Technische Hochschule Ingolstadt

Technische Hochschule Ingolstadt

Faculty of Mechanical Engineering

Master of Science in Renewable Energy Systems

Master's Thesis

Comparison of SCADA Systems and Their Effect on Predictive Maintenance Strategies in Large PV Power Plants

Submitted by

Rezzan Güzel

Issued on: 22.12.2022

Submitted on: 09.01.2023

First Examiner: Prof. Dr.-Ing. Daniel Navarro Gevers

Second Examiner: Prof. Dr.-Ing. Tobias Schrag

Declaration of Independence

I hereby declare that this thesis is my own work, that I have not presented it elsewhere for examination purposes and that I have not used any sources or aids other than those stated. I have marked verbatim and indirect quotations as such.

Ingolstadt, 16.12.2022

Rezzan Güzel

Abstract

Large PV power plant investments have accelerated in the last decade due to economic and political developments in the world. These investments have led to the technological development and maturity of all components used in a large PV power plant. However, large PV power plants still face various failures at the system and component levels. Detecting these failures instantly or even in advance will minimize the operation and maintenance costs and maximize the system reliability and safety of large PV power plants. As SCADA is still the most mature control and monitoring technology in the industry, PV power plant designers and operators can benefit from its functionalities to develop predictive maintenance strategies and control the large PV power plant built in a wide area. The thesis aims to conduct research on available SCADA infrastructure technologies for large PV power plants and compare them in a technical aspect to understand their effect on developing predictive maintenance strategies.

Developing predictive maintenance strategies for large PV power plants depends on the correct, coherent, and reliable data provided by the SCADA system. All the necessary SCADA components that generate and transmit the operational data are identified. Standard SCADA communication protocols and topologies are compared in terms of their applicability to large PV power plants and their positive contribution to developing predictive maintenance strategies. Key performance indicators are identified for generating the actual operational trends, which are compared to expected or designed operating trends in the scope of predictive maintenance strategy. To improve the reliability and safety of the SCADA system, cyber-attack risk mitigation methods and redundancy instruments are presented.

By considering all the findings and results of the multifaceted comparisons, a SCADA infrastructure system is designed as a case study to improve the effectiveness of the predictive maintenance strategy of a gold mine PV + BESS hybrid power plant in Burkina Faso.

Keywords: Photovoltaics, Large PV power plant, SCADA system, SCADA standard communication protocols, SCADA system network topologies, Predictive maintenance strategy, Key performance indicator, SCADA cyber-security

Acknowledgement

I would like to express my gratitude and indebtedness to Prof. Dr.-Ing. Daniel Navarro Gevers, for his inspiring guidance, and valuable suggestions during the thesis.

I am blessed to have received unconditional love and support from my parents, sisters, and girlfriend. Without their trust and endless encouragement, it would have been impossible to pursue my dreams.

Rezzan Güzel

Ingolstadt, December 2022

Table of Contents

List of Figures

List of Tables

List of Abbreviations

5G	5th Generation
AC	Alternating Current
ASCII	American Standard Code for Information Interchange
BESS	Battery Energy Storage System
BOOL	Boolean
CCTV	Closed-Circuit Television
CIP	Critical Infrastructure Protection
CPR	Corrected Performance Ratio
DC	Direct Current
DINT	Double Integer
DNP	Distributed Network Protocol
EPA	Enhanced Performance Architecture
GHI	Global Horizontal Irradiance
HMI	Human-Machine Interface
HV	High Voltage
I/O	Input/Output
IEC	International Electrotechnical Commission
IED	Intelligent Electronic Device
IEEE	Institute of Electrical and Electronics Engineers
INT	Integer
IoT	Internet of Things
IP	Internet Protocol
ISO	International Standard Organization
KPI	Key Performance Indicators
LAN	Local Area Network
LV	Low Voltage
MAC	Media Access Control
MBWA	Mobile Broadband Wireless Access
MTU	Master Terminal Unit
MV	Medium Voltage
NCSC	National Cyber Security Centre
NERC	North American Electric Reliability Corporation
NIST	National Institute of Standards and Technology
NNSA	National Nuclear Security Administration
O&M	Operation and Maintenance
OSI	Open System Interconnection
PLC	Programmable Logic Controller
PV	Photovoltaic
RS	Recommended Standard
RTU	Remote Terminal Unit
SCADA	Supervisory Control and Data Acquisition
SCL	Substation Configuration Language
SMV	Sampled Measured Values
TCP/IP	Transmission Control Protocol/Internet Protocol
TLS	Transport Layer Security
TS	Transformer Station
UBYTE	Unsigned Byte
UDINT	Unsigned Double Integer

UINT	Unsigned Integer
US	United States
WiMAX	Worldwide Interoperability for Microwave Access
ZigBee	Zonal Intercommunication Global standard

Symbols

°C	Centigrade
C	Temperature correction coefficient
E_{out}	AC energy generation
G_{ref}	Reference irradiance
H	Total irradiation on the plane of array
ha	Hectar
Hz	Hertz
km	Kilometer
kV	Kilovolt
kW	Kilowatt
kWh	Kilowatt-hour
kWh/kW_{peak}	Specific energy yield
kWp	Kilowatt-peak
m	Meter
MW	Megawatt
MW_{peak}	Megawatt-peak
MWh	Megawatt-hour
P_0	Rated DC power of the subjected PV array
$PR_{expected}$	Expected performance ratio
T_{module}	Temperature of the module
$T_{reference}$	Temperature under the standard test conditions
V	Volts
Y_E	Expected AC energy generation
Y_F	Energy yield
Y_r	Refence yield
γ	Coefficient of the maximum-power temperature

1 Introduction

In the last 50 years, increasing energy demand with economic and political events dictated the industry and countries to operate power plants with the highest level of efficiency. This demand for high efficiency can only be achieved by real-time monitoring and control of the power plant. Supervisory control and data acquisition (SCADA) is a system structure that enables the power plant operator to monitor, control, obtain and save the data of the power plant operation by using a combination of different hardware and software elements.

Each power plant type has its characteristic operation. Therefore, the SCADA system that will be implemented in each power plant type also needs to be unique to the operation type. The diversification and change of energy production sources, such as solar photovoltaic (PV), wind, and biomass energy generation, have forced SCADA systems to be upgraded, as well. Consequently, deciding on the proper SCADA system selection for large power plants requires an inclusive knowledge of SCADA systems and their power plant idiosyncratic advantages and disadvantages.

Although photovoltaic (PV) systems are one of the most mature renewable energy systems, which have shown a sharp increase over the past decades, they still face various failures at the system and component levels. The significant failures and unexpected events in large PV power plants can be classified as geological instability issues such as severe weather events, string fuses failures, overvoltage, inverter failures, substation over temperature, low DC insulation, MV/LV transformer failures and module failures[1].

Due to these failures and unexpected events, minimizing operation and maintenance costs and maximizing the system reliability and safety of large PV power plants remain major concerns for investors and grid and power plant operators[2].

Moreover, due to unstandardized operation and maintenance (O&M) practices and approaches, companies develop various proprietary strategies, which eventually may cause a decrease in revenue and an increase in project cost, as well as the efficiency of the PV power plant.

As a result, PV power plant maintenance strategies are gaining interest among companies and research institutions in order to develop a best practice for PV power plant generation processes[3]. The maintenance procedures of large PV power plants include the following types:

Preventive Maintenance: This type of maintenance is scheduled by operators or PV power plant designers by considering environmental conditions, type, and warranty terms of the components.

Corrective Maintenance: Usually, it refers to replacing the failed components to repair the damage. It is advantageous for the PV power plant to have a remote reset feature to minimize the damaged component effect on other components during the corrective maintenance event. If the component requiring corrective maintenance is not so vital for the process, this kind of maintenance can be combined with scheduled maintenance tasks.

Condition-Based or Predictive Maintenance: This type of maintenance activity can be achieved by using the real-time data or historical data logged in the data logger or historian and comparing them with usual process data or design specifications to predict failures or identify them at earlier chapters before becoming vital to PV power plant process. This kind of maintenance is essential to prevent unexpected failures, which may cause penalties from grid operators (for on-grid systems) or revenue loss.

As it is explained above, maintenance activities can be planned or scheduled according to various parameters, such as checking the actual energy generation and comparing it with the design specifications or noticing the sudden failures via real-time data. For a PV power plant operator to be able to schedule the maintenance task, there is a need for an advanced data acquisition, monitoring, and control system.

This study aims to research available SCADA systems in the market for large PV power plants, compare them in technical aspects, and understand the effect on predictive maintenance strategies. To achieve this aim, available SCADA components, topologies, communication protocols, and data types are being defined and compared in different aspects in order to conclude their adequacies, inadequacies and effects on the predictive maintenance strategies. These multifaceted comparisons were used to design the SCADA infrastructure for a hybrid renewable power plant project in Burkina Faso as a case study to improve the effectiveness of predictive maintenance activities.

2 SCADA Systems Evolution and Their Functionalities in Large PV Power Plants

2.1 SCADA System Evolution

All initial SCADA systems were based on the simple wired topologies at the very beginning of their existence. However, due to major technological developments and rising concerns regarding the increased operational and maintenance costs, more advanced wireless topologies have been integrated to SCADA. As a consequence of this situation, SCADA systems had a continuous development[4].

Integrating the old SCADA topologies and components with the new technology component and topologies brought more vulnerability to the systems in terms of security. Multiple monitoring schemes and changing data computation areas from physical devices to Web or cloud-based systems consequently reduce the system security.

Nevertheless, the decision on the most proper SCADA system infrastructure is highly dependent on the use case, critical system parameters, and technical and financial barriers. Since energy security is highly critical for countries and companies, designing the necessary SACADA infrastructure requires the utmost attention and understanding of the components and topologies.

The following sections examine the SCADA system evaluation under four categories: classical wire-based SCADA systems, distributed SCADA systems, network-based SCADA systems and IoT-cloud-based SCADA systems.

2.1.1 Classical Wired-Based SCADA Systems

These centralized SCADA systems are standalone control and monitoring structures. Classical wired-based monolithic SCADA systems were not utilizing open-source software but producer-specific software, which increased dependency on the producer company. Although the communication between the remote terminal units (RTUs) and master terminal unit (MTU) was carried out by the wires, wide area network protocol used for the communication between the RTUs[5]. Figure 1 illustrates the representative architecture of wire-based classical SCADA systems.

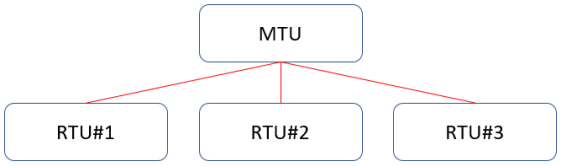

Figure 1: General architecture of wire-based classical SCADA system

2.1.2 Distributed SCADA Systems

The primary benefit distributed SCADA systems brought, compared to classical wired-based SCADA systems, was the first utilization of the local area network (LAN) protocols. However, the major drawback of earlier SCADA generations, producer-specific protocols and software, remains the same in this generation of SCADA systems as well. Although this system increased the redundancy and reduced the cost of the system, the operation and maintenance costs due to proprietary components still were relatively high [6]. Figure 2 shows a representative architecture of distributed SCADA systems.

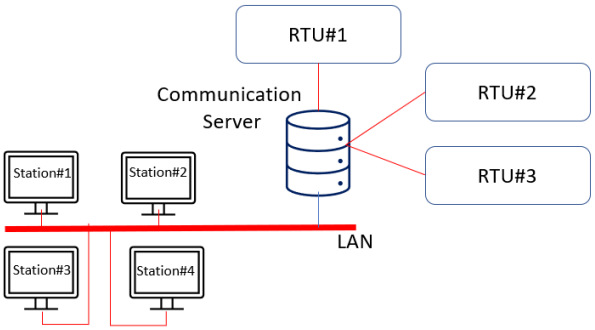

Figure 2: General architecture of distributed SCADA system

2.1.3 Network-based SCADA Systems

Compared to the first- and second-generation SCADA systems, network-based SCADA systems had a game-changer functionality, utilising open-source software and standard protocols. This functionality reduced the initial investment, operation, and maintenance costs significantly. Moreover, being independent of the producer, increased flexibility and redundancy and facilitated the future expansion of the system. Due to the fact that many SCADA components are reachable from the Internet, the vulnerability of the system to cyber-attacks has increased [7]. Figure 3 depicts the basic structure of network-based SCADA systems.

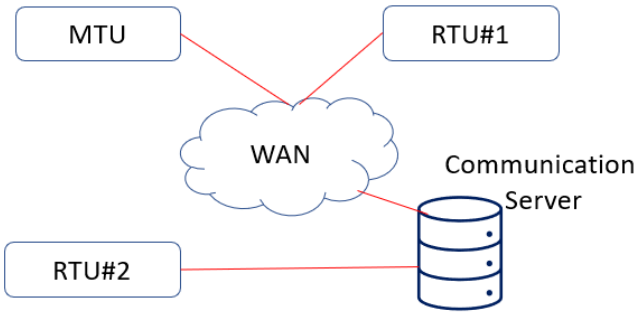

Figure 3: General architecture of network-based SCADA system

2.1.4 IoT-Cloud Based SCADA Systems

IoT-cloud-based SCADA systems are widely used nowadays. The main reason for this popularity is that IoT-cloud-based SCADA systems significantly facilitate the integration and maintenance of the SCADA system. All the data storage and computation are being executed in the cloud system. Compared to the classical SCADA systems, IoT-cloud-based SCADA systems are more specific targets for cyber-attacks [4]. Figure 4 depicts the basic structure of IoT-cloud based SCADA systems. Nevertheless, considering the flexibility, redundancy, accessibility, and scalability they bring to the system and the studies in the cyber security area, these systems will find more place for themselves both in the industry and energy areas.

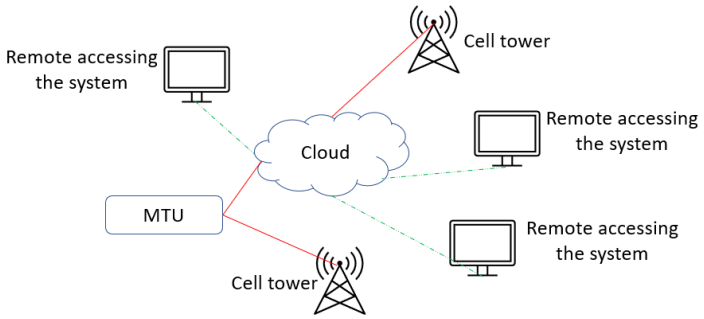

Figure 4: General architecture of IoT-cloud based SCADA system

2.2 SCADA Functionality in Large PV Power Plant

SCADA is a system of data obtaining and interaction process by linking hardware and software components to enable the SCADA operator to monitor, acquire real-time data and control the operation through a human-machine interface (HMI) and store the data in a log file. This outstanding functionality brings many benefits to PV power plant operations, such as enabling the operator to analyse PV power plant performance and to understand, if it meets the expectation, to define the malfunctioning components and even control the power plant generation output to minimize the impact on the grid.

By analysing and comparing real-time data with the PV power plant design data, which usually includes nominal operation conditions of the PV power plant components, maintenance periods can be defined earlier to prevent unnecessary system shutdowns and loss of money.

Large PV power plants contain various components located on an enormously large area. In order to achieve necessary communication between these different components, one or a combination of a few advanced communication technologies is needed. As SCADA has been flexible in terms of communication protocols and used in disturbed systems such as electrical transmission and distribution systems for decades, it may provide certain techno-economic advantages to large power plants [8].

3 SCADA Components in Large PV Power Plants

SCADA systems in large PV power plants have different components located in a wide area. Each of these components fulfils the requirements of different SCADA functionalities, namely controlling, monitoring, data collection and storage. Generally, SCADA system components can be classified under three major subsystems: hardware components, software components, and PV power plant field control and measurement components.

3.1 Hardware Components

Hardware components are the physical elements of the PV power plant SCADA systems and can be categorised as master station hardware, field control units (remote terminal units) and field devices based on their location and functionalities [9].

3.1.1 Master Terminal Units

Master terminal units (MTU) or SCADA server can simply be defined as central unit or central computer of the SCADA system. Depending on the PV power plant size, they can be located in one or more places [10] [11]. The major functionalities of the MTUs are:

- Commanding to remote terminal units (RTUs), programmable logic controllers (PLCs) and intelligent electronic devices (IEDs),
- Requiring and collecting data from RTUs/PLCs/IEDs,
- Processing the acquired data,
- Storing the acquired data,
- Displaying the received data in the form of tables, graphs, illustrations, pictures, and curves via human-machine interface to the operator.

The bidirectional communication between MTU and RTUs/PLCs/IEDs is carried out by the MTU programs, which can be either triggered manually by the system operator or automatically defined earlier. In general, most of the instructions are automatically triggered. As MTU is the master in association with RTUs/PLCs/IEDs, whenever MTU require the data, the slave units react immediately to provide it. Figure 5 illustrates the typical inputs and outputs of the MTUs.

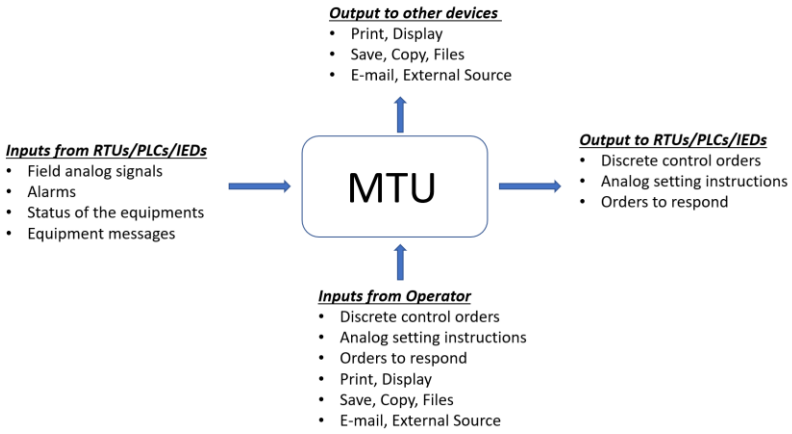

Figure 5: MTUs' inputs and outputs from and to other devices

MTUs enable one or more operators to follow the process, set the inputs and change the existing system settings simultaneously.

3.1.2 Remote Terminal Unit

A remote Terminal Unit (RTU) transmits and controls the data collected in a remote field to a central SCADA system or central station using different communication protocols. The most crucial feature of RTUs is that they can work as master and slave at the same time in master-slave architecture. [12]

The Digital and analogue input/output (I/O) capacity of RTUs are expandable. This feature enables the PV power plant to integrate the future extensions to their SCADA systems without changing the whole existing infrastructure.

It is possible to create logic diagrams and mathematical calculations in RTUs. The received data can be processed through these logic diagrams and mathematical calculations to create output signals to control the field devices and send data to the central station. Figure 6 shows the typical inputs and outputs of the RTUs.

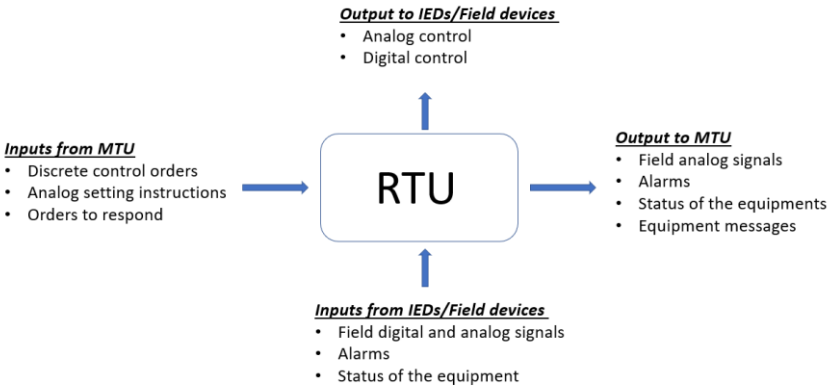

Figure 6: RTUs' inputs and outputs from and to other devices

3.1.3 Programmable Logic Controller

Programmable Logic Controllers are microcomputers which can collect data, send data to a central station, process the received data via logic diagrams and mathematical calculations and generate outputs for field devices. Unlike conventional computers, they are designed with multiple inputs and outputs.

3.1.3.1 Comparison of RTUs and PLCs

In the industry, the usage of RTUs and PLCs has slowly gotten on the same pathway in the last decades. Although PLCs have similar functionalities to RTUs, they have significant differences, which may affect the decision-making during the SCADA design of the large PV power plants.

In practice, PLCs are more suitable for local controls because of communication protocol limitations. Due to communication protocol flexibility, RTUs have been used more for the purpose of remote control tasks. [13]

Table 1 shows the comparison of the PLCs and RTUs regarding the vital decision parameters.

Table 1: PLCs and RTUs comparison table

Parameters of Comparison	PLCs	RTUs
I/O Sizing	In general, less I/O numbers per module	High I/O number
SCADA requirement	May function also without SCADA	Cannot function without SCADA connection
Control area	Local control area	Remote control area, wide geographical telemetry
Operation type	Cyclical	Event driven
Data transmission	All data is being transmitted as per programming (Possibly also unnecessary data)	Changes only (Only requested data, very efficient)
Environments and Temperature	Stable environments (Operation temperature usually between -20°C and +60°C)	High tolerance to environment conditions (Operation temperature usually between -40°C and +70°C)
Power requirement	Depends on permeant power supply	Batteries or charging units
Communication protocols	Modbus, DNP3 and DF-1(on serial ports)	Modbus RTU, TCP/IP, Serial ports (RS-232/485), MDLC, IEC60870)
Integrated radio communication	No	Yes
Remote changes	No	Yes

3.1.4 Intelligent Electronic Devices

Intelligent Electronic Devices (IEDs) can be simply defined as intelligent sensors or actuators. They are capable to carry out the data acquisition, communication, control, and process tasks. Their deployment in the SCADA system of the large PV power plants is usually in the substations, which contain digital meters and protection relays [11].

3.2 Software Components

All supervisory and human machine interfaces, including workstations, are part of the SCADA software components. SCADA software components can be classified under two sub-groups, which are proprietary and open software[9].

Proprietary Software: They are patented software and developed by individual companies to communicate with their own hardware. The main drawback of this kind of software is that it increases the dependency on the manufacturer[9].

Open Software: This type of software is supported by multiple companies. They became popular due to the flexibility that they bring to the system [9].

3.2.1 Historian

All the data collected by MTUs is stored or logged in the historian. This allows the operator to examine the historical data and compare it with the real-time data when necessary. The collected data can either be stored on-site in the local hard drives or off-site on internet-based servers. [14]

Each stored data has a tag (data ID), value and time stamp. These data are categorized according to context, specific events, alarms, notifications, dates, etc. Here data can be transformed into meaningful information by basic calculations.

3.2.2 Human-Machine Interface

Human-machine interface (HMI) is a combination of hardware and software. It enables humans or operators to visually monitor the received data from the controlled process, make necessary adjustments on the system settings and fast respond to the emergency.

The HMI can monitor real-time and historical data, alarms, equipment messages, the status of the equipment and process, and other required information. Moreover, the operator may configure set points and control algorithms to adapt the system to internal and external changes. [11]

3.3 PV Power Plant Field Control and Measurement Components

PV power plant field control and measurement components are the first groups of elements in the SCADA hierarchy which provide necessary measurement data and react to the control signals received from MTUs and RTUs/PLCs/IEDs. These components consist of string boxes, solar trackers, weather stations and various field sensors such as temperature sensors, wind sensors or anemometers.

3.3.1 String Box

In large PV power plants, PV strings are combined by string boxes which are then connected to power conditioning systems or, in other names, inverters. Almost 2% of the PV modules create failures in the first five years of their installation [15]. Theoretically, these failures can be identified by examining the output generation change of inverters. However, considering the number of modules connected per inverter in large PV power plants, these changes in the energy generation can be very little and cannot be differed from the weather effect on the generation. As PV modules do not have electronic components to provide individual data, the data generated by the string boxes are essential for the PV power plant operator to detect damaged PV modules.

3.3.2 Solar Tracker

A solar tracker is a 1- or 2- motorised axis system that positions the PV modules to receive maximum irradiation from the Sun. By developing interfaces between solar trackers, HMIs and central stations, solar trackers can be automatically controlled via the SCADA system or manually positioned by the PV power plant operator. [16]

3.3.3 Weather Station

The weather station is a combination of data loggers and sensor units which sense the weather information to convert meaningful data to be displayed by the SCADA system [17].

Depending on the size of the PV power plant, the number of weather stations may be more than one, in order to monitor the entire PV power plant. Weather stations may or may not have wind sensors or anemometers, which provide necessary data for control and safety for the trackers to be protected against strong wind or storm. Weather stations are capable of providing the following data to the SCADA system:

- Solar irradiation,
- Air pressure,
- Humidity,
- Ambient temperature,
- Speed and direction of the wind.

4 SCADA Communication Protocols in Large PV Power Plants

Large PV power plants consist of various components with different functionalities from different manufacturers. The success of the SCADA system is highly dependent on the smooth communication between these components.

In order to achieve this smooth communication, SCADA systems have their own standardized communication protocols. The communication protocols regulate the bidirectional data exchanges through a communication link [18]. In other words, a PV power plant component must speak the same language or use the same communication protocol as other elements of PV power plants to be understood. Without these protocols, the data will remain coarse data that is not structured and safe. All the standardized protocols are based on the ISO (International Standard Organization) seven-layer ‘’open system interconnection (OSI)’’ or, ‘’transport control protocol/internet protocol (TCP/IP)’’ or ‘’Enhanced Performance Architecture (EPA)’’ model [19].

The main target of the OSI model is to create a framework to enable all network and system to be connected and be able transfers data to each other. The model offers independence to the manufacturer company in terms of communication. The basic structure of modes is shown in Table 2.

Table 2: Open System Interconnection (OSI) layer by ISO

OSI Layer name	OSI Layer level
Application	7
Presentation	6
Session	5
Transport	4
Network	3
Data Link	2
Physical	1

Layer 7 is the "application layer" where network service is provided to the end user, such as the PV power plant operator or grid operator. Layer 6, or the "presentation layer", is where syntax tasks are carried out. At this layer, data are being converted from one form to another in order to be readable by the end user. Controlling the ports and maintaining necessary connections are being executed at layer 5. Layer 4 is responsible for coordinating the amount and speed of the data that need to be sent through network connections. The data is being directed at layer 3. Layer 2 is the most complex layer. When this layer receives the data, it checks the errors and then divides the bits into data frames. Layer 1 is the "physical layer" which consists of network cables, wireless radio frequencies, connectors, etc.

In earlier 1970s, Advanced Research Projects Agency Network developed TCP/IP [20]. While the OSI model was designed for the new communication protocols to be built on, the TCP/IP was created on the existing communication protocols. However, the creator of the model keeps adopting the model to the industrial requirements, which is the reason that TCP/IP is still widely used model [21]. The other major

difference between TCP/IP and the OSI model is that it has only four layers, as shown in Table 3.

Table 3: Layers of TCP/IP model

TCP/IP Layer name	TCP/IP Layer level
Application	4
Transport	3
Internet	2
Network Interface	1

In this model, layer 4 enables the end-user or operator to visualize the received data the same as with OSI layer 7. However, the TCP/IP model does not require the presentation and session layers to carry out application tasks. Again, similar to OSI model, transportation layer is responsible for data flow and reliability. Layer 3 regulates the data-sending tasks. Moreover, all necessary IP addresses are assigned for the data's final destination. The network interface layer is a physical layer, and its primary duty is to enable the data exchange between the SCADA components, which are located on the same network. Ethernet protocol is being utilized at this level in order to identify how data must be transmitted. Each network components have its own unique MAC address, which enables ethernet protocol to ensure data has been sent to the correct destination.

Another network model on which communication protocols have been based on is Enhanced Performance Architecture (EPA) model. The main advantage of this model is that it has less additional load compared to OSI and TCP/IP models and only uses direct communication links [22]. This model only uses three layers of the OSI model, as it is depicted in Table 4.

Table 4: Layers of EPA model

EPA Layer name	EPA Layer level
Application	3
Datalink	2
Physical	1

There are many SCADA communication protocols. This paper will explain the communication protocols which are usually being used in large PV power plants. The SCADA communication protocols of large PV power plants can be classified under two major headings, namely wired and wireless communication protocols.

4.1 Wired SCADA Communication Protocols in Large PV Power Plants

Wired communication is a system where data exchange is executed through wire-based technologies such as fibre optic cables, ethernet cables, copper cables, coaxial cables, etc. There are many standardized SCADA communication protocols used in PV power plants which utilize wired communication technologies, such as IEC 61850, IEC 60870-5 and DNP3. Apart from these standardized SCADA communication protocols, there are other vendor-specific communication

protocols, such as Modbus RTU and Profibus. Although Modbus RTU and Profibus are vendor-specific, they have been widely recognized. There are other vendor-specific protocols. However, since they are not usually used by PV power plants due to their high vendor dependency, this paper has yet to present their functionalities.

4.1.1 IEC 61850

Usually, generated energy from large PV power plants is transmitted to load nodes at medium or high voltage levels in order to decrease the power losses during the distribution or transmission. However, the power generation source of PV power plants is the PV modules, and the generated power is in low voltage and direct current (DC) form. This DC power can neither be utilized by traditional electrical devices nor transmitted via existing AC transmission lines. Therefore, with the help of inverters, this DC power is converted to AC power. However, this converted power again has a low voltage form which is usually less than 1kV. In order to regulate the low voltage to medium or high voltage, electrical substations are necessary. Electrical substations in large PV power plants consist of step-up transformers, MV or HV switchgear, circuit breakers, isolators, insulators, surge arresters and metering systems [23]. Moreover, all the energy metering tasks are carried out in the electrical substation. Therefore, PV power plant electrical substation is vital and needs to be monitored and controlled with utmost care.

Generally, PV power plant electrical substations use industrially accepted ethernet-based IEC 61850 standards for communication with other components of the SCADA system [18].

In addition to the define how data will be exchanged, by the virtualized model of logical devices, all the services, data, the way that devices should behave and where to store data can be identified in the IEC 61850 protocol. Every data generated by the components is tagged with a descriptive string which consists of a register number, index number and storage destination in this protocol [24].

Another main future of IEC 61850 is that it uses a standardized language, namely substation configuration language (SCL), as is explained in the IEC 61850 part 6 [25]. Each SCL file has a header which defines SCL configuration, a description of the substation, an IED description, a description of the communication system and a definition of logical nodes [26].

As it is mentioned in the earlier chapters, open and standardized communication protocols bring the benefit of flexibility to the operator and system for the future needs of system extensions and reduce the dependency on a specific vendor. Since IEC 61850 provides predetermined data and device names, PV power plant operators can understand what received data means without checking defined maps.

Other key advantages of the IEC 61850 communication protocol are as follows [27]:

- Procurement uncertainty can be eliminated by using SCL, which enables the user to define exact expectations from each device.

- IEC 61850 enable the devices to transmit data over the same physical network, which reduces the additional cost and consequently decreases the wiring costs.
- A single transducer backing sampled measured values (SMV) can transmit numerous signals to multiple instruments. Therefore, there is no need signal specific transducers.
- Due to standardized data type and device addresses, IEC 61850 almost does not need manual configuration.

Even though IEC 61850 was initially developed for substation applications, it is a very comprehensive communication protocol at the moment. Many aspects of the implementation of the IEC 61850 to PV power plants, such as data modelling for the inverters, are already identified in IEC 61850-7-420 [28]. However, there is still time to identify IEC 61850 as an entirely suitable communication protocol for all PV power plant processes [27].

4.1.2 IEC 60870-5

IEC 60870-5 is a group of standards developed for supervisory control and data acquisition by the International Electrotechnical Commission. This communication protocol is based on the three-layer EPA model [29]. The protocol was initially developed for electrical power systems and has been widely used in European countries.

This paper will focus on the two sub-group communication protocols of the IEC 60870-5, namely IEC 60870-5-101 and IEC 60870-5-104, which are being widely used in solar applications.

4.1.2.1 IEC 60870-5-101

IEC 60870-5-101, or by another name, IEC101, is a sub-branch communication protocol of IEC 60870-5 communication protocol which standardizes the bi-directional telecontrol among the devices. This protocol supports multiple SCADA topologies such as point-to-point, multiple point-to-point and star.

In this protocol, information objects are created by categorized data, and an individual address is assigned to each information object. It supports two different data transmission processes, namely unbalanced and balanced transmission. In an unbalanced system, data is transmitted from controlled stations to the controller as an answer to a request from the control station [30]. Figure 7 represents a simple topology of unbalanced data transmission prosses.

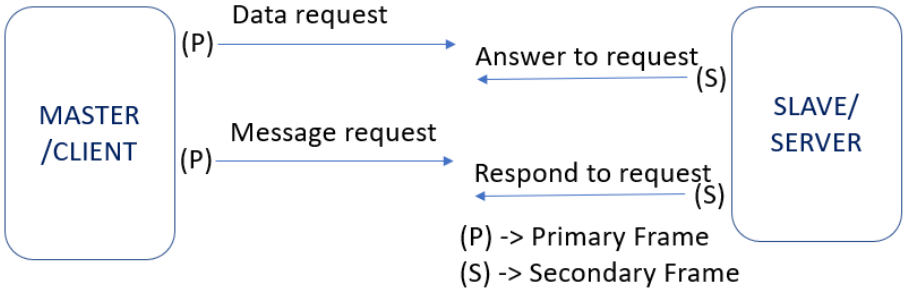

Figure 7: Unbalanced data transmission process

However, in a balanced transmission process, all stations can create and transmit the message, which decreases the transmission frequency. Figure 8 is an illustration of the simple structure of the balanced data transmission process. This type of transmission can be only implied to point-to-point and multiple point-to-point topologies [30].

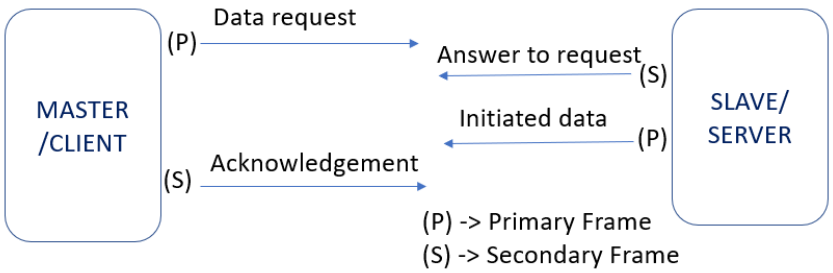

Figure 8: Balanced data transmission process

The acquired data can be categorized under different groups critical (high priority) and non-critical (low priority). This enables the system operator to monitor and control the vital components of the PV power plant more accurately.

Data transmission is only possible with a serial link between the devices. While this feature brings more robustness to the system, the need for point-to-point connection brings additional wiring costs. However, as it is specified in IEC 60870-5, this serial communication can be converted to TCP/IP communication via IEC 60870-5-104. The process of converting between this subgroup of IEC 60870-5 is bi-directional [31]. With this unique functionality, the system may be more flexible in terms of data communication.

4.1.2.2 IEC 60870-5-104

IEC 60870-5-104 is an addition to IEC 60870-5-101 in order to extend the network access capacity of IEC 60870-5-101 by combining all the functionalities and enabling it to run over TCP/IP. The primary purpose was to extend the functionality of this protocol beyond the substation level [29]. Thoracically, IEC 60870-5-104 can support both balanced and unbalanced data transmission. However, since TCP usually creates point-to-point connections, the most used case for this protocol is the balanced mode.

The controlled stations or slave devices in the master-slave hierarchy transmit different data, such as changes in the trend, regular cyclical data transmission, answers to commands and read requests. The operator or automized control centre can require real-time data from the controlled centres at any time.

IEC 60870-5-104 has the ability to generate and transmit data with their timestamp, which allows the operator to recognize when the data has been received, which component has sent the data and what is the data quality. Moreover, thanks to its buffer functionality, IEC 60870-5-104 can store the data in the event of communication failures and send the stored data when the communication problem is being solved.

As the primary purpose of development was to extend the functionality of the IEC 60870-5-101 beyond the substation level, it supports many different types of objects. This allows SCADA systems to receive more quality information.

4.1.3 DNP3

Like other standard protocols, the DNP3 communication protocol was also developed for the electrical power industry. However, today it is being used in a wide range of industries like water treatment centres, the oil & gas industry, and renewable energy plants.

It is designed with the ability to transmit large but fewer data packages. Although the initial version of the DNP3 protocol was only used between RTUs and IEDs, today, it supports additional topologies, such as point-to-point and multiple point-to-point topologies [32].

DNP3 and IEC 60870-5-101 were created from the exact origin. Nevertheless, they have gained significant acceptance on different sides of the world. While IEC 60870-5-101 is usually preferred in European countries, especially in the electrical power industry, DNP3 was the dominant protocol in North America in different industries. DNP3 is being supported by the DNP3 Users Group and Technical committee actively. Thus, DNP3 will be adopted for future advancement of system requirements [33].

Like the IEC 60870-5 protocol, the DNP3 protocol is also based on the EPA model and can categorize the data according to their priority. This protocol supports a balanced data transmission method which means controlled centres can transmit their messages without request. Multiple different data types can be transmitted in a single message [34].

4.1.4 Modbus

Modbus serial open communication protocol, located at the OSI model's application level, was initially designed for the PLCs [35]. However, it has been almost an industry-standard protocol for transmitting data between control and industrial electronic devices nowadays [32].

Modbus-supported devices may transmit data only in a master-slave hierarchy and on an unbalanced data transmission process where data is transmitted from slave to master as per the request of the master. Master devices have the ability to identify the individual slave device data to be transmitted or can generate a message for all the devices [36].

As mentioned above, Modbus is an application protocol, and it is nothing but a simple messaging structure which identifies the rules for data structure and data interpretation. This reduces the implantation time.

There are three types of Modbus protocols being used in the industry, namely, Modbus ASCII, Modbus RTU and Modbus TCP/IP. For a better understanding of these different Modbus functionalities, serial communication protocols which are used by Modbus ASCII and RTU should be explained.

In general, standard protocols use either TCP/IP or serial communication protocols to be connected to the SCADA systems. This connection can be physical, which is being utilized by serial communication

protocols or over ethernet (TCP/IP). There are two major physical and electrical connections, namely RS-232 and RS-485.

RS-232 is developed to organize data exchange between the transmitter and receiver devices. It is still widely used in SCADA devices for data-transmitting tasks. The most favourable feature of this protocol is that only by using two wires, full duplex (data sending and receiving at the same time) data transmission is possible. However, the major drawback is that the data exchange distance is only limited to approximately 15-20 meters. The other disadvantage is that data exchange is possible only in a point-to-point topology [37].

RS-485 serial communication protocol was developed to compensate for the disadvantage of the RS-232 in order to meet industry requirements. Data transfer distance is increased to approximately 1200m with this protocol. The limitation of data exchanges only between one driver and receiver in RS-232 was also eliminated with the RS-485 protocol, which enables data transmission between 32 drivers and 32 receivers. Due to this feature, this protocol supports multi-point-to-point topology beside point-to-point topology [38].

4.1.4.1 Modbus ASCII

Modbus ASCII, which stands for American Standard Code for Information Interchange, is one of the two serial protocols of Modbus communication protocol. It is the first developed Modbus, and it can only run on two physical layers, which are RS-232 and RS-485. It can transmit data only on an unbalanced data process, and there is only one master in the system [39].

4.1.4.2 Modbus RTU

The other Modbus, based on serial communication, is Modbus RTU. While Modbus ASCII frames the code with ASCII characters, Modbus RTU uses binary data. Modbus RTU requires almost half the number of bytes than Modbus ASCII in order to transfer the same information. Thus, Modbus RTU is faster in terms of data transmission. It is the most widely used protocol in the industry [40].

4.1.4.3 Modbus TCP/IP

Modbus TCP/IP is the latest developed Modbus protocol. Although it has almost the same abilities as Modbus RTU, the major difference is that Modbus TCP/IP runs over TCP/IP and ethernet in order to carry the data [41]. Same as with other Modbus protocol families, it is open software. It is just only a data transmission protocol; in contrast to IEC protocols, it does not identify data or how to store data.

4.1.4.4 Comparison of Modbus protocols

The Modbus communication protocol family is one of the most matured and widely used communication protocols in industrial and PV power plant SCADA systems. Selection of the most applicable Modbus communication protocol mostly depends on the component requirements and system application requirements.

As Modbus ASCII is an ancient and complex protocol, the usage of this protocol is very limited nowadays. In contrast to Modbus ASCII, Modbus RTU and Modbus TCP/IP are simple, and it is possible to use with all types of PLCs and almost all PV power plant field control and measurement components.

Since Modbus TCP/IP can run over the TCP/IP and Ethernet network protocols and is consequently supported by wired and wireless media, it is more flexible and faster compared to Modbus RTU. Another advantage of Modbus TCP/IP over the Modbus RTU is that while there can be only one master in Modbus RTU protocol, Modbus TCP/IP supports the systems which have multiple masters.

Table 5 presents the different aspects of the functionalities of Modbus communication protocols for a better understanding.

Table 5: Comparison of Modbus communication protocols

Feature	Modbus ASCII	Modbus RTU	Modbus TCP/IP
Transmission efficiency	Low	Medium	High
Data structure	ASCII	Binary	Binary
Data Transmission Process	Unbalanced	Unbalanced	Unbalanced
Master/Client Structure	Only one master/client	Only one master/client	Multiple master/client
Used network	RS-232, RS-485	RS-232, RS-485	TCP/IP & Ethernet
Supported topologies	Star and bus	Star and bus	Star, mash, ring, tree, bus
Supported media	Wired	Wired	Wired / Wireless
Number of components	Less components in the system (Only one master, limited number of slaves)	Less components in the system (Only one master, limited number of slaves)	More components in the system (Only one master, limited number of slaves)
Usage in PV power plant components	N/A	Inverters, trackers, weather stations, string boxes, thermal sensors	Inverters, trackers, weather stations, string boxes, thermal sensors

4.1.5 Profibus

Large power plants and factory process requires more sub-control centres. In these operations, the limited number of masters which can be used in Modbus protocols may not be sufficient. In order to eliminate these limitations, Modbus protocol brought and generated the solution to the needs of the industry in the very late 20. Century Profibus, which stands for Process Field Bus, communication protocol was developed [42].

Profibus communication protocol was created based on the OSI seven-layer model [43]. This feature ensures that different devices from different suppliers can communicate with each other without additional effort. Profibus communication protocol only supports the bus topology, where all master and slave devices are linked to a central bus.

Same as the Modbus communication protocol, Profibus is also based on the master(client)/slave(server) hierarchy. In contrast to Modbus ASCII and Modbus RTU, it supports multiple master structures as well. Same as Modbus, Profibus also run on the RS-485 transmission media. Another difference is that Profibus does not use the exact RS-485 specification, but instead, it uses the specification with additions, such as the allowed number of devices is increased to 126 regardless of if it is master or slave [40].

Table 6 shows the major differences between the Modbus and Profibus protocols which may help the decision-making for the most appropriate communication protocol selection for the system.

Table 6: Comparison of Profibus and Modbus protocols

Feature	Profibus	Modbus
Transmission efficiency	Higher	Lower
Data structure	Binary	Binary
Data transmission process	Unbalanced	Unbalanced
Master/Client structure	Multiple master/client	Mono master/client (Except Modbus TCP/IP)
Used network	RS-485	RS-232, RS-485- TCP/IP
Supported topologies	Bus	Star and bus (Star, mash, ring, tree, bus for Modbus TCP/IP)
Supported media	Wired	Wired (Wired and wireless for Modbus TCP/IP)
Number of components	126 (Regardless of if it is master or client)	32 (Usually 1 master)
Model based on	OSI	EPA
PV power plant applications	Large PV power plants	Small PV power plants

Profibus is more robust compared to the Modbus communication protocol. Therefore, it is utilization with specific PV power plant components such as PV-trackers are more reliable [44].

4.2 Wireless SCADA communication protocols in large PV power plants

Wireless communication is a system where data exchange is executed without wire infrastructure. Compared to wired systems, the usage of wireless systems is very new in SCADA systems. They are mostly utilized by power grid, water, wastewater, oil & gas and telemetry applications.

Although wireless communication technologies have some benefits compared to wired communication protocols, such as decreased installation costs and fewer earthworks, wired technologies remain more advantageous in terms of reliability and stability [45].

Wireless communication technologies have been widely used in factory SCADA processes due to relatively short distances. However, when it comes to critical operations such as power grid management and power generation, there are still some concerns to use with vital components. Nevertheless, with technological development such as 5G, there is no doubt that wireless communication technologies will

be utilized more by SCADA systems.

This section will cover the explanation of different communication technologies and compare them in terms of their applicability to large PV power plants by considering the predictive maintenance activities.

4.2.1 Wireless LAN

Wireless LAN is based on the IEEE 802.11 communication protocol, and it can support point-to-point and multiple point-to-point topologies [46]. Wireless LAN has been used by grid operators in order to monitor and control distributed energy resources such as conventional power plants and renewable power plants and their substations to provide the required generation and consumption balance [47].

As it is mentioned earlier, IEC 61850 communication protocol runs over ethernet and TCP/IP. Due to the fact that Wireless LAN technologies are also considered ethernet-based technologies, the applications which use IEC 61850 communication protocol, such as substations, can get the benefit of Wireless LAN functionalities. Moreover, Wireless LAN can be used to increase the redundancy of data transmission by applying parallel to wired technologies like a fibre optic [48].

The major drawbacks of Wireless LAN technologies are as follows:

- Availability and reliability problems,
- Electro-magnetic disturbances due to High Voltage environment, may affect the data exchange speed,
- Radiofrequency disturbances caused by other wireless components may cause malfunctioning,
- There are not enough number of Wireless LAN components in the market [49].

4.2.2 WiMAX

WiMAX is based on the IEEE 802.16 communication protocol and stands for ‘’Worldwide inter-operability for Microwave access [50]. Currently, many grids benefit from WiMAX applications for power failure detection, remote metering, and real-time energy pricing with different time-dependent tariffs. However, the high cost of the WiMAX tower is still a vital parameter to consider [49]. This technology can be utilized by grid operators for remote PV power plant metering applications or for data transfer to remote O&M centres.

4.2.3 Cellular

Communication via Cellular wireless technology highly depends on cellular service availability. Cellular communication only supports point-to-point topology [51]. Apart from obtaining the information from ethernet and serial protocols, this communication technology can transfer this information to another communication protocol. This functionality makes cellular technology unique, among others. Recent developments in 5G technology may further advance this technology enormously. Considering the high monthly fees, this technology might not be economically feasible for continuous data transfer in large PV power plants.

Following wireless technologies have relatively short coverage areas for a large PV power plant, or they do not have wide usage in SCADA systems yet like Digital microwave technology:

- Mobile Broadband Wireless Access (MBWA)(Based on IEE 802.20 communication standard)
- Digital Microwave Technology
- Bluetooth
- ZigBee (Based on IEEE 802.15.14 communication standard)

Table 7 shows the comparison of wireless SCACA technologies and their possible application in large PV power plants.

Table 7: Comparison of wireless SCADA technologies in large PV power plants

Wireless Communication Technology	Rate of Data	Coverage area (Approximately)	Potential PV power plant Application	Limitation
Wireless LAN	Range between 1 and 54Mbps	Up to 100m	Increase the redundancy of wired based Ethernet and TCP/IP protocol	Low reliability and availability Electro-magnetic disturbance Radio frequency disturbance Not widespread
WiMAX	Up to 75Mbps	Up to 48km	PV power plant metering Data transfer to remote O&M centres	Expensive WiMAX tower Probably require a license
Cellular	Range between 60 and 240Kbps	Range between 10 and 50km	Data transfer to remote O&M centres	Expensive for continues data transfer
Bluetooth	Up to 720Kbps	Range between 1 and 100m	Local applications such data transfer from field sensors to RTU	Low data rate and short coverage area
Digital Microwave	Up to 155Kbps	Up to 60km	Data transfer to remote O&M centres	Low data rate
ZigBee	Range between 20 and 250Kbps	Range between 1 and 100m	Local applications such data transfer from field sensors to RTU	Low data rate and short coverage area

4.3 Comparison of SCADA Communication Protocols for Large PV Power Plants

Considering the availability, reliability, and low data rate features of the wireless SCADA communication technologies, they might need to be more favourable for critical operations and operations where real-time data is the priority. However, as wireless communication topologies are a very favourite topic of research and development studies and especially development in the area of 5G communication technologies, they might change SCADA communication infrastructure design in the near future.

Profibus and Modbus are serial communication protocols. Therefore, a SCADA communication design of the whole PV power plant cannot be based on either of them. Moreover, both communication protocols have been used by factories and process automation applications for years. They have matured protocols for these applications. Their utilization by power applications is very limited compared to factory and process automation applications. Nevertheless, as they are elementary and open protocols, they still can find a place for themselves in large PV power plant applications, such as for the communication between RTU and inverter.

IEC 61850 is the most matured and industrially accepted communication protocol for substations. Due to the advantages, such as eliminating the procurement uncertainties, wire cost reduction, almost no need for manual configuration and decreased number of transducers, that IEC 61850 brings to investors and designers, it is almost a monopoly of substation communication protocol. Therefore, they are being used in the substation of large PV power plants worldwide.

IEC 60870-5-101, IEC 60870-5-104 and DNP3 are based on the initial IEC 870-5 document, and they have similarities. However, in detail, they have significant differences which affect the appropriate communication protocol selection. IEC 60870-5-101 is a serial-based communication protocol, while IEC 60870-5-104 and DNP3 are TCP/IP-based protocols.

Both DNP3 and IEC 60870-5-104 have advantageous features such as timestamping, time synchronization and user-definable objects for O&M operations of PV power plants. These functionalities allow the operators to identify the error location, the time of the event and data quality. DNP3 is a multi-purpose communication protocol being utilized widely, especially in North America, by the water, wastewater, and oil & gas industry. IEC 60870-5-104 is more specific to energy projects, and it is the dominant communication protocol in Europe for energy projects. Moreover, IEC 60870-5-104 offer more user-definable objects, which may increase the data quality.

5 SCADA Communication Topologies in Large PV Power Plants

SCADA communication topologies are basically network topologies which enable physical and logical SCADA components to link with each other. OSI model has been introduced in the earlier section of this study. While the physical topology of the system is defined by the first layer of the OSI model, logical topology is defined by the data link layer [52].

There are three main SCADA communication topologies, namely, point-to-point, point-to-multi-point, and point-to-point-to-point.

Point-to-point or peer-to-peer is the simplest SCADA system communication structure. The information is transferred between only two points or stations. Nevertheless, this structure increases the total cost of infrastructure due need for individual wiring or communication channels between each point. Depending on the communication protocol and the SCADA hierarchy structure, data can be transmitted simultaneously or regularly. Moreover, both devices can be master or one master and one slave [53]. Figure 9 shows a simple structure of point-to-point topology.

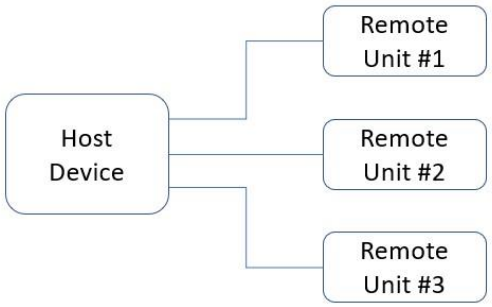

Figure 9: Point-to-Point topology

Point-to-multi-point or multi-point-to-point is a more complex structure than point-to-point topology due to the increased number of devices. Usually, there is at least one master in this type of topology. The major handicap of this structure is the bandwidth, and all the slaves receive the same massage generated by the master unit [53]. Figure 10 illustrates a simple structure of point-to-multi-point topology.

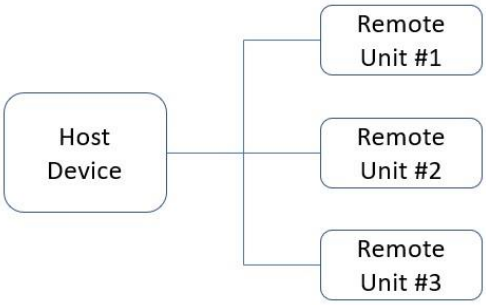

Figure 10: Point-to-Multi-Point topology

The other basic topology is point-to-point-to-point, where data is transmitted from field units to the central unit or vice versa through individual communication channels. While this type of communication brings more reliability and redundancy to the system, it increases the price [53]. Figure 11 depicts a simple structure of point-to-point-to-point topology.

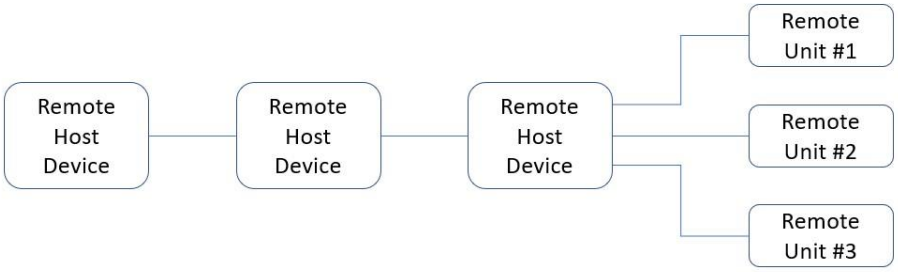

Figure 11: Point-to-Point-to-Point topology

However, in order to create the necessary flexibility, redundancy and robustness for complex systems such as large PV power plants, SCADA systems may utilize the combination of these protocols. Consequently, there are other SCADA topologies with system-specific advantages and disadvantages. The following chapter introduces these topologies and compares them with different aspects of their usage in large PV power plants.

As SCADA systems are very vital components of PV power plants, and they are usually the target of cyber-attacks, the information related to their topologies is not available to the public. In order to visualize the topologies for a better understanding of the topic, the following topologies are generated as representative SCADA structures. To make the created topologies as real scenarios, various PV power plant components like PV strings, string boxes, central inverters, substations, RTUs, PLCs, MTUs, weather stations, historians, SCADA centres, and a remote centre have been used in the topologies. Generally, a complex system can only be structured by a combination of topologies. Therefore, red connection lines have been used to show the explained topology, and blue has been used for only point-to-point structure.

5.1 SCADA Mesh Network Topology

SCADA Mash network topology refers to where RTUs/PLCs/IEDs have fully or partially connected to each other and to MTU. Although this type of topology brings a high redundancy and robustness to the system, it increases the investment, installation, commissioning, operation, and maintenance costs. The system needs to be more flexible in terms of future extensions due to high cabling costs [54]. However, partially meshed topologies are being preferred by large PV power plants to increase the redundancy at RTU or control level in order to increase data reliability and fault identification.

Figure 12 depicts a representative partially meshed SCADA network topology. By creating an additional physical layer between RTUs, the data exchange among the RTUs will not need to be carried out through MTU. The same logic can be applied to the control side of the SCADA topology where the historian,

weather station, SCADA centre and remote centre are connected to each other to MTU. Generally, this is not a structure preferred by large PV power plants.

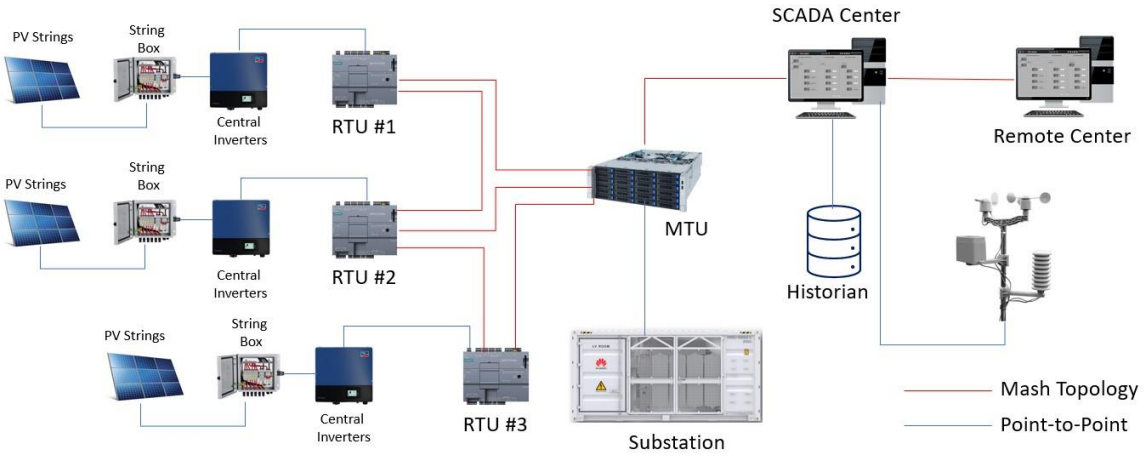

Figure 12: Representative partially meshed SCADA network topology

5.2 SCADA Star Network Topology

In SCADA star network topology, each component of the PV power plant SCADA system is connected to MTU with a point-to-point connection. All the data exchange traffic goes through the MTU, and this makes star topology very poor in terms of redundancy and robustness [55]. As each device only require one cable to transmit the data, the cable infrastructure cost is low. Less effort for fault detection and future extensions are two main advantages of this topology.

IEC 608-70-5-101, IEC 608-70-5-104, and DNP3 standard protocols are capable of running over this topology [56]. By an additional repeater between MTU and MTU-connected devices, the redundancy of this topology can be increased. This addition makes star topology favourable for large PV power plant applications. Figure 13 illustrates a representative SCADA star network topology at the control and RTU levels.

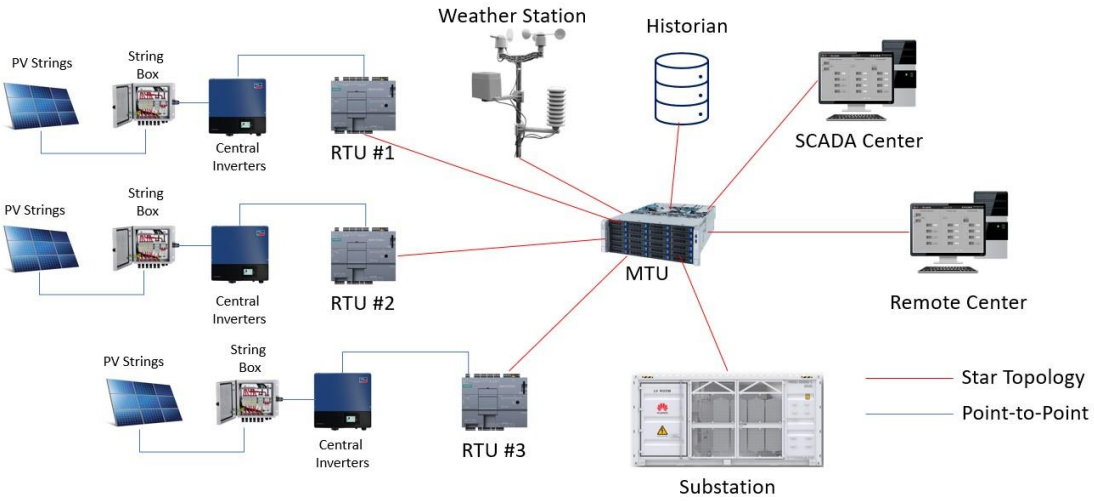

Figure 13: Representative SCADA star network topology

5.3 SCADA Ring Network Topology

Ring network topology has been one of the most popular network strategies in traditional SCADA design due to it is the cheapest way to increase redundancy. Usually, in large PV power plants, the fibre network is divided into loops; the number of loops is highly dependent on the size of the plant, the shape of the site, and how critical the operation is. IP protocols such as TCP/IP can be efficiently utilized through ring topology. Thus, standard SCADA protocols, which are IEC61850, IEC 608-70-5-101, IEC 608-70-5-104 and DNP3, can be used in this topology.

SCADA ring network topology has low Installation operation and maintenance costs. However, designing is a complex process due to the need for the arrangement of IP addresses, configuration, and data object classifications [57]. Figure 14 shows a representative SCADA ring network topology at the RTU level.

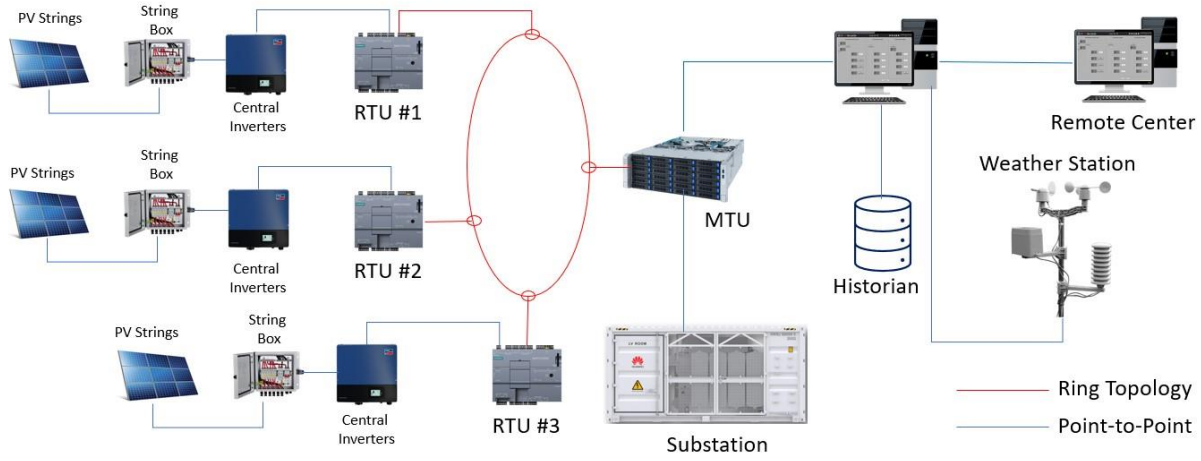

Figure 14: Representative SCADA ring network topology

5.4 SCADA Bus Network Topology

SCADA Bus network topology was the most famous before the switches were applied to the SCADA system. This is the simplest and cheapest way to enable the system to exchange data. It is still being utilized, especially for temporary work. However, lack of redundancy and robustness make bus topology an unfavourable SCADA network topology for large PV power plants. Nevertheless, this topology with additional switches can still be preferred only at the individual level of SCADA design, such as at the information bus, control bus or field bus. Figure 15 shows a possible PV power plant network topology where the information bus is designed with a bus network topology.

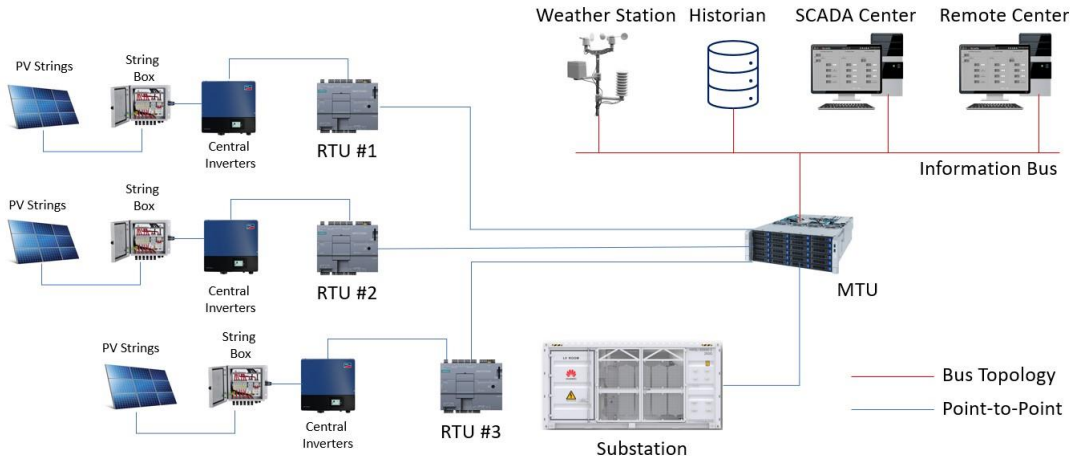

Figure 15: Representative SCADA bus network topology

5.5 SCADA Tree Network Topology

SCADA tree network topology is a hierarchical structure which combines star and bus topologies. Therefore, it brings the advantages and disadvantages of these topologies to the system. Due to it is hierarchical and straightforward design, it is applicable to SCADA applications. Moreover, the system may benefit from an independent point-to-point structure for future expansions and localize the failure not to affect the rest of the system. However, any failure in the bus may affect the whole system [58].

As all standard SCADA communication protocols can run over this topology and with its topology-specific advantages, bus topology is widely being used by PV power plants. However, this might only be favourable if additional redundancy precautions, such as repeaters at the backbone busses, are applied. Figure 16 illustrates a representative SCADA tree network topology where PV field control and measurement devices are linked to remote input and output (RIO) units. RIOs transfer the data through PLCS/IEDs, RTUs, and MTU and to the final destination, which is the SCADA centre or remote SCADA centre.

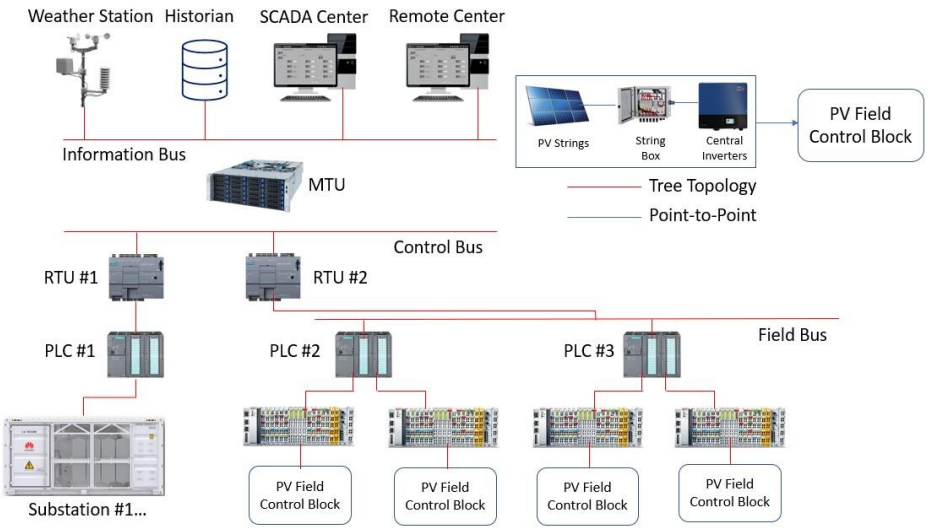

Figure 16: Representative SCADA tree network topology

5.6 Comparison of SCADA Network Topologies in Large PV Power Plants

Decision-making of the suitable network topology for PV power plant SCADA system depends on many parameters. The use case of the PV power plant, like grid support, off-grid power supply, backup power, and hydrogen generation, is the main base for decision-making.

From a technical perspective, as mesh topology is the most reliable, robust, and redundant structure, it can be preferred for the SCADA network. However, due to intense data traffic jams, this might not be favourable for grid support applications where real-time data is vital.

From an economic perspective, bus and tree topologies might be suitable due to their simplicity and, consequently, their low price. However, very low redundancy and robustness might bring a considerable risk to the PV power plant.

Star and ring topologies are the most techno-economic solution for effectiveness improvement of predictive maintenance strategies in large PV power plants. Considering the robustness, flexibility and high data transmission period that star topology may bring and the redundancy of the ring topology, PV power plants may utilize one or a combination of these topologies.

Table 8: Comparison of SCADA network topologies

	SCADA Network Topologies				
Parameter	**Mesh**	**Star**	**Ring**	**Bus**	**Tree**
Supported Standard Communication Protocols	IEC61850, IEC 608-70-5-101, IEC 608-70-5-104, DNP3	IEC61850, IEC 608-70-5-101, IEC 608-70-5-104, DNP3	IEC61850, IEC 608-70-5-101, IEC 608-70-5-104, DNP3	IEC61850, IEC 608-70-5-101, IEC 608-70-5-104, DNP3	IEC61850, IEC 608-70-5-101, IEC 608-70-5-104, DNP3
Configuration	Difficult	Easy	Difficult	Easy	Easy
Installation Cost	Expensive	Expensive	Moderate	Cheap	Cheap
Robustness	High	High	Moderate	Low	Low
Redundancy	High	Moderate	Moderate	Low	Low
Future Extension	Difficult	Easy	Moderate	Easy	Easy
Data Transmission Time	Moderate	High	Moderate	High	Moderate
Suitability to large PV power plants	Partially	Yes	Yes	No / Partially	Yes

6 SCADA Data Types and Cyber Security in Large PV Power Plants

Predictive maintenance strategies depend on the data types, data acquisition and storage strategies, and cyber security of the overall SCADA system. Each selected component, communication protocol and topology have its unique data types, data storage and cyber security requirement.

This chapter introduces the basic data types which are being utilized by all the communication protocols, data acquisition methods and periods, selection of the key performance indicators (KPIs) and cyber security requirements for large PV power plant SCADA systems.

6.1 SCADA Data & Signal Types

The acquired data and their quality and quantity are the sine qua non of the predictive maintenance strategies in large PV power plants. The data types, formats and frames are usually dependent on the selected communication protocols. However, they are all based on the basic programming data structure. Moreover, RTUs/PLCs/IEDs receive and transmits signals, and consequently, even these signals have their own data types and structure.

As all the acquired data is calculated through logic models to visualize the raw data, the type and format of the data have a correlation between the time for acquisition and the area that they occupy in the data storage area [59].

The following signals and data types are widely being utilized by RTUs/PLCs/IEDs signals and communication protocols.

6.1.1 Digital Signals

Digital signals are usually used by RTUs/PLCs/IEDs as input and output to identify the status of the component and systems. These signals are based on the Boolean data types, which only have a one-bit place, and this one-bit can be either 1 or 0. As these signals occupy the least place in the data storage area and are very useful in order to define the status of the devices, such as on, off, and fail, it is an industry standard for all the standard communication protocols.

6.1.2 Analog signals

Achieving comprehensive monitoring and controlling of the PV power plant SCADA system requires not only the status of the device and system but also dynamic values from both sensors and components. Analogue signals are integer data-based signals whereby defining a range matching with sensor and component sense values can be converted to readable and meaningful values. The conversation range usually highly depends on the manufacturer company. For example, a temperature sensor can produce an electrical signal between 4mA to 20mA, and this range can be converted to a range of between 0 to 65535 possible data to be able to sense by the RTUs/PLCs/IEDs. Another manufacturer of temperature sensors can use a range of 1V to 10V. However, the logic of the conversation will remain the same.

6.1.3 Bit

The general format of every data is based on the bit regardless of the signal type and communication protocol. This format consists of either 1 or 0. Usually, 1 refers to the on, and 0 refers to the off. However, there are minor cases where this situation is vice versa. In SCADA systems, these types of data can also be named Boolean.

6.1.4 Byte

The combination of 8 bits creates a byte which means there can be only 28 possible expressions of byte format. Byte formats can be divided into sub-format systems, namely signed and unsigned. A signed byte means the possible byte expression can be either positive or negative; an unsigned byte refers only to the possible positive data.

Following equation shows the signed and unsigned data format range:

Signed Byte data: $Min = -1x\left[2^{(n-1)}\right], \quad Max = \left[2^{(n-1)}\right] - 1$ (1)

Unsigned Byte data: $Min = 0, \quad Max = \left[\left(2^{(n)}\right) - 1\right]$ (2)

6.1.5 Other Common Data Types in SCADA System

Bits and Bytes are fundamental to all the data frames. The other data frames, which are still may be counted as the basic frames, are derived from the bits and bytes and are utilized by almost all signals and standard communication protocols. Table 9 illustrates the other basic data types, their description, size, and range values:

Table 9: Common data types in SCADA system

Data type	Description	Size (In bits)	Range
BOOL	Boolean	1	No range, either 0 or 1
BYTE	Byte	8	-127 to +127
UBYTE	Unsigned Byte	8	0 to 256
INT	Integer	16	-32767 to 32767
UINT	Unsigned Integer	16	0 to 65536
DINT	Double Integer	32	-2147483647 to +2147483647
UDINT	Unsigned Double Integer	32	0 to 4294967296
Real	Real	32	-3,4028235E+38 to +3,4028235E+38 (Difference between Real and DINT, Real data can be defined as decimal value.)
String	String text	N/A	up to 250 ASCII characters

6.2 Data Acquisition

Correct Data acquisition and sampling time is the initial condition of feasible and sustainable PV power plant maintenance strategies. Large PV power plants can be divided into two major subsystems in terms of data generation, namely electrical components of the PV power plant such as active energy, reactive energy, frequency, and power factor etc. and satellite or weather stations' climate data such as irradiance, ambient temperature, photovoltaic temperature, humidity, wind speed and wind direction [60].

IEC 61724-1:2021 explain the general terminology and required equipment for monitoring and analysing the PV power plant performance [61]. IEC 61724-1:2021 classify the monitoring systems under three subsystems, which are Class A, Class B and Class C. These classifications groups refer highest, moderate, and lowest levels of data monitoring for Class A, Class B and Class C, respectively. The number of PV power plant monitoring systems is also defined in this standard, and it is dependent on the PV power plant size. Table 10 illustrates the minimum number of the monitoring system for different PV power plant capacities as per IEC 61724-1:2021.

Table 10: Minimum number of monitoring systems by plant capacity as per IEC 61724-1 [61]

Plant Capacity (in AC)	Number of the monitoring systems
Until 5MW	1
From 5WMW to 40MW	2
From 40MW to 100MW	3
From 100MW to 200MW	4
From 200MW to 300MW	5
From 300MW to 500MW	6
From 500MW to 750 MW	7
From 750MW to up	8

Moreover, IEC 61724-1:2021 standardize the minimum parameters to be measured for each class of data monitoring level, as it is shown in Table 11. However, these parameters are the minimum requirements, and usually, large PV power plants need more parameters to be measured in order to generate sufficient data for creating feasible, efficient and sustainable predictive maintenance strategies.

Table 11: Minimum parameters to be measured for each class of data monitoring as per IEC 61724-1:2021 [61]

	Class A	Class B	Class C
Data of electrical components			
Allowed maximum uncertainties	0%	3%	N/A
Power factor at AC side	✓	✓	✓
Energy generation	✓	✓	✓
AC Power generation	✓	✓	✓
AC array power generation	✓	✓	✓
AC array voltage and current generation	✓	✓	
DC array voltage and current generation	✓		
DC array power generation	✓		
Data of Satellite or Weather station(s)			
Allowed maximum uncertainties	3%	8%	N/A
Global horizontal irradiance (GHI)	✓	✓	
In-pane irradiance	✓	✓	✓
Ambient air temperature	✓	✓	✓
PV panel temperature	✓	✓	
Humidity	✓	✓	
Wind speed	✓	✓	
Wind direction	✓		

To ensure the correct and reliable data-collecting process, the sampling time for each data type and an evaluation process should be defined. The sampling frequency for the data can be assigned according to their dynamic structure and criticalities. For example, as irradiance values are highly dynamic, it is better to assign a very short time sampling period, such as 1 second. However, the transfer switch status is relatively static; it is better to assign a sampling period whenever the event occurs.

In the process of data evaluation, all the acquired raw data should go through a data quality check process where all data synchronization, repetitive data, time shifting, data shifting, missing data and calibration problems will be identified and corrected. Another essential step of the data evaluation process is the time stamping check and synchronization before data is utilized in key performance indicator calculations.

6.3 Key Performance Indicators (KPIs)

Key performance indicators are the measured and calculated values that demonstrate the actual performance of the PV power plant. Defining KPIs and their detection and calculation methods are the keys to creating proper predictive strategies. These KPIs may be related to technical or financial parameters.

As clearly explained in the earlier chapters, the basis of creating the predictive maintenance strategies for large PV power plants is the comparison of the fundamental operational values with the past operational data or the designed values. Apart from providing information for these comparisons, KPIs

enable the operator to compare the PV power plant with other PV power plants under similar environmental and climatic conditions [62].

As per IEC 61724-1:2021 following KPIs are defined for performance analyses of PV power plants.

6.3.1 Reference Yield

This term refers to the corresponding hours under the reference irradiance. It can be calculated by the following formula [63]:

$$Y_r = \frac{H}{Gref} \quad (3)$$

Where:

- Y_r is the refence yield
- H is the total irradiation on the plane of array
- $Gref$ is the reference irradiance

As it can be seen from the formula, the unit of the reference yield is the hour (h), and consequently, it is a function of site location, direction and inclination angle of the PV modules, and the weather conditions.

6.3.2 Energy Yield

Energy yield refers to the needed PV array operation hours at the plate power of PV modules in order to generate the same amount of energy. It can be calculated by the following formula [64]:

$$Y_F = \frac{E_{out}}{P_0} \quad (4)$$

Where:

- Y_F is the energy yield
- E_{out} is the AC energy generation
- P_0 is the rated DC power of the subjected PV array

Similar to reference yield, the unit of the energy yield is also hour (h) or kWh/kW. This value is important for the designers and operators to normalize the PV power plant energy generation with respect to overall PV power plant size.

6.3.3 Expected Energy Yield

Expected energy yield can be defined as the expected energy generation of a PV power plant in the certain time period, and it can be calculated by the following formula [64]:

$$Y_E = \frac{PR_{expected}}{Y_r} \quad (5)$$

Where:

- Y_E refers expected AC energy generation

- $PR_{expected}$ refers the expected performance ratio by considering the real system and environment parameters

The unit of the expected energy yield is kWh/kW_{peak}. As this indicator is only based on expectation the accuracy of the value is dependent on the calculation and simulation methods.

6.3.4 Performance Ratio

Performance ratio refers to the performance of the PV power plant under real environmental conditions. This indicator is the most famous value for PV power plant reliability identification due to it consists of losses, mismatches, and natural climatic conditions. The performance ratio can be calculated by the following formula [64]:

$$PR = \frac{Y_f}{Y_r} = \frac{\frac{E_{out}}{P_0}}{\frac{H}{G_{ref}}} \quad (6)$$

PR does not have a unit and is usually calculated on a monthly or yearly basis. However, to identify the failing or malfunctioning components it is better to calculate daily or weekly basis.

6.3.5 Weather Corrected Performance Ratio

The temperature-corrected performance ratio is another value defined in IEC 61724-1:2021 in order to increase the PR accuracy by considering the weather variations. It can be calculated by following formula [64]:

$$CPR = \frac{Y_f}{Y_r} = \frac{\frac{E_{out}}{P_0}}{\frac{HxC}{G_{ref}}} \quad (7)$$

Where:

C is the temperature correction coefficient and calculated as following [64]:

$$C = 1 + \gamma\, x\, (T_{module} - T_{reference}) \quad (8)$$

Where:

- T_{module} is the temperature of the module
- $T_{reference}$ is the temperature under the standard test conditions which is 25°C
- γ is the coefficient of the maximum-power temperature

6.3.6 Soiling Ratio

The term of soiling ratio refers to the ratio of the actual power generation of the PV array under the existing soiling conditions to the power generation that would occur if the PV array were clean. Soiling is dependent on the site location and conditions. There is not any standardized formula to calculate the soiling ratio. However, the general acceptance is to measure the maximum power generation of the soiled PV array in proportion to the clean PV array power generation or the datasheet value of the PV array.

Apart from the KPIs identified by the IEC 61724-1:2021, two more indicators need to be considered during the predictive maintenance strategy creation: performance loss rate and availability.

As PV power plant components consist of different electrical and electronic components, it is expected that they will have performance reduction during this time. The performance degradation of the components is dependent on many parameters such as used materials, the technology that has been used, environmental conditions, the preferred balance of system components and etc. This value can be calculated as the ratio between the daily/weakly/monthly/yearly PR to two or more following years' values [65].

There are two main types of availability measurements which are energy-based availability and time-based availability in large PV power plants. Energy-based availability refers to the ratio between the actual energy generation to the reference energy yield. Time-based availability is the ratio between the time that a PV power plant generates a certain amount of energy and the expected time, according to design specifications, to produce the same amount of energy [64].

6.4 SCADA Cyber Security

The communication protocols, which run over the open Transmission Control Protocol / Internet Protocol (TCP/IP), have been the standard for conventional and renewable power plants for the last two decades. This brings the benefit of not only enabling the communication between the components and systems in the power plant but also necessary communication between the power plant to other network systems such as grid networks or other power plants [6]. However, these advantages also bring some risks to the overall SCADA system, such as remote cyber-attacks on the system due to plant SCADA system is accessible from the internet.

SCADA system cyber-security is vital both for operational and maintenance activities. Especially as the creation of predictive maintenance activities is highly dependent on the data acquisition and data quality received from the SCADA system, all the necessary preventive measures should be taken during the design phase. To eliminate cyber security risks on the SCADA system, the system-specific threats and vulnerabilities should be identified.

This chapter introduces the SCADA system threats, SCADA system vulnerabilities and the options to mitigate the cyber security risks.

6.4.1 SCADA System Cyber Threats

SCADA systems have been the target of cyber-attacks from different types of attackers. These attacks were not only attacks from remote locations but also from insiders or by attack campaigns or using malware [66].

6.4.1.1 Terrorist Attacks

Terrorist attacks on the SCADA system are a typical case for critical government infrastructure. These attacks usually target alarm systems and closed-circuit television systems and try to control whole monitoring and control systems. During the literature research, any significant terrorist attack on renewable power plants and specifically on large PV power plants, has not been founded.

6.4.1.2 Insider attacks

Unfortunately, cyber-attacks or SCADA systems damages are not always from remote locations. There have been many attacks or damages from the insiders to the SCADA system of different industries, such as the sewer system in Maroochy, the National Nuclear Security Administration (NNSA) of the US Department of Energy and a massive blackout in London [67]. During the literature research, any significant insider attack on renewable power plants and specifically on large PV power plants, has not been founded.

6.4.1.3 Hacker attacks

SCADA systems are mostly being targeted by hackers. There is different motivation, such as controlling the whole SCADA system to create irreversible damage, collecting data or manipulating the data that the SCADA system provide. These attacks target different network levels of the SCADA system. The first publicly known attack on a PV power plant was also made by hackers. The firewall of a utility in the western United States was targeted by hackers in May 2019. The main reason for this vulnerability was the unpatched Cisco brand firewall which enabled hackers to crash the components. This attack destroyed the paramount communication between the generation units, and eventually, this situation impacted 500MW of PV power plant generation [68].

6.4.1.4 Malware Attacks

The term Malware stands for malicious software. Malware software are secretly injected into a file or an application, and they perform activities in order to steal or manipulate data, slow down system operation or attack the whole system to damage [67]. Malwares are in different variety of shapes such as spyware, viruses and trojan horses, and ultimately target confidentiality, integrity and availability of the system [69].

6.4.2 SCADA System Vulnerabilities

When initially SCADA systems were developed, security was not a big concern due to stand-alone communication networks. However, over time these systems become connected to the internet and to each other over internet protocols. Since standard SCADA communication protocols running over TCP/IP models became the industry standard in large PV power plants, these systems share the same vulnerabilities with the information technologies vulnerabilities [67]. To detect the vulnerabilities from a broad perspective, not only SCADA system-specific vulnerabilities considered but also the vulnerability that information technologies bring to the system should also be considered.

The main vulnerability of the traditional system is very little implementation of encryption and authentication security mechanisms [70]. Although encryption needs might be very system specific, the encryption of data integrity, especially in large PV power plants, is a necessity in order to make sure that data quality cannot be manipulated by malicious attacks.

Authentication has traditionally been another vulnerability for SCADA systems. Usually, producers either support weak authentication or do not support authentication at all. Even the support for weak authentication is very limited, with recommending the use of the same authentication for all components.[71]

As explained earlier, the failure at the cyber security of SCADA systems has essential consequences which make these initial systems targets. Therefore, the attention to the weakness of the standard communication protocol that is being utilized by the SCADA system has been increased. Traditionally SCADA systems are being designed in a way to minimize the effort to debug systems. However, this same functionality makes easy data interception and data manipulation. In addition to these vulnerabilities, the lack of authentication and encryption only makes SCADA systems easier targets.

The standard SCADA communication protocols namely IEC 61850, IEC60870-5-101, IEC60870-104, DNP3 and Modbus are lack of proper authentication, authorization, integrity and confidentiality features [70].

6.4.3 Cyber-Security Risk Mitigation Methods

Awareness of SCADA system cyber security risks has increased in the last decades. This situation brought more interest to this vital topic and made industries, institutions, and governments develop risk mitigation tools against it. To standardize the cyber security process for SCADA systems, several standards have been created and involved in the SCADA system design. The standards and regulations that may be applicable to the large PV power plant SCADA system are as follows:

- NERC 1300 - CIP-002-3 - CIP-009-3, Critical Cyber Asset Identification,
- ISO/IEC 27001:2013 - Information technology - Security techniques - Information security management systems – requirements,
- ISO/IEC 27032:2012 - Information technology - Security techniques - Guidelines for Cybersecurity,
- NIST - 800-12-13-26-82, An Introduction to Information Security,
- National Cyber Security Centre (NCSC) - Cyber Essentials.

Moreover, to make sure that the designed SCADA infrastructure has the necessary security protection, the following cyber-attack risk mitigation methods should be considered by PV power plant investors and designers, producers, and aggregators:

- The very initial step for mitigating the cyber-attack risks is to separate the internal and external communication of the large PV power plant. After that, by creating proper secure ports,

configured zones, subnets, context-based firewalls and gateways, the access between isolated communication systems can be designed. This will ensure that internal communication will only handle the contact among site instruments, controllers and the SCADA system. Moreover, the external communication system will only control communication at the internet level.

- Identifying role-based authentication to ensure the relevant person has access to identified components or software. Authentication can be defined for different vital functions such as deleting and manipulating the data, data creation and HMI access. This will help to mitigate the risks of insider attacks on the system.

By integrating transport layer security (TLS) into the SCADA system, data integrity, authentication, and encryption can be secured. This functionality will help to mitigate the risk of remote hacker attacks [68].

- By integrating antivirus software into the SCADA system, the risk of malware software attacks can be mitigated.
- The backup data can also be stored in a physical location where external access is prohibited. This physical backup data storage system should periodically be tested in order to make sure that the stored data still has the same quality and can be recovered at any requirement.
- Apart from the network and component level mitigation methods, the physical security of the PV power plant also can be increased by CCTV systems, access control systems, security alarm systems and fences in order to secure vital SCADA components from unauthorized physical actions.

7 Case Study: SCADA Infrastructure Design of a Gold Mine Hybrid Renewable Power Plant, By Considering the Predictive Maintenance Activities

In the earlier sections, SCADA systems in large PV power plants have been discussed and compared from different perspectives to explain the prioritization of the different communication protocols, topologies, key performance indicators, security, and components for creating and increasing the effectiveness of predictive maintenance strategies in large PV power plants.

In light of the findings and conclusion of the earlier sections, a SCADA infrastructure was designed as a case study for a gold mine in Burkina Faso to enable the investor and system operator to increase the effectiveness of predictive maintenance activities during the operation.

Mining operations are very critical in terms of continuity, and any disturbance or blackout in the system might cause irreversible technical and financial losses. Especially for the system with high inverter-based renewable energy generation penetration, grid stability and power quality are the parameters that need the utmost attention. Due to these reasons, implying predictive maintenance strategies to the operation scheme is vital.

7.1 General Description of the Project

The gold mine is located in Burkina Faso, operating 365 days and 8760 hours during the year. Heavy fuel oil generators have supplied the necessary power requirements for years. However, the company aimed to decrease its decarbonization which appears to be due to power generation. To achieve this aim, the mining company decided to build a hybrid PV power plant and battery energy storage systems that will operate along with the existing minimum number of heavy fuel oil generators.

The gold mine site has the following environmental and electrical grid conditions:

- Maximum ambient temperature is 50°C
- Yearly average ambient air temperature is 28°C
- Minimum ambient air temperature is 1°C
- Elevation above the sea is less than 1000m
- The electrical system is an off-grid system
- Nominal system voltage is 33kV
- Maximum system voltage is 36kV
- Electrical frequency is 50Hz
- The voltage level for the auxiliary systems is 230/400 V_{AC} and110 V_{DC}

The gold mine owner has performed a detailed feasibility study to assess different options to integrate the PV power plant and battery energy storage system (BESS) into the existing mining operation. The configuration determined by the owner is as follows:

- A solar PV power plant with $32MW_{AC}$ and 40MWp capacity
- A battery energy storage system with 16MW/9MWh capacity

The new inverter-based power plant will integrate into the existing mining electrical system via step-up transformers and MV switchgear, and the generated energy will be transmitted by using a 33kV double circuit overhead transmission line, as is shown schematically in Figure 17.

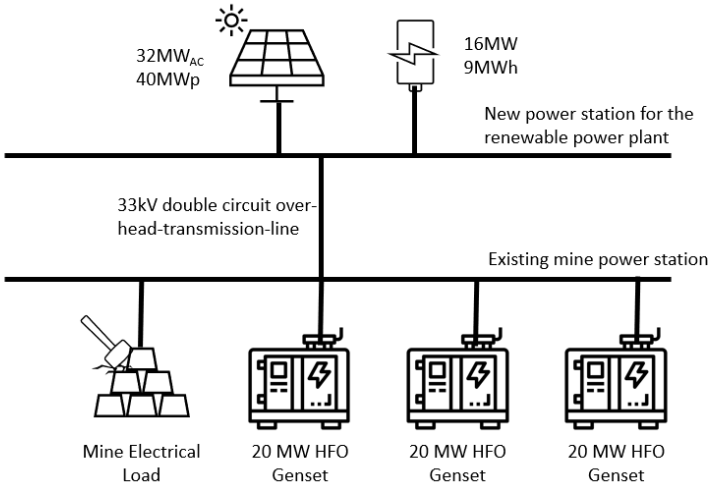

Figure 17: Schematic electrical layout of the mining project

To be able to provide necessary contingency, redundancy and reliability requirement, PV generation decided to be evacuated through five identical compact substations, namely TS1, TS2, TS3, TS4, and TS5, which consists of the followings:

- 1 piece of 3 winding 6500kVA 0,8/0,8/33kV step-up transformer
- 1 piece of 33kV ring main unit which includes 1 piece of transformer protection switchgear, 1 piece of incomer switchgear and 1 piece of outgoing switchgear

The BESS charging events are also designed in a way that the system can be charged or discharged through 2 identical substations, namely BESSA and BESSB, and each one of them consists of the followings:

- 3 pieces of 2 winding 3000kVA 0,71/33kV step-up transformer
- 1 piece of 33kV ring main unit which includes, 3 piece of incomer switchgear and 1 piece of outgoing switchgear

Overall, these substations will transmit the generated power to the new 33kV power station of the hybrid renewable power plant via a 33kV double circuit over-head-transmission-line. Figure 18 illustrates the schematic single-line diagram of the PV power plant and battery energy storage system electrical infrastructure.

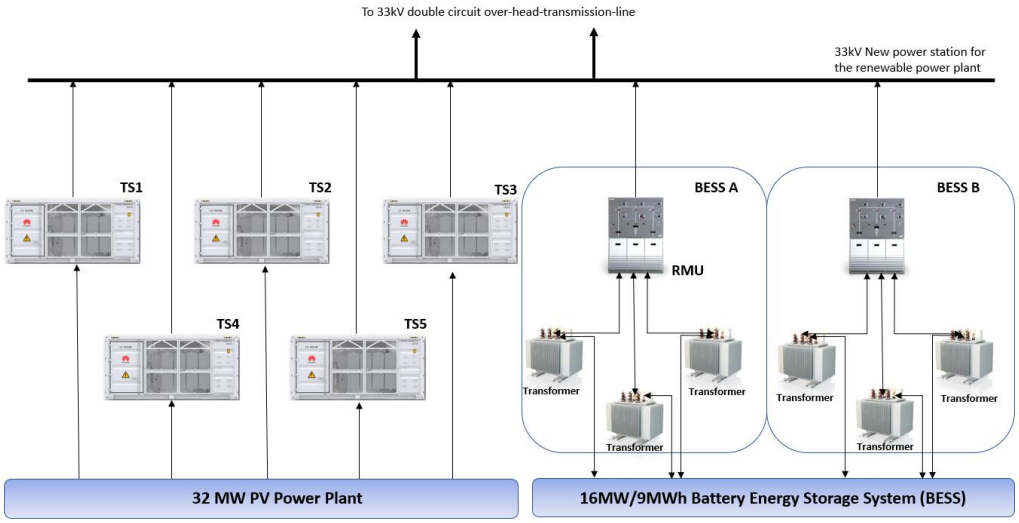

Figure 18: Schematic single line diagram of the PV and BESS plant

7.2 SCADA Infrastructure Design of the Project

7.2.1 SCADA System Components of the Project

The project subjected to the case study has 32MW$_{AC}$ PV and 16MW/9MWh BESS capacity, which means the hybrid renewable power plant will be located on an extensive area. This reality is considered during the defining of the SCADA system components and their numbers.

PV power plant has five substations, as it is depicted in Figure 18. 130 pieces of 250kW string inverters are considered in the scope of the PV power plant design to convert the generated DC power by the PV modules to AC power which will be regulated to a medium voltage level by the substations. These string inverters collect several DC inputs, which are connected to the PV strings. Due to reason that all DC inputs are being monitored by the inverters, this design assure that the necessary data related to PV strings will be supplied by these inverters to the PV power plant SCADA system.

PV power plant compact substations can be defined as an integrated solution for inverters, transformers, MV switchgear and associated equipment to be controlled. Each compact substation system generates several digital and analogue signals, such as transformer and MV switchgear status, the temperature of the substation container and fire alarm signals. As these compact substations are crucial components of the overall PV power plant system, all the data and signals should be monitored by SCADA systems. Moreover, a dry contact relays provide internal pressure, alarm temperature and trip temperature status to the SCADA system. The compact substation will consist of different types of circuit breakers and switches. To enable the SCADA system to monitor the open, close and trip status of these elements, related relays should provide the signals to the SCADA system. Each compact substation container should also have intrusion and fire alarm sensors to provide necessary information against intrusion, fire, and vandalism via the substation centre provider to the SCADA system.

A data concentrator considered for each PV power plant compact substation to provide the data regarding string inverters, transformers, MV switchgear, and fire and intrusion signals. All these will be gathered directly on the data concentrator.

An RTU is considered for the control and monitoring of the inverters groups in each substation unit, and another RTU is located at the battery group of each RMU unit to enable data transmission to the SCADA system and enable the operator to monitor and control these devices.

To collect all the data at the 33kV new power station of the hybrid renewable power plant, an RTU was considered in the SCADA system design. This component will collect all the data regarding circuit breakers, switchgear, and energy analyser. This way, the SCADA operator will be able to acquire all the data from one source and be able to monitor and control the 33kV power station.

An MTU unit is considered to provide the necessary flexibility to the operator in terms of the monitoring and control of all the site devices. By defining the behaviour and parameters according to design and connection conditions, for all components, the MTU can ensure the system is running automatically as desired. As grid stability and power quality are two essential parameters of the system, it is considered that MTU should have the following control modes:

- Active and reactive power control,
- Power factor control,
- Voltage and frequency control,
- Ramp rate control,
- Connect/disconnect PV power plant,
- Connect/disconnect BESS.

Due to the fact that the PV power plant capacity is $32MW_{AC}$, the minimum number of the weather monitoring system should be two, as per IEC 61724-1:2021. These weather stations are considered to provide data directly to the MTU unit in order to provide quick access for the operator and also can be used in the necessary control diagrams. The weather stations are considered to provide the following data for being utilized in the key performance indicators which are used for the predictive maintenance strategies:

- Ambient temperature,
- Module temperature,
- Irradiation,
- Barometric pressure,
- Relative humidity,
- Wind speed and direction.

Apart from the main components, the SCADA system should have other necessary supportive components to be able to run without any significant issues and provide the necessary data for creating the predictive maintenance strategy. These components are electrical cabinets, rack cabinets, monitoring system cabinets, uninterruptable power supply units, industrial servers, workstations or operator stations, media converters, industrial ethernet switches and data loggers.

7.2.2 SCADA Communication Protocol Selection and Topology Design of the Project

A physical cable network has the advantages such as availability and reliability over wireless systems. Therefore, a physical cable infrastructure is considered in the project's design. A physical cable network is a communication road where data is transmitted from one component to another. Therefore, the necessary cable type and sections are selected according to communication protocol and topology.

Utilizing the same communication protocol for each part and component of the SCADA system may cause data mismatches, data delay and increased installation, operation, and maintenance costs. Instead of that, as explained in the earlier sections, using different communication protocols that can run together and bring maximum data efficiency at both component and system levels is a better approach for supplying the necessary data for predictive maintenance strategy design. Therefore, considering each advantage and disadvantage of standard communication protocols, the following parameters have been considered during the SCADA infrastructure design.

33kV new power station, PV power plant and BESS substations are critical components of hybrid renewable power plant projects due to the data they provide to the SCADA system is essential in terms of the continuity pf the system. Therefore, IEC 61850 is considered an internal communication protocol for these components due to advantages such as eliminating procurement uncertainties, wire cost reduction, almost no need for manual configuration, decreased number of transducers, and being the most matured and industrially proven communication protocol for the substations.

The length between string inverters and RTU units is considerably short. Due to the fact that there will be only one master in the relationship between the RTU unit and inverter and considering the data transmission velocity, Modbus RTU is considered the best communication option between inverters and RTUs.

Each BESS system is designed to communicate with each other and the RTU unit. However, this is not a hierarchical process; instead, it is a process where each BESS system and RTU is a master and battery pack, and the devices in the battery container are the slaves. Thus, to enable the system to have this type of hierarchical structure, the communication between BESS and RTU units is designed as Modbus TCP.

The main logic behind creating the predictive maintenance strategy is to acquire reliable and quality data and use them to calculate the key performance indicators and compare them with trends and design specifications. The condition of reliable and quality data is to use a communication protocol that can

define different objects, provides necessary time stamps for acquired data and flexibility in filling the data gaps, and makes it easy to fix the data mismatches. Thus, IEC 60870-5-104 is considered to be the communication protocol between the RTUs, substations and MTU unit due to being an open protocol, ability of integrating the different protocols with interfaces, being a dominant protocol for the energy projects in Europe and Africa.

IEC 60870-5-104 communication protocol runs over TCP/IP. To minimize the data losses due to long-distance cable connections and considering the nature of the TCP/IP model, optical fibre cables are the best physical transmitter for the IEC 60870-5-104 communication protocol. The optical fibre cables are considered to be shielded in order to avoid electromagnetic interfaces, which may eventually cause the wrong data or data loss.

The data transmission can be handled over the ethernet for the sensors, such as weather stations and pyranometers. The distance between the sensor stations and MTUs designed to be less than 1000m, where CAT6 cables are considered physical data transmission roads.

To assure that the SCADA system has redundancy, the topology between the RTUs, substations and MTU unit, fibre optic cable infrastructure was designed as a ring topology. This topology will bring redundancy, more flexibility and high tolerance to a physical disconnection of the devices to the overall SCADA system.

Star topology is considered where ethernet CAT6 cables are being utilized, due to the reason that these cables are supplying the necessary connection between the small measurement devices to the master units such as RTUs and MTUs. This will bring fast data transmission, robustness, and flexibility for future extensions to the SCADA system.

All the acquired data is designed to be stored in a physical server at the site where the utmost physical securities ensured. The necessary storage time for each data type is explained in further sections. The overall SCADA infrastructure for supporting the predictive maintenance strategy design is schematical showed in the Figure 19.

Overall, the SCADA system is designed to monitor all devices in the hybrid renewable power plant and supply the necessary data for creating the predictive maintenance strategy. The system provides access from local and remote locations, demonstrating information from the hybrid power plant, component status, energy generation, and weather data conditions, among others. This enables the system to convert these data into a powerful tool to take preventive measures.

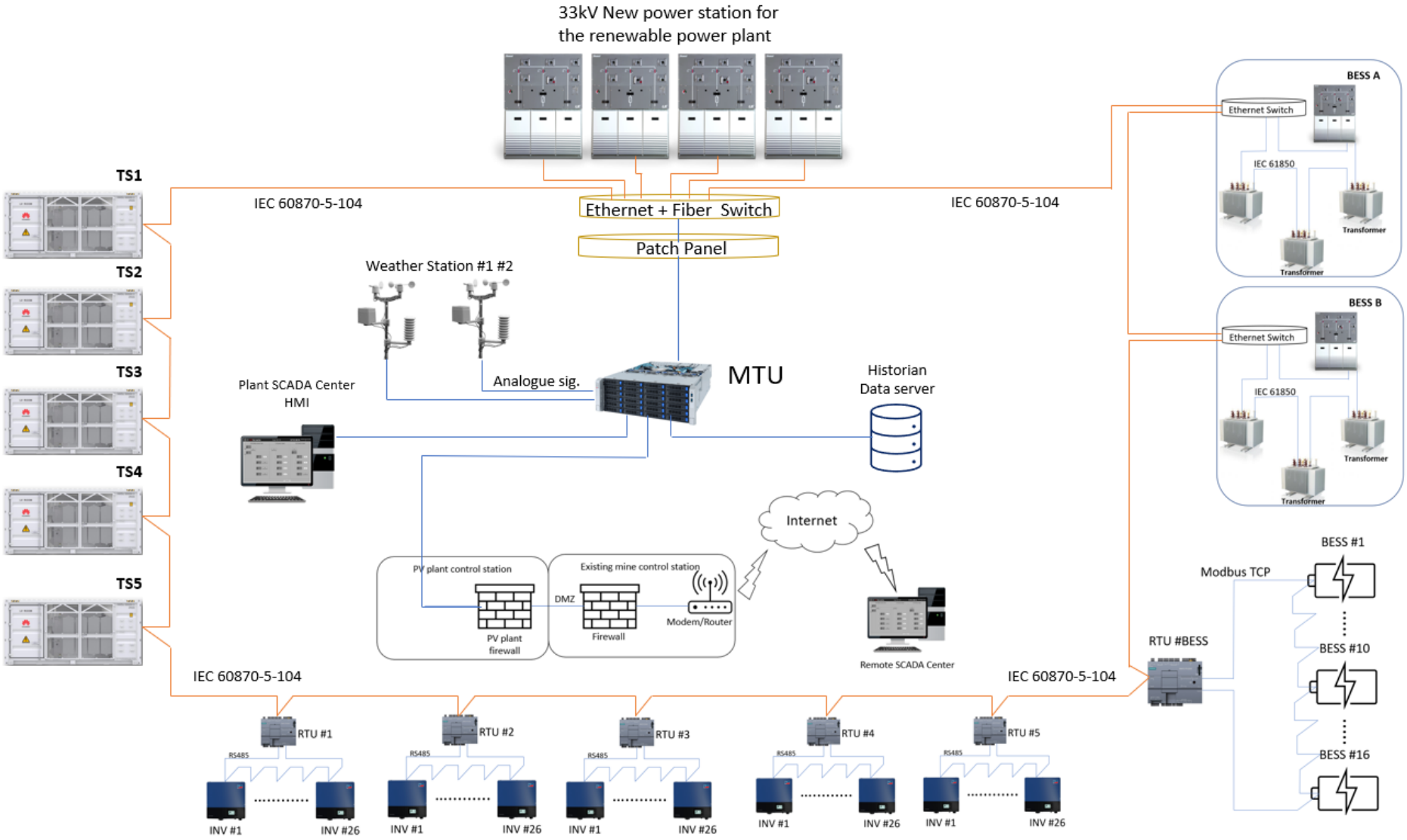

Figure 19: Schematic design of PV power plant and BESS SCADA topology

7.2.3 SCADA Data Acquisition, Storage, and Variable Calculations

Correct, coherent and reliable data acquisition is mandatory to build a predictive maintenance strategy. In order to provide the quality data, this section provides information regarding data acquisition and storage as well as how variables and KPIs should be calculated and structured.

7.2.3.1 Data Acquisition

Data acquisition sampling time is vital to ensure that any critical trend or status has not been neglected. Although data acquisition sampling time can be easily defined for most variables based on industry standards, some variables require a higher frequency of data acquisition to provide correct and coherent information to predictive maintenance measures.

The other important parameter is providing data to the control stations without gaps. Therefore, the data acquisition period of all dataloggers and monitoring components is designed in a way that they can provide data without time gaps. All the data loggers and monitoring devices are considered to have the capability of backfilling the missing data without losing any timestamps in case of a communication failure.

Table 12: Data acquisition frequency and types

Component	Variable	Data sampling time interval
Inverter	Power & Energy related data	1 min.
	Errors, events, and alarms	Whenever event occurs
Power Analyser	Power & Energy related data	1 sec.
	Errors, events, and alarms	1 min.
Substation I/O Signals	All substation I/O signals	Whenever event occurs
Power Plant Controller	Power & Energy related data	1 sec.
Protection relay	Power & Energy related data	1 min.
	Errors, events, and alarms	Whenever event occurs
Battery Energy Storage System	Power & Energy related data	1 min.
	Errors, events, and alarms	Whenever event occurs
Weather Station and Sensors	Irradiance	1 sec.
	Ambient temperature	1 min.
	PV module temperature	10 sec.
	Humidity	10 sec.
	Pressure	1 min.
	Wind direction	1 min.
	Wind speed	1 min.

In order to minimize synchronization failures, all the components are considered to be set to local time. While Table 12 shows variable acquisition frequency at the system level, Appendix A provides a piece of detailed information regarding each variable and the data acquisition period. Additionally, all the

status data, such as errors, alarms and events considered to be acquired whenever they occur.

7.2.3.2 Data Storage

Data storage is a mandatory functionality in order to provide the base trends and information to compare with future operation data. All the data considered to be classified as per IEC 61850 and IEC 60870-5-104 requirements. Table 13 shows a simple list of acquired, calculated, and correlated data and their sampling period to be stored in the data storage.

Table 13: Data storage period, method, and type

Parameter	Period (Minutes)	Value
Power	During every 10 minutes with 1-minute intervals.	Mean
		The highest
		The lowest
Energy	At the end of every 10 minutes	Cumulative
PV module temperature	During every 10 minutes with 1-minute intervals.	Mean
		The highest
		The lowest
Irradiance	During every 10 minutes with 1-minute intervals.	Mean
		The highest
		The lowest
Ambient temperature	During every 10 minutes with 1-minute intervals.	Mean
		The highest
		The lowest
Humidity	At the end of every 10 minutes	Mean
Wind speed	During every 10 minutes with 1-minute intervals.	Mean
		The highest

The database is designed in a way that all kinds of data can be exported to third-party software such as excel or CSV. The data and variables will be stored in a Raw data format and as well as in the form of calculation results. It is important to note that there are some parameters, such as energy, which are cumulative values. Calculation methods assigned to obtaining the values are depicted in Appendix B. All these calculations are based on the formulation of KPIs and industry standards. Statistical calculations are considered to be done on the RAW data, and each of them allows the distinction between RAW and calculated data.

A regular data backup task is also considered to prevent data losses. The necessary data backup is considered to be hosted on the internal hard drive. However, in order to increase the data redundancy, an external two terabyte was also designed in the scope of the project to back up the acquired data.

7.2.4 SCADA System Software of the Project

As discussed in the earlier sections, the SCADA system software is designed to be open software where the dependency on the producer can be eliminated, enabling variable changes as well as allowing sufficient data integrity. Another criterion of SCADA system software selection was flexibility which enables future extensions without any technical and financial constraints. Selected SCADA software service life is higher than the hybrid renewable power plant life expectancy. Moreover, selected SCADA software does not require any customer licences.

The SCADA system software has ability of visualizing historical data and create trends from these data. The additional selected SCADA system software has the functionality of monitoring the following real-time data:

- All real time data provided by the SCADA components such as energy meters, analysers, inverters, and substations,
- Meteorological data from the weather station,
- Power plant performance and malfunctioning components,
- AC and DC power generation,
- Accumulated energy generation.

For better visualization of the predictive maintenance needs, graphical and informative windows are also integrated into SCADA software. SCADA software is designed to show the data by means of the following screens:

- Plant information,
- Dashboard,
- Daily generation,
- Monthly generation,
- Annual generation,
- Performance ratio,
- Availability,
- Alarms and events.

The software is designed so that the operator or client can select different data and variables from different screens to check historical data or compare them with real-time data.

7.2.5 KPIs of the Project

To be able to execute the predictive maintenance measurements of the hybrid renewable power plant, the key performance indicators were selected as per IEC 61724-1:2021 and with the utmost care. The selection of the KPIs is affecting the data storage capacity as well as the effort to catch the malfunctioning or extreme generation capacity degradations.

The main philosophy of predictive maintenance measurements is to compare the power plant's design conditions or historical data to real-time data. Therefore, KPI selection, detection and calculation are the critical parameters of reliable predictive maintenance strategies. Therefore, the following KPIs are considered to be the most feasible values in terms of reliability and storage occupancy:

- Reference yield,
- Energy yield,
- Expected energy yield,
- Performance ratio,
- Weather corrected performance ratio,
- Soiling ratio.

7.2.6 Cyber-Security of the Project SCADA System

Data quality and real-time data acquisition are vital parameters for creating predictive maintenance strategies. After comparing the availability, reliability, and data quality features of the wired and wireless SCADA communication technologies in the earlier sections, it was considered to design PV power plant SCADA infrastructure with wired-based technologies. However, as the client also wants the data to be available to third-party clouds to be able to monitor the site remotely, an internet interface was also designed. As it is explained in the earlier chapters, to provide necessary simplicity and security, the internet connection architecture designed in a logic arrangement where the PV power plant isolated from the existing system and connected to the internet via firewalls through the existing gold mine control centre, as it is shown Figure 20.

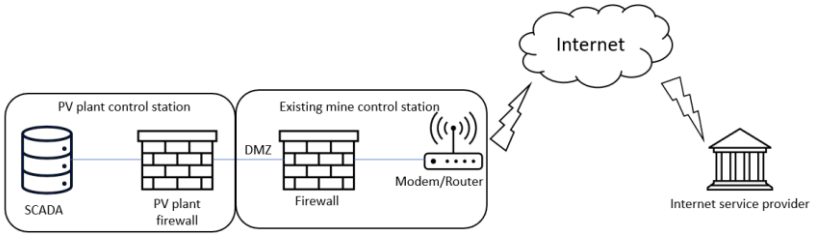

Figure 20: Schematic of the internet service architecture

Additionally, to provide a necessary buffer between the PV power plant and external connection, the ‘’demilitarized zone (DMZ)’’ should be prepared for the SCADA system at the PV site firewall.

To mitigate the cyber-attack risks, the following baseline functionalities also are implied in the SCADA design:

- The firewall is configured in a way to deny all the data traffic apart from the data traffic, that is explicitly necessary for the continued system operation,
- All the SCADA servers, workstations, MTUs, RTUs and CCTV, selected to be with the malware protection,

- All the patches were being tested before they are applied to the live systems,
- SCADA system software and hardware have the ability to detect and prevent intrusion into the network system,
- Cyber key threats, which are explained above, were considered during the risk assessment process,
- The designed SCADA system has a comprehensive and documented backup regime for the purpose of disaster recovery, business continuity and incident response,
- The designed SCADA system has an encyclopaedic and verified backup regime for disaster recovery, business continuity and incident reaction,
- All the connections to the hybrid renewable power plant from the outside or third-party systems were designed in a way that it can only occur with SCADA operator authorization verification,
- All the user accounts were authenticated with passwords which align with the at least above-mentioned password requirements,
- All the user accounts were authenticated with passwords aligned with the earlier-mentioned password requirements.

8 Conclusion

The thesis aims to conduct research on available SCADA infrastructure technologies for large PV power plants, compare them in a technical aspect to understand their effect on developing predictive maintenance strategies and, in the light of these understandings, design a SCADA system infrastructure for a gold mine PV power plant in Burkina Faso as a case study. As a result, the following conclusions and inferences summarized from the study:

1. Apart from the traditional SCADA system hardware and software components such as RTUs, PLCs and MTUs, the alternative components such as weather stations, pyranometers and inverter monitoring relays should also be integrated into the SCADA system infrastructure to ensure that each component of the large PV power plant is optimally controlled and monitored. These additional components will provide more data to create accurate trends to compare with the expected operation trends in order to detect the failures or malfunctioning components in advance.
2. The successes of the SCADA system and, consequently, the predictive maintenance strategy to be developed by the acquired data from SCADA are highly dependent on the smooth communication between components of the large PV power plant. Considering the necessity of security, reliability and flexibility, it concluded that open and wired communication protocols have significant advantages over proprietary and wireless communication protocols. Among the open and wired communication protocols, DNP3 and IEC 60870-5-104 are standard communication protocols that meet the control and monitoring requirements for developing predictive maintenance strategies in large PV power plants. However, due to IEC 60870-5-104 is more specific to energy projects, offers more user-definable objects and is the dominant power plant communication protocol in Europe, it may be preferred especially for the large PV power plants located in Europe and Africa.
3. As the redundancy and robustness of the system and data transmission speed highly depend on the SCADA network topology, decision-making on the topology is a vital parameter for developing predictive maintenance strategies. After a techno-economic comparison of the SCADA network topologies, a combination of a star and ring network topology is considered the most feasible infrastructure for increasing the effectiveness of predictive maintenance strategies. This combination has robustness, flexibility, high data transmission speed of the star topology, and high redundancy of the ring topology.
4. Defining KPIs and identifying the data subjected to KPIs calculations are the key parameters for generating large PV power plant operation trends. Therefore, KPIs such as reference yield, energy yield, expected energy yield, performance ratio, weather-corrected performance ratio and soiling ratio should be included in the trends to be developed to compare with the usual process data or design specifications.

5. In order to meet the grid code or operator requirements, the SCADA system utilized in large PV power plants usually has an internet connection. This makes PV power plant SCADA systems a potential target for cyber-attacks. To mitigate the cyber-security risk, internal and external communication should be separated by firewalls, role-based authentication should be implied into the system, TLS systems should always be considered in SCADA systems design, and the acquired data should always be stored in a physical location as a backup.
6. The SCADA infrastructure of the gold mine hybrid renewable power plant was designed in a way that covers the earlier findings and provides the necessary precondition for developing predictive maintenance strategies.

8.1 Future Work

This study examined the SCADA technologies for large PV power plants from an infrastructure point of view and compared them for their optimum usage for the predictive maintenance strategy development activities. Future studies can investigate the effect of SCADA data processing and logic programming methods on developing predictive maintenance activities. This can provide a comprehensive SCADA hardware and software infrastructure guidance and increase the effectiveness of the predictive maintenance strategies for large PV power plants.

References

[1] G. D. Lorenzo, R. Araneo, M. Mitolo, A. Niccolai, F. Grimaccia. Review of O&M Practices in PV Plants: Failures, Solutions, Remote Control, and Monitoring Tools. IEEE J. Photovoltaics 2020;10(4):914–26. https://doi.org/10.1109/JPHOTOV.2020.2994531.

[2] Liu Q, Zhao Y, Zhang Y, Kang D, Lv Q, Shang L. Hierarchical context-aware anomaly diagnosis in large-scale PV systems using SCADA data. In: 2017 IEEE 15th International Conference on Industrial Informatics (INDIN). IEEE; 2017, p. 1025–1030.

[3] Walker HA. Best Practices for Operation and Maintenance of Photovoltaic and Energy Storage Systems; 3rd Edition; 2018.

[4] A. Sajid, H. Abbas, K. Saleem. Cloud-Assisted IoT-Based SCADA Systems Security: A Review of the State of the Art and Future Challenges. IEEE Access 2016;4:1375–84. https://doi.org/10.1109/ACCESS.2016.2549047.

[5] Zerdazi I, Fezari M, Mokhtar B, Bayart M. Evolution and Vulnerability in SCADA Systems. In: ; 2019.

[6] Er. K. B. MOHD. UMAR ANSARI. An Analysis and Study on SCADA. IJAREEIE 2020;9(5).

[7] S. V. B. Rakas, M. D. Stojanović, J. D. Marković-Petrović. A Review of Research Work on Network-Based SCADA Intrusion Detection Systems. IEEE Access 2020;8:93083–108. https://doi.org/10.1109/ACCESS.2020.2994961.

[8] Enescu F, Bizon N. SCADA Applications for Electric Power System. In: ; 2017, p. 561–609.

[9] Narayan Lal Purohit A. Data Acquisition of Solar Power Plant Using Scada System. IJETT 2015;23(4):189–94. https://doi.org/10.14445/22315381/IJETT-V23P237.

[10] Khatri P. Review of SCADA system for photovoltaic power plant, April 2018 2019.

[11] Alade A. A., Ajayi O. B., Okolie S. O., Alao D. O. Overview of the Supervisory Control and Data Acquisition (SCADA) System. International Journal of Scientific & Engineering Research 2017;8(10).

[12] M. M. Ahmed, W. L. Soo. Supervisory Control and Data Acquisition System (SCADA) based customized Remote Terminal Unit (RTU) for distribution automation system. In: 2008 IEEE 2nd International Power and Energy Conference; 2008, p. 1655–1660.

[13] Motorola I. SCADA Systems: A Comparision of RTUs and PLCs 2007.

[14] Eren H, Yee I. Data Historian. In: Instrument Engineers' Handbook: Fourth Edition; 2012, p. 454–464.

[15] Goto Tetsuo, Morishita Yuki, Take Masayuki, Asao Yoshihisa, Shimoguchi Takefumi, Matsushita Tomohisa. String Monitoring Unit for Megawatt Solar Power Plants. In: ; 2017.

[16] Robalo BM, Figueiredo JM. Supervisory Control developed for a Solar Tracking Prototype based on PV-Technology. IFAC Proceedings Volumes 2010;43(1):291–6. https://doi.org/10.3182/20100329-3-PT-3006.00053.

[17] Karthik Krishnamurthi, Suraj Thapa, Lokesh Kothari, Arun Prakash. Arduino Based Weather Monitoring System. International Journal of Engineering Research and General Science 2015;3(2).

[18] Yadav G, Paul K. Architecture and security of SCADA systems: A review. International Journal of Critical Infrastructure Protection 2021;34:100433. https://doi.org/10.1016/j.ijcip.2021.100433.

[19] Rao Kalapatapu. SCADA PROTOCOLS AND COMMUNICATION TRENDS. ISA 2004.

[20] G. Strawn. Masterminds of the Arpanet. IT Professional 2014;16(3):66–8. https://doi.org/10.1109/MITP.2014.32.

[21] Mohammed M. Alani. Guide to OSI and TCP/IP Models. In: SpringerBriefs in Computer Science; 2014.

[22] A. Bani-Ahmed, L. Weber, A. Nasiri, H. Hosseini. Microgrid communications: State of the art and future trends. In: 2014 International Conference on Renewable Energy Research and Application (ICRERA); 2014, p. 780–785.

[23] Hindocha K, Shah S. Design of 50 MW Grid Connected Solar Power Plant. International Journal of Engineering Research and 2020;V9. https://doi.org/10.17577/IJERTV9IS040762.

[24] C Ralph Mackiewicz. Technical Overview and Benefits of the IEC 61850 Standard for Substation Automation. In:

[25] Y. Rangelov, N. Nikolaev, M. Ivanova. The IEC 61850 standard — Communication networks and automation systems from an electrical engineering point of view. In: 2016 19th International Symposium on Electrical Apparatus and Technologies (SIELA); 2016, p. 1–4.

[26] Gaviano A, Weber K, Dirmeier C. Challenges and Integration of PV and Wind Energy Facilities from a Smart Grid Point of View. Energy Procedia 2012;25:118–25. https://doi.org/10.1016/j.egypro.2012.07.016.

[27] Ahmed Elgargouri. IMPLEMENTATION OF IEC 61850 IN SOLAR APPLICATIONS [Master´s thesis for the degree of Master of Science in Technology]: FACULTY OF TECHNOLOGY TELECOMMUNICATION ENGINEERING; 30th of 2012.

[28] Hussain S, Ali I. IEC 61850-7-420 based Communication Configuration to Integrate DER to Distribution System. IEEE Standards Education - Student Application Papers Applying Industry Standards 2015. https://doi.org/10.13140/RG.2.2.17656.03846.

[29] Matoušek Petr, doc. Ing., Ph.D., M.A. Description and analysis of IEC 104 Protocol. Brno, Czech Republic; 2017.

[30] IEC 60870-5-101. Telecontrol Equipment And Systems - Part 5-101:Transmission Protocols -

Companion Standard For Basic Telecontrol Tasks. 2nd ed.: IEC; 2015. [October 11, 2022].

[31] Advantech Czech s.r.o. Protocol IEC101/104 User Module: APPLICATION NOTE. Czech Republic; 2020.

[32] Kang DJ, Robles R. Compartmentalization of Protocols in SCADA Communication 2010.

[33] Clarke G, Reynders D, Wright E (eds.). Practical Modern SCADA Protocols. Oxford: Newnes; 2003.

[34] Triangle MicroWorks, Inc. DNP3 Overview. 1st ed. Raleigh, North Carolina; 2002.

[35] Modbus-IDA. MODBUS APPLICATION PROTOCOL SPECIFICATION V1.1b: Modbus-IDA; 2006; Available from: https://modbus.org/docs/Modbus_Application_Protocol_V1_1b.pdf. [October 25, 2022].

[36] ACROMAG INCORPORATED. INTRODUCTION TO MODBUS TCP/IP. U.S.A; 2005.

[37] Mark E. Hazen. Understanding Some Basic Recommended Standards for Serial Data Communications - A comparison of RS-232, RS-422 and RS-485: High Performance Analog; 2003.

[38] Harkishen Singh, & Gavin Mangeni. RS Based SCADA System for Longer Distance Powered Devices. International Journal of Electrical, Electronic and Communication Sciences.

[39] Sena Technologies. Introduction to MODBUS: Technical Tutorial; 2002.

[40] James Powell PE. Profibus and Modbus: a comparison. Germany; 2013.

[41] M. Moghimi, C. Bennett, D. Leskarac, S. Stegen, J. Lu. Communication architecture and data acquisition for experimental MicroGrid installations. In: 2015 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC); 2015, p. 1–5.

[42] C. Zebermann, L. Kaster and H. Rake. PROFIBUS - OPEN ON-SITE COMMUNICATION. Achen; 1990.

[43] PROFIBUS Nutzerorganisation e. V. PROFIBUS System Description: Technology and Application. Germany; 2016.

[44] Igor Petrović MV. USAGE AND ADVANTAGES OF PROFIBUS COMMUNICATION PROTOCOL FOR INDUSTRY. Technical journal 8. Technical college in Bjelovar 2014.

[45] Willig A, Matheus K, Wolisz A. Wireless Technology in Industrial Networks. Proceedings of the IEEE 2005;93:1130–51. https://doi.org/10.1109/JPROC.2005.849717.

[46] Shyam Parekh. IEEE 802.11 Wireless LANs.

[47] P. M. Kanabar, M. G. Kanabar, W. El-Khattam, T. S. Sidhu, A. Shami. Evaluation of communication technologies for IEC 61850 based distribution automation system with distributed energy resources. In: 2009 IEEE Power & Energy Society General Meeting; 2009, p. 1–8.

[48] G. Thonet, B. Deck. A new wireless communication platform for medium-voltage protection and control. In: IEEE International Workshop on Factory Communication Systems, 2004. Proceedings; 2004, p. 335–338.

[49] P. P. Parikh, M. G. Kanabar, T. S. Sidhu. Opportunities and challenges of wireless communication technologies for smart grid applications. In: IEEE PES General Meeting; 2010, p. 1–7.

[50] Felician A. The WiMAX Technology. Oeconomics of Knowledge 2010;2.

[51] Azeem S, Sharma S. Wireless Cellular Technologies and Convergence 2017;5:766. https://doi.org/10.17762/ijritcc2321-8169.1705148.

[52] Espina J, Falck T, Mülhens O. Network Topologies, Communication Protocols, and Standards. In: Body Sensor Networks; 2007, p. 145–182.

[53] Keith Stouffer, Victoria Pillitteri, Suzanne Lightman, Marshall Abrams, Adam Hahn. Guide to Industrial Control Systems (ICS) Security: Supervisory Control and Data Acquisition (SCADA) Systems, Distributed Control Systems (DCS), and Other Control System Configurations such as Programmable Logic Controllers (PLC). NIST Special Publication 800-82; 2015.

[54] Nivedita Bisht, Sapna Singh. ANALYTICAL STUDY OF DIFFERENT NETWORK TOPOLOGIES. International Research Journal of Engineering and Technology (IRJET) 2015;2(01).

[55] John Peter, Timo Perttunen. NETWORK TOPOLOGIES; 2014.

[56] Clarke G, Reynders D, Wright E. Practical Modern SCADA Protocols: DNP3, 60870.5 and Related Systems; 2004.

[57] Gary W. Scheer. Comparison of Fiber-Optic Star and Ring Topologies for Electric Power Substation Communications. In: ; 1999.

[58] VANDANA NIGAM. TOPOLOGIES; 2017.

[59] K.SARASWATHI. PLC AND DATA ACQUISITION SYSTEMS. Enathur, Kanchipuram; 2017.

[60] Lindig S, Louwen A, Moser D, Topic M. Outdoor PV System Monitoring—Input Data Quality, Data Imputation and Filtering Approaches. Energies 2020;13. https://doi.org/10.3390/en13195099.

[61] IEC 61724-1:2021. Photovoltaic system performance - Part 1: Monitoring(2): IEC; 2021.

[62] Oprea S, Adela B. Key Technical Performance Indicators for Power Plants. In: ; 2017.

[63] Jahn U, Nasse W. Performance analysis and reliability of grid-connected PV systems in IEA countries. Proceedings of 3rd World Conference on Photovoltaic Energy Conversion 2003;3. https://doi.org/10.1109/WCPEC.2003.1305009.

[64] SERENDI-PV Consortium. Smooth, Reliable and Dispatchable Integration of PV in EU Grids:

Key Performance Indicators (KPIs) on state of the art of PV reliability, performance, profitability and grid integration; 2020.

[65] Leloux J, Taylor J, Moretón R, Narvarte L, Trebosc D, Desportes A. Monitoring 30,000 PV systems in Europe: Performance, Faults, and State of the Art; 2015.

[66] Hemsley KE, E. Fisher R. History of Industrial Control System Cyber Incidents. United States; 2018.

[67] F. Daryabar, A. Dehghantanha, N. I. Udzir, Nor Fazlida binti Mohd Sani, S. bin Shamsuddin. Towards secure model for SCADA systems. In: Proceedings Title: 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec); 2012, p. 60–64.

[68] Walker A, Desai J, Saleem D, Gunda T. Cybersecurity in Photovoltaic Plant Operations. United States; 2021.

[69] F. Daryabar, A. Dehghantanha, N. I. Udzir. Investigation of bypassing malware defences and malware detections. In: 2011 7th International Conference on Information Assurance and Security (IAS); 2011, p. 173–178.

[70] A. Volkova, M. Niedermeier, R. Basmadjian, H. de Meer. Security Challenges in Control Network Protocols: A Survey. IEEE Communications Surveys & Tutorials 2019;21(1):619–39. https://doi.org/10.1109/COMST.2018.2872114.

[71] Timothy M. Yardley. SCADA: Issues, Vulnerabilities, and FutureDirections. login Usenix Mag 2008;33.

Appendix A: Data Acquisition Frequency and Types

Component	Variable	Data sampling time interval
Inverter	DC current/voltage/power input	1 min.
	Output current per phase	1 min.
	Output voltage per phase	1 min.
	Output power per phase	1 min.
	Total AC Power output	1 min.
	Power factor	1 min.
	Energy produced daily	10 min.
	Total energy produced	10 min.
	Local network frequency	1 min.
	Operating time	Once in a day
	Inverter Serial Number	Once in a day
	Grounding Resistance	1 min.
	Inverter Temperature	1 min.
	Status/ Alarms	Whenever event occurs
	Grounding failure	Whenever event occurs
	Inverter internal failure	Whenever event occurs
	Temperature protection activation	Whenever event occurs
	Inverters On/Off switch indication	Whenever event occurs
Power Analyser	Active and reactive energy	1 sec.
	Active and reactive power	1 sec.
	Grid voltage V1, V2, V3	1 sec.
	Grid Voltage V12, V23, V31	1 sec.
	Current phase I1, I2, I3	1 sec.
	Grid frequency	1 sec.
	Power factor	1 sec.
Energy Meter	Active and reactive energy	1 min.
	Active and reactive power	1 min.
	Grid voltage V1, V2, V3	1 min.
	Grid Voltage V12, V23, V31	1 min.
	Current phase I1, I2, I3	1 min.
	Grid frequency	1 min.
	Power factor	1 min.
	Alarms	1 min.
MV Feeders Energy Analyser	Active and reactive energy	1 min.
	Active and reactive power	1 min.
	Voltage V1, V2, V3	1 min.
	Voltage V12, V23, V31	1 min.
	Current phase I1, I2, I3	1 min.
	Grid Frequency	1 min.

	Power Factor	1 min.
	Alarms	1 min.
Building I/O Signals	Circuit Breaker status	Whenever event occurs
	Line Disconnector switch	Whenever event occurs
	Earth Switch status	Whenever event occurs
	Switchgear pressure status	Whenever event occurs
Power Plant Controller	Active and reactive energy	1 min.
	Active and reactive power	1 sec.
	Grid voltage V1, V2, V3	1 sec.
	Grid Voltage V12, V23, V31	1 sec.
	Current phase I1, I2, I3	1 sec.
	Grid frequency	1 sec.
	Operation/Controlling mode	1 min.
	Operation/Commands	1 min.
	Operation/Set-points	1 min.
	Plant status	1 min.
	Alarms	1 min.
	Status	1 min.
Protection Relay	Active and reactive power	1 min.
	Grid voltage V1, V2, V3	1 min.
	Grid Voltage V12, V23, V31	1 min.
	Current phase I1, I2, I3	1 min.
	Grid frequency	1 min.
	Power Factor	1 min.
	Local/remote status	Whenever event occurs
	27, 50, 50N, 59, 81R, 81U/O protection status	Whenever event occurs
	Trip and/or alarm reset	Whenever event occurs
	Alarms and errors	Whenever event occurs
Battery Energy Storage System	AC input voltage	1 min.
	AC input current	1 min.
	DC output voltage	1 min.
	DC output current	1 min.
	DC output power	1 min.
	Battery level	1 min.
	Internal Temperature	1 min.
	Main Status	Whenever event occurs
	Faults status	Whenever event occurs
	Alarms	Whenever event occurs
	Alarms	Whenever event occurs
Weather Station and Sensors	Irradiance	1 sec.
	Ambient temperature	1 min.
	Photovoltaic module temperature	10 sec.

Humidity	10 sec.
Barometric pressure	1 min.
Wind speed	1 min.
Wind direction	1 min.

Appendix B: Calculation Method and Storage of Acquired Variables

Type	Value	Formula	Screening
$val_{minute}T$	Sum	$val_{sum} = val_{minute}T - val_{minute0}$	Cumulative values
$val_{avg.}$	Average	$val_{avg} = \frac{\sum_{i=1}^{10} val\_minute(i)}{10}$	During every 10 minutes with 1-minute intervals
val_{max}	Maximum	$val_{max} = Max([val_{minute1}; val_{minute10}])$	During every 10 minutes with 1-minute intervals
val_{min}	Minimum	$val_{min} = Min([val_{minute1}; val_{minute10}])$	During every 10 minutes with 1-minute intervals
$val_{avg.}10$	Average	$val_{avg} = \frac{\sum_{i=00:10}^{01:10} val\_10minute(i)}{6}$	During every 10 minutes with 1-minute intervals
$val_{avg.}day$	Average	$val_{avg} = \frac{\sum_{i=01:00}^{00:00} val\_hour(i)}{24}$	During every day with 1-day intervals
$val_{avg.}day$	Sum	$val_{sum} = val_{minute}n - val_{minute1}$	During every day with 1-minute intervals
$val_{avg.}month$	Average	$val_{avg} = \frac{\sum_{i=1}^{31} val_{day}(i)}{31}$	During every month with 1-day intervals
$val_{avg.}year$	Average	$val_{avg} = \frac{\sum_{i=1}^{12} val_{month}(i)}{12}$	During every year with 1-month intervals