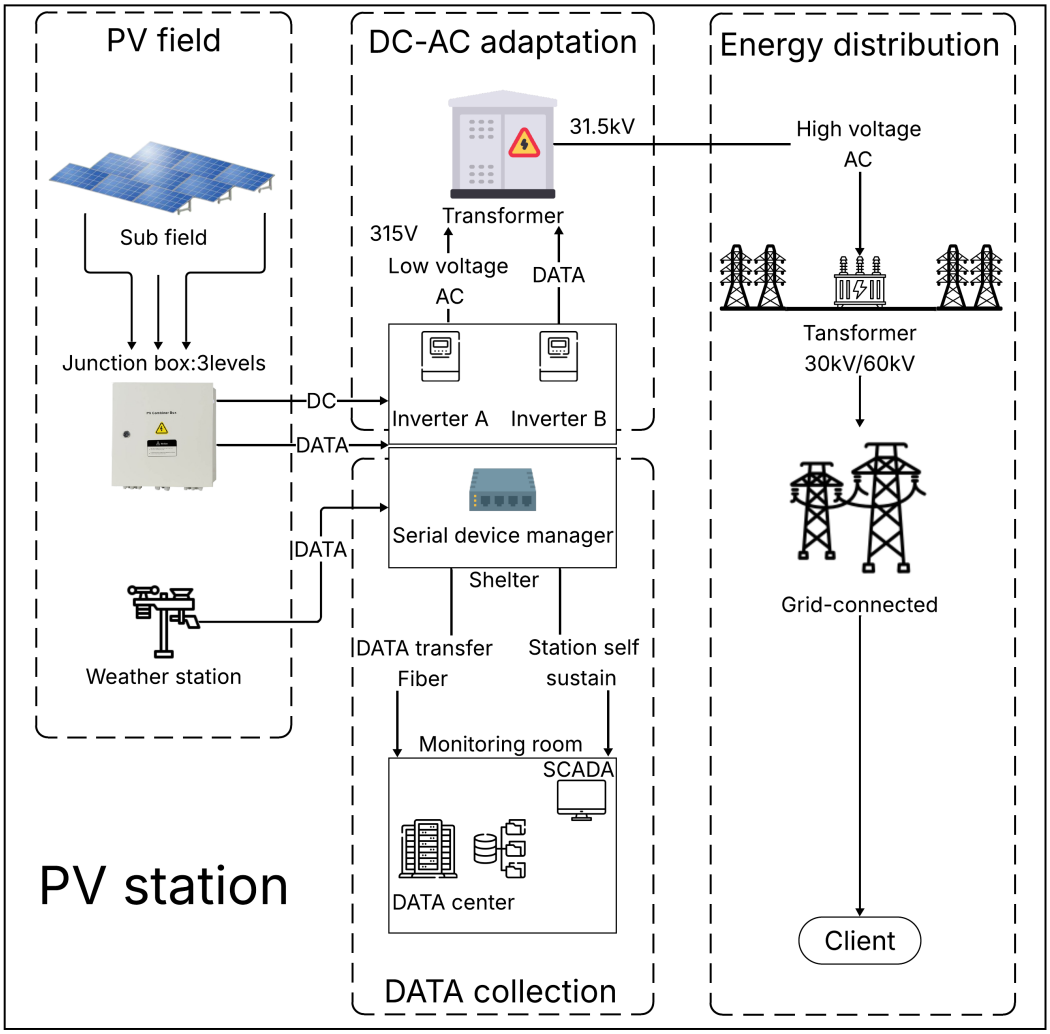


Figure 2.3: General layout of the PV Power Plant

These zones are designed to operate in an integrated manner, ensuring both operational efficiency and seamless communication. The following sections trace the respective trajectories of energy and data through these zones, highlighting their specific roles and interactions.

2.2.2.1 Stages of Energy Conversion and Flow

The energy life cycle within a PV station encompasses three principal stages: generation, conversion, and distribution. Each phase contributes to the efficient transformation and transmission of solar energy to the utility grid.

- **PV Generation Field**
  This zone hosts the primary energy source, where solar radiation is converted into direct current (DC) electricity by photovoltaic modules. Sensors and a weather station are also deployed here to collect environmental and electrical data. Both the generated energy and acquired data are routed through multi-level junction boxes and delivered via underground conduits to the shelter.

- **Power Conversion Zone**
  The shelter serves a dual purpose. In its upper compartment, inverters convert DC energy into alternating current (AC). The AC output is then routed through low-voltage transformers and loop cells, which prepare the power for elevation to medium-voltage levels. These components comprise the adaptation stage, bridging the generation and distribution phases.

- **Energy Distribution Segment**
  In this final stage, the adapted energy reaches the substation, where it is stepped up to grid-compatible voltages (30 kV/60 kV). It is subsequently injected into the national utility grid via a transmission line for end-user consumption.

Figure 2.4 illustrates the energy flow pathway from the photovoltaic module array to the substation, along with the structural layout of the associated sub-fields[28].

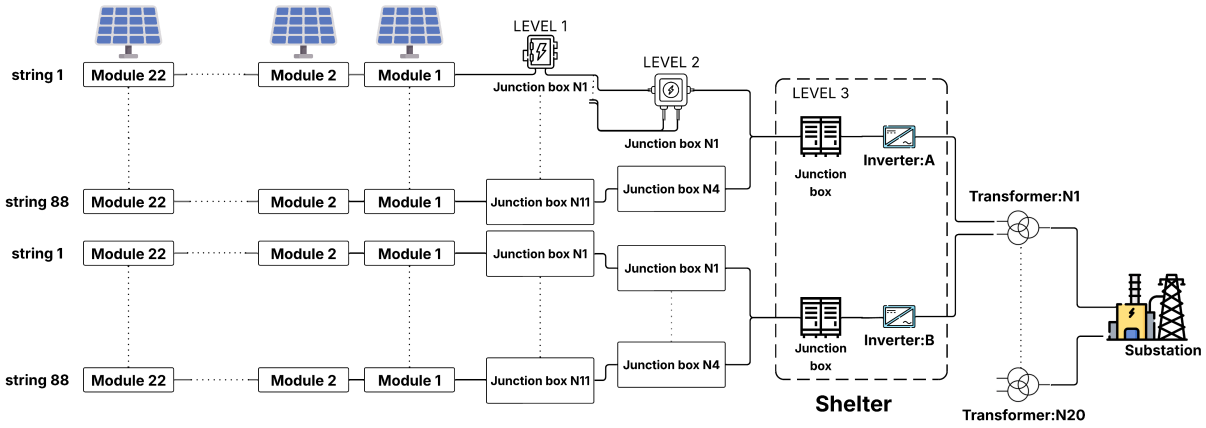


Figure 2.4: Energy flow: from field generation to grid injection via shelter and substation.

2.2.2.2 Data Acquisition and Transmission

The data pathway in a photovoltaic station is responsible for acquiring, transmitting, and processing operational and environmental information necessary for real-time supervision, control, and long-term performance evaluation. It consists of two principal components: data sources and the communication network.

- **Data Sources**

  Multiple subsystems generate data across the PV station, each contributing to the comprehensive monitoring of the plant[28]:

  - **Environmental Sensors (Weather Station)**
    - Pyranometer for plane-of-array irradiance.
    - Ground humidity and precipitation sensors
    - Module surface temperature sensors
    - Five-element meteorological mast measuring: ambient temperature, wind direction and speed, atmospheric pressure, and relative humidity,
  - **Junction-Box Processing Units**
    Measure electrical properties of individual PV strings, including string current and voltage.
  - **Equipment Status Reports**
    Inverters and transformers generate internal data on operational states, fault diagnostics, alarms, and other performance indicators.

- **Communication network**

Parallel to the energy flow, data is transmitted from field-level sensors and junction boxes via RS485 serial communication to a serial device server located in the lower section of the shelter. Subsequently, this data is relayed via fiber-optic and Ethernet connections to the monitoring room, where a SCADA (Supervisory Control and Data Acquisition) system enables real-time supervision, control, and long-term performance archiving.

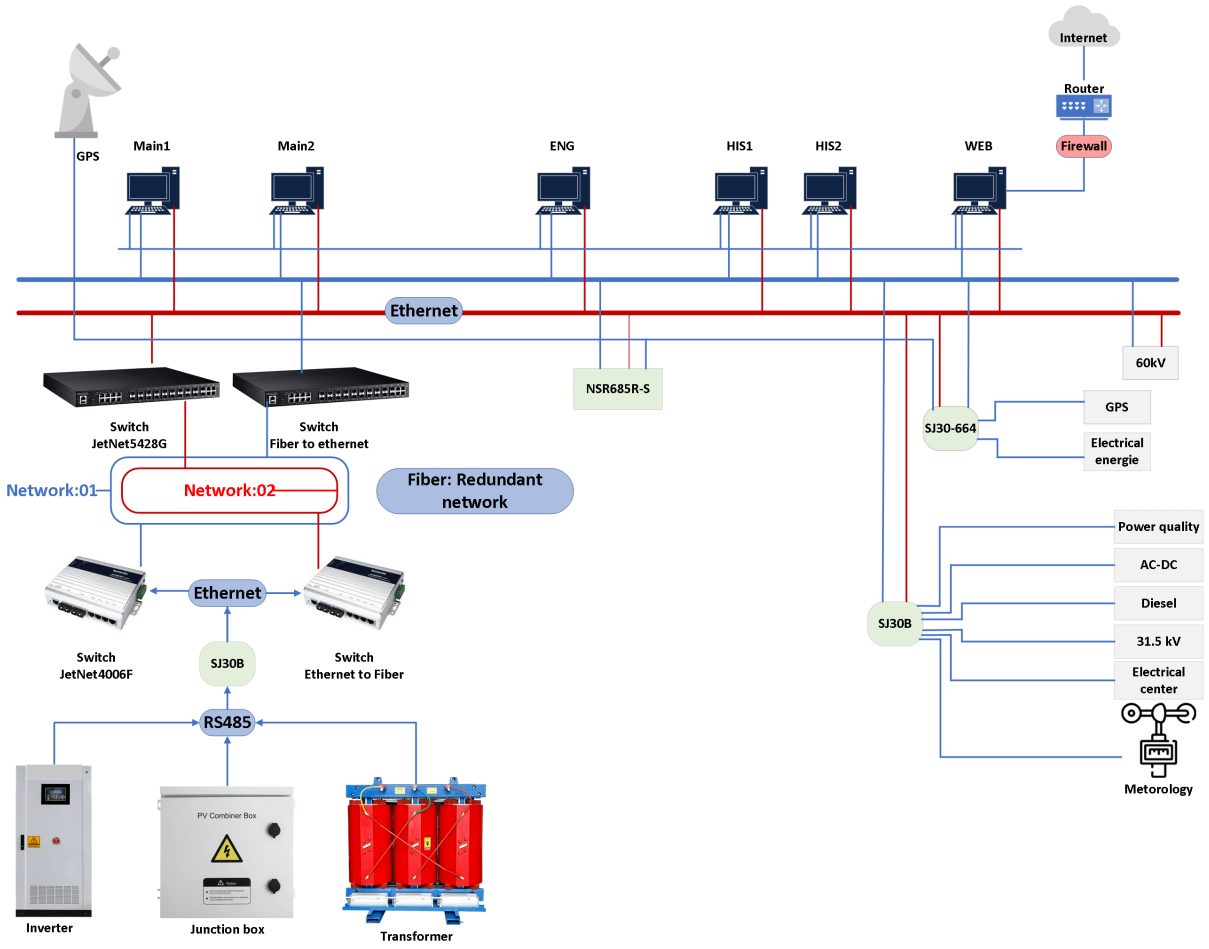


Figure 2.5: SCADA network

- **Subsystem Layout**

To further contextualize the station's infrastructure, Figure 2.6 presents a technical overview of key subsystems[28]. It includes real-world images of critical components(such as the photovoltaic array, inverter shelter, transformer unit, and substation configuration), corresponding to the operational zones previously described. This visual supplement reinforces the conceptual distinction between energy and data pathways by linking them directly to their physical installations.



Figure 2.6: Site technical equipment

2.3 Existing SCADA Interface

2.3.1 Interface Pages

The current SCADA platform provides a modular, real-time supervisory view of the photovoltaic installation. It enables operators to observe key operational parameters, monitor the status of critical components, and navigate between subsystems through a structured set of graphical interface pages.

To ensure clarity and operational efficiency, the interface adopts a hierarchical structure. Each page is dedicated to a specific monitoring function, allowing tasks to be organized across dis-

tinct views. This structure supports focused observation and streamlined system navigation. The underlying organizational logic is illustrated in Figure 2.7, which presents the navigational hierarchy and functional segmentation of the SCADA interface.

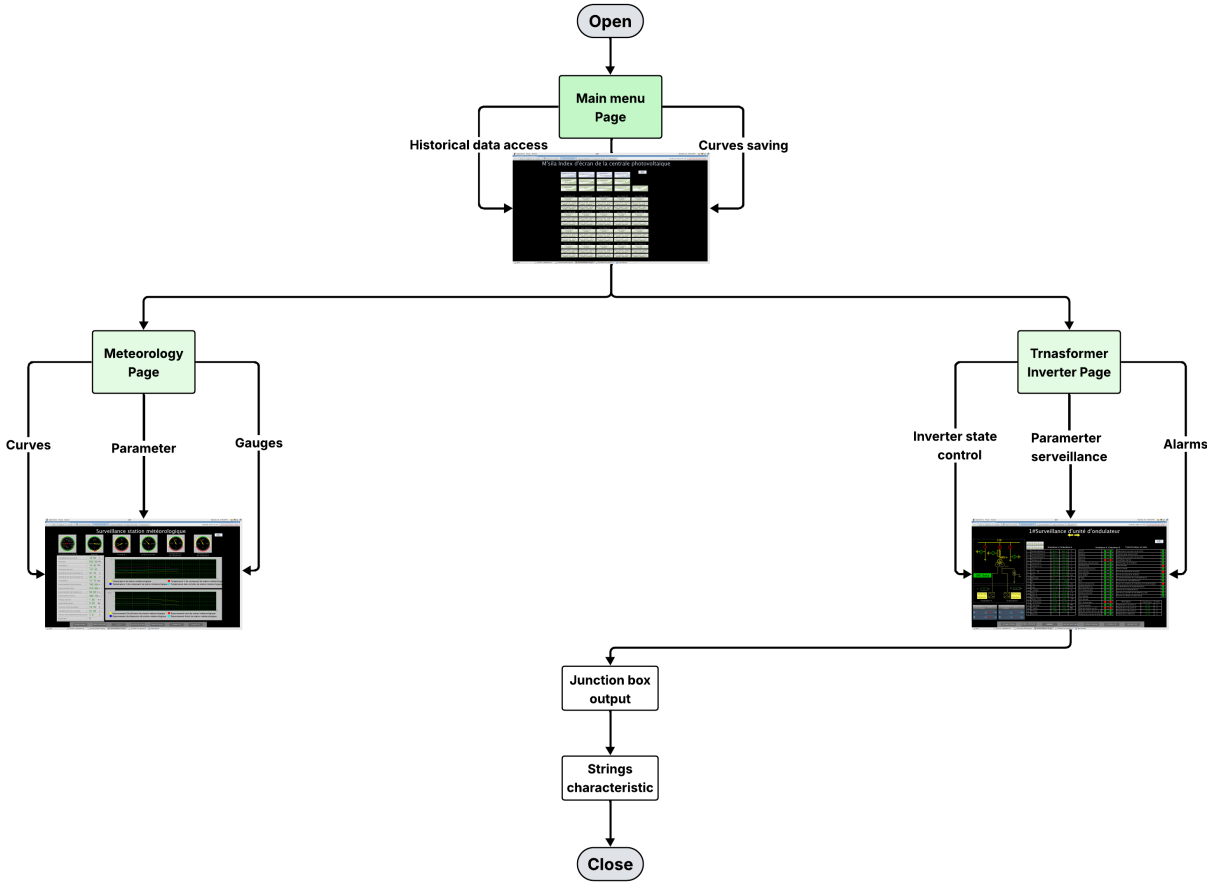


Figure 2.7: SCADA menu

The following pages represent the primary monitoring screens of the SCADA system. Their layout and role correspond to the aforementioned hierarchical organization, with each page offering dedicated insight into specific plant subsystems.

- **Main Menu**

The main menu serves as the central navigation hub, providing access to subsystem-specific pages, including subfield monitoring at the transformer-inverter and junction box levels, various system surveillance schematics, and report export functionalities. Its primary role is to streamline user interaction and facilitate task-oriented workflows.



Figure 2.8: SCADA mainMenu

- **Inverter and transformer**

This screen presents real-time electrical parameters for inverter units and transformer sections, including voltage, current, power output, and internal temperature. Its primary role is to enable operational tracking of the energy conversion chain and to support early detection of component anomalies.

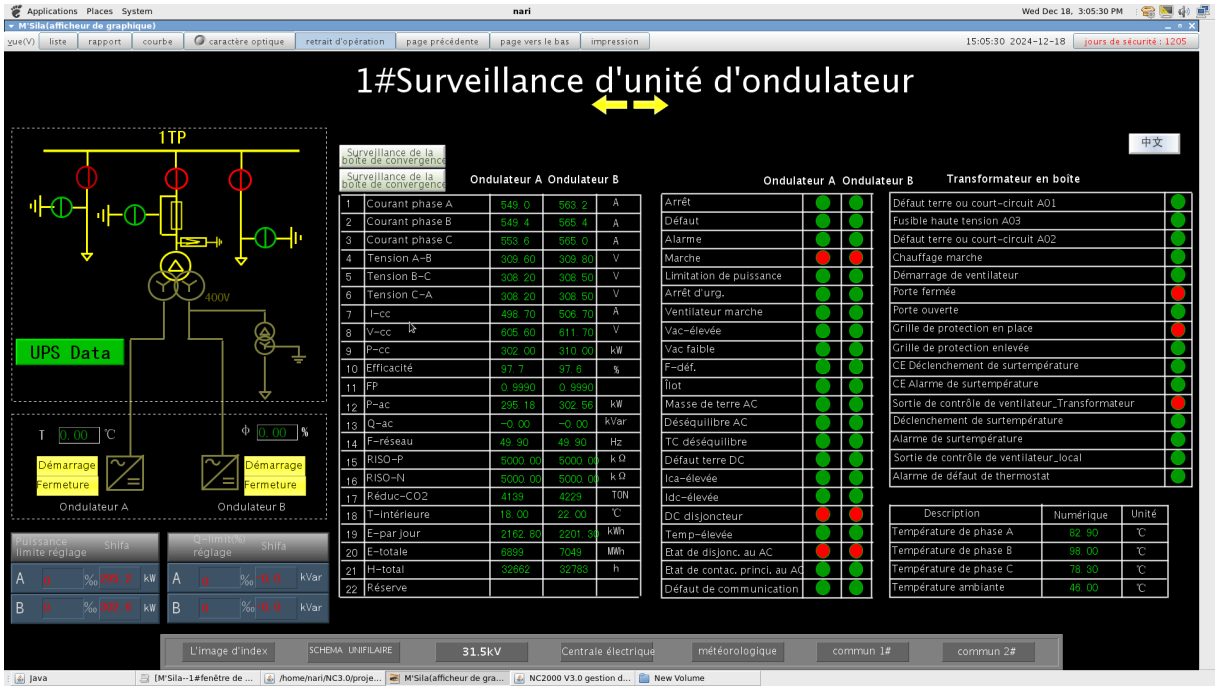


Figure 2.9: SCADA inverter surveillance

- **Meteorology**

Displays essential environmental data including solar irradiance, ambient and panel temperatures, and wind speed. This page supports performance contextualization by correlating environmental conditions with electrical output trends.

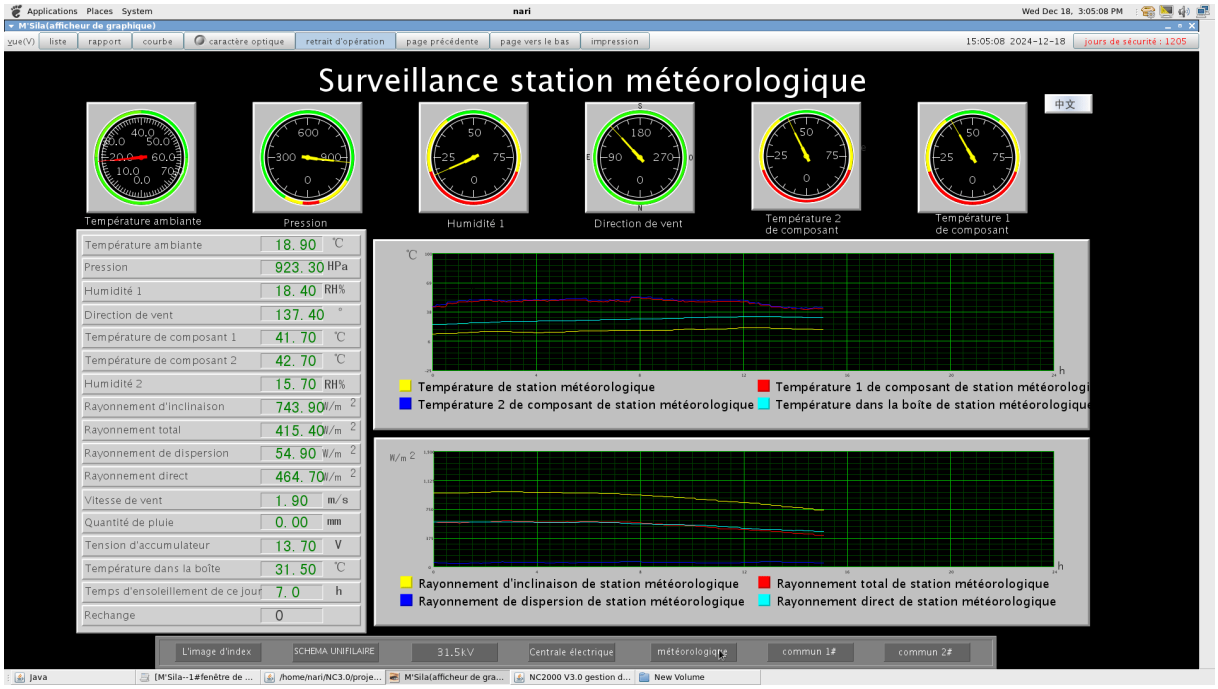


Figure 2.10: SCADA weather monitoring

- **Existing SCADA Interface Limitations**

While functionally operational, these interface pages exhibit notable limitations in visualization richness, analytical depth, and performance assessment metrics. The absence of predictive modeling, temporal analytics, and component□specific diagnostics constrains the system's ability to support proactive maintenance or intelligent supervision. These deficiencies underscore the need for a more advanced, insight□driven interface—a direction pursued in Chapter III.

2.3.2 Requirements and Assessment

To evolve from passive monitoring toward proactive fault prevention, the interface must shift from static data visualization to intelligent, insight-driven supervision. A modern supervisory platform should extend beyond basic observation to incorporate advanced capabilities that enable operators to interpret real-time conditions, assess component performance, and respond preemptively to emerging issues. To fulfill this role, several key enhancements are required:

- **Time-Series Data Visualization:** Interactive, temporal plots make it possible to detect periodic trends, identify slow performance drifts, and recognize subtle behavioural anomalies that static numeric displays may conceal.

- **Performance and Loss Analysis:** Quantitative insights into specific losses improve system-level understanding and facilitate relative performance assessment.

- **Predictive Performance Modeling:** Real-time comparison with expected voltage, current, and power trajectories helps define a meaningful performance baseline.

- **Fault Detection:** Deviations from expected performance profiles enable the early identification of anomalous behaviour across modules, strings, or inverters.

- **Diagnosis:** Observing the severity, duration, and distribution of these deviations supports fault classification and localization, guiding appropriate corrective actions. replacement.

2.4 Conclusion

The station's instrumentation and SCADA interface reliably capture key parameters; however, their basic displays and absence of performance-rating metrics constrain proactive monitoring. This project addresses those limitations by delivering advanced graphical analytics, predictive output curves, and comprehensive fault-detection diagnostics, thereby providing operators with a deeper, more actionable understanding of the plant's health and performance.

Chapter 3

SCADA Interface Development

3.1 Introduction

This chapter presents a comprehensive methodology for developing a Supervisory Control and Data Acquisition (SCADA) interface for large-scale photovoltaic (PV) power plant monitoring systems. SCADA systems are critical for achieving efficient and reliable PV system operation by enabling real-time monitoring and control. The development process begins with exploring the communication architecture, which forms the backbone of data transmission in such systems. This architecture uses a hierarchical network topology to incorporate field instruments (such as RS-485/Modbus RTU devices) with serial-to-Ethernet gateways and a fibre optic/Ethernet infrastructure, with everything converging on a centralised control server. This configuration ensures robust data acquisition, transmission, and processing across the entire PV infrastructure network. The second phase involves a detailed analysis of the architectural design and implementation of the SCADA server. Particular attention is devoted to the frameworks that govern data processing, storage, and visualisation. The aim is to develop an enhanced Human-Machine Interface (HMI) that improves usability and workflow to optimise operational efficiencies.

3.2 Communication Architecture

As shown in Figure 3.1, the hierarchical network architecture incorporates RS-485/Modbus RTU-compliant field devices at the terminal level. These devices communicate with the core infrastructure via serial-to-Ethernet gateways, which connect them to a high-bandwidth fibre-optic/Ethernet backbone. The network ends at centralized control servers, which enable system-wide data aggregation and real-time monitoring.

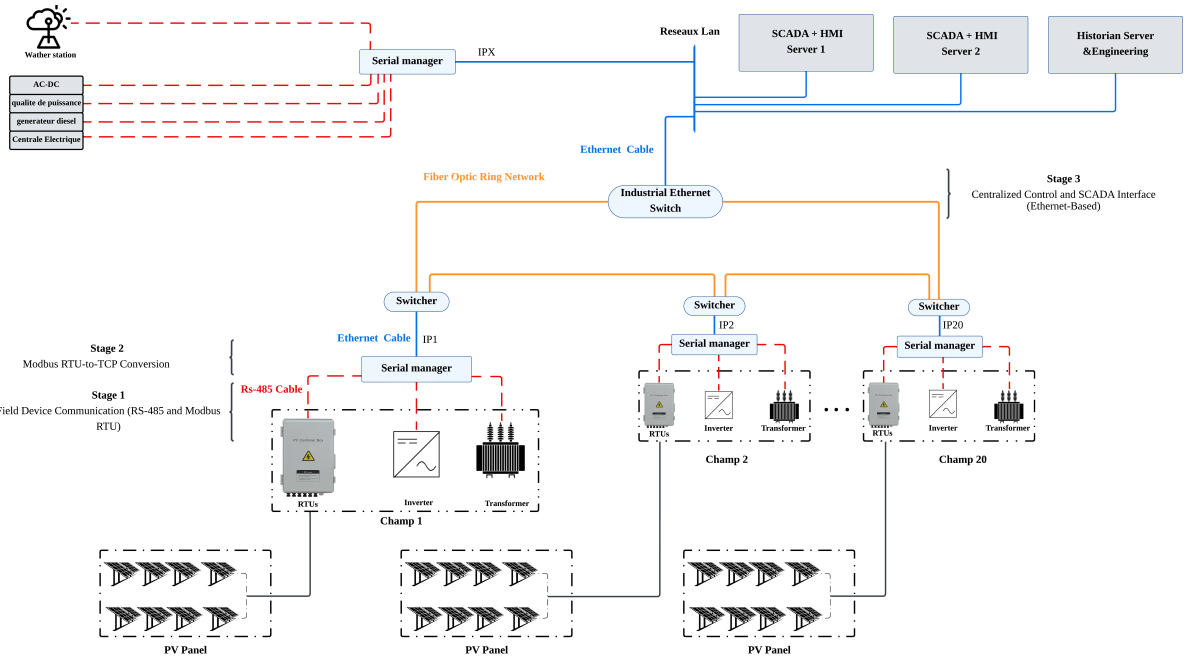


Figure 3.1: Schematic Architecture of Data Transmission in a Large-Scale PV Grid System

The first layer of this architecture is centred on field-level communication protocols, leveraging the robust characteristics of the RS-485 physical layer implementation combined with the Modbus RTU application layer protocols. This combination provides deterministic data exchange capabilities, even in the electrically challenging environments that are typically found in large-scale PV installations. The intermediate layer incorporates protocol conversion mechanisms that bridge the gap between legacy serial communication systems and modern Ethernet-based networks. Finally, the supervisory layer implements comprehensive data management, visualization, and control functionalities via a centralized SCADA server architecture. The data transmission process is structured into three distinct stages, ensuring reliable communication from field devices to the control centre:

3.2.1 Stage 1 Modbus RTU protocol

In the first stage, field devices such as remote terminal units (RTUs), inverters, transformers, and environmental sensors are interconnected via RS-485 cabling to a centralized serial device manager. The Modbus RTU protocol, widely adopted in such systems[29], is utilized over the RS-485 physical layer. As depicted in Figure 3.2, the Modbus RTU frame structure functions as the fundamental communication protocol, governing the formatting and transmission of data packets across the serial bus[24].

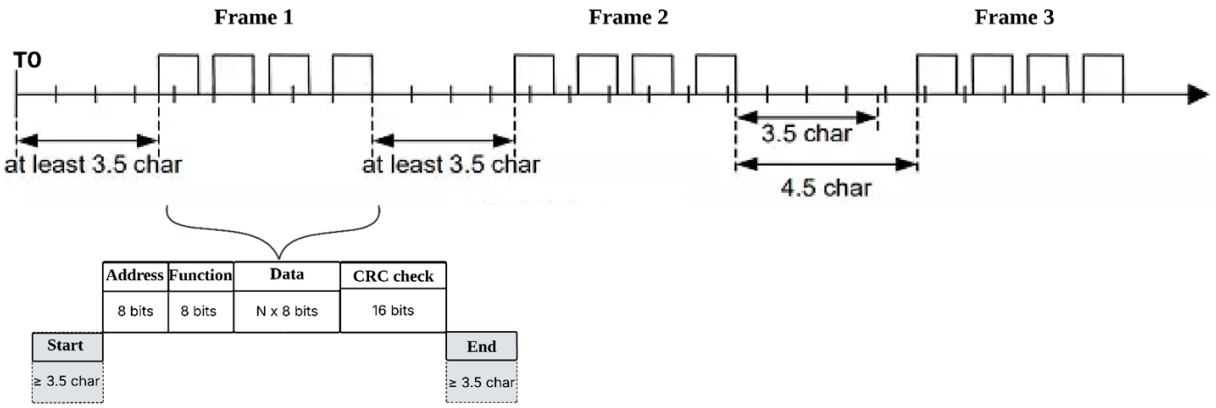


Figure 3.2: MODBUS message RTU framing

The Modbus RTU frame structure ensures reliable communication between devices and comprises:

- **Slave Address:** Identifies the target device (1-247) that should respond to the request.
- **Function Code:** Defines the specific operation to be performed, such as reading or writing registers (refer to Table 3.1 for common function codes).
- **Data:** Includes the parameters necessary for the operation, such as register addresses and values.

- **(Cyclic Redundancy Check):** A two-byte checksum that allows the receiving device to verify the integrity of the transmitted data, detecting any errors that may have occurred during transmission.

To clarify Modbus protocol operations, Table 3.1 provides the commonly used function codes and their corresponding descriptions[30].

Table 3.1: Common Modbus Function Codes

Function Code (Hex)	Function Name	Description
0x01	**Read Coils**	Reads the status of discrete (binary) output coils. Returns ON/OFF values for each addressed coil.
0x02	**Read Discrete Inputs**	Reads the status of discrete (binary) input contacts (e.g., digital sensors or safety switches). Returns ON/OFF values for each addressed input.
0x03	**Read Holding Registers**	Reads one or more 16-bit holding registers (read/write registers). Commonly used to retrieve analog data (e.g., voltage, current, power values).
0x04	**Read Input Registers**	Reads one or more 16-bit input registers (read-only registers). Typically used to retrieve sensor measurements or status information.
0x05	**Write Single Coil**	Writes a single discrete output (coil). Used to force one binary output ON (0xFF00) or OFF (0x0000) at a specified coil address.
0x06	**Write Single Register**	Writes a single 16-bit holding register at a specified address. Used to change one setpoint or configuration parameter (e.g., change operating mode).
0x0F	**Write Multiple Coils**	Writes values to multiple consecutive discrete output coils in a single transaction. The request includes a byte count and coil value bitmap.
0x10	**Write Multiple Registers**	Writes values to multiple consecutive holding registers. The request includes a byte count followed by the list of register values to be written.
0x16	**Mask Write Register**	Updates specific bits within a single holding register using an AND-mask and an OR-mask. Useful for modifying individual control or status bits.
0x17	**Read/Write Multiple Registers**	Combines a read and a write operation in one atomic transaction. The first part reads N registers; the second part writes M registers starting at another address.

3.2.2 Stage 2 Modbus RTU-to-TCP Conversion (Serial Device Manager)

In PV SCADA architecture, the serial device manager serves as an essential gateway between the Level 1 RS-485/Modbus RTU network and the Level 2 Ethernet network, as shown in Figure 3.3. Its main function is to convert Modbus RTU message frames, which were originally

designed for serial communication, into Modbus TCP packets suitable for transmission over an IP network.

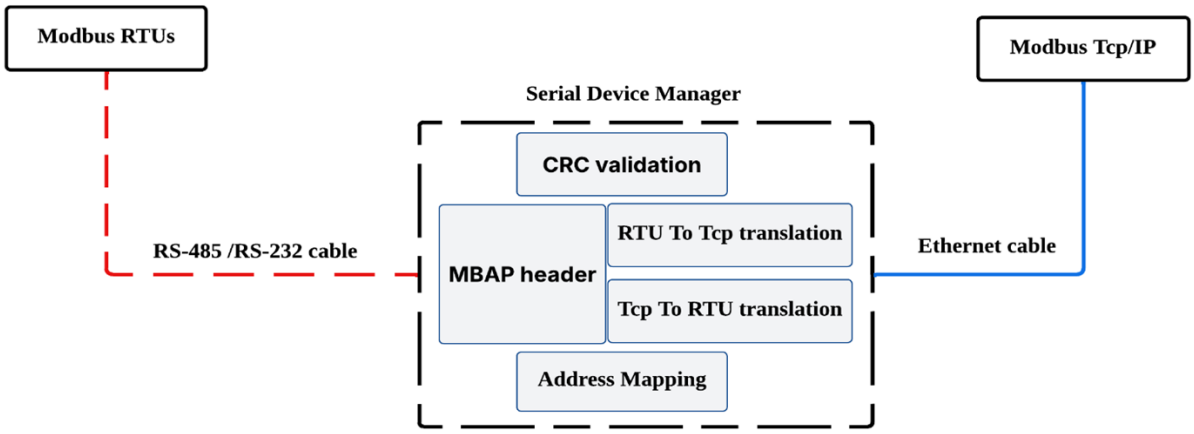


Figure 3.3: Modbus RTU-to-TCP Conversion

As illustrated in Figure 3.3, the conversion system consists of three components, namely: Modbus RTU field devices, a central Serial Device Manager, and the Modbus TCP/IP network. The Serial Device Manager shown in the dashed box serves as a Modbus gateway, performing the following functions to bridge serial and Ethernet communication:

- **CRC Validation:** The purpose of this operation is to verify the integrity of incoming Modbus RTU messages. This is performed to detect any potential transmission errors.
- **RTU to TCP Translation:** The process entails the removal of the RTU frame, which comprises the slave address, function code, data, and CRC. Subsequently, the Protocol Data Unit (PDU) is encapsulated within a Modbus Application Protocol (MBAP) header. Finally, it is packaged for transmission over TCP/IP.
- **TCP to RTU Translation:** The system under consideration receives Modbus TCP requests from the Ethernet side, removes the MBAP header, reassembles the PDU into an RTU frame, and appends the CRC for delivery over the serial link.
- **Address Mapping:** It is imperative to maintain a consistent mapping of Modbus addresses between the serial network and the IP network. This is necessary to ensure that each field device and TCP client retains a unique, correct identifier.

The Modbus TCP/IP is situated on the right side. This network functions as a communication platform for supervisory control systems (SCADA interfaces and HMI). The network establishes a connection to the Serial Device Manager via an Ethernet cable (indicated by the solid blue line). This connection facilitates the exchange of IP-based data with supervisory systems; therefore, establishing a thorough end-to-end pathway that extends from field sensors to high-level control interfaces. The original Modbus RTU payload is encapsulated and transmitted through the OSI layers[31], as demonstrated in Figure 3.4:

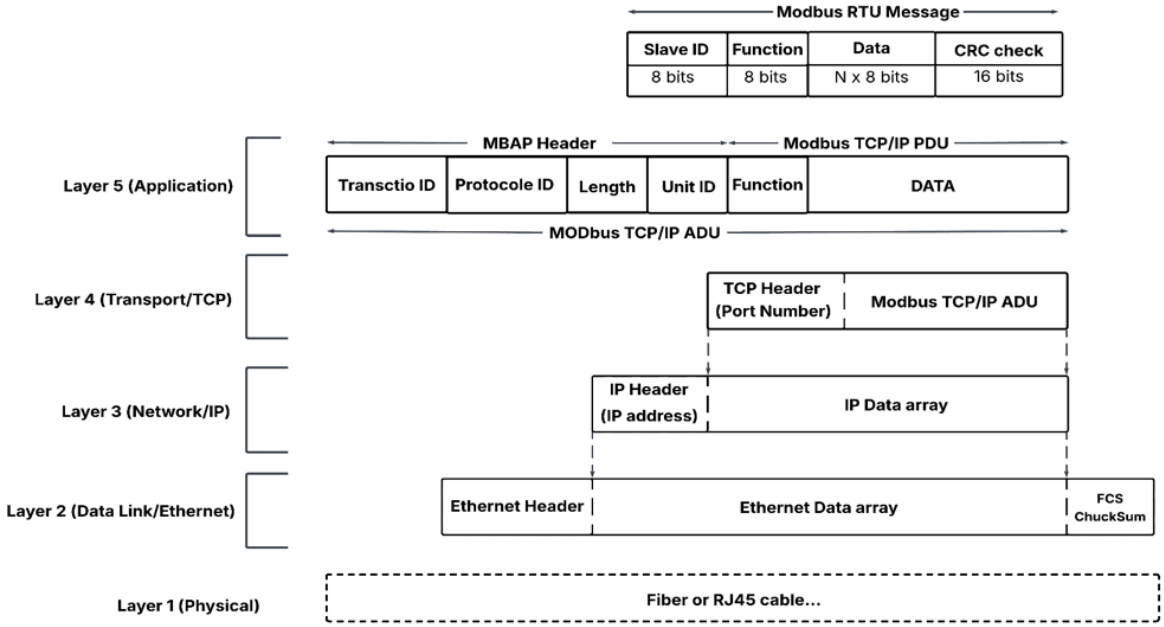


Figure 3.4: Modbus RTU Payload Encapsulation into a TCP/IP–Ethernet Frame

The original Modbus RTU payload (function code and data) is encapsulated in successive protocol headers before transmission over the physical layer(e.g., fiber-optic or RJ45 cable), as shown in Figure 3.4 and detailed below:

- **Application Layer (MBAP Header)**
  A 7-byte MBAP header (Transaction ID, Protocol ID, Length, Unit ID) is prepended to the RTU PDU (Function Code + Data).

- **Transport Layer (TCP Segment)**
  The MBAP + PDU is encapsulated in a TCP segment, which adds source and destination ports (typically port 502), sequence and acknowledgment numbers, and a TCP checksum to ensure reliable delivery.

- **Network Layer (IP Packet)**
  The TCP segment is wrapped in an IP packet, including source and destination IP addresses, TTL, a protocol field set to TCP, and an IP header checksum for correct routing.

- **Data Link Layer (Ethernet Frame)**
  The IP packet is placed inside an Ethernet frame, which adds source and destination MAC addresses, an Ether-type, and a Frame Check Sequence (CRC-32) for link-level integrity.

- **Physical Layer**
  Finally, the complete Ethernet frame is converted into a bitstream and transmitted over the chosen medium (e.g., fiber-optic cable or twisted-pair copper via RJ45).

3.2.3 Stage 3 Request-Response Communication Mechanisms

During each polling cycle, the SCADA server generates a Modbus TCP request directed at the target device (see Figure 3.5). This request includes critical identifiers such as Transaction ID, Protocol ID, Unit ID, Function Code, and Register Addresses, all encapsulated within a Modbus Application Data Unit (ADU). Upon receiving the request, the device processes these fields and responds with a structured data frame (Figure 3.6). The response contains the requested data in hexadecimal format, which the SCADA server decodes using the device's specific register map. This request-response mechanism is fundamental to the systematic and reliable collection of data from field devices.

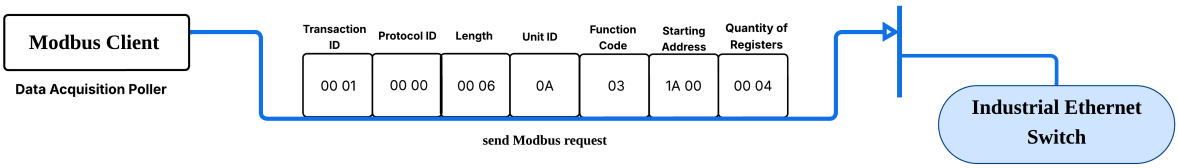


Figure 3.5: Sending Modbus TCP Requests

The target device responds with a Modbus TCP frame :

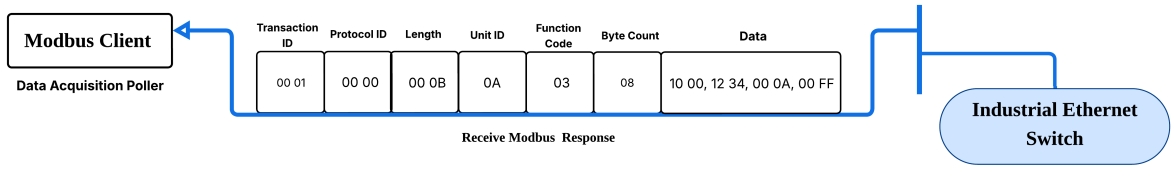


Figure 3.6: Receive Modbus TCP Response

Once received, the responses are decoded and normalized before being transmitted to the SCADA server. At this stage, the data has been structured and stored within the database. The processed measurements are then made available through the Front-End interface, enabling real-time monitoring, historical analysis, and fault detection.

This progression from low-level protocol communication to high-level data visualization is essential for ensuring complete system observability and operational intelligence in industrial PV monitoring architectures.

3.3 Design and Implementation of a Real-Time SCADA Server

After examining the low-level data transmission from RS-485/Modbus RTU devices via the Ethernet network, this section explores the high-level architecture of the SCADA server. The Back-End Data Processing module manages data acquisition, storage, and control functions

over the Ethernet network. The Front-End Visualisation module, realised as a human-machine interface (HMI), provides operators with tools for real-time monitoring and control. Figure 3.7 illustrates the overall SCADA server architecture.

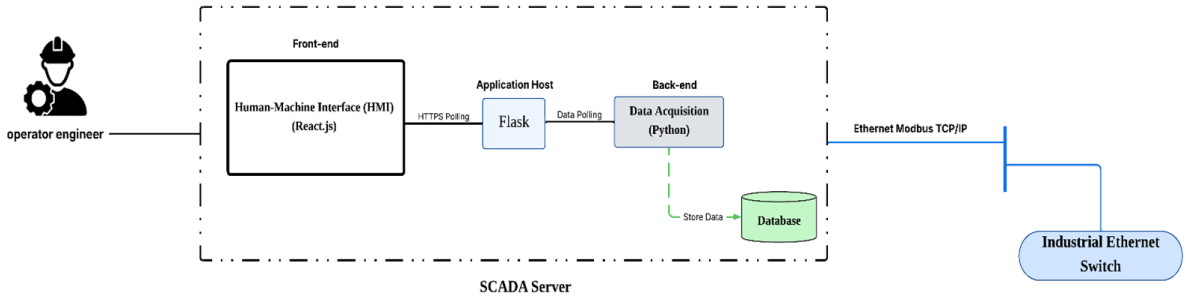


Figure 3.7: SCADA Server Architecture and Network Connectivity

The system architecture, depicted in Figure 3.7, presents a SCADA server configuration employing Modbus TCP/IP protocol for Ethernet-based communication with field devices. The server's operational framework comprises four principal functional modules:

- **Back-End Data Processing:** This layer handles core logic tasks, including Modbus read-/write operations, data validation, configuration management, and formatting data for storage and analysis.
- **Data Storage:**Incoming data from the Back-End is systematically recorded and organized within the chosen database system.
- **Application Server (Flask):** Serving as an intermediary, Flask exposes RESTful API endpoints that facilitate data exchange between the Back-End processing and the HMI. It is also responsible for authentication, request routing, and applying business logic.
- **Front-End Visualization:** The HMI provides real-time dashboards, trend charts, and alarm panels, allowing operators to monitor system status and respond to events promptly.

The following sections describe the design principles and implementation steps for each of these components in the SCADA server.

3.3.1 Back-End Data Processing

The Back-End Data Processing component handles all core logic tasks, including Modbus read-/write operations, data validation and normalization, configuration management, and data formatting for storage and analysis. This component operates continuously in the background to maintain real-time data acquisition regardless of Front-end activity.

- **Configuration File**

The configuration file serves as a centralized repository for all device-specific parameters. Each field device is identified by its Modbus TCP/IP network parameters, primarily its IP address. Additionally, each logical tag is mapped to its corresponding register address, data type, scaling coefficient, and unit of measurement. By consolidating these definitions in a single file (see Figure 3.8), the system becomes more maintainable and flexible. For instance, incorporating a new inverter or updating the serial device manager involves only modifying the configuration file, without necessitating changes to the core code base. This centralization ensures consistency in register mappings and facilitates version control, thereby enhancing the overall reliability of the system.

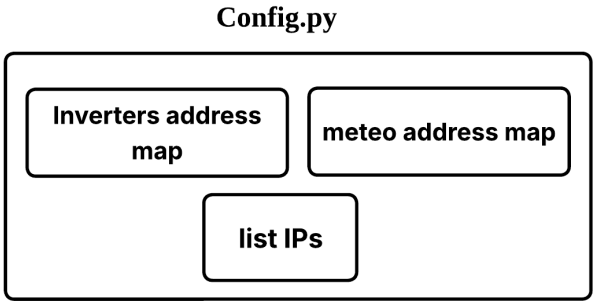


Figure 3.8: Config file

- **Data Normalization**

Figure 3.9 illustrates the data normalization process, where floating-point measurements from RTUs and PLCs are typically transmitted as integer values spread across one or more 16-bit registers. To make this data usable, it is essential to convert these raw integer values into their corresponding engineering units, such as volts, amps, or degrees Celsius. This conversion ensures that the data is accurately interpreted and analysed within the SCADA system. It is worth mentioning that all factors and unit conversions are sourced from the inverter's technical data sheet.

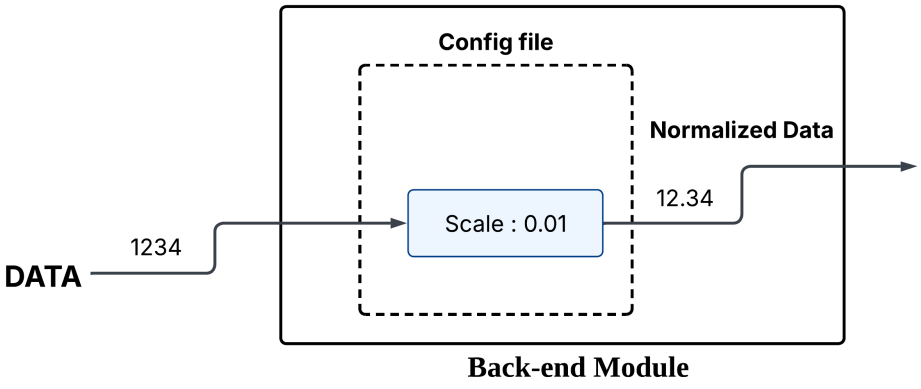


Figure 3.9: Data Normalization Process

3.3.1.1 Python-Based Data Acquisition Module

The back-end functionality of the SCADA server is divided into separate Python scripts, each of which handles specific tasks related to data acquisition, storage, performance calculation, and analytics. The read/write modules are among the most important of the key modules. In this module, two Python scripts are designed for reading and writing incoming data by referencing the centralized configuration file to identify each target device, namely:

- **Inverter.py**: Responsible for polling photovoltaic data inverters.
- **station_meteo.py**: Handles data collection from meteorological sensors.

Figure 3.10 illustrates the system structure and operational workflow, showing the implemented modules executing the following sequence:

- Loading configuration parameters (IP address, Modbus unit ID, registers, polling interval).
- Formulating and transmitting Modbus protocol requests.
- Apply scaling factors to convert 16-bit register values to engineering units (floating point).
- Save normalized measurements to the database.

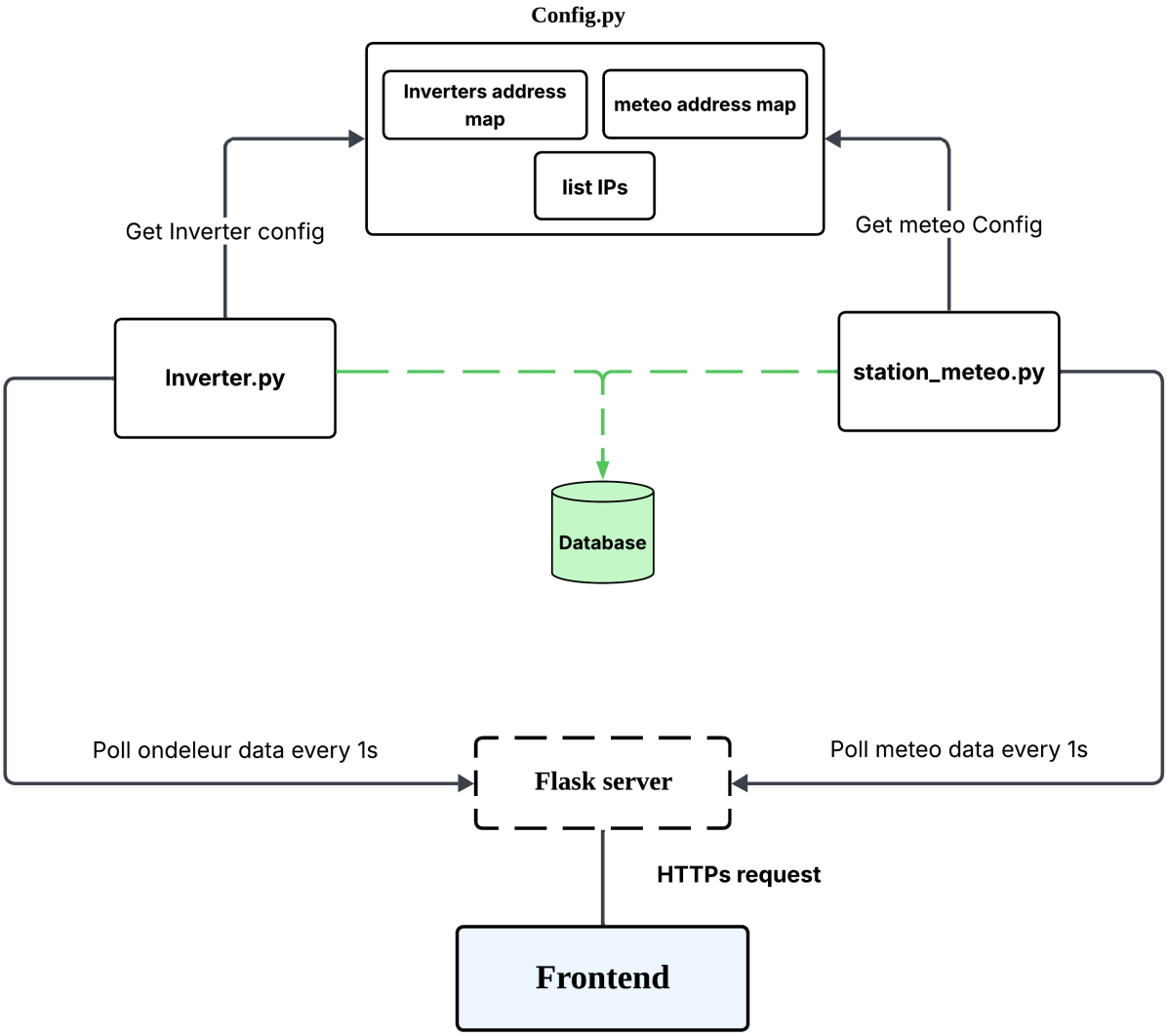


Figure 3.10: Architecture of Polling Modules

This modular approach ensures that each component focuses on its specific function, thereby enhancing clarity, maintainability, and ease of debugging.

- **Data Export to Excel**

The Data Export to Excel feature enables operators to extract historical inverter and meteorological data from the database into a spreadsheet format. This functionality supports offline analysis, reporting, and long-term archiving without the need for external tools. Figure 3.11 illustrates how this feature is integrated within the Back-End scripts and how it interacts with the HMI to provide user-driven access to historical measurements.

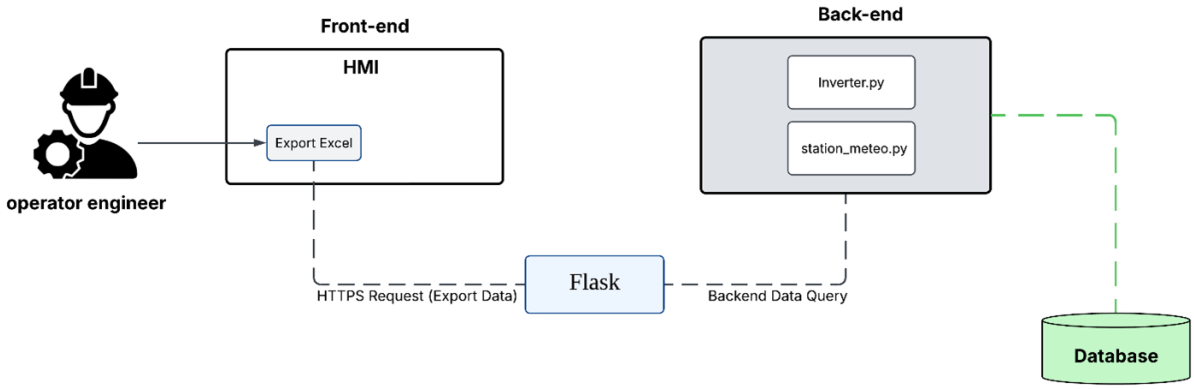


Figure 3.11: Data Export and Back-End Processing System

As illustrated in Figure 3.11, a dual-level architecture including a front-end human-machine interface (HMI) and a back-end processing layer is highlighted. When the engineer operator selects 'Export Excel' on the HMI, an HTTPS request is sent to the Flask server. The server translates this into a data request and forwards it to the 'Inverter.py' and 'Station_Meteo.py' modules. These scripts retrieve the requested historical measurements from the database and format them into a spreadsheet. The Flask server then returns this spreadsheet to the HMI, allowing the operator to download it directly.

3.3.1.2 Parameter Calculation Module

After storing real-time measurements, a supplementary Python module was developed to compute higher-order performance metrics for each inverter (see Figure 3.12). This module calculates the following performance indicators:

- Yield Total (Yr)
- Actual Yield (Ya)
- Firm Yield (Yf)
- Performance Ratio (PR)

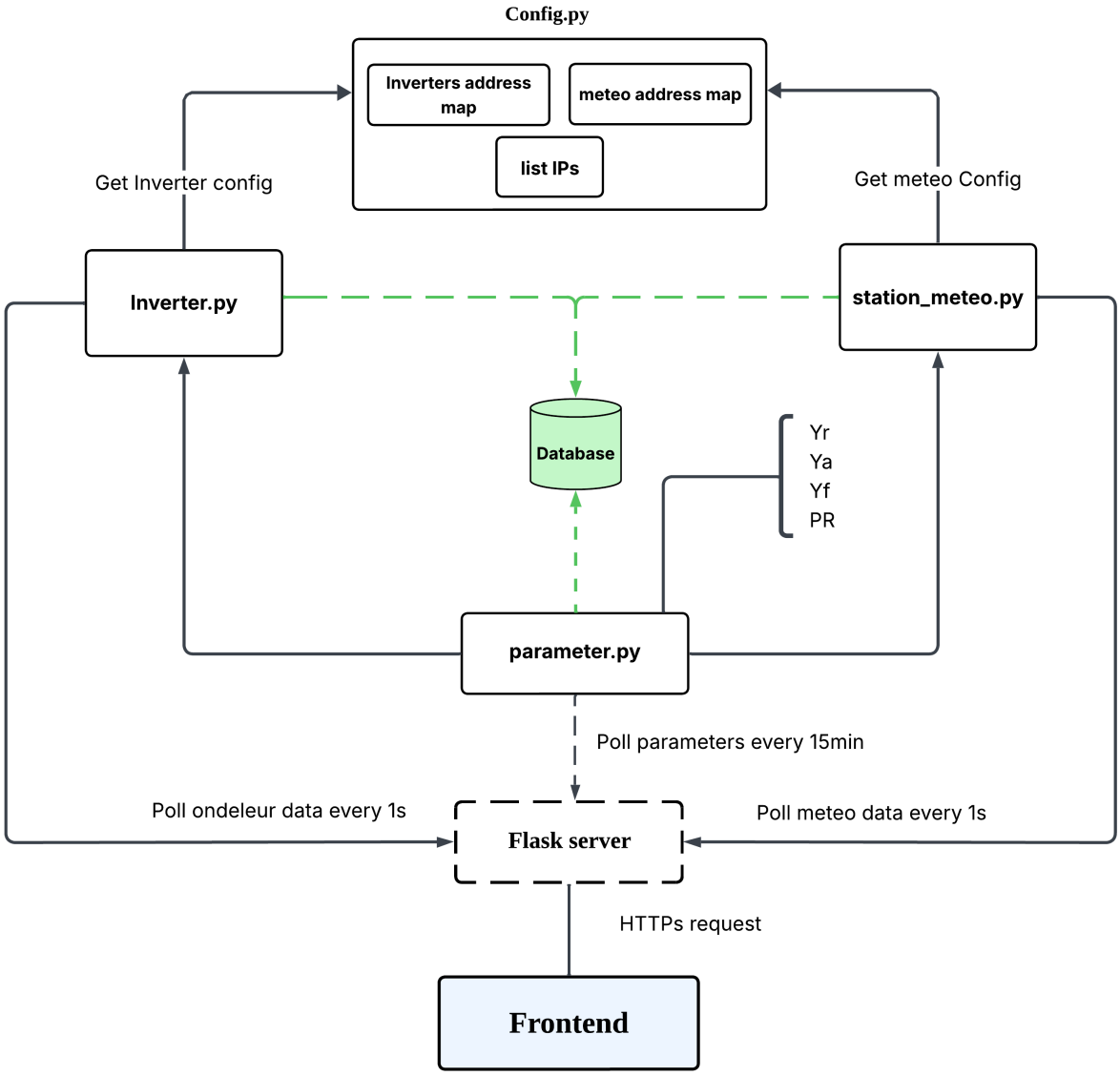


Figure 3.12: Flowchart with Parameter Module

As can be seen in Figure 3.12 of the provided diagram, the “parameter.py” function is designed to obtain real-time data from the “Inverter.py” and “Station_Meteo.py” modules in order to calculate the following variables: Yr, Ya, Yf, and PR.

3.3.1.3 Predictive Module

The function “prediction_model.py” script is to predict the inverter DC power output. The system integrates processed feature inputs from station_meteo.py (meteorological data) and actual measurements from inverter.py. Subsequently, it calculates a predictive target, designated Pdc_predict. This target is stored in the MongoDB database and subsequently displayed on the interface (Front-End).

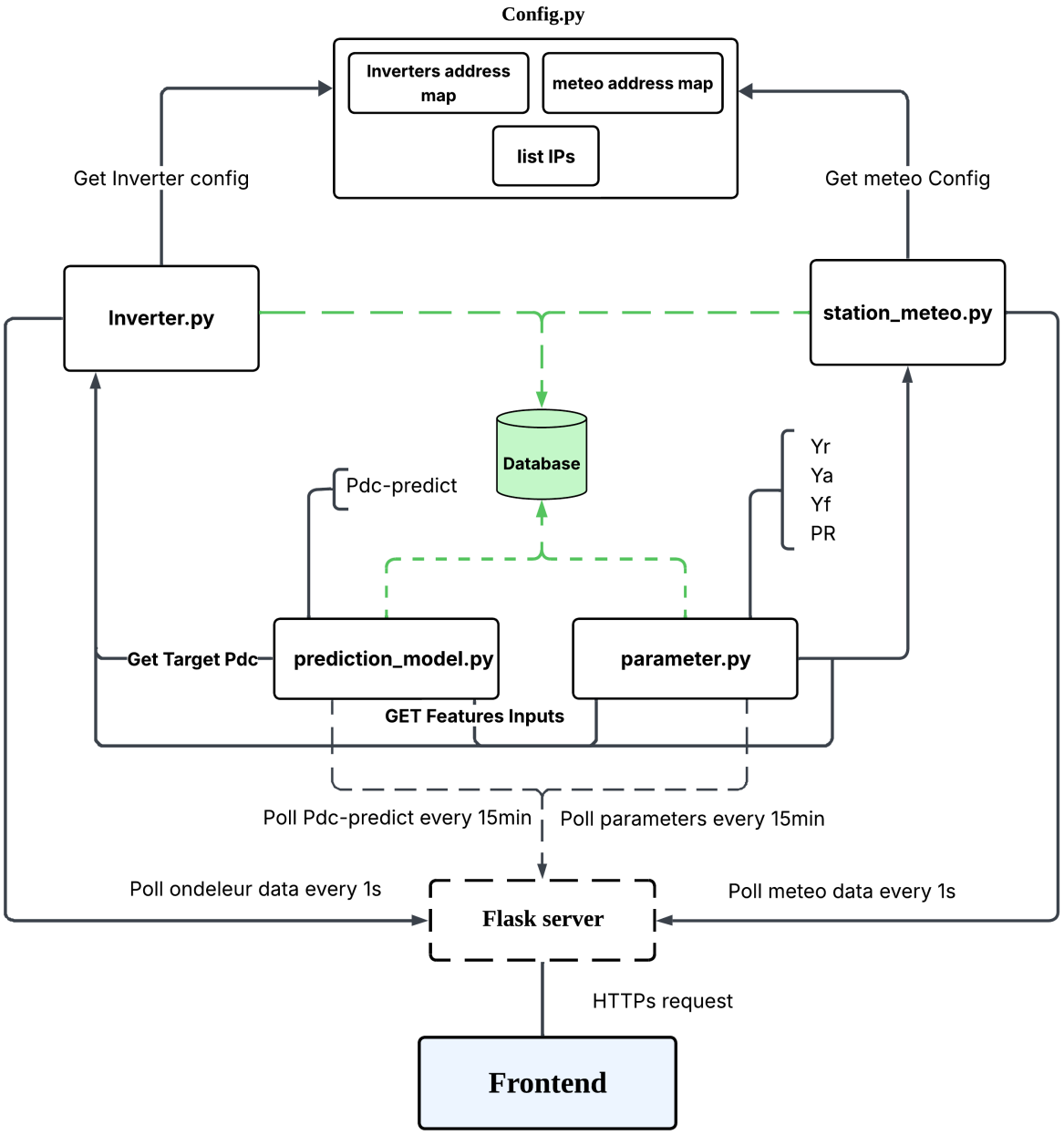


Figure 3.13: Flowchart with Analytic Module

In this Figure 3.13, the “Prediction_model.py” script gets data input features from the “station_meteo.py” script and the target measurement (PDC) from “inverter.py” in real-time data polling. The model is pre-trained using ML techniques such as Random Forest or the proposed CNN-LSTM. After that, it stores the outputs as "Pdc_Prediction" in MongoDB.

3.3.2 Data Storage with MongoDB

In a data-driven environment such as PV system monitoring, it is crucial to be able to store, manage, and process large volumes of time-series data efficiently. MongoDB, a document-oriented NoSQL database, offers the flexibility, scalability, and performance required for this application. Its JSON-like structure allows complex and evolving data to be accommodated without the need for rigid data migration. Furthermore, MongoDB's horizontal scaling (sharding) and built-in time-to-live (TTL) indexes enable the ingestion of thousands of data points per second[32]. High-resolution data is retained temporarily, while older records are automatically archived or removed[33]. Figure 3.14highlights the structure of the used database.

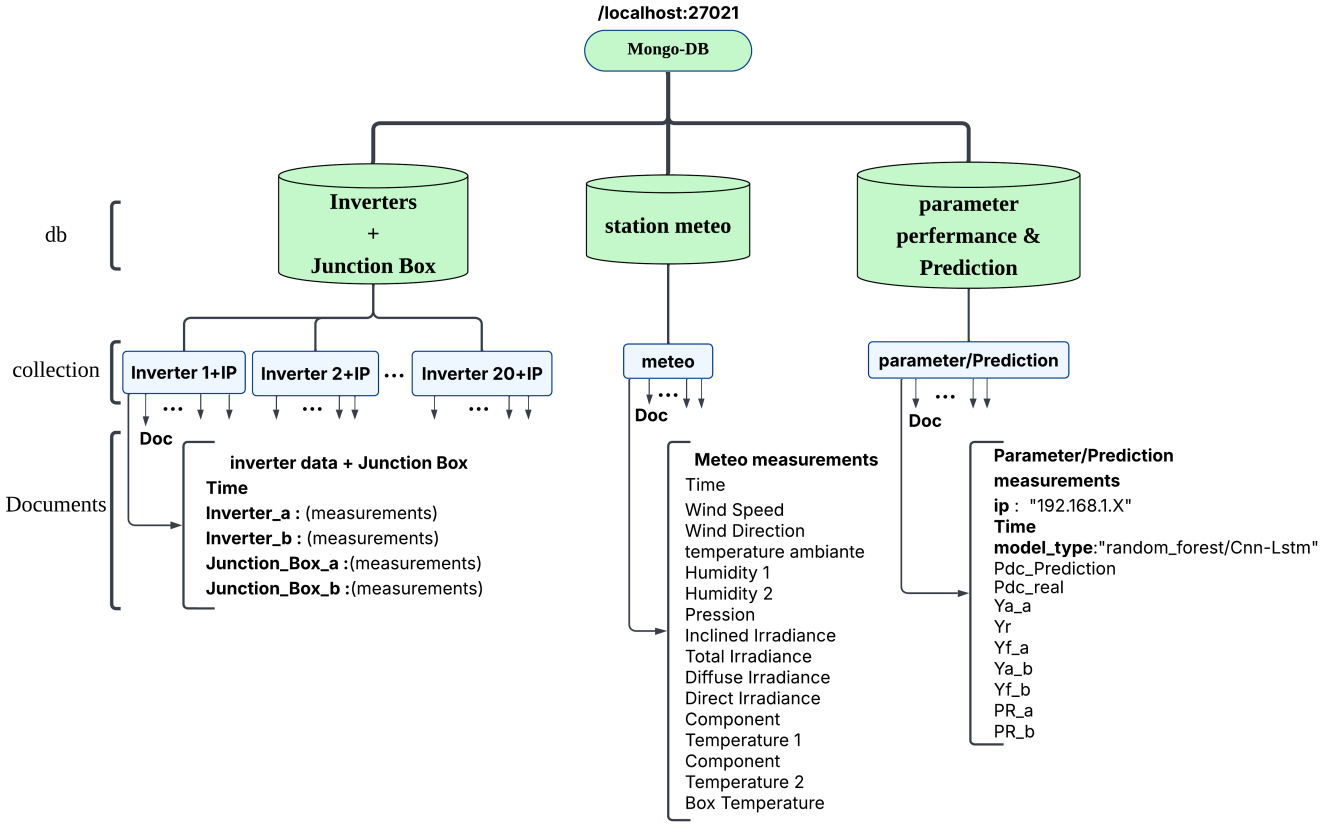


Figure 3.14: Database Architecture

The database is logically divided into three primary collections:

- **Inverters Collection:** Stores real-time and historical data from inverters and junction boxes.
- **Meteo Collection:** Captures environmental data from meteorological stations.
- **Performance and Prediction Model Collection:** This collection aggregates performance metrics (YR, Ya, Yf, and PR) and prediction model outputs. This structure facilitates efficient data retrieval and supports predictive analytics and fault detection systems. Table 3.2

presents a comparative analysis of MongoDB (NoSQL) and traditional relational (SQL) databases.

Table 3.2: Comparative Analysis of MongoDB (NoSQL) and Relational SQL Databases

Criterion	MongoDB (NoSQL)	Relational SQL
Storage Footprint	Compact storage with BSON compression; built-in TTL purges old data automatically	Generally larger due to normalized tables and indexes; requires manual archival scripts.
Schema Flexibility	Schema-less, can add fields without migrations	Rigid schema; altering table structures often requires migrations
Performance	Optimized for high-frequency writes (thousands of docs/sec); fast time-range queries via compound indexes	Good for transactional workloads; large bulk inserts and time-series queries can degrade performance
Ideal Use Case	High-throughput telemetry, evolving data models, and real-time analytics	Structured, transactional data with stable schemas (e.g., financial records)

3.3.3 Application Server (Flask)

The Flask application server acts as middleware, facilitating data exchange between the Back-End processing layer and the Front-End HMI. It receives HTTP requests from the Front-End, processes them, and communicates with Back-End services to retrieve or update data. This architecture enhances scalability and simplifies system maintenance by allowing independent updates to the Front-End without affecting the Back-End logic. Figure 3.15 presents the architectural role of Flask in orchestrating HTTP request handling and mediating data exchange between front-end and back-end system components.

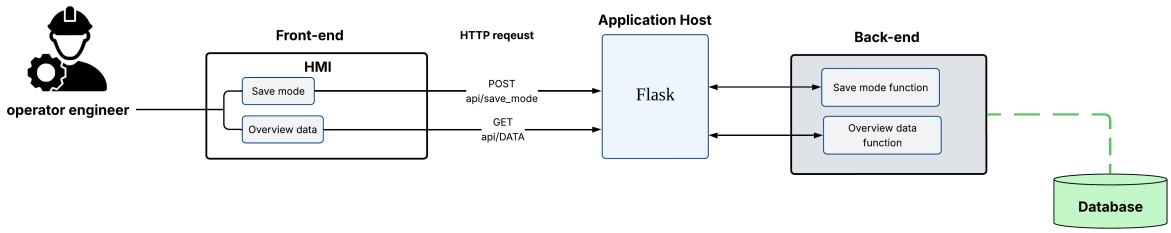


Figure 3.15: Flask as a Middleware: Managing HTTP Requests and Data Flow

Specifically, Flask functions as a server that receives HTTP requests from the Front-End dashboard, processes them, and communicates with the Back-End services to retrieve the necessary data. These responses are relayed back to the Front-End, where they are either visualized or interacted with by the user. This approach enhances scalability and simplifies system upgrades.

3.3.4 Front-End and Visualization

The Front-End HMI is designed to provide operators with an intuitive interface for monitoring and controlling the PV system. It is structured into three main pages:

- **Overview:** Displays system status and aggregated data.
- **Details:**: Allows operators to access specific inverter and junction box information.
- **Report:**Enables data export and trend visualization.

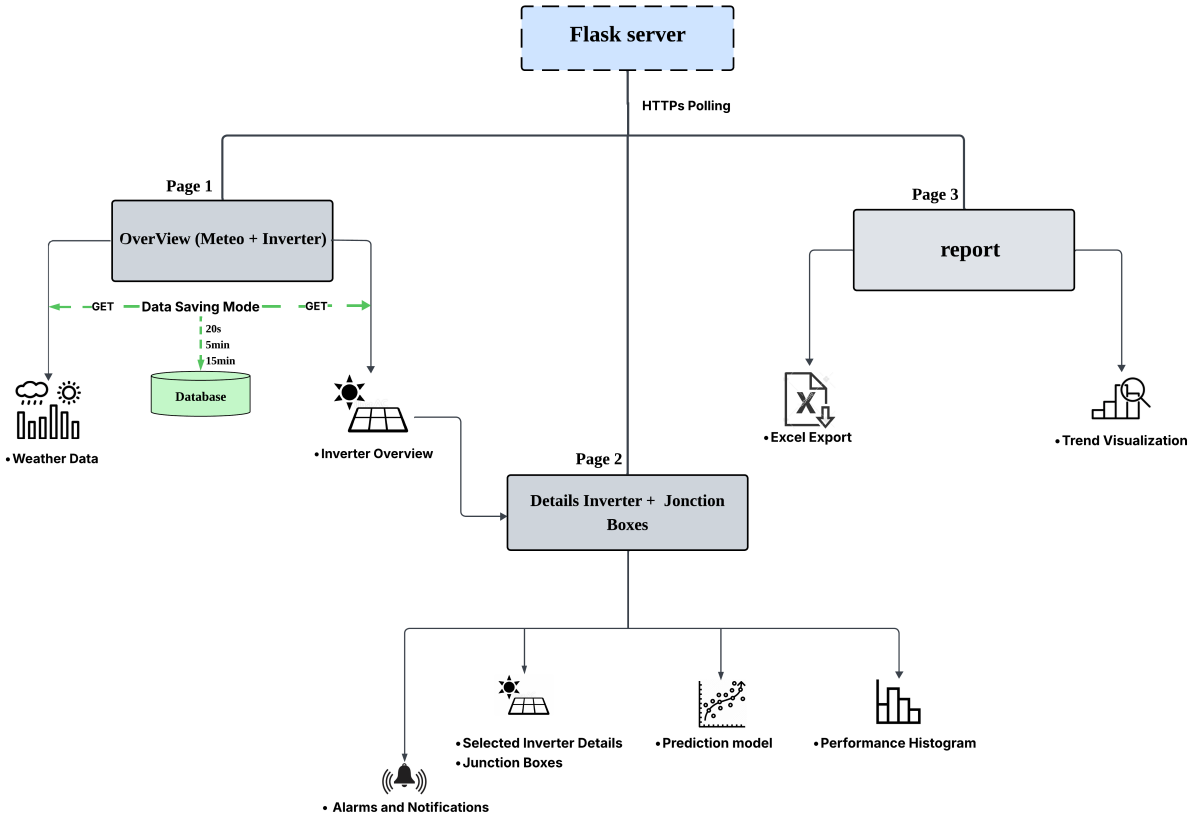


Figure 3.16: Front-End HMI Architecture and Data Flow

The HMI interacts with the Back-End via HTTPS polling, ensuring real-time updates and seamless operation.

Essentially, the Front-End serves as the central point of interaction for operators, as shown in Figure 3.16. It abstracts the complexity of Back-End data sources and presents actionable insights and control options for the PV system.

To provide an overview of the entire system, the following global architecture diagram, Figure 3.17 consolidates the key components and interactions discussed throughout this work. It

illustrates the complete data flow from user interaction at the Front-End interface, through the Flask-based middleware layer, to the Back-End processing and data storage components.

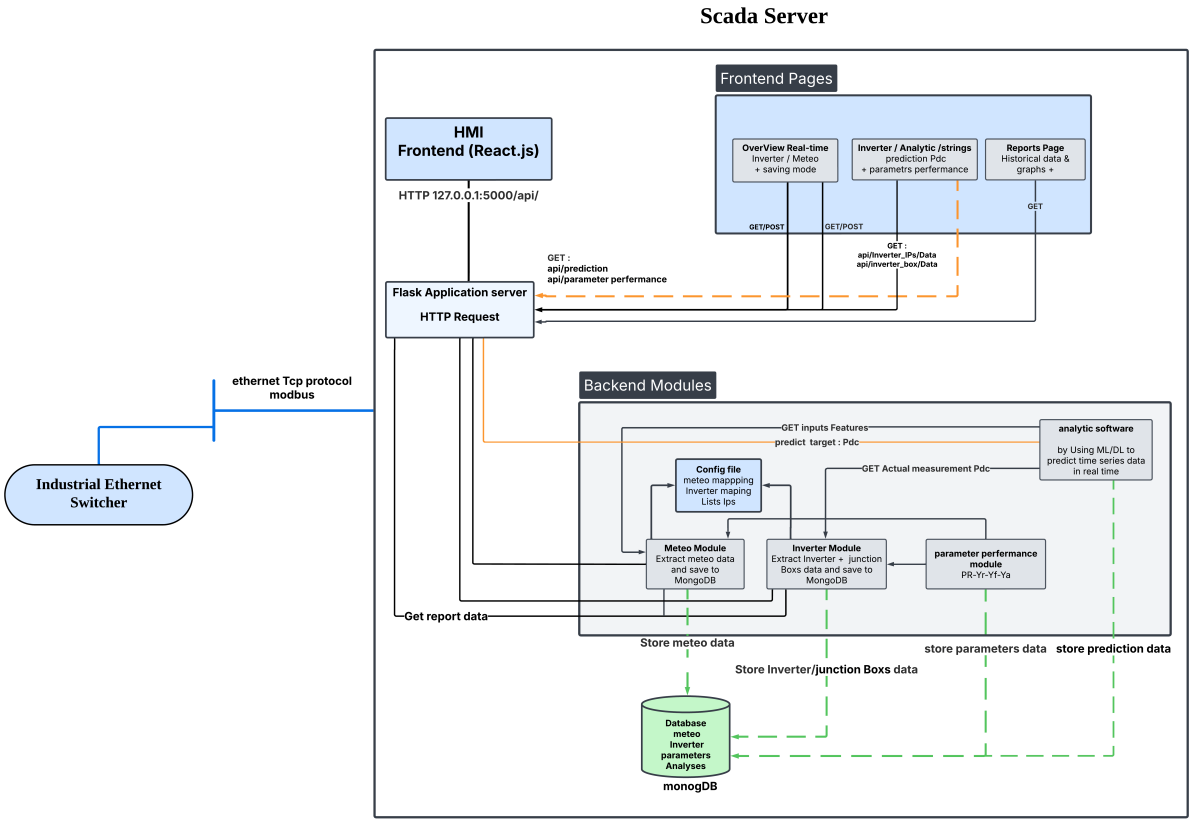


Figure 3.17: Full-Stack Web Application Architecture

3.4 Data Server Emulation

The development of a Supervisory Control and Data Acquisition (SCADA) system for large-scale photovoltaic (PV) power plants requires a rigorous and iterative testing phase. Given the complexity and real-time nature of such systems, continuous testing under realistic conditions is essential to validate performance, ensure reliability, and fine-tune both the backend and frontend components.

However, direct access to a live PV plant during development is often impractical due to operational constraints, safety concerns, or logistical limitations. To overcome this, system emulation emerges as a critical methodology. By simulating the behaviour of the PV plant and its data flow, developers can carry out comprehensive validation, fault injection, and performance monitoring in a controlled environment.

The **Raspberry Pi 4 Model** is proposed as a *low-cost, flexible, and efficient emulation platform*. Leveraging its robust processing capabilities, network interfaces, and Linux-based environment,

the Raspberry Pi is configured to act as a **virtual server**, mimicking the behaviour of the principal data acquisition switch in a real PV monitoring system, as shown in figure 3.18.

Specifically, the Raspberry Pi is programmed to:

- Replay a **historical dataset spanning two years**, representing key operational variables such as solar irradiance, ambient temperature, voltage, current, and power output.
- Emulate **real-time data transmission** using standard industrial communication protocols `Modbus TCP`.
- Maintain consistent controllable sampling intervals (e.g., every 5 or 10 seconds), simulating the temporal dynamics of a real PV installation.

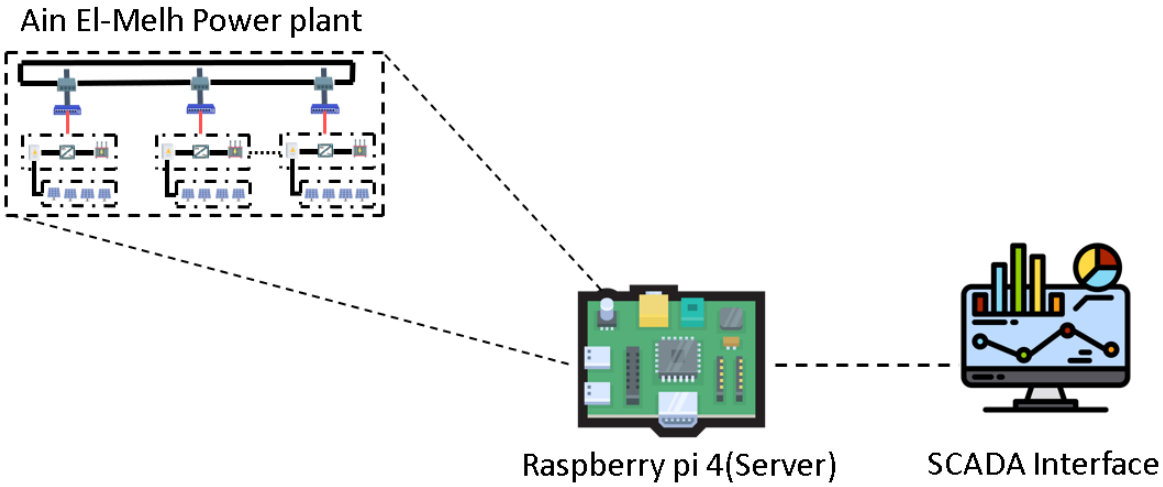


Figure 3.18: Raspberry pi4 server

This approach allows for full-scale system testing, including:

- Backend ingestion and processing performance under sustained data flow.
- Frontend responsiveness and visualization accuracy.
- Real-time fault injection for evaluating diagnostic algorithms.
- System stability, error handling, and logging mechanisms.

The emulation process implemented on the Raspberry Pi 4 follows a structured methodology designed to replicate the behaviour of a real PV monitoring system. This approach ensures accurate data representation and seamless integration with the SCADA interface, figure 3.19 shows the data sending emulation. The following steps summarize the procedure:

- **Initial Configuration:** An IP address and a slave ID are assigned to the Raspberry Pi to emulate the identity of a field device within a Modbus-based or IP-addressable network.

- **Data Preparation:** A two-year historical dataset, including key parameters such as irradiance, temperature, voltage, current, and power, is stored locally on the Raspberry Pi. This dataset serves as the source for the emulated data stream.
- **Data Conversion:** Since many industrial protocols transmit data in raw integer form, the floating-point measurements in the dataset are converted into integer format. This ensures compatibility with the SCADA system and allows realistic simulation of sensor output.
- **Data Mapping and Addressing:** Each subfield is associated with a specific IP address. Within each virtual device, the corresponding measurement types are encapsulated into specific Modbus address registers within specified register types, such as handling and coil registers, also the number of registers addressed is assigned to each type of data, allowing the SCADA system to interpret them accordingly.
- **Transmission Parameters:** The size of each data packet and the sampling interval (e.g., every 5 or 10 seconds) are configured to mimic real-time operational behaviour. These parameters are adjustable to simulate different network loads and measurement frequencies.
- **Server Initialization:** The emulation server script is launched remotely via SSH connection to the Raspberry Pi. This script initializes the configured transmission parameters and begins serving data.
- **Data Transmission:** Data is transmitted from the Raspberry Pi to the SCADA backend via an Ethernet connection. The SCADA system receives, stores, and visualizes this data as though it originated from real hardware components in a live PV plant.

This emulation methodology enables robust system testing under realistic and reproducible conditions. It also facilitates the validation of data ingestion, processing pipelines, and visualization tools within the SCADA environment prior to deployment in a live setting.

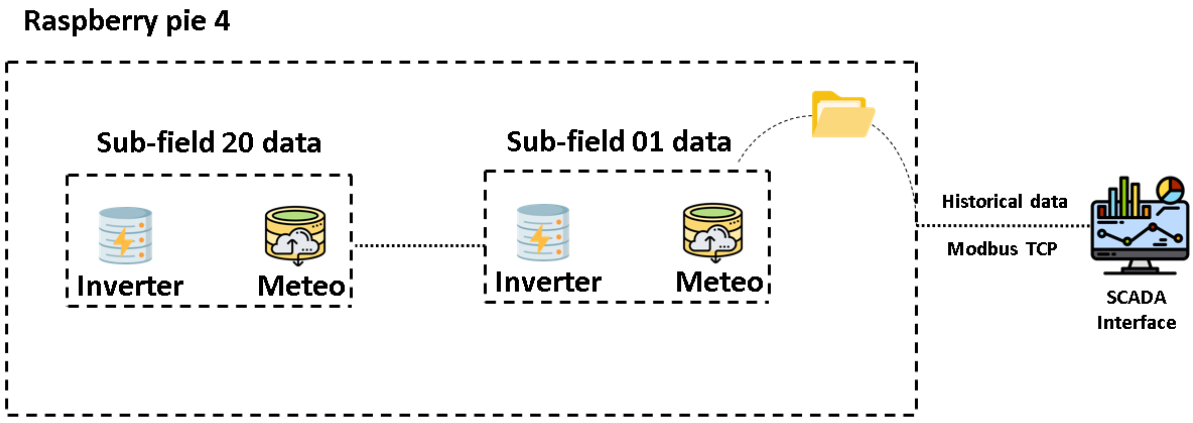


Figure 3.19: Raspberry pi4 data sending emulation

The Raspberry Pi thus serves not only as a **data source emulator**, but also as a **critical integration tool**, ensuring that the SCADA system can operate reliably once deployed in a real-world environment. Its use accelerates development, supports modular testing, and allows reproducible experiments, making it an indispensable component of the proposed smart PV monitoring solution.

3.5 Conclusion

In this chapter, a complete framework was presented for the development and implementation of a SCADA interface tailored for large-scale photovoltaic (PV) power plant monitoring. The development process was divided into two core domains: the communication architecture and the SCADA server architecture.

Firstly, we examined the hierarchical communication network, which begins with RS-485 / Modbus RTU field devices transmitting data to a serial device manager. This intermediary performs real-time protocol conversion to Modbus TCP, enabling integration with Ethernet-based infrastructure. We outlined the encapsulation of Modbus RTU payloads into TCP/IP packets and described how these are transmitted through the OSI layers across physical media such as fiber-optic and RJ45 Ethernet. A robust request-response communication mechanism was illustrated, detailing how data is reliably polled from field devices and delivered to the SCADA server for further processing.

Secondly, the SCADA server architecture was detailed, encompassing the Back-End data processing modules, Flask application server, MongoDB storage system, and Front-End visualization layer. Specific focus was placed on modular Python-based scripts for real-time data acquisition from inverters and meteorological sensors. The configuration management strategy was explained, along with the normalization of raw register values into engineering units. Additional features such as Excel export functionality, performance metrics, and prediction module calculation were implemented to support operational analytics.

Thirdly, the role of MongoDB as the data storage layer was discussed, highlighting its advantages in handling large-scale time-series data, scalability through sharing, and automatic data lifecycle control via TTL indexing. The overall system design enables scalable, modular, and real-time monitoring of PV systems while supporting future integration of intelligent analytics and diagnostic modules.

Finally, to enable continuous testing and validation of the SCADA system in the absence of a live PV environment, a full emulation mechanism was implemented using the Raspberry Pi 4. Acting as a virtual data acquisition unit, the Raspberry Pi was configured to replay a two-

year historical dataset over Ethernet using Modbus TCP or HTTP REST protocols. This data was preprocessed, converted to raw integer format, encapsulated into protocol-specific address registers, and transmitted at predefined intervals. The Raspberry Pi thus served as a critical tool for simulating real-time field behavior, ensuring reliable backend integration, and verifying end-to-end system functionality. This approach not only accelerated development but also enhanced the robustness and adaptability of the SCADA architecture.

Chapter 4

Novel Deep Learning Prediction Model, Fault Detection, and Diagnosis in Grid-Connected Photovoltaic Systems

4.1 Introduction

This chapter presents a comprehensive and innovative methodology for performance prediction, fault detection, and diagnosis in large-scale photovoltaic (PV) systems. We introduce a hybrid deep learning model specifically designed to capture the complex spatio-temporal patterns in PV array data, followed by an integrated approach for automated fault detection and diagnosis. The final section describes the implementation of these advanced techniques within a standalone offline application, providing an accessible and effective tool for PV system evaluation and maintenance.

4.2 Methodology

The methodology is structured into three main stages, each building upon the previous one to create a robust and fully integrated solution.

First, we design and implement a novel hybrid deep learning model that combines a one-layer Convolutional Neural Network (CNN) with Long Short-Term Memory (LSTM) units. This architecture is trained on historical PV generation data from Ain El Melh, Algeria, after undergoing rigorous preprocessing steps, including data cleaning, normalization, feature selection, and partitioning into training and testing sets. The hybrid model is chosen for its superior ability to learn both spatial and temporal dependencies inherent in PV output data[34].

Second, we introduce an automated offline fault detection technique. This method continuously monitors each inverter's Corrected Performance Ratio (CPR) and employs adaptive statistical thresholds to identify anomalies. When a deviation is detected, a complementary diagnosis algorithm utilizes pattern recognition and statistical analysis on historical records to classify and pinpoint the root cause of the fault.

Finally, the optimized prediction model and the fault detection & diagnosis routines are encapsulated within a standalone application. This software processes archived PV system data to generate automated fault alerts, performance predictions, and detailed diagnostic reports, all within an intuitive offline environment.

4.3 Proposed Hybrid Deep Learning Model (CNN-LSTM)

4.3.1 The Dataset

The dataset utilized in this research comes from a large grid-connected, ground-mounted photovoltaic (PV) system situated in Ain El-Melh, Algeria. The values (min,max,avg) are introduced

to provide a quick overview of the data range and typical behaviour of each input feature, as described in the table 4.1 :

Table 4.1: PV plant monitoring data

Feature	Description	Minimum	Maximum	Average
Tp	Panel temperature (°C)	−2.5000	74.8000	27.8800
Gtilt	Tilt irradiation (W/m²)	0	1565.7	308.2515
Gtot	Total irradiation (W/m²)	0	1395.6	237.9987
Gdisp	Dispersion irradiation (W/m²)	0	648	73.6811
Gdirect	Direct irradiation (W/m²)	0	1364.3	234.4052
WS	Wind speed (m/s)	0	22.2	3.7939
H	Humidity (%)	0	71.6000	35.2273
P	Pressure (Pa)	0	927	911.3952
VDC	Voltage (V)	0	780.4000	323.4112
IDC	Current (A)	0	944.3000	182.6364
PDC	Power (kW)	0	569.4410	107.6481

4.3.2 Data Processing

To ensure the reliability of our model, data preprocessing is essential[35]. This involves:

- **Cleaning and Imputation:** Missing data points were addressed through interpolation for minor gaps, while severely erroneous or outlier values were removed to maintain dataset integrity.
- **Normalization:** All features were normalized to a common scale to mitigate biases due to differing units and magnitudes, facilitating efficient model convergence.
- **Feature Selection:** Correlation analysis and importance ranking techniques identified the most influential variables impacting PV output, including module temperature and solar irradiance.
- **Data Partitioning:** The refined dataset was split into training and testing subsets using an 80/20 ratio to enable unbiased evaluation of model performance.

4.3.3 CNN-LSTM Hybrid Model

The proposed CNN-LSTM structure, as it is shown in Figure 4.1. consists of one CNN layer,one LSTM layer, and a fully connected layer. The input of the neural network has nine variables, eight of which are features such as temperature, irradiance, global radiation, and others (see the

section Dataset for more information about the variables), while the ninth variable is the target (in this schematic the target is set as PV power generated Pdc).

First, the upper layer of CNN-LSTM consists of CNN. The CNN layer can receive various variables that affect PV output. On the other side, the dataset is separated into two parts: 80% for training the model and 20% to validate the results. In the validation part, we used data from Ain Melh Bosaada in 2023, as shown in Table 4.1. The CNN consists of an input layer that accepts sensor variables as inputs, an output layer that extracts features to LSTMs, and several hidden layers. The hidden layers used consist of: a convolution layer, a ReLU layer, an activation function, or a pooling layer. From the presented data in the input, a unique output pattern is generated. The CNN extracts the local features and the LSTM temporal part. With this structure, the neural network "learns" for every input a weight that determines a specific output.

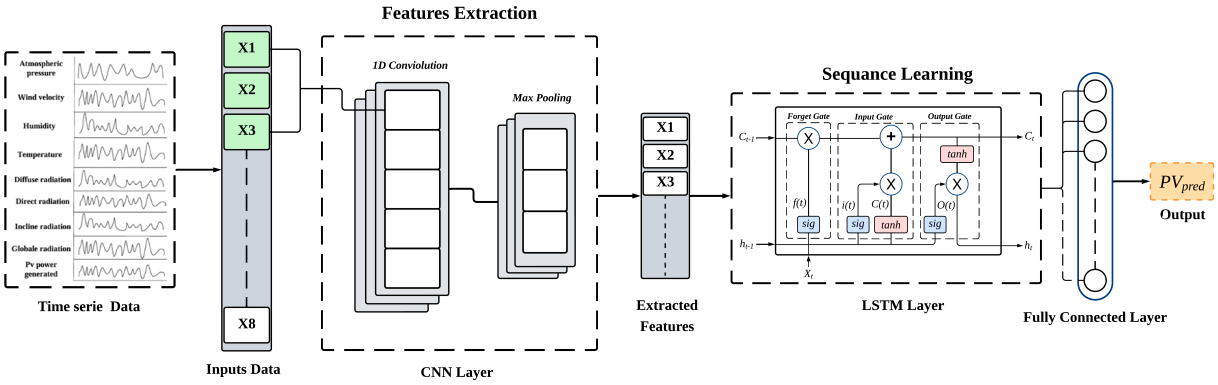


Figure 4.1: The proposed CNN-LSTM Architecture

The hybrid CNN-LSTM architecture employs a single convolutional layer with ReLU activation, followed by a pooling layer. The LSTM component utilizes one hidden layer, while a fully-connected output layer (2048 neurons) incorporates a 0.1 dropout rate to prevent overfitting. Detailed parameter specifications are provided in Table 4.2.

Hyperparameters were optimized via trial-and-error: Initial values were set, models were trained on 80% of the data and validated on 20% across seasonal subsets spanning two years. Parameters were iteratively adjusted until optimal performance was achieved.

To validate model efficacy, PV output predictions were evaluated under distinct weather scenarios (cloudy/clear skies) and benchmarked against: LSTM, Random Forest, and 5D CNN-LSTM

A comprehensive analysis of comparative results is presented in subsequent sections.

Table 4.2: Layer configuration of the proposed model

Layer	**Configuration**	**Activation Function**
Convolution 1	Filters = 64; kernel size = 3	ReLU
Max-pooling 1	Kernel size = 2; stride = 2	--
LSTM 1	Units = 64	Tanh, Sigmoid
Fully Connected	Neurons = 2048	--
Dropout	Dropout rate = 0.1	--
Training Parameters	Loss function: MSE, Optimizer: Adam, Batch size: 32	

4.3.4 Performance Evaluation Metrics

To evaluate the model's performance several standard metrics are used, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2), which collectively provide a robust measure of prediction accuracy and reliability[36].

$$\text{MAE} = \frac{1}{n}\sum_{i=0}^{n-1}|y_i - \hat{y}_i| \tag{4.1}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=0}^{n-1}(y_i - \hat{y}_i)^2} \tag{4.2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n-1}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n-1}(\bar{y} - y_i)^2} \quad \text{with} \quad \bar{y} = \frac{1}{n}\sum_{i=0}^{n-1} y_i \tag{4.3}$$

4.3.5 Prediction Results

- **DC Power prediction**

Figure 4.2 compares the predicted versus actual daily PV output for two distinct test scenarios, demonstrating the superior accuracy of the proposed CNN–LSTM model. Across diverse weather conditions, the coefficient of determination (R^2) consistently exceeds 0.9992 (see Table 4.3). Under cloudy conditions, the Mean Absolute Error (MAE) measures 2.4475 kW, which improves to 1.3795 kW on clear days. Similarly, the Root Mean Square Error (RMSE) decreases from 3.8901 kW in overcast scenarios to 2.1774 kW under clear skies. The enhanced performance under clear-sky conditions is attributed to the diminished variability in irradiance and module temperature, yielding more stable power generation throughout the day.

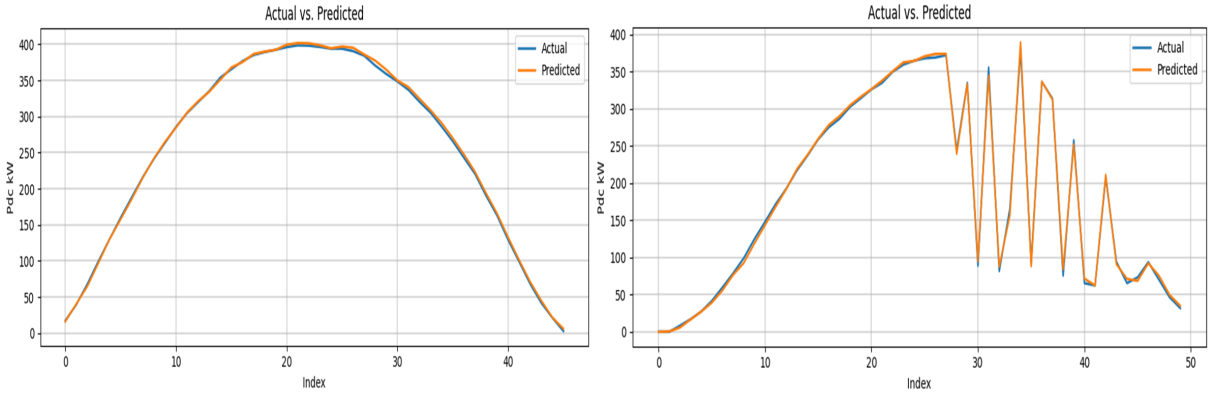


Figure 4.2: Predicted and measured power (a) clear sky day (b) cloudy day

Table 4.3: Performance metrics under different weather conditions

Condition	R^2	RMSE (kW)	MAE (kW)
Clear skies	0.9998	2.1774	1.3795
Cloudy	0.9992	3.8901	2.4475

- **DC Current prediction**

Two distinct test scenarios were also employed to evaluate the model's ability to predict the PV system's DC current. As shown in Figure 4.3, the CNN–LSTM predictions closely track the measured values, with coefficients of determination (R^2) exceeding 0.9991 in both cases (Table 4.4). The Mean Absolute Error (MAE) spans from 2.2967 A to 5.1988 A, while the Root Mean Square Error (RMSE) ranges between 3.5286 A and 8.2184 A. These results confirm the model's robust accuracy in capturing DC current dynamics under varying daily conditions.

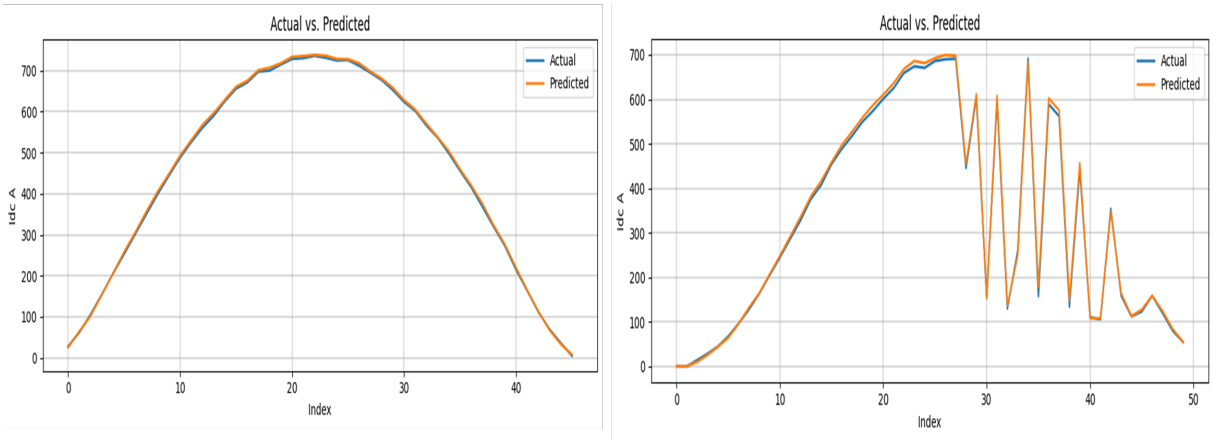


Figure 4.3: Predicted and measured current (a) clear sky day (b) cloudy day

Table 4.4: Performance metrics of current prediction under different weather conditions

Condition	R^2	RMSE (A)	MAE (A)
Clear skies	0.9998	3.5286	2.2967
Cloudy	0.9991	8.2184	5.1988

- **DC Voltage prediction**

Figure 4.4 presents the predicted DC voltage generated by the CNN-LSTM model alongside the measured values under both clear and cloudy sky conditions. As shown, the predicted voltage closely matches the actual measurements. This alignment is supported by the regression results in Table 4.5, where the coefficient of determination (R^2) reaches approximately 0.9660, indicating an excellent fit. These results confirm the model's robustness and adaptability, particularly under stable weather conditions.

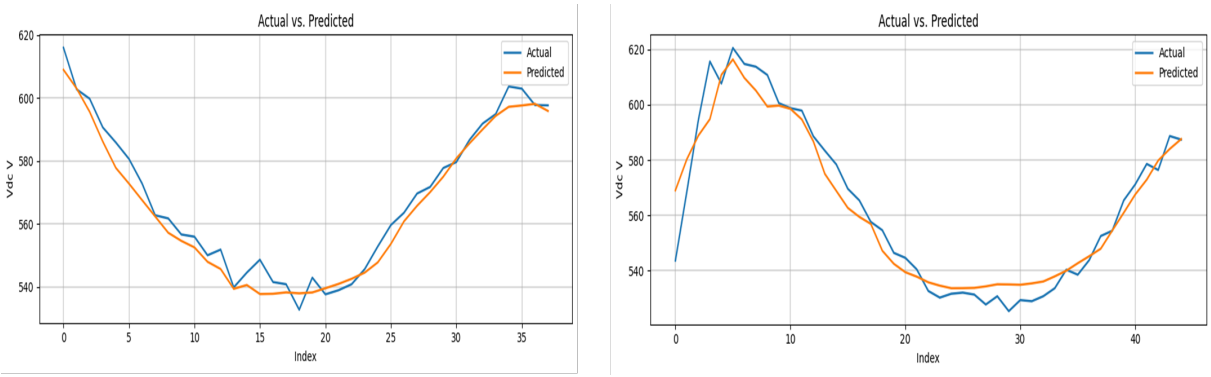


Figure 4.4: Predicted and measured voltage (a) clear sky day (b) cloudy day

Table 4.5: Performance metrics of voltage prediction under different weather conditions (Alternative Model)

Condition	R^2	RMSE (A)	MAE (A)
Clear skies	0.9660	4.3260	3.5461
Cloudy	0.9414	7.1951	5.3641

- **Performance Comparison with Other Models**

The comparative performance of the proposed one-layer CNN-LSTM hybrid model was evaluated against standard baseline architectures (a standalone LSTM, a five-layer CNN-LSTM, and a Random Forest), all trained on the same dataset. Table 4.6 (reproduced from our master's research work) summarizes key metrics: coefficient of determination (R^2), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), along with the total number of trainable parameters and model size.

Table 4.6: Comparative Evaluation of Proposed CNN-LSTM and Benchmark Models

Model	R^2	RMSE	MAE	N Parameters
CNN-LSTM	0.9998	3.5286	2.2967	169,793 (663.25 KB)
LSTM	0.9957	11.9234	7.9842	51,821 (202.43 KB)
5CNN-LSTM	0.9975	8.9992	5.4994	1,089,713 (4.16 MB)
Random Forest	0.9680	21.0200	7.4000	**Non-parametric**

From the table, the CNN-LSTM model achieves the best results compared to the other methods. This is because CNN is used to extract important features, while LSTM handles the prediction. This explains why it outperforms the standalone LSTM model.

However, when the number of CNN layers increased (as in the 5CNN-LSTM), the model returned a much larger number of parameters. This could reduce accuracy when working with real-time data, and it requires more hardware resources to run. It may also suffer from instability or forgetting over time.

4.4 Proposed Fault Detection and Diagnosis Methodology

4.4.1 Fault Detection Technique

The proposed fault detection approach efficiently identifies anomalies in large-scale photovoltaic (PV) systems by analysing each inverter's Corrected Performance Ratio (CPR) using statistical outlier detection combined with temporal fault tracking[37]. This methodology integrates three core components:

- **Corrected Performance Ratio (CPR)**

$$\mathrm{CPR} = \frac{\mathrm{PR}}{1 - \delta \times (T_p - T_{p\_STD})} \tag{4.4}$$

$$\mathrm{PR} = \frac{Y_r}{Y_f} \tag{4.5}$$

Where:

- T_p: Cell temperature computed from measured meteorological data (°C)
- T_{STD}: Cell temperature under standard test conditions (typically 25°C)
- δ: Temperature coefficient for power (%/°C, negative in sign) that corresponds to the installed modules
- Y_r: Reference Yield
- Y_f : Final Yield

This correction ensures that CPR reflects true performance degradation, independent of natural fluctuations in temperature.

- **Abnormal Behaviour Detection – Hampel Identifier**

The Hampel identifier serves as a robust statistical framework for detecting photovoltaic (PV) inverter faults by identifying outliers in Corrected Performance Ratio (CPR) values. An inverter 'Inv' at time t is flagged as faulty if its CPR (CPRt, Inv) falls below an adaptive bound. This dynamic thresholding approach reliably distinguishes between inherent operational variability and significant performance deviations indicative of faults. Figure 4.5 illustrates this concept.

$$\mathrm{CPR}_t^{\mathrm{inv}} \leq \text{bound} \tag{4.6}$$

$$\text{bound} = \mathrm{mean}(\mathrm{CPR}_t) - \alpha \times S \tag{4.7}$$

$$S = \frac{1}{0.6745} \times \mathrm{median}\left|\mathrm{mean}(\mathrm{CPR}_t) - \mathrm{CPR}_t^{\mathrm{inv}}\right| \tag{4.8}$$

- α: Threshold constant
- S: Median Absolute Deviation (MAD) scale estimator

To prevent instability caused by the vanishing of the scale parameter S during low inverter activity periods, such as sunrise, sunset, or limited inverter operation, a minimum threshold of 0.07 was imposed. Thus, S is redefined as max(S,0.07), ensuring numerical stability and consistent model performance under low-irradiance conditions.

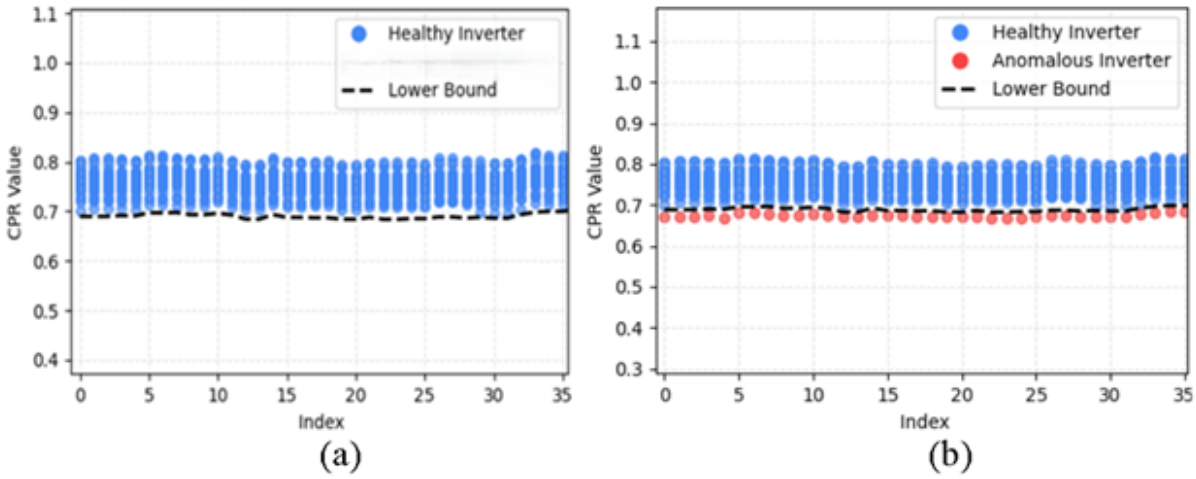


Figure 4.5: Detection of PV Inverter Faults Using the Hampel Identifier

In Figure 4.5(a) displays normal inverter operation, where CPR values fluctuate above the bound.While Figure 4.5(b) shows a fault condition, where the inverter's CPR falls below the bound, signaling a performance anomaly indicative of a fault.

- **Fault Detection Frequency**

While individual inverter CPR outliers may signal faults, transient anomalies caused by environmental fluctuations, shading, or measurement noise can also produce similar deviations. To distinguish isolated events from persistent faults, the Anomaly Frequency Index (AFI) is introduced. The AFI quantifies the occurrence and consistency of CPR outliers over a specified time window, enabling more reliable fault detection.

$$AFI_i = \frac{1}{M} \sum_{t=1}^{M} H\left(\text{bound}_t - CPR_t^i\right)$$

where The function H is Heaviside step function :

$$H(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases}$$

The figure 4.6 illustrates the overall fault detection framework, detailing the sequential processing steps from data acquisition through CPR of the Inverter, adaptive thresholding via the Hampel filter, and fault classification based on AFI evaluation.

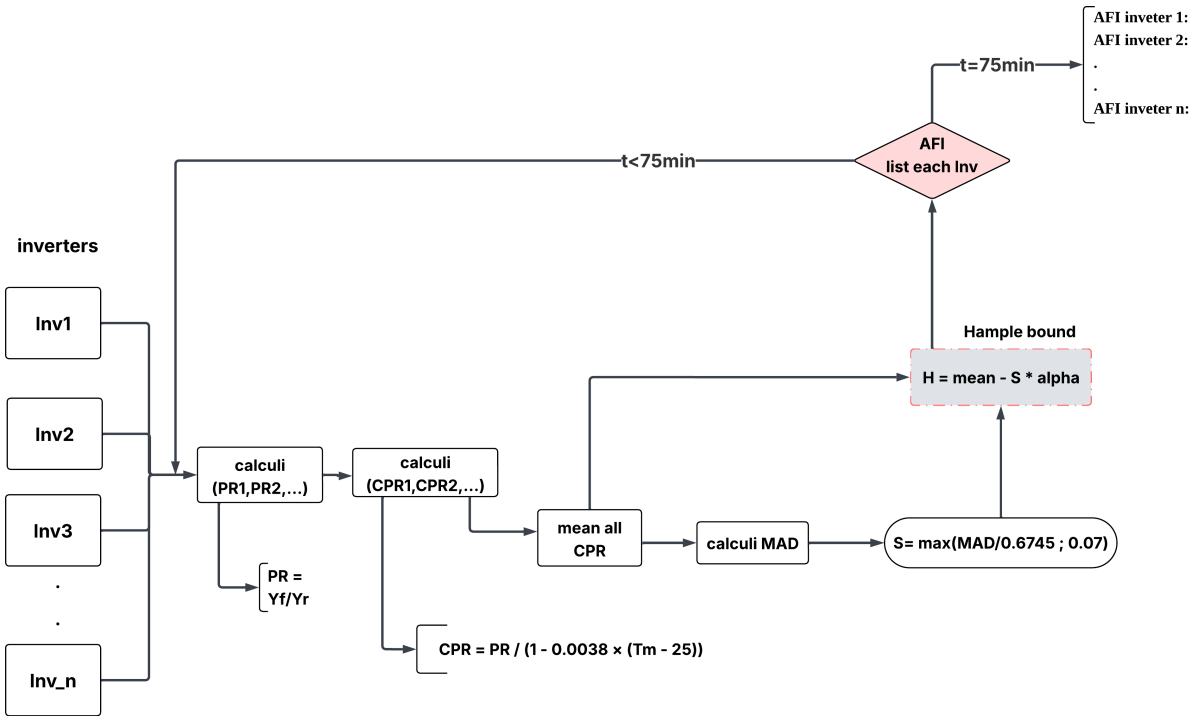


Figure 4.6: Fault Detection Technique Based on CPR Metric and Hampel Identifier

To illustrate the process, the Algorithm (Algorithm 1) presents a concise step-by-step outline of the fault detection method.

Algorithm 1: Fault Detection Technique

Input :

- Current matrix $I_{M\times N}$ (M: period of detection, N: inverters)
- Voltage matrix $V_{M\times N}$
- Inclined irradiance vector G_{din}
- Cell temperature vector T_p

Output: A flag for each inverter indicating whether it is faulty

for *each* $T \leq M$ **do**
  for *Inv* $\leq N$ **do**
    Compute mean CPR across inverters: $CPR_{t,\text{Inv}}$;
Compute mean CPR across inverters:

$$\text{mean}(CPR_t) = \frac{1}{N}\sum_{\text{Inv}=1}^{N} CPR_{t,\text{Inv}}$$

;
for *Inv* $\leq N$ **do**
  Calculate absolute deviation:

$$MAD_{\text{Inv}} = |\text{mean}(CPR_t) - CPR_{t,\text{Inv}}|$$

Compute scale parameter:

$$S = \frac{1}{0.6745} \times \text{median}(MAD_{\text{Inv}})$$

Enforce $S \leftarrow \max(S, 0.07)$;
Set detection boundary: $\text{bound}_t = \text{mean}(CPR_t) - \alpha S$;
for *Inv* $\leq N$ **do**

$$AFI_{\text{Inv}} = \frac{1}{M}\sum_{t=1}^{M} H(\text{bound}_t - CPR_{t,\text{Inv}})$$

Where $H(x)$ is the Heaviside step function
for *Inv* $\leq N$ **do**
  for *in* AFI_{Inv} **do**
    if $AFI_{Inv} \geq 0.6$ **then**
      Mark the inverter as faulty;

4.4.2 Diagnosis Approach

After fault detection flags anomalous behaviour via the Anomaly Fault Indicator (AFI), signaling a potential issue, the next task is pinpointing its origin. Our innovative approach hinges on crafting adaptive boundaries for DC current (Idc) and voltage (Vdc). These boundaries stem from predictions by a trained CNN-LSTM proposed model, which predicts the expected Idc and Vdc under normal, fault-free conditions for that day. Using these predictions and defined equations, we establish Dynamic boundaries that sharply distinguish normal operations from various fault states, enabling accurate fault type classification.

- **Dynamic Boundaries for Fault Diagnosis**

To construct the dynamic boundaries for fault diagnosis, a hybrid deep learning model (CNN-LSTM) is used to predict the expected direct current $I_{\mathrm{DC_{pred}}}$ and voltage $V_{\mathrm{DC_{pred}}}$ values under fault-free operating conditions. These predictions represent the baseline behaviour of the photovoltaic system in normal operation and serve as reference points for fault classification[38].

To account for natural variability and prediction uncertainty, the standard deviations of the model's prediction errors, denoted as σ_I for current and σ_V for voltage, were computed from the training dataset. These statistical measures quantify the deviation between the model's outputs and the actual measured values during periods confirmed to be free of faults (healthy system).

An adjustable parameter p is introduced to define the confidence level of the boundary thresholds, with $p = 1$ corresponding to approximately 68.2% confidence. This parameter allows the dynamic boundaries to be scaled appropriately to balance fault sensitivity and robustness against false alarms.

Figure 4.7 illustrates the concept of standard deviation spread in a normal distribution, demonstrating how the choice of $p = 1$ defines the boundary limits that capture the majority of normal behaviour while allowing detection of anomalies beyond these limits.

By integrating the model's predictions with these statistically derived thresholds, the proposed method facilitates accurate, adaptive monitoring of system behaviour, thereby enhancing fault detection and diagnosis performance.

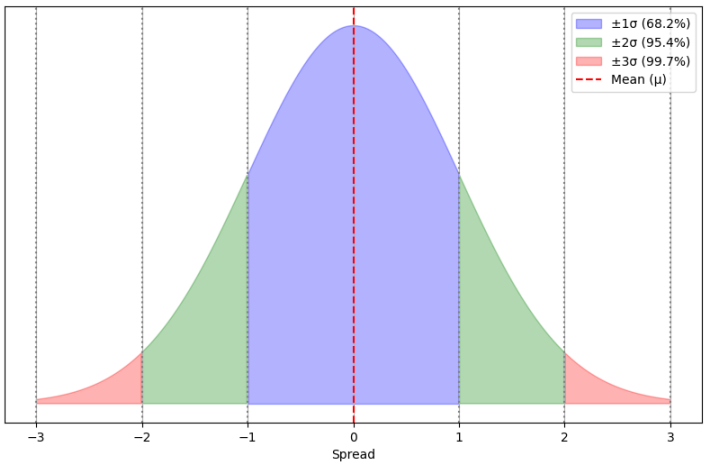


Figure 4.7: Standard Deviation Spread in a Normal Distribution

- **Dynamic Lower Bounds**

The dynamic lower boundaries for DC current and voltage are defined as:

$$I_{\text{DC,lower}} = I_{\text{DC,pred}} - p\sigma_I \tag{4.9}$$

$$V_{\text{DC,lower}} = V_{\text{DC,pred}} - p\sigma_V \tag{4.10}$$

If the real-time measured current or voltage falls below these thresholds, it suggests abnormal behaviour. Mathematically, a fault condition is identified if:

$$I_{\text{DC,measured}} < I_{\text{DC,lower}} \tag{4.11}$$

$$V_{\text{DC,measured}} < V_{\text{DC,lower}} \tag{4.12}$$

- **Dynamic upper Bounds**

To detect overvoltage conditions, an upper threshold is defined exclusively for the DC voltage. Within the system's design and operational framework, overcurrent conditions are not considered critical; therefore, no upper bound is enforced for the DC current.

$$V_{\text{DC,upper}} = V_{\text{DC,pred}} + p\sigma_V \tag{4.13}$$

A real-time measurement is considered to exceed the upper dynamic boundary when the following conditions are satisfied:

$$V_{\text{DC,measured}} > V_{\text{DC,upper}} \tag{4.14}$$

To ensure the system is operating normally, all real-time measurements must remain within their corresponding bounds:

$$I_{\text{DC,lower}} < I_{\text{DC,meas}} \quad \text{and} \quad V_{\text{DC,lower}} < V_{\text{DC,meas}} < V_{\text{DC,upper}} \tag{4.15}$$

Measurements outside these boundaries trigger fault diagnosis processes and help categorize fault types with improved precision.

4.4.3 Fault Diagnosis Scenarios

This approach enables precise differentiation between normal and faulty conditions. The overall architecture of the diagnosis process, including fault detection and classification via adaptive boundaries, is illustrated in Figure 4.8, which provides a clear visual representation of the step-by-step decision-making flow from anomaly detection to fault classification.

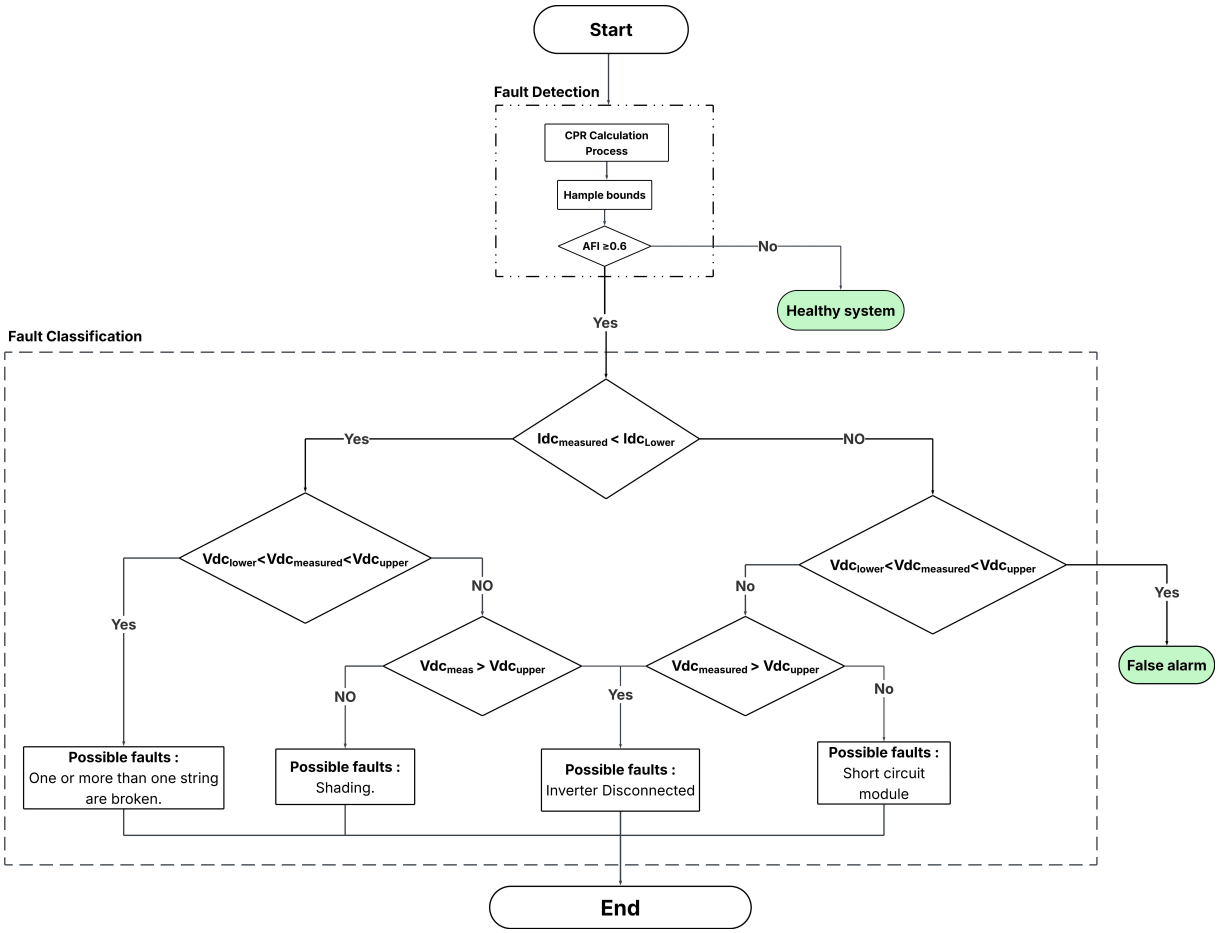


Figure 4.8: Fault detection and diagnosis flowchart of the proposed method

In this flowchart, once a fault within the photovoltaic (PV) system has been confirmed, the approach systematically isolates three potential fault sources by evaluating two adaptive-boundary indicators derived from predicted current and voltage values. The following scenarios are employed to identify specific fault types:

- **Scenario 1:** As depicted in Figure 4.9, both the measured DC current and voltage exceed their respective thresholds. This condition is indicative of partial shading, which diminishes the irradiance received by the PV modules, thereby causing simultaneous reductions in both current and voltage.

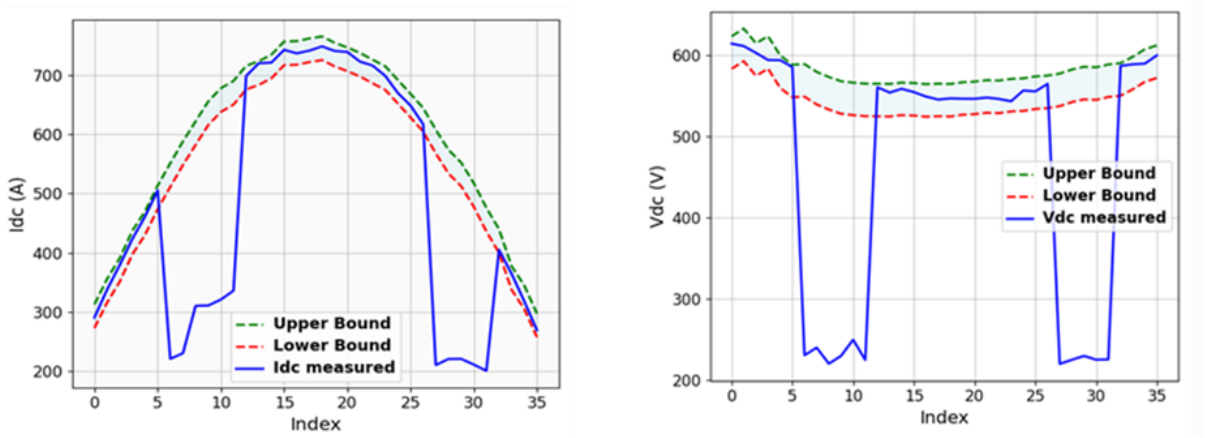


Figure 4.9: Partial Shading or Soiling Effects on PV Module

- **Scenario 2:** Illustrated in Figure 4.10, the measured DC current remains within its threshold band, while the voltage exceeds its threshold. This scenario suggests the presence of a short-circuit fault within the PV modules, where the current is unaffected but the voltage drops significantly.

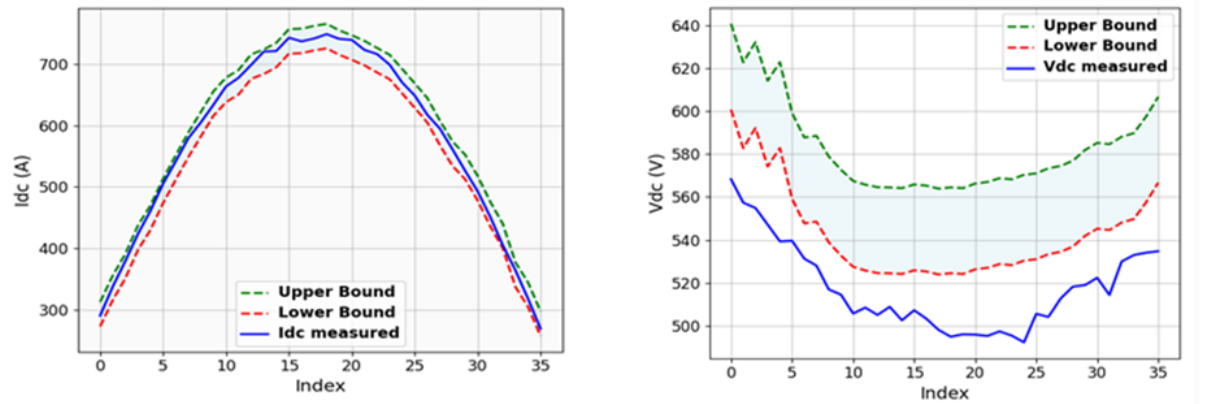


Figure 4.10: Short-Circuit Fault in PV Modules

- **Scenario 3:** As shown in Figure 4.11, the measured DC current exceeds its threshold, whereas the voltage remains within its designated bounds. This situation points to an issue within the PV strings, such as a disconnection or degradation, which primarily affects the current output.

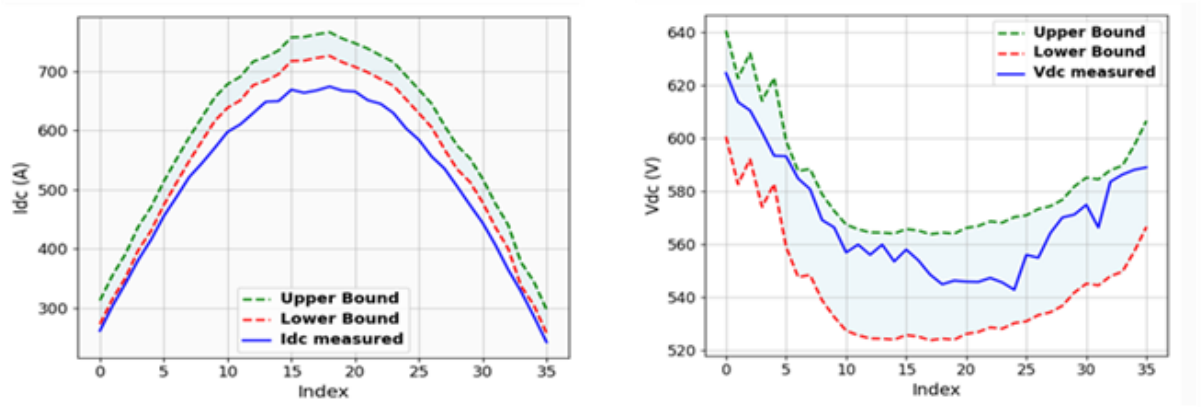


Figure 4.11: String-Level Fault in PV System

4.5 Standalone Python GUI for Prediction and Fault Diagnosis

This section outlines the design and implementation of a standalone graphical user interface (GUI) application developed for power prediction, fault detection, and diagnosis in photovoltaic systems (see Figure 4.12). The application features an intuitive, user-friendly interface that integrates seamlessly with the proposed CNN–LSTM model, enabling users to:

- Input relevant system and environmental data.
- Configure model prediction parameters.
- Visualize predicted values and fault detection.
- Perform key parameter calculations.
- Automatically detect and diagnose system faults.

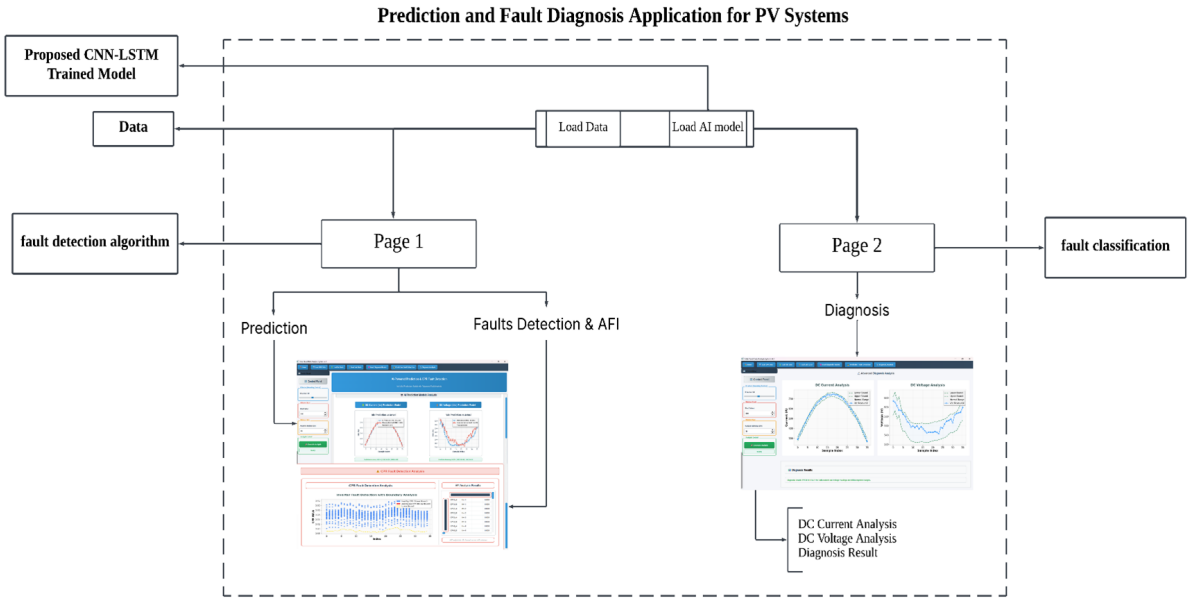


Figure 4.12: Standalone Offline Application for PV System prediction, Fault Detection and Diagnosis

The system architecture is structured around two core functional modules: prediction and anomaly detection (Page 1), and detailed fault diagnosis (Page 2).

4.5.1 Page 1 - Prediction and Fault Detection Module

This page leverages a pre-trained CNN-LSTM deep learning model to predict outputs of the PV system, such as DC current and voltage, based on environmental and operational inputs (e.g., irradiance, temperature). The input data and trained model are loaded through dedicated interface controls. Once data processing, the system applies a fault detection algorithm, and Deviations are quantified using the Anomaly Fault Index (AFI), computed via an adaptive thresholding method. This process enables identification of abnormal operating behaviour, which is visually represented through interactive plots and status indicators on the GUI.

4.5.2 Page 2 - Fault Diagnosis Module

After detecting an anomaly, this page presents detailed visualizations of DC current and voltage profiles over time, enabling users to examine deviation patterns. The system applies predefined diagnostic criteria to classify fault types, distinguishing among various failure modes. Results display in an intuitive format, supporting rapid interpretation and decision-making to ensure both early detection and precise classification, thereby enhancing PV system reliability and maintainability.

4.6 Conclusion

In this chapter, we have developed and validated a comprehensive methodology for PV system performance prediction, fault detection, and diagnosis. By employing a novel CNN-LSTM hybrid model, we achieved highly accurate predictions of PV output under diverse environmental and operating conditions. Our fault detection and diagnosis techniques, based on robust statistical analysis and adaptive performance boundaries, ensure reliable identification and classification of system anomalies.

The implementation of these models in a standalone offline application demonstrates their practical value for PV system monitoring and maintenance. Looking ahead, we plan to integrate this intelligent diagnostic tool into the SCADA interface developed in Chapter 3. This integration will enable real-time fault prediction and automated performance analysis within the operational monitoring platform, enhancing the overall responsiveness, efficiency, and intelligence of PV plant supervision.

Chapter 5

Field Experimental Validation

5.1 Introduction

This final chapter presents the results obtained from the operational deployment of the developed SCADA interface within a grid-connected photovoltaic (PV) power plant. The data and visualizations discussed are drawn from a representative experimental day, selected to demonstrate the system's monitoring, analytical, and logging capabilities under real-world conditions.

Rather than focusing solely on interface design, the discussion emphasizes how each functional layer contributes to effective supervision and decision-making. The chapter highlights the integration of real-time visualization with advanced diagnostics, positioning the interface as a central tool for managing and optimizing PV plant performance. The following sections illustrate the system's hierarchical and modular structure, demonstrating how it translates complex operational behaviour into intuitive, actionable insights.

5.2 SCADA Interface Presentation

The developed SCADA solution comprises two complementary interfaces designed to address distinct operational needs within the photovoltaic plant supervision framework.

The first interface functions in real time, providing continuous monitoring of system parameters alongside historical data logging capabilities. It offers multi-level visualization, from overall plant performance down to granular details such as individual strings and junction boxes, thereby enabling operators to maintain situational awareness and perform immediate assessments during plant operation.

The second interface operates offline, dedicated exclusively to advanced fault detection and diagnosis. Leveraging predictive models and diagnostic algorithms, this standalone application supports detailed analysis and interpretation of system anomalies, facilitating preventive maintenance and reliability assurance without impacting real-time monitoring activities.

Together, these interfaces form a cohesive supervisory environment, balancing the demands of live operational oversight with comprehensive fault analytics. Each interface and its functionalities are elaborated in the following sections, illustrated with representative figures derived from actual implementation.

5.3 Monitoring Interface

The monitoring interface provides real-time visualization of PV system performance metrics at varying granularities:system-wide level, inverter-level, and string-level. Divided into multiple pages, the interface enables a comprehensive understanding of operational behaviour. and serves as the foundation for operational oversight by offering an immediate depiction of current system status and localized component behaviour.

5.3.1 General Monitoring Page

The main monitoring page consolidates multiple sub-levels of supervision, including inverters, junction boxes, and strings, into a cohesive layout. This hierarchical design allows operators to transition from general overviews to highly specific parameter inspections seamlessly. Figures 5.1 to 5.10 display real-time data collected during the experimental day, highlighting the interfaces functionality under actual operating conditions.

Figure 5.1 presents the main monitoring interface (upper section), which displays key meteorological parameters using intuitive gauge indicators. These parameters include solar irradiance, ambient temperature, and wind speed, providing real-time environmental context. Additionally, station-wide statistics, such as daily energy production and peak power output, are summarized to offer a comprehensive overview of the plant's performance at a glance.

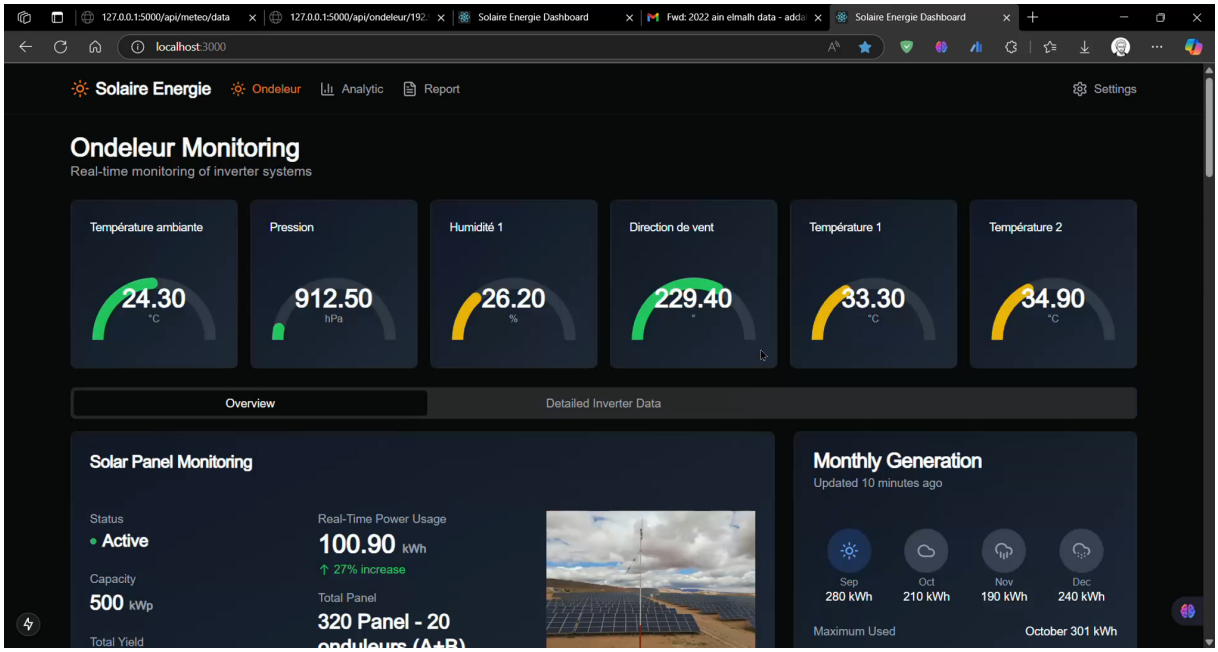


Figure 5.1: Inverter Gauges

Figures 5.2 and 5.3 illustrate the temporal evolution of key meteorological parameters through interactive line plots. These visualizations display the daily progression of variables such as solar irradiance and ambient temperature, enabling users to track environmental conditions in real time and identify trends that may impact system performance.

For instance, users can selectively toggle the display of individual parameters, such as direct irradiance—on or off using the interactive legend (as shown in Figure 5.3). This functionality helps reduce visual clutter and allows users to focus on specific data of interest.

Additionally, each graph can be expanded into full-screen mode to facilitate detailed examination of fluctuations, supporting more precise technical analysis and interpretation.

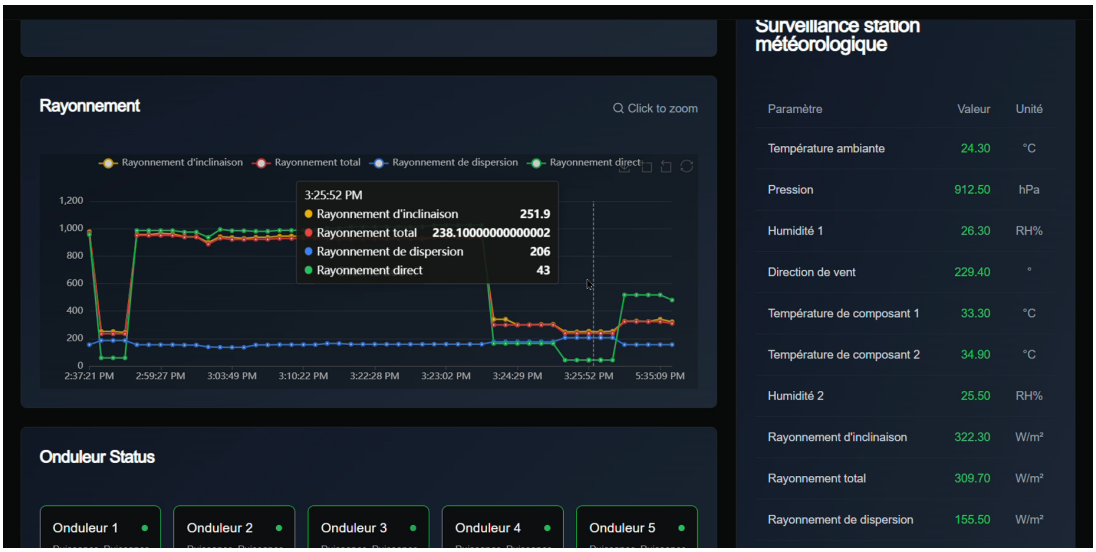


Figure 5.2: Parameters curve representation

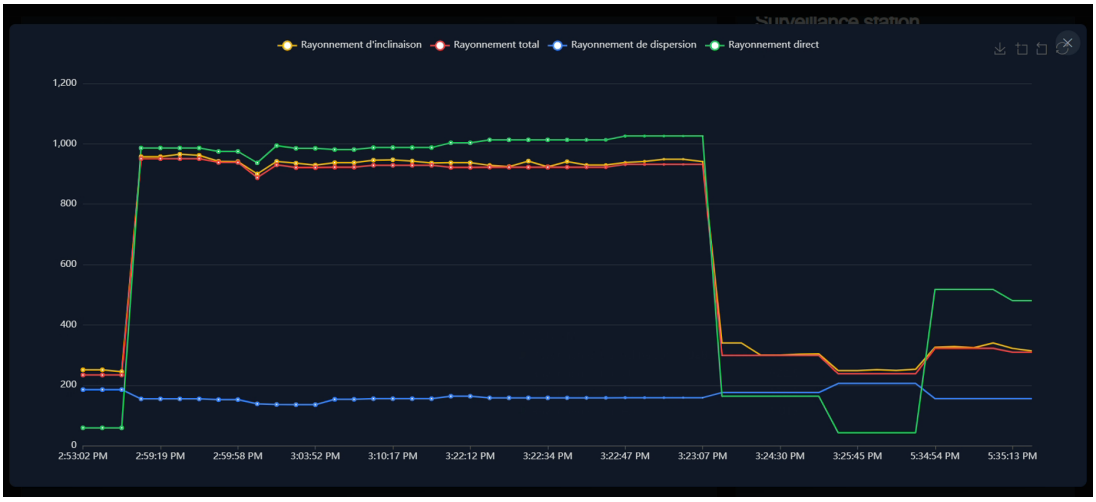


Figure 5.3: Curve full size

To monitor the photovoltaic field at a surface level, the interface provides independent, real-time supervision of each subfield by individually tracking the operational status of its corresponding inverter. Real-time visualization of critical electrical parameters—such as DC voltage, DC current, internal temperature, and AC power output—is provided for each unit. This enables targeted diagnostics and performance evaluation at the subfield level, ensuring localized supervision and rapid fault identification.

On the lower side of the interface, a list of all inverters is provided, allowing users to access detailed data for each unit by simply clicking on its identifier. General power-related metrics are also displayed in this section. On the right side, supplementary meteorological data, such as humidity and wind speed, are presented to complete the environmental context. This functionality is illustrated in Figure 5.4.

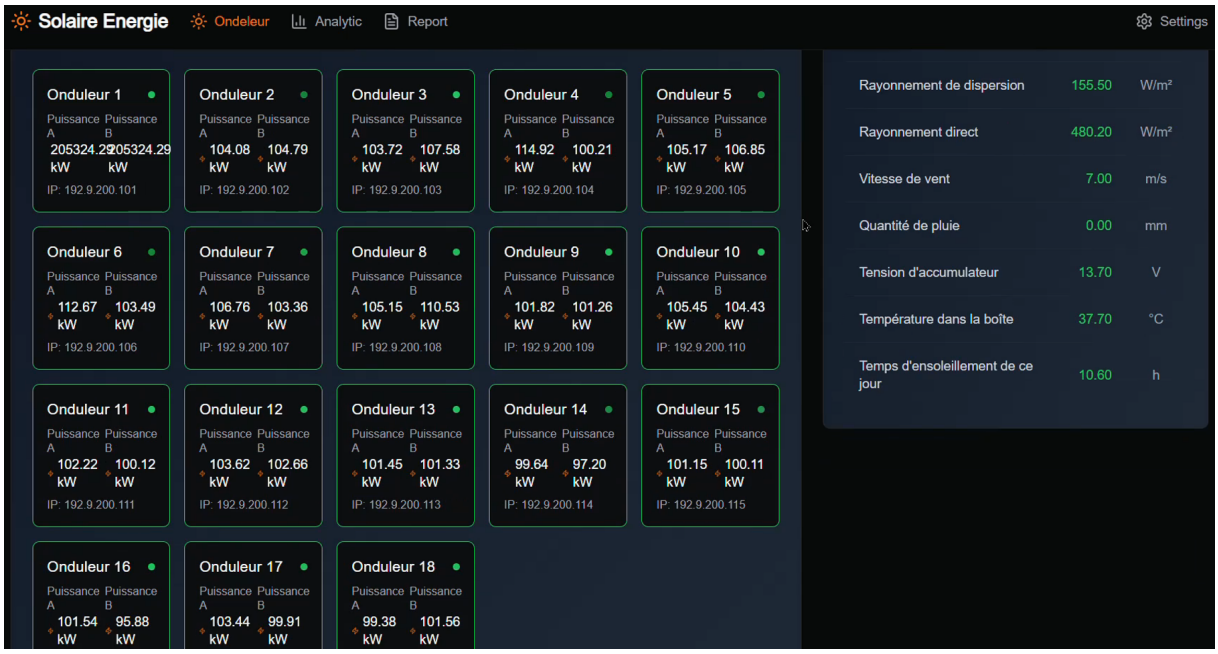


Figure 5.4: Inverters surface level monitoring

5.3.2 Inverter Monitoring Page

At a more detailed level, each inverter can be accessed individually to display its corresponding measurements, operational indicators, and key performance characteristics. The interface presents a consolidated information block that includes system-wide metrics, such as AC power output and hourly energy production, for both Side A and Side B, along with detailed DC and AC input/output values. In addition, the most important operational characteristics are provided, such as inverter frequency and temperature. This functionality is illustrated in Figure 5.5.

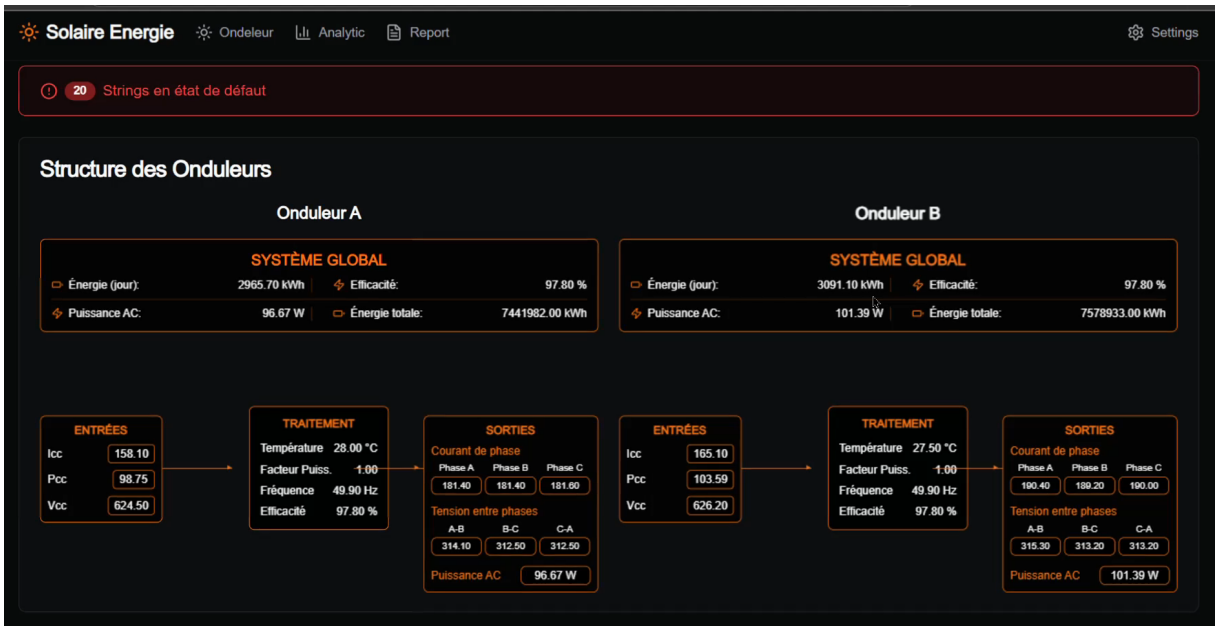


Figure 5.5: Inverter monitoring

- **Inverter Alarms System**

An alarm system is implemented across multiple levels of the monitoring interface, from inverters to junction boxes and down to individual strings. This layered approach ensures that faults can be detected and localized with high precision, enabling timely and targeted maintenance interventions.

At the inverter level, an alarm flag is raised when one or more connected strings exhibit abnormal behaviour. The interface not only signals the presence of a fault but also indicates the number of affected strings, allowing operators to quickly assess the extent of the issue without navigating away from the summary view.

This functionality is demonstrated in Figure 5.6, where inverter-level alerts highlight the number of strings currently flagged as faulty.

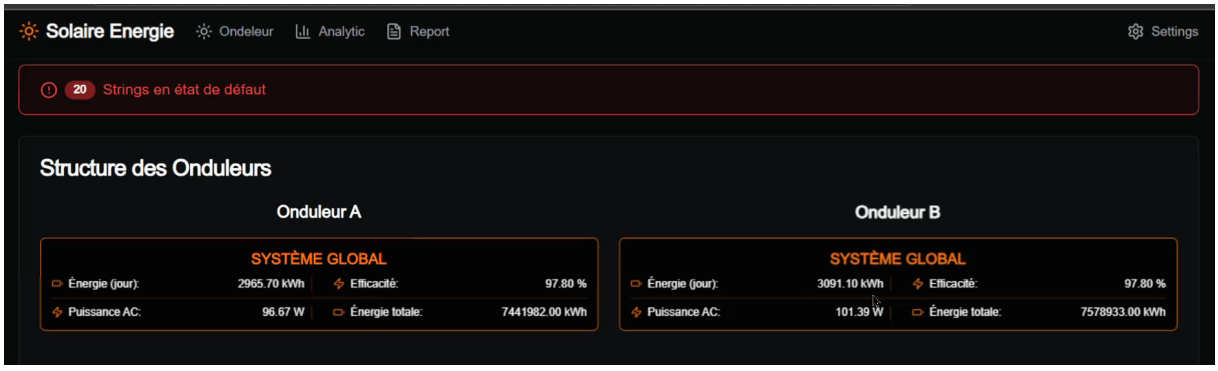


Figure 5.6: Inverter alarms

- **Performance Parameters**

This module presents a set of computed performance indicators for each inverter (Sides A and B), including the Reference Yield (Yr), Array Yield (Ya), and Yield Factor (Yf), which represent the expected solar energy input, the actual energy output from the array, and the normalized energy output, respectively. These indicators are visualized using bar charts to support intuitive comparison across inverter units and time intervals.

To assess system performance, the Performance Ratio (PR) is calculated by comparing actual energy output to expected theoretical values. This indicator provides a clear measure of how efficiently the system is operating under current conditions, making it a critical input for fault detection.

The module provides operators with a structured overview of real-time, daily, and cumulative performance, facilitating continuous monitoring and performance evaluation. This functionality is illustrated in Figure 5.7.



Figure 5.7: Performance parameters

- **Prediction Module**

Real-time environmental inputs are continuously fed into the prediction module, enabling it to estimate future values of key electrical parameters such as power, voltage, and current. These predicted values are visualized alongside actual real-time measurements in the form of progressive curves, allowing operators to observe deviations, assess forecast accuracy, and anticipate system behaviour. By anticipating short-term system behaviour and revealing deviations from expected patterns, and early indications of expected performance trends.The predictive insights support operators with informed decision-making and responding proactively.
An illustration of the prediction interface is presented in Figure 5.8.

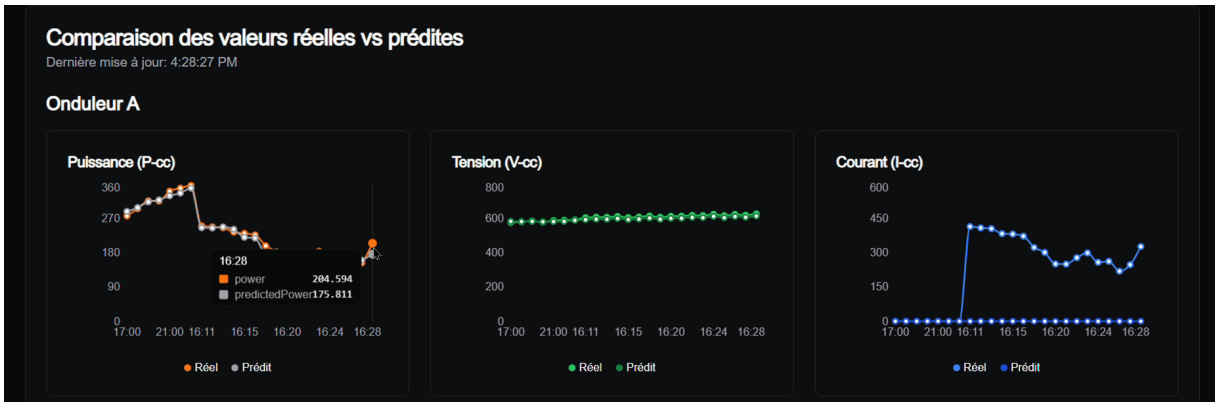


Figure 5.8: Prediction module

5.3.3 String Monitoring Page

The junction box level represents the most detailed stage of PV field monitoring, offering precise electrical insight into the operation and health of the photovoltaic array. This interface enables fine-grained supervision by presenting electrical measurements at progressively localized levels within each subfield.

Three levels are defined:

- **Level 1:**Show each inverter's junction boxes.
- **Level 2:**Collect all inverter strings of each junction box side to side.
- **Level 3:**Isolate a single junction box and display its strings.

- **Collective Junction Box**

At the top of this hierarchy (Level 1), the system evaluates each inverter's Side A and Side B by monitoring their respective junction boxes. For each junction box, the interface displays the aggregated DC current and power from all connected strings, as well as the shared voltage common to those strings. This representation provides a structured overview of subfield behaviour, facilitating direct comparison between inverter sides and supporting early detection of potential irregularities. Additionally, string-level alarms are integrated within each junction box, allowing for the prompt identification of localized faults. This functionality is illustrated in Figure 5.9.

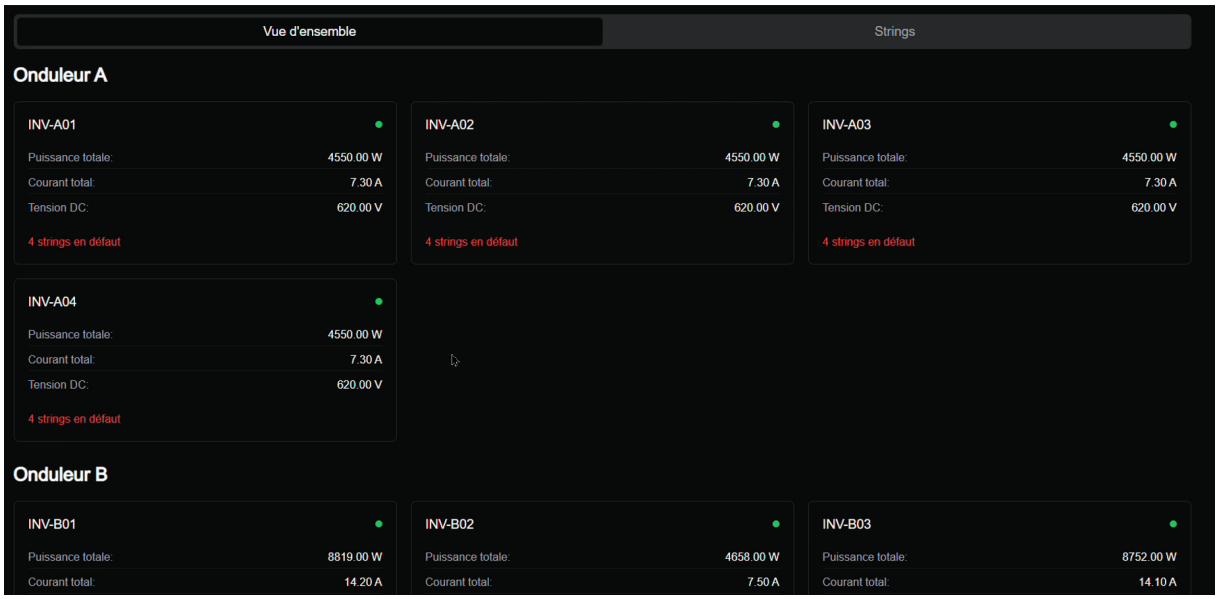
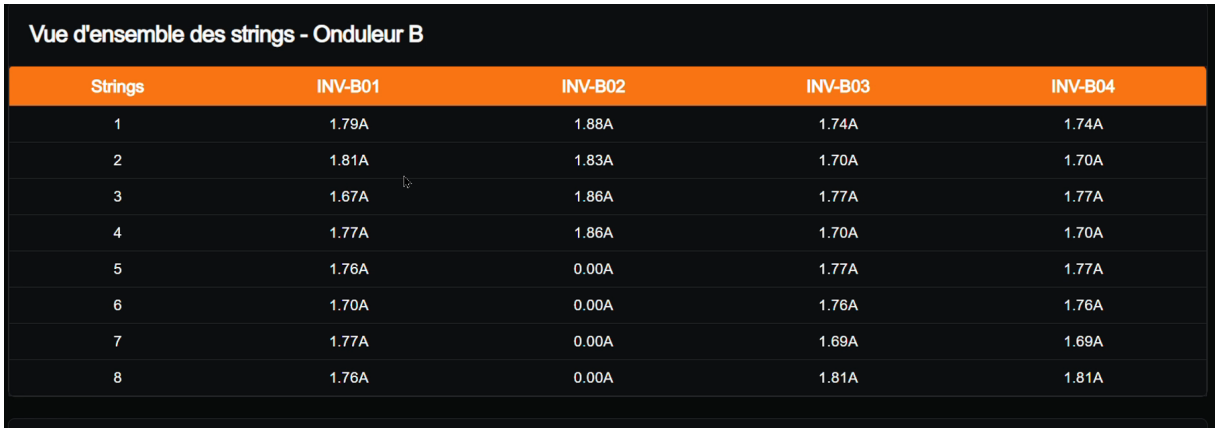


Figure 5.9: Inverter's junction boxes

- **Collective Strings**

Descending to Level 2, the interface focuses on one side of a selected inverter, presenting the full assembly of its associated junction boxes along with their constituent PV strings. Each junction box serves as an aggregation point, and the interface displays both the box-level electrical characteristics, such as total voltage and current, and the individual string currents feeding into it. This arrangement allows users to assess the collective behaviour of grouped strings while retaining visibility into individual contributions, thereby enhancing the detection of subsystem irregularities or mismatches. This functionality is depicted in Figure 5.10.



Vue d'ensemble des strings - Onduleur B

Strings	INV-B01	INV-B02	INV-B03	INV-B04
1	1.79A	1.88A	1.74A	1.74A
2	1.81A	1.83A	1.70A	1.70A
3	1.67A	1.86A	1.77A	1.77A
4	1.77A	1.86A	1.70A	1.70A
5	1.76A	0.00A	1.77A	1.77A
6	1.70A	0.00A	1.76A	1.76A
7	1.77A	0.00A	1.69A	1.69A
8	1.76A	0.00A	1.81A	1.81A

Figure 5.10: Inverter's strings

- **Single Junction Box Strings**

Finally, at the most granular scale (Level 3), the interface provides detailed monitoring within a selected junction box, where each individual PV string is displayed separately. Real-time readings of DC voltage and current are presented for each string, enabling localization performance assessment. In addition, the system highlights detected anomalies at the string level, allowing for immediate identification and targeted intervention. This functionality is depicted in Figure 5.11.

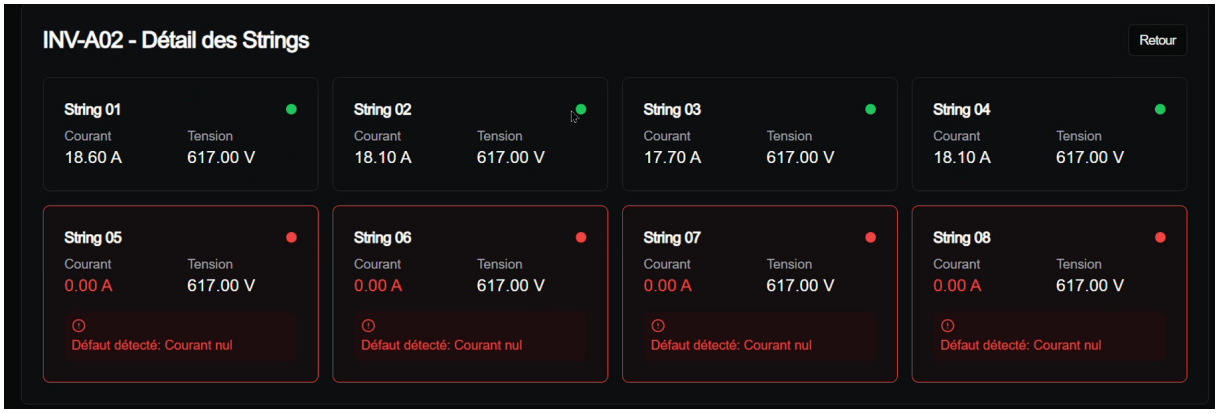


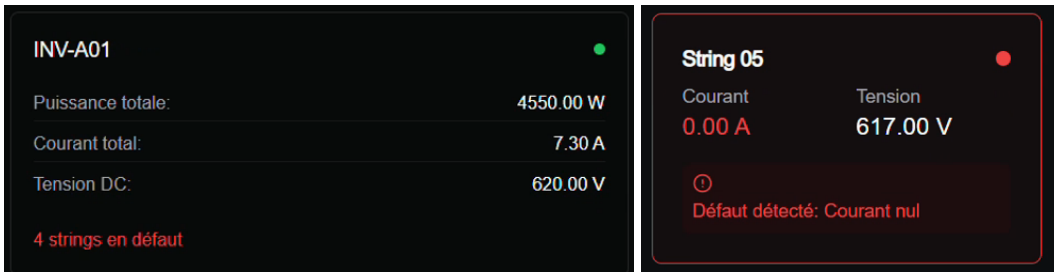
Figure 5.11: Isolated junction box level 1

- **Strings Alarms System**

Continuing with the alarm system, fault detection extends beyond the inverter level to both junction boxes and individual strings.

At the junction box level, each unit displays its alarm flag when one or more connected strings are in fault. The interface also shows the number of affected strings, helping operators quickly narrow down the location of the issue without needing to inspect each string individually. This functionality is illustrated in Figure 5.12a.

At the string level, each string is shown with its own alarm indicator. A fault is raised when the current value is null, typically signaling a disconnection or malfunction. The interface marks these cases with both a visual flag and a current value of zero, enabling quick identification of faulty strings. This behaviour is demonstrated in Figure 5.12b.



(a) Junction box alarms (b) String alarms

Figure 5.12: Alarms on junction box and string level

5.3.4 Historical Logging Page

This section outlines the long-term data management capabilities embedded within the SCADA interface. In addition to supporting real-time monitoring and analytical functionalities, the system ensures persistent logging of both raw measurements and derived performance indicators. These archival features are essential for facilitating trend analysis, regulatory compliance, and comprehensive performance audits over time.

The interface supports flexible configurations for data sampling rates, saving intervals, and storage formats, allowing adaptation to various operational and reporting requirements.

- **Data Saving Configuration**

  The interface provides flexible logging configurations, enabling operators to tailor data acquisition strategies based on specific operational needs. Users can target particular components, such as individual inverters or junction boxes, for focused monitoring and recording. Key logging parameters are fully customizable, including:

  - **Sampling interval:** Defines how frequently measurements are recorded.
  - **Inverter selection:** Enables logging of only selected metrics, reducing unnecessary data storage.

This configurability ensures efficient use of storage resources while maintaining high-resolution data for critical components. Figure 5.13 illustrates the data saving configuration interface.

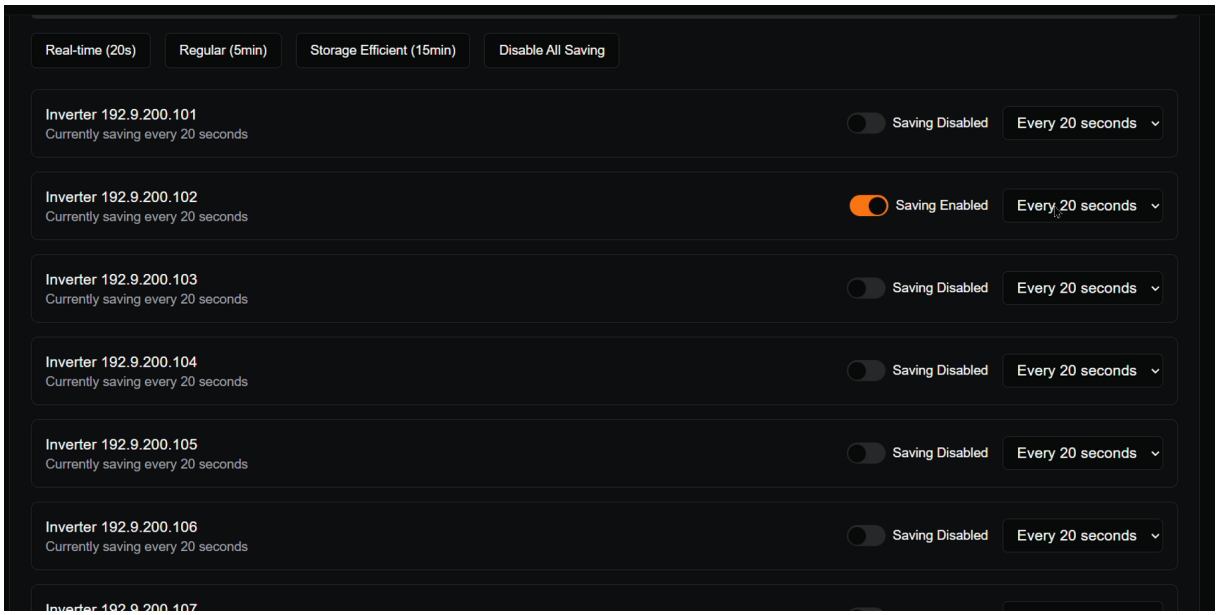


Figure 5.13: Data saving parameters

- **Historical Data Visualization and Report**

The reporting page enables visualization of historical data through time-series plots. By selecting a specific date, users can review previously recorded system behaviour in the form of performance curves, providing contextual insights into past operational conditions. In addition to visualization, the interface supports customized data export: operators can select which inverters to include and optionally choose whether to incorporate junction box-level data. The exported file is generated in an Excel-compatible format, facilitating further offline analysis and reporting. This functionality is demonstrated in Figure 5.14.



Figure 5.14: Historical report

- **Data Exported File**

  The exported files are generated in an Excel-compatible format and are structured to ensure both readability and traceability. Each file begins with general metadata, including the station name, the selected date, and the identifier of the associated subsystem, such as the meteorological station name, the inverter number and side designation (A or B), or the corresponding junction box reference.

  The main body of each file is organized according to the nature of the subsystem:

  – **Meteorological Station File:** This file contains environmental parameters recorded by the weather station, including irradiance, ambient temperature, wind speed, atmospheric pressure, and module temperature. An example of this layout is shown in Figure 5.15.

  – **Junction Box File:** Each junction box is identified by the inverter and side to which it is connected (e.g., "Inverter 10 – Side A"). The recorded parameters focus on string-level characteristics, such as voltage and current per string. An example layout that includes junction box data is shown in Figure 5.16.

  – **Inverter File:** Data are structured by inverter, with separate sections for Sides A and B. Each section presents key electrical parameters—voltage, current, power, and energy—recorded over time. An example without junction box data is shown in Figure 5.17.

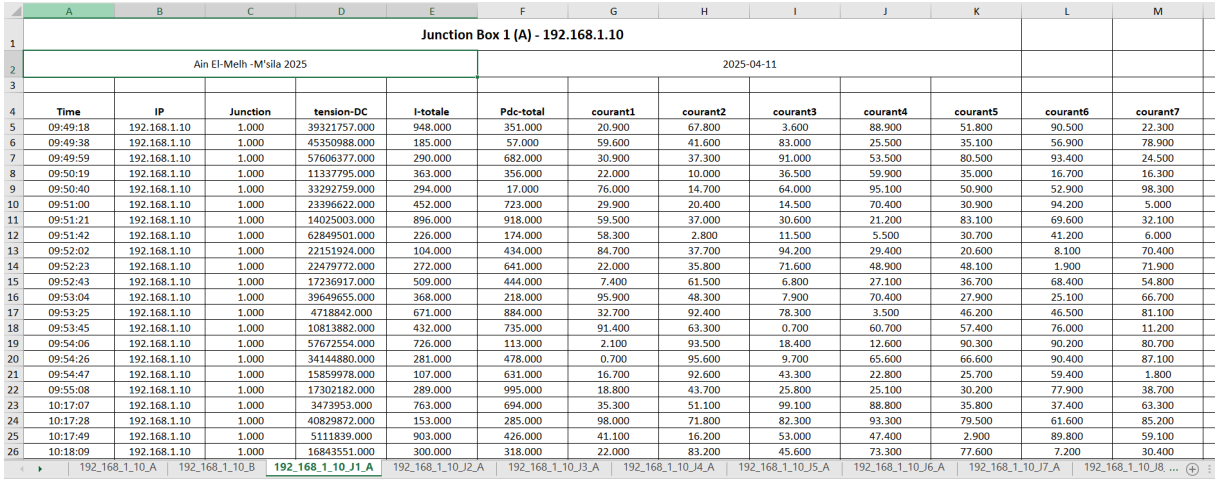
For all three file types, each row of data includes a precise timestamp, allowing accurate alignment of measurements with their corresponding acquisition time. This standardized structure supports offline analysis, formal reporting, and long-term data archival.

A1 | Energie&Meteorologiquie

	A	B	C	D	E	F	G	H	I	J	K	L
1	Energie&Meteorologiquie											
2	Ain El-Melh -M'sila 2025						2025-02-20					
3												
4	Temps	Temperature ambiante (°C)	Temperature panneau (°C)	Rayonnement inclinaison	Rayonnement total (W/m²)	Rayonnement direct (W/m²)	Rayonnement diffuse (W/m²)	Vitesse de vent (m/s)	Direction de vent (°)	Pression (hPa)	Humidite (%)	Yr (h)
5	15:43:54	70.000	5.300	73.000	24.000	6.800	2.300	39.000	3.000	89.000	36.000	0.000
6	15:44:14	56.000	8.000	20.000	57.000	9.300	5.600	46.000	75.000	47.000	48.000	0.000
7	15:44:34	27.000	2.000	83.000	65.000	0.600	9.600	83.000	72.000	59.000	56.000	0.000
8	15:44:54	21.000	0.000	25.000	28.000	8.700	3.300	70.000	14.000	33.000	55.000	0.000
9	15:45:15	87.000	0.300	54.000	12.000	7.500	8.700	2.000	67.000	56.000	53.000	0.000
10	15:45:35	73.000	9.200	99.000	28.000	6.500	6.700	22.000	0.000	64.000	54.000	0.000
11	15:45:55	27.000	0.600	74.000	79.000	3.700	0.800	93.000	6.000	5.000	46.000	0.000
12	15:46:15	51.000	8.500	49.000	20.000	6.700	10.000	87.000	23.000	39.000	52.000	0.000
13	15:46:36	73.000	1.700	13.000	8.000	2.600	8.900	68.000	80.000	84.000	23.000	0.000
14	15:46:56	41.000	5.500	76.000	8.000	0.900	4.700	26.000	76.000	2.000	36.000	0.000
15	15:47:16	38.000	3.100	91.000	84.000	9.400	9.300	98.000	49.000	95.000	34.000	0.000
16	15:47:36	84.000	8.700	24.000	52.000	2.300	1.000	72.000	72.000	35.000	26.000	0.000
17	15:47:56	72.000	0.500	92.000	57.000	2.900	0.600	87.000	27.000	3.000	56.000	0.000
18	15:48:17	92.000	9.500	30.000	52.000	5.000	7.900	58.000	86.000	63.000	14.000	0.000
19	15:48:37	4.000	3.100	34.000	82.000	0.200	0.000	67.000	56.000	36.000	19.000	0.000
20	15:48:57	72.000	8.800	43.000	16.000	4.400	0.000	91.000	29.000	4.000	23.000	0.000
21	15:49:17	96.000	3.800	80.000	98.000	2.700	1.200	71.000	75.000	30.000	29.000	0.000
22	15:49:38	55.000	7.000	79.000	80.000	9.000	4.300	31.000	25.000	83.000	39.000	0.000
23	15:49:58	49.000	7.800	77.000	91.000	2.100	7.700	40.000	13.000	66.000	38.000	0.000
24	15:50:18	100.000	8.900	11.000	39.000	5.900	4.500	74.000	92.000	100.000	26.000	0.000
25	15:50:38	55.000	3.600	16.000	36.000	1.800	0.000	19.000	29.000	41.000	39.000	0.000
26	15:50:59	19.000	7.200	98.000	20.000	3.600	4.600	55.000	94.000	66.000	38.000	0.000

Energie&Meteorologiquie

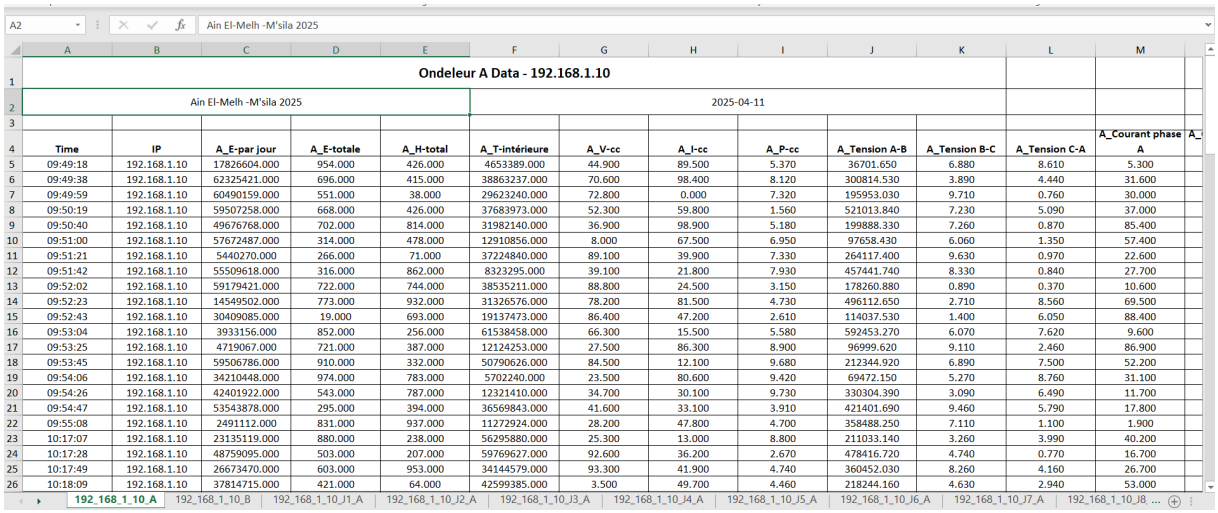
Figure 5.15: meteo Excel file example



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Junction Box 1 (A) - 192.168.1.10												
2	Ain El-Melh -M'sila 2025					2025-04-11							
3													
4	Time	IP	Junction	tension-DC	I-totale	Pdc-total	courant1	courant2	courant3	courant4	courant5	courant6	courant7
5	09:49:18	192.168.1.10	1.000	39321757.000	948.000	351.000	20.900	67.800	3.600	88.900	51.800	90.500	22.300
6	09:49:38	192.168.1.10	1.000	45350988.000	185.000	57.000	59.600	41.600	83.000	25.500	35.100	56.900	78.900
7	09:49:59	192.168.1.10	1.000	57606377.000	290.000	682.000	30.900	37.300	91.000	53.500	80.500	93.400	24.500
8	09:50:19	192.168.1.10	1.000	11337795.000	363.000	356.000	22.000	10.000	36.500	59.900	35.000	16.700	16.300
9	09:50:40	192.168.1.10	1.000	33292759.000	294.000	17.000	76.000	14.700	64.000	95.100	50.900	52.900	98.300
10	09:51:00	192.168.1.10	1.000	23396622.000	452.000	723.000	29.900	20.400	14.500	70.400	30.900	94.200	5.000
11	09:51:21	192.168.1.10	1.000	14025003.000	896.000	918.000	59.500	37.000	30.600	21.200	83.100	69.600	32.100
12	09:51:42	192.168.1.10	1.000	62849501.000	226.000	174.000	58.300	2.800	11.500	5.500	30.700	41.200	6.000
13	09:52:02	192.168.1.10	1.000	22151924.000	104.000	434.000	84.700	37.700	94.200	29.400	20.600	8.100	70.400
14	09:52:23	192.168.1.10	1.000	22479772.000	272.000	641.000	22.000	35.800	71.600	48.900	48.100	1.900	71.900
15	09:52:43	192.168.1.10	1.000	17236917.000	509.000	444.000	7.400	61.500	6.800	27.100	36.700	68.400	54.800
16	09:53:04	192.168.1.10	1.000	39649655.000	368.000	218.000	95.900	48.300	7.900	70.400	27.900	25.100	66.700
17	09:53:25	192.168.1.10	1.000	4718842.000	671.000	884.000	32.700	92.400	78.300	3.500	46.200	46.500	81.100
18	09:53:45	192.168.1.10	1.000	10813882.000	432.000	735.000	91.400	63.300	0.700	60.700	57.400	76.000	11.200
19	09:54:06	192.168.1.10	1.000	57672554.000	726.000	113.000	2.100	93.500	18.400	12.600	90.300	90.200	80.700
20	09:54:26	192.168.1.10	1.000	34144880.000	281.000	478.000	0.700	95.600	9.700	65.600	66.600	90.400	87.100
21	09:54:47	192.168.1.10	1.000	15859978.000	107.000	631.000	16.700	92.600	43.300	22.800	25.700	59.400	1.800
22	09:55:08	192.168.1.10	1.000	17302182.000	289.000	995.000	18.800	43.700	25.800	25.100	30.200	77.900	38.700
23	10:17:07	192.168.1.10	1.000	3473953.000	763.000	694.000	35.300	51.100	99.100	88.800	35.800	37.400	63.300
24	10:17:28	192.168.1.10	1.000	40829872.000	153.000	285.000	98.000	71.800	82.300	93.300	79.500	61.600	85.200
25	10:17:49	192.168.1.10	1.000	5111839.000	903.000	426.000	41.100	16.200	53.000	47.400	2.900	89.800	59.100
26	10:18:09	192.168.1.10	1.000	16843551.000	300.000	318.000	22.000	83.200	45.600	73.300	77.600	7.200	30.400

192_168_1_10_A | 192_168_1_10_B | 192_168_1_10_J1_A | 192_168_1_10_J2_A | 192_168_1_10_J3_A | 192_168_1_10_J4_A | 192_168_1_10_J5_A | 192_168_1_10_J6_A | 192_168_1_10_J7_A | 192_168_1_10_J8. ...

Figure 5.16: Junction box inverter 10 side A Excel file example



A2 | Ain El-Melh -M'sila 2025

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Ondeleur A Data - 192.168.1.10												
2	Ain El-Melh -M'sila 2025					2025-04-11							
3													
4	Time	IP	A_E-par jour	A_E-totale	A_H-total	A_T-intérieure	A_V-cc	A_I-cc	A_P-cc	A_Tension A-B	A_Tension B-C	A_Tension C-A	A_Courant phase A	A_
5	09:49:18	192.168.1.10	17826604.000	954.000	426.000	4653389.000	44.900	89.500	5.370	36701.650	6.880	8.610	5.300	
6	09:49:38	192.168.1.10	62325421.000	696.000	415.000	38863237.000	70.600	98.400	8.120	300814.530	3.890	4.440	31.600	
7	09:49:59	192.168.1.10	60490159.000	551.000	38.000	29623240.000	72.800	0.000	7.320	195953.030	9.710	0.760	30.000	
8	09:50:19	192.168.1.10	59507258.000	668.000	426.000	37683973.000	52.300	59.800	1.560	521013.840	7.230	5.090	37.000	
9	09:50:40	192.168.1.10	49676768.000	702.000	814.000	31982140.000	36.900	98.900	5.180	199888.330	7.260	0.870	85.400	
10	09:51:00	192.168.1.10	57672487.000	314.000	478.000	12910856.000	8.000	67.500	6.950	97658.430	6.060	1.350	57.400	
11	09:51:21	192.168.1.10	5440270.000	266.000	71.000	37224840.000	89.100	39.900	7.330	264117.400	9.630	0.970	22.600	
12	09:51:42	192.168.1.10	55509618.000	316.000	862.000	8323295.000	39.100	21.800	7.930	457441.740	8.330	0.840	27.700	
13	09:52:02	192.168.1.10	59179421.000	722.000	744.000	38535211.000	88.800	24.500	3.150	178260.880	0.890	0.370	10.600	
14	09:52:23	192.168.1.10	14549502.000	773.000	932.000	31326576.000	78.200	81.500	4.730	496112.650	2.710	8.560	69.500	
15	09:52:43	192.168.1.10	30409085.000	19.000	693.000	19137473.000	86.400	47.200	2.610	114037.530	1.400	6.050	88.400	
16	09:53:04	192.168.1.10	3933156.000	852.000	256.000	61538458.000	66.300	15.500	5.580	592453.270	6.070	7.620	9.600	
17	09:53:25	192.168.1.10	4719067.000	721.000	387.000	12124253.000	27.500	86.300	8.900	96999.620	9.110	2.460	86.900	
18	09:53:45	192.168.1.10	59506786.000	910.000	332.000	50790626.000	84.500	12.100	9.680	212344.920	6.890	7.500	52.200	
19	09:54:06	192.168.1.10	34210448.000	974.000	783.000	5702240.000	23.500	80.600	9.420	69472.150	5.270	8.760	31.100	
20	09:54:26	192.168.1.10	42401922.000	543.000	787.000	12321410.000	34.700	30.100	9.730	330304.390	3.090	6.490	11.700	
21	09:54:47	192.168.1.10	53543878.000	295.000	394.000	36569843.000	41.600	33.100	3.910	421401.690	9.460	5.790	17.800	
22	09:55:08	192.168.1.10	2491112.000	831.000	937.000	11272924.000	28.200	47.800	4.700	358488.250	7.110	1.100	1.900	
23	10:17:07	192.168.1.10	23135119.000	880.000	238.000	56295880.000	25.300	13.000	8.800	211033.140	3.260	3.990	40.200	
24	10:17:28	192.168.1.10	48759095.000	503.000	207.000	59769627.000	92.600	36.200	2.670	478416.720	4.740	0.770	16.700	
25	10:17:49	192.168.1.10	26673470.000	603.000	953.000	34144579.000	93.300	41.900	4.740	360452.030	8.260	4.160	26.700	
26	10:18:09	192.168.1.10	37814715.000	421.000	64.000	42599385.000	3.500	49.700	4.460	218244.160	4.630	2.940	53.000	

192_168_1_10_A | 192_168_1_10_B | 192_168_1_10_J1_A | 192_168_1_10_J2_A | 192_168_1_10_J3_A | 192_168_1_10_J4_A | 192_168_1_10_J5_A | 192_168_1_10_J6_A | 192_168_1_10_J7_A | 192_168_1_10_J8. ...

Figure 5.17: inverter 10 side A Excel file example

While the diversity of monitoring views enables comprehensive supervision of both macro- and micro-level dynamics within the PV field, these capabilities alone are not sufficient. The complexity of large-scale photovoltaic systems necessitates the integration of analytical models to support deeper interpretation, performance assessment, and predictive fault detection.

5.4 Analysis Interface

Advancing from surface-level monitoring to in-depth interpretation, the analysis interface enables higher-order supervision by transforming raw data into actionable insights through prediction algorithms and fault detection and diagnostic procedures. This analytical layer is designed to enhance situational awareness by identifying performance deviations, projecting future system states, and recognizing early signs of failure.

Each sub-module is organized to support a comprehensive analytical workflow, beginning with predictive model visualization, followed by fault detection, and concluding with detailed fault diagnosis.

5.4.1 Home Page

To enhance system reliability and operational safety, a standalone application was developed to integrate predictive modeling, fault detection algorithms, and diagnostic evaluation into a unified interface. The objective was to provide a practical, user-friendly tool that enables early identification of performance deviations and supports informed maintenance actions.

The application is structured to guide users through a complete diagnostic workflow, from setting detection parameters to interpreting prediction results and finally the system health assessment. As shown in Figure 5.18, the landing page offers centralized access to key components of the application, including AI model loading, fault detection configuration, and diagnostic execution, while maintaining process traceability through real-time system status updates.

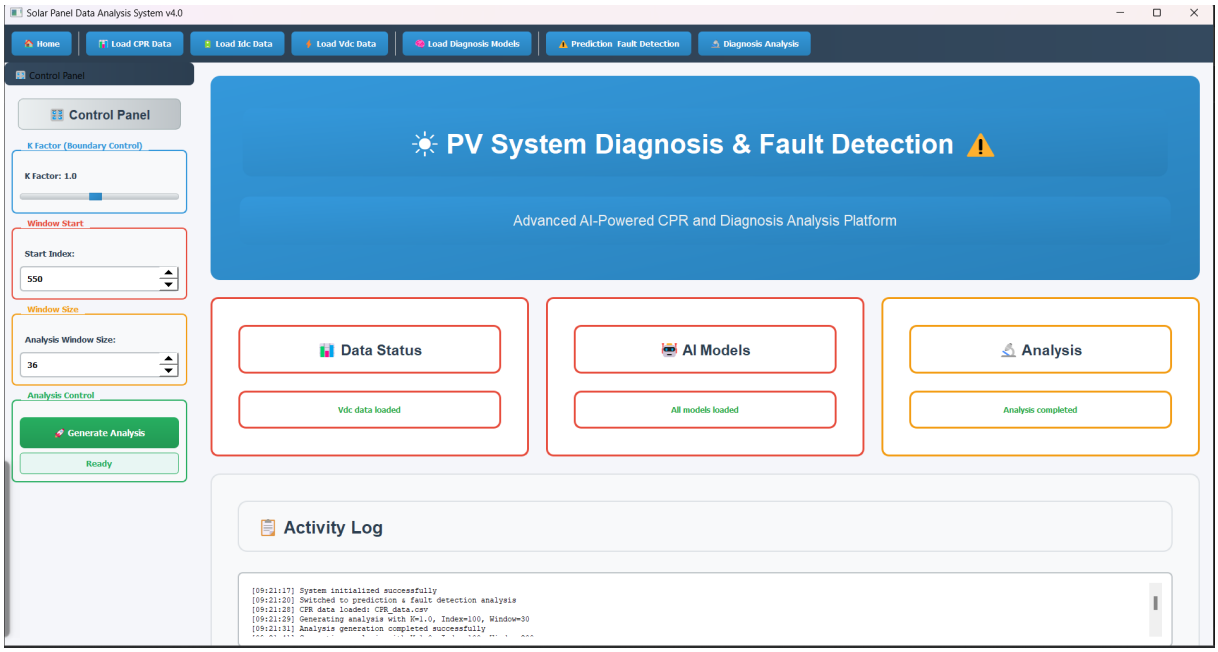


Figure 5.18: Home page PV System Diagnosis and Fault Detection

5.4.2 Prediction and Fault Detection page

The prediction module, as shown in Figure 5.19, presents a visual comparison between predicted and measured values of DC voltage and current. This allows for the validation of offline-trained models under actual operational conditions. By plotting both real-time and predicted signals on the same axes, the interface helps operators assess the accuracy and reliability of prediction models and detect early performance shifts behaviour.



Figure 5.19: Prediction Page

The fault detection procedure as shown in figure 5.20, is based on a Corrected Performance Ratio (CPR) classification scheme, which defines diagnostic boundaries to assess inverter health. Inverters are considered healthy when their observed CPR values fall within CPR-defined limits, and anomalous when deviations are detected when values drop below lower thresholds. This method provides a structured way to evaluate performance against expected

To complement this classification, the system calculates the Anomaly Frequency Index (AFI) for each inverter, a quantitative measure of how frequently an inverter deviates from normal operating conditions. These values are presented in an AFI matrix, offering a comprehensive view of anomaly consistency and distribution across the plant. Together, the CPR-based classification and AFI quantification form a robust framework for early fault detection.

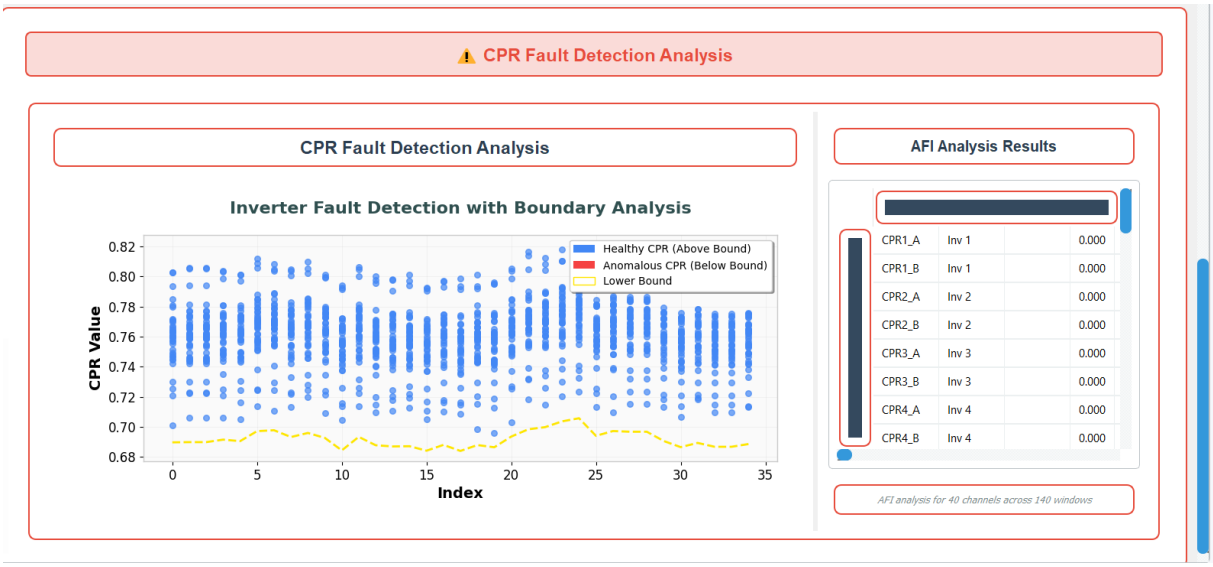


Figure 5.20: Faults Detection Page

5.4.3 Diagnosis page

The final diagnosis interface (Figure 5.21) provides a concise summary of system health through comparative analysis of predicted versus measured DC voltage and current. When both metrics remain within expected thresholds, the system is labelled as healthy. In more critical scenarios, this module can pinpoint the nature and extent of inverter malfunctions.

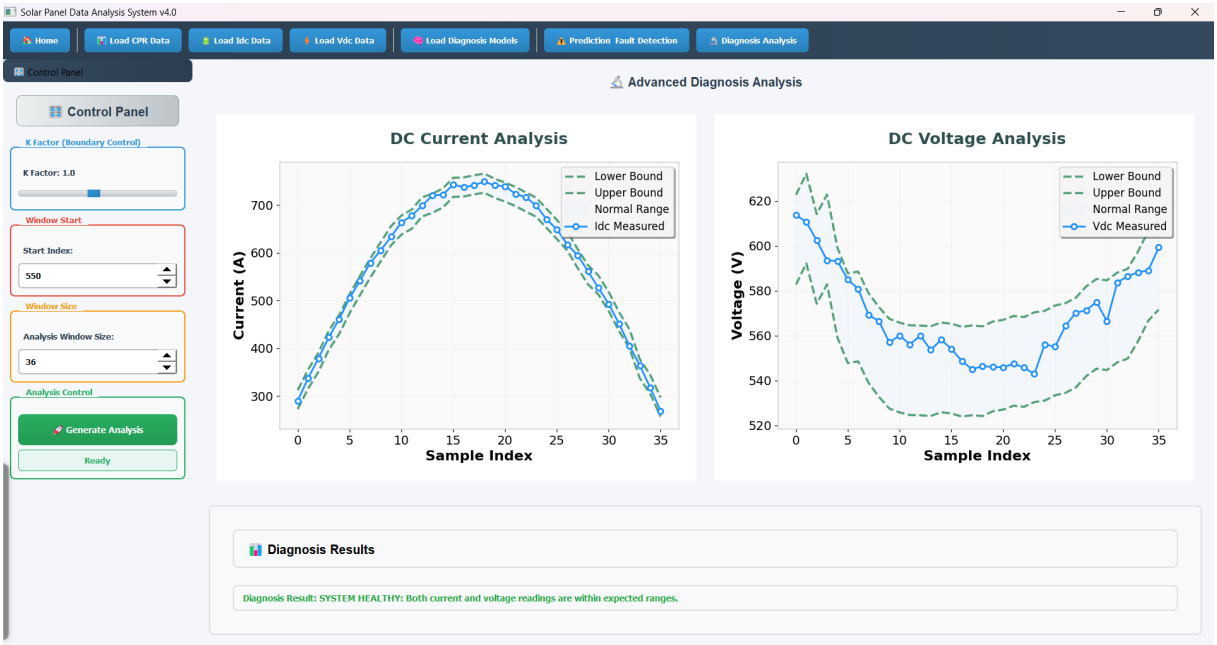


Figure 5.21: Diagnosis Analysis Page

Collectively, these functionalities reflect a shift from traditional threshold-based alarms to data-driven predictive diagnostics, improving fault localization, reducing response time, and supporting preventive maintenance strategies within PV plant operations.

5.5 System Evaluation

To evaluate the effectiveness of the proposed SCADA system, a comparison is conducted with the previous interface, highlighting improvements in usability, data resolution, visualization, and diagnostics. Figure 5.1 summarizes the key differences.

Table 5.1: comparison summary of old and new SCADA interface

Feature	Old SCADA System	New SCADA Interface (This Work)	Improvement / Added Value
User Interface (UI)	Basic, non-intuitive layout	Modern, user-friendly dashboard with modular design	Improved user experience and navigation
Visualization (Graphs/Charts)	Static and minimal	Dynamic, customizable charts with zoom and layer selection	Enhanced data interpretation
Alarm System	Unavailable	Multilevel alerts with history and logs	Better diagnostics and faster operator response
Data Interchange	Basic, limited access	Exportable formats (CSV, Excel) with enhanced access	Improved data analysis and reporting
Performance Factors	Unavailable	Standardized PV parameters	Component-level and global system efficiency tracking
Prediction Module	Unavailable	Deep learning hybrid model	Comparison of supposed vs. actual values
Fault Detection & Diagnosis	Unavailable	Not reliant on historical data	Enables predictive maintenance
Customization Options	Very limited	Personalized layout saving and display preferences	Better adaptability to user needs
Open-Source Architecture	Not applicable	Fully robust and extensible software architecture	Enables cost-effective deployment, integration flexibility, and long-term scalability

These enhancements collectively improve the system's responsiveness, clarity, and scalability.

5.6 Conclusion

The developed SCADA interface meets the core objectives of this project, delivering a robust, modular, and user-oriented platform for supervising grid-connected photovoltaic power plants. Its layered architecture enables comprehensive operational oversight through real-time monitoring, analytical interpretation, and historical data management.

Throughout this chapter, the interface was validated under real operating conditions, demonstrating its capacity to transform raw operational data into meaningful insights. The results

confirm that the system enables not only continuous supervision but also informed interpretation and timely decision-making. The structured integration of monitoring, analysis, and data saving reflects a practical and scalable approach to photovoltaic plant management, aligning with the project's overarching goals of reliability, usability, and operational transparency.

Together, these results highlight the system's responsiveness, flexibility, and practical applicability, confirming its suitability as a complete SCADA solution for photovoltaic plant supervision.

General Conclusion

This final year project focused on the comprehensive redesign of the SCADA interface for a photovoltaic delivery substation, leveraging a modern approach built on advanced web development technologies. The project successfully transformed an outdated, rigid, and limited interface into a flexible, user-centred, interactive, and scalable solution that meets current requirements for data supervision and industrial visualization.

The work began with an in-depth analysis of the existing system, identifying functional and visual limitations. Guided by the insights gained from this evaluation, a new architecture was proposed. It features a clear separation between the frontend developed for visualization and the back-end implemented for logic. Additionally, a dedicated communication layer ensures smooth interaction with field devices via the Modbus protocol. A robust database was also integrated to store real-time measurements and equipment states efficiently.

The resulting interface allows real-time monitoring of electrical parameters and operating statuses of inverters and transformers. It also introduces advanced features such as energy trends prediction, performance analysis, and early fault detection and diagnostic capabilities. Moreover,the use of interactive visualization tools significantly enhances data interpretation through dynamic and interactive graphs, and supports more informed decision-making by operators.

In summary, this work demonstrates the feasibility and practical benefits of integrating modern web technologies into industrial control systems, resulting in enhanced robustness, performance, and scalability. This redesign further establishes a solid foundation for future enhancements, including the integration of artificial intelligence for advanced predictive analytics and the deployment of mobile platforms to enable remote monitoring capabilities. Ultimately, the system is designed to operate as an integral part of the photovoltaic power plant and to enable future interoperability with national grid dispatching systems, aligning with the broader vision of smart energy infrastructure.

Bibliography

[1] INTERNATIONAL ENERGY AGENCY. Snapshot of global pv markets 2023. Technical Report, 2023. URL: https://iea-pvps.org/wp-content/uploads/2023/04/IEA_PVPS_Snapshot_2023.pdf.

[2] Bhuwan Pratap Singh, Sunil Kumar Goyal, and Prakash Kumar. Solar pv cell materials and technologies: Analyzing the recent developments. *Materials Today: Proceedings*, 43:2843--2849, 2021. CRMSC-2021. URL: https://www.sciencedirect.com/science/article/pii/S2214785321000146, doi:10.1016/j.matpr.2021.01.003.

[3] Daniel de B. Mesquita, João Lucas de S. Silva, Hugo S. Moreira, Michelle Kitayama, and Marcelo G. Villalva. A review and analysis of technologies applied in pv modules. In *2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*, pages 1--6, 2019. URL: https://ieeexplore.ieee.org/abstract/document/8895369, doi:10.1109/ISGT-LA.2019.8895369.

[4] Y. Zhu and W. Xiao. A comprehensive review of topologies for photovoltaic i–v curve tracer. *Solar Energy*, 196:346--357, 2020. URL: https://www.sciencedirect.com/science/article/pii/S0038092X19312344, doi:10.1016/j.solener.2019.12.020.

[5] S. Sorrentino and G. De Carne. Assessment of the usability and accuracy of the simplified one□diode models for photovoltaic modules. *Energies (MDPI)*, 2017. URL: https://www.mdpi.com/1996-1073/10/11/1717.

[6] University of Central Florida FSEC. Cells, Modules, Panels and Arrays. Technical Guide. Technical report, Florida Solar Energy Center, University of Central Florida, 2025. URL: https://energyresearch.ucf.edu/consumer/solar-technologies/solar-electricity-basics/cells-modules-panels-and-arrays/.

[7] Wenham Xiao. *Photovoltaic Power System: Modeling, Design, and Control*. Wiley, Hoboken, NJ, 2017. URL: https://dokumen.pub/photovoltaic-power-system-modelling-design-and-control-9781119280408-111928040html.

[8] V. Karthikeyan, S. Rajasekar, Vipin Das, P. Karuppanan, and Asheesh Kumar Singh. *Grid-Connected and Off-Grid Solar Photovoltaic System*, pages 125--157. Springer International Publishing, 2017. doi:10.1007/978-3-319-50197-0_5.

[9] J Sreedevi, N Ashwin, and M Naini Raju. A study on grid connected pv system. In *2016 National Power Systems Conference (NPSC)*, pages 1--6, 2016. URL: https://ieeexplore.ieee.org/abstract/document/7858870, doi:10.1109/NPSC.2016.7858870.

[10] Rabindra Satpathy and Venkateswarlu Pamuru. Off□grid solar photovoltaic systems. In *Solar PV Power: Design, Manufacturing and Applications from Sand to Systems*, chapter 7. Elsevier, 2021. URL: https://sdgresources.relx.com/book-chapters/solar-pv-power-chapter-7-grid-solar-photovoltaic-systems.

[11] Zhang Yijie, Ma Tao, and Yang Hongxing. A review on capacity sizing and operation strategy of grid-connected photovoltaic battery systems. *Energy and Built Environment*, 5(4):500--516, 2024. Open Access PDF. URL: https://www.sciencedirect.com/science/article/pii/S2666123323000156, doi:10.1016/j.enbenv.2023.04.001.

[12] Amal Herez, Rania Khaled, Haitham Abu-Rub, Nouredine Ghaffour, and Ahmad El Hajj. A review on the classifications and applications of solar photovoltaic technology. *AIMS Energy*, 11(6):1102--1130, 2023. URL: https://www.aimspress.com/article/doi/10.3934/energy.2023056, doi:10.3934/energy.2023056.

[13] International Electrotechnical Commission. Photovoltaic system performance– part1: Monitoring. Standard IEC 61724□1:2021, IEC, 2021. URL: https://cdn.standards.iteh.ai/samples/103495/100a1ea212914dd38ba72a9138af4ce1/IEC-61724-1-2021.pdf.

[14] Daryl R. Myers. Evaluation of the performance of the pvusa rating methodology applied to dual junction pv technology. Technical Report NREL/TP-6A2-45376, National Renewable Energy Laboratory, Golden, CO, July 2009. URL: https://www.nrel.gov/docs/fy09osti/45376.pdf.

[15] INTERNATIONAL ENERGY AGENCY PVPS. Analysis of photovoltaic systems. Technical Report, 2020. URL: https://iea-pvps.org/wp-content/uploads/2020/01/rep2_01.pdf.

[16] Sarah Kurtz. Weather corrected performance ratio. Technical Report, 2013. URL: https://www.nrel.gov/docs/fy13osti/57991.pdf.

[17] João Figueiredo and José Sá da Costa. A scada system for energy management in intelligent buildings. *Energy and Buildings*, 49:85--98, 2012. URL: https://www.sciencedirect.com/science/article/pii/S0378778812000722, doi:10.1016/j.enbuild.2012.01.041.

[18] ScienceDirect Topics. Supervisory control and data acquisition system. https://www.sciencedirect.com/topics/engineering/supervisory-control-and-data-acquisition-system, 2025. Accessed June 2025.

[19] Vincenzo Pagano, Christian Cianfrini, and Velio Casella. A study on an iot-based scada system for photovoltaic utility plants. *Electronics*, 13(11):2065, 2023. URL: https://www.mdpi.com/2079-9292/13/11/2065, doi:10.3390/electronics13112065.

[20] Hu Guozhen, Cai tao, Chen Changsong, and Duan Shanxu. Solutions for scada system communication reliability in photovoltaic power plants. In *2009 IEEE 6th International Power Electronics and Motion Control Conference*, pages 2482--2485, 2009. URL: https://ieeexplore.ieee.org/abstract/document/5157821, doi:10.1109/IPEMC.2009.5157821.

[21] Ioan Cristian Hoarcă. Energy management for a photovoltaic power plant based on scada system. In *2021 13th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, pages 1--9, 2021. URL: https://ieeexplore.ieee.org/abstract/document/9515136, doi:10.1109/ECAI52376.2021.9515136.

[22] A. Setiawan, Sugeng, K. I. Koesoema, S. Bakhri, and J. Aditya. The scada system using plc and hmi to improve the effectiveness and efficiency of production processes. In *IOP Conference Series: Materials Science and Engineering*, volume 550, page 012008, 2019. URL: https://iopscience.iop.org/article/10.1088/1757-899X/550/1/012008/pdf, doi:10.1088/1757-899X/550/1/012008.

[23] V. Ramnarayan, V. Srividya, and R. Lakshmipathi. Scada protocols and communication trends. In *ISA EXPO 2004 Technical Conference*. ISA (International Society of Automation), 2004. Accessed June 2025. URL: https://d1wqtxts1xzle7.cloudfront.net/36429091/TP04ISA048-SCADA-protocols-and-communications-trends-libre.pdf.

[24] Shreyas Sharma. Modbus rtu: A comprehensive guide to understanding and implementing the protocol. https://fr.scribd.com/.../Modbus-RTU-Guide, 2025. Accessed June 2025.

[25] Antonio Armenta. Differences between a programmable logic controller (plc) and a remote terminal unit (rtu). https://fr.scribd.com/document/753802059/

Differences-Between-a-Programmable-Logic-Controller-PLC-and-a-Remote-Terminal-
2022. Accessed June 2025.

[26] Ramanjaneyulu Marni, Anirudh Reddy R, Ramchander Nirudi, and R. Harshavardhana Rao Y. Accurate string measurements for solar pv power plants to improve the performance and reliability. *International Journal of Engineering Research & Technology (IJERT)*, 5(10):1--5, 2016. URL: https://www.ijert.org/accurate-string-measurements-for-solar-pv-power-plants-to-improve-the-performa doi:10.17577/IJERTV5IS100018.

[27] Hongwen Song, Yong He, and Xuefang He. Grid-connected photovoltaic systems: An overview of recent research and emerging pv converter technology. *International Journal of Emerging Electric Power Systems*, 17(6):597--614, 2016. URL: https://www.degruyter.com/document/doi/10.1515/ijeeps-2016-0078/html, doi:10.1515/ijeeps-2016-0078.

[28] Moadh KICHENE. Contribution to the development of new approaches to monitoring, fault detection and diagnostics for grid-connected photovoltaic systems. Mémoire de fin d'étude, USTO-MB, Oran, 2024.

[29] MODBUS Organization, Inc. Modbus over serial line specification and implementation guide. https://modbus.org/docs/Modbus_over_serial_line_V1_02.pdf, 2002. Accessed: 2025-02-20.

[30] MODBUS Organization, Inc. Modbus application protocol specification. https://modbus.org/docs/Modbus_Application_Protocol_V1_1b3.pdf, 2006. Accessed: 2025-02-20.

[31] MODBUS Organization, Inc. Modbus messaging on tcp/ip implementation guide. https://modbus.org/docs/Modbus_Messaging_Implementation_Guide_V1_0b.pdf, 2006. Accessed: 2025-02-20.

[32] GE Vernova. M6xxm modbus protocol manual. https://www.gevernova.com/grid-solutions/sites/default/files/resources/products/manuals/m6xxm_en_m_b.pdf. Accessed: 2025-02-20.

[33] MongoDB Inc. Time series collections. https://www.mongodb.com/docs/manual/core/timeseries-collections/, 2025. Accessed: 2025-02-20.

[34] Su-Chang Lim, Jun-Ho Huh, Seok-Hoon Hong, Chul-Young Park, and Jong-Chan Kim. Solar power forecasting using cnn-lstm hybrid model. *Energies*, 15(21):8233, 2022. URL: https://www.mdpi.com/1996-1073/15/21/8233, doi:10.3390/en15218233.

[35] Ahmad Abu Shanab, Taghi M. Khoshgoftaar, Randall Wald, and Jason Van Hulse. Data quality measures and data cleansing for data mining: A case study. *International Journal of Business Intelligence and Data Mining*, 7(2):116--134, 2012. doi:10.1504/IJBIDM.2012.048730.

[36] Victor Hugo C. de Albuquerque, Pethuru Raj, and Satya Prakash Yadav. *Smart Systems: Innovations, Security and Applications*. De Gruyter, 2024. doi:10.1515/9783111323749.

[37] Ao Dong, Yingying Zhao, Xiwei Liu, Li Shang, Qi Liu, and Dahai Kang. Fault diagnosis and classification in photovoltaic systems using scada data. In *2017 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC)*, pages 117--122, 2017. doi:10.1109/SDPC.2017.31.

[38] Y. H. Chen, R. Liang, Y. Tian, and F. Wang. A novel fault diagnosis method of pv based on power loss and i–v characteristics. In *IOP Conference Series: Earth and Environmental Science*, volume 40, page 012022. IOP Publishing, 2020. URL: https://iopscience.iop.org/article/10.1088/1755-1315/40/1/012022, doi:10.1088/1755-1315/40/1/012022.