Performance Analysis of Grid-Connected Photovoltaic Systems

by

Walid Omran

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2010

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Solar energy is one of the most promising renewable resources that can be used to produce electric energy through photovoltaic process. A significant advantage of photovoltaic (PV) systems is the use of the abundant and free energy from the sun. However, these systems still face major obstacles that hinder their widespread use due to their high cost and low efficiency when compared with other renewable technologies. Moreover, the intermittent nature of the output power of PV systems reduces their reliability in providing continuous power to customers. In addition, the fluctuations in the output power due to variations in irradiance might lead to undesirable performance of the electric network. The support of governments, electric utilities, researchers and consumers is the key to overcoming the aforementioned obstacles and enhancing the maturity of the technology in this field.

The primary objective of the research proposed in this thesis is to facilitate increasing the penetration levels of PV systems in the electric network. This can be achieved by quantifying and analyzing the impacts of installing large grid-connected photovoltaic systems on the performance of the electric network accurately. To achieve this objective, the development of a new and intelligent method is introduced. The method utilizes the available data efficiently to produce accurate realistic results about the performance of the electric network without overestimating or underestimating the impacts of the PV system. The method utilizes historical environmental data collected over a number of years to estimate the profile of the output power of the PV system. In addition, the method considers the actual data of the electric network. Hence, the interaction between the output power of the PV system and the electric network components can be simulated to identify the possible operational problems.

After identifying the operational problems that might arise due to installing PV systems, especially due to power fluctuations, different strategies that can mitigate these problems are studied in detail. These strategies include installation of energy storage devices, use

of dump loads, and operation below the maximum power point. Upon studying the mitigation strategies, their economical aspects are investigated. The economical aspect is crucial for PV systems because of their high cost, which is reflected on the price of the energy produced by them.

The presented research integrates techniques from different fields of engineering such as data mining, mathematical optimization and power systems. This research is expected to contribute to the advancement of PV technology by introducing methods that will help in carrying out in-depth evaluation of the performance of PV systems and providing feasible solutions to the operational problems that might arise from the installation of these systems.

Acknowledgements

Thank God for helping me achieve this work.

I then would like to thank my supervisors, Professor Mehrdad Kazerani and Professor Magdy Salama for their continuous guidance throughout the period of my PhD studies. Their valuable suggestions and discussions were always helpful and inspiring. Also, their support and encouragement were my greatest motive to aim for the best.

I would also like to thank my Ph.D. committee members: Professor Jatin Nathwani, Professor Siva Sivoththaman, Professor Kankar Bhattacharya, and the external examiner Dr. Josep Guerrero from Universitat Politecnica de Catalunya.

My appreciation is also extended to Professor Mohamed Kamel, Professor Miguel Anjos, and Professor Ramadan A El Shatshat for their useful discussions. Also, many thanks to my colleagues in the power group for their help and support.

I would like to show my deepest gratitude and respect to my family, especially my parents, the ones to whom I owe all the success in my life. No words can express my gratitude to them, but I pray God to bless them and reward them.

Million thanks to my little son who, despite his young age, accepted trading our playing time together with the research time. I hope I will be able to make it up to him.

A final word to my wife; without you I could have never been able to achieve this work. Your patience and encouragement were always a source of strength for me. You are the shining moon that lightens my life.

Table of Contents

List of Figures

List of Tables

List of Symbols

$a_{i,j}$ The i^{th} transformation coefficient for the j^{th} segment

$\boldsymbol{A}$ projection matrix that projects the original data onto the new basis

Aa Area of the PV array

Aa_i Average distance between the i^{th} segment and all other segments in the same cluster

APE Absolute Percentage Error

Ba_i Average distance between the i^{th} segment and all other segment not included in the same cluster

BS Battery Storage

C_E Capital cost related to the battery energy capacity

C_P Capital cost related to the battery power capacity

C_i Centroid of the i^{th} cluster

Ca_i Average distance between all segments in cluster i and their centroid

$CC_{i,j}$ Distance between the centroids of clusters i and j

CF Capacity factor of the PV system

$\boldsymbol{Cov}$ Covariance matrix of the Pattern matrix $\boldsymbol{X}$

d Integer representing the day of the year

d_r Average annual market discount rate

DBI The Davis-Bouldin index

$Dist(i,k)$ Euclidean distance between the i^{th} and k^{th} segments

DL Dump Load

E_B Energy discharged from the battery

$E_{B\,max}$ Maximum energy storage capacity of the battery

E_{loss} Energy loss

Et Equation of time

$\boldsymbol{F}$ Matrix containing the mean of each feature

$f(t)$ Mean of the samples at each time t

f_{ij}	j^{th} feature of the i^{th} segment
G_b	Direct (beam) irradiance received by a horizontal surface
G_{bn}	Direct (beam) normal irradiance received by a surface perpendicular to the sun rays
$G_{bn\text{-}m}$	Measured direct (beam) normal irradiance on a surface perpendicular to the sun rays
G_d	Diffuse irradiance received by the horizontal surface
G_g	Global irradiance received by a horizontal surface
$G_{g\text{-}m}$	Measured global irradiance on a horizontal surface
G_o	Extraterrestrial irradiance received on a horizontal surface
G_{on}	Normal extraterrestrial irradiance
G_{sc}	The solar constant
G_{tg}	Global irradiance received by a tilted surface
$G_{tg\text{-}m}$	Measured global irradiance received by a tilted surface
$\overline{G}_{tg-m}$	Measured average daily irradiance
G_{tb}	Beam irradiance received by a tilted surface
G_{td}	Diffuse irradiance received by a tilted surface
G_{tr}	Reflected irradiance or albedo
K_T	Thermal derating coefficient
L	Number of significant principle components
L_{st}	Standard meridian for the local time zone
LA	Lead Acid
M	Number of data points in each pattern
MAE	Mean Absolute Error
$MAPE$	Mean Absolute Percentage Error
MBE	Mean Bias Error
$MPPT$	Maximum power point tracking
m_x	Rows of the Pattern matrix $\boldsymbol{X}$ (features)
N	Number of PV power patterns in each year
N_B	Number of times the battery will be replaced during period T ($N_B = {}^{T}/_{T_B}$)

N_i Number of intervals

n_C Total number of clusters

n_f Number of features

n_s Total number of segments

n_x Columns of the Pattern matrix $\boldsymbol{X}$ (patterns)

NaS Sodium Sulfur

NOCT Nominal operating cell temperature

O_E Annual operating cost related to the battery energy storage capacity

O_P Annual operating cost related to the battery power capability

P_B Power generated from the battery, +ve: discharging & -ve: charging

$P_{B\,max}$ Maximum power capability of the battery

P_C Power curtailed from the PV system

P_D Power dumped from the PV system

P_{Dmax} Power rating of the dump load

P_G Power injected into the grid

P_{Gmax} Maximum power injected into the grid

P_{MPP} Maximum DC power generated from the PV system

P_{PV} AC Power generated from the PV system

$P_{PV\text{-}freq}$ Frequency of occurrence of the PV power in the i^{th} interval calculated from the original clusters

$P^*_{PV\text{-}freq}$ Frequency of occurrence of the PV power in the i^{th} interval calculated from the virtual clusters

$P_{X\text{-}Y\_av}$ Average active power flowing in Branch X-Y

$P_{X\text{-}Y\_freq}$ Frequency of occurrence of active power flowing in Branch X-Y

$P_{X\text{-}Y\_max}$ Maximum active power flowing in Branch X-Y

PC Principal component

PCA Principal component analysis

PCU Power Conditioning Unit

PV Photovoltaic

$Q_{X\text{-}Y\_av}$ Average reactive power flowing in Branch X-Y

$Q_{X\text{-}Y\_freq}$ Frequency of occurrence of reactive power flowing in Branch X-Y

$Q_{X\text{-}Y\_max}$ Maximum reactive power flowing in Branch X-Y

R_{PV} Profit gained from the PV system over period T

$R_{PV/BS}$ Profit gained from the PV/BS system over period T

$R_{PV/BS\text{-}PC}$ Profit gained from the PV/BS system while curtailing the power over period T

$R_{PV/DL}$ Profit gained from the PV/DL system over period T

$R_{PV/PC}$ Profit gained from the PV system over period T while curtailing the power

$RMAE$ Relative Mean Absolute Error

$RMBE$ Relative Mean Bias Error

S_E Salvage value related to the battery energy storage capacity ($/kWh)

s_k k^{th} PV power segment

S_P Salvage value related to the battery power capability ($/kW)

SC Partition Index

SI Silhouette index

Sil_i Silhouette coefficient of the i^{th} segment

SOC_{min} Minimum allowable state of charge

SSE Sum of Squared Errors

STC Standard Test Conditions

T Number of years studied

T_a Ambient Temperature

T_B Lifetime of the battery (years)

T_c Cell temperature

x Lower limit for power fluctuations (kW)

$\boldsymbol{X}$ Pattern matrix

$\boldsymbol{X}'$ Reconstructed pattern matrix

$x'_j(t)$ j^{th} reconstructed segment

y Upper limit for power fluctuations (kW)

V_{Z_av} Average voltages at Node Z

V_{Z_freq} Frequency of occurrence of the voltage at Node Z

V_{Z_max} Maximum voltage at Node Z

$v_i(t)$ i^{th} principal component

$\boldsymbol{Z}$ Pattern matrix after removing the mean

α Price of energy sold from the PV system ($/kWh)

α_s Solar altitude

β Present worth related to the battery power capability ($/kW)

γ Present worth corresponding to the battery energy storage capacity ($/kWh)

δ Present worth corresponding the dump load power capacity ($/kW)

δ_s Declination angle

ΔR Percentage change in profit

Δt Minute resolution of data points (minute)

η_{ch} Energy efficiency of the battery

η_{DCloss} 1 – the fractional power loss in the DC-side

η_{dust} 1 – the fractional power loss due to dust on the PV array

η_{inv} Inverter efficiency,

η_m PV module efficiency

$\eta_{mismatch}$ 1 – the fractional power loss due to module parameter mismatch

η_{MPP} 1 – the fractional power loss due to DC current ripple and MPPT algorithm error caused by the switching converter

η_o Rated module efficiency at Standard Test Conditions

η_{PB} Power efficiency of the battery charger

θ_i Angle of incidence

θ_Z Zenith angle

λ_i i^{th} eigenvalue of the covariance matrix $\boldsymbol{Cov}$

μ_{ik}	Fuzzy membership
ρ	Ground reflectance
τ	Tilt angle of the PV array
φ_{lat}	Latitude of the location
φ_{long}	Longitude of the location
ω	Hour angle

Chapter 1

Introduction

1.1 General

The aim of this chapter is to present the motivation behind the work done in this thesis. The chapter also provides the main objectives of the research as well as the thesis organization.

1.2 Motivation

In the near future, the demand for electric energy is expected to increase rapidly due to the global population growth and industrialization. This increase in the energy demand requires electric utilities to increase their generation. Recent studies predict that the world's net electricity generation is expected to rise from 17.3 trillion kilowatt-hours in 2005 to 24.4 trillion kilowatt-hours (an increase of 41%) in 2015 and 33.3 trillion kilowatt-hours (an increase of 92.5%) in 2030 [1]. Currently, a large share of electricity is generated from fossil fuels, especially coal due to its low prices. However, the increasing use of fossil fuels accounts for a significant portion of environmental pollution and greenhouse gas emissions, which are considered the main reason behind the global warming. For example, the emissions of carbon dioxide and mercury are expected to increase by 35% and 8%, respectively, by the year 2020 due to the expected increase in

electricity generation [2]. Moreover, possible depletion of fossil fuel reserves and unstable price of oil are two main concerns for industrialized countries.

To overcome the problems associated with generation of electricity from fossil fuels, renewable energy sources can be participated in the energy mix. One of the renewable energy sources that can be used for this purpose is the light received from the sun. This light can be converted to clean electricity through the photovoltaic process. The use of photovoltaic (PV) systems for electricity generation started in the seventies of the 20th century and is currently growing rapidly worldwide. In fact, many organizations expect a bright future for these systems. For example, the European Photovoltaic Industry Association (EPIA) expects that the global cumulative PV capacity will reach 200 GW by the year 2020 and 800 GW by the year 2030 [3], as illustrated in Figure 1-1. Also, a recent survey presented by the International Energy Agency (IEA) reports that over 2.26 GW of PV capacity was installed during the year 2007 in the IEA reporting countries. This shows an increase of more than 50 % over the previous year, and brings the total installed capacity in these countries to about 7.8 GW [4].

PV systems are usually used in three main fields: 1) satellite applications, where the solar arrays provide power to satellites, 2) off-grid applications, where solar arrays are used to power remote loads that are not connected to the electric grid, and 3) on-grid, or grid-connected applications, in which solar arrays are used to supply energy to local loads as well as to the electric grid. Grid-connected PV systems currently dominate the PV market, especially in Europe, Japan and USA. For example, in 1990 only 27% of the cumulative installed PV capacity was connected to the grid, but by the end of 2007 this percentage increased to more than 90%, as illustrated in Figure 1-2.

Grid-connected PV systems can be installed on the facades and rooftops of buildings, on the shades of parking lots. They can also be installed as power plants that aim to inject all their produced power into the grid. Table 1-1 shows a variety of grid-connected PV projects that have been recently installed or are planned to be installed in different countries around the world.

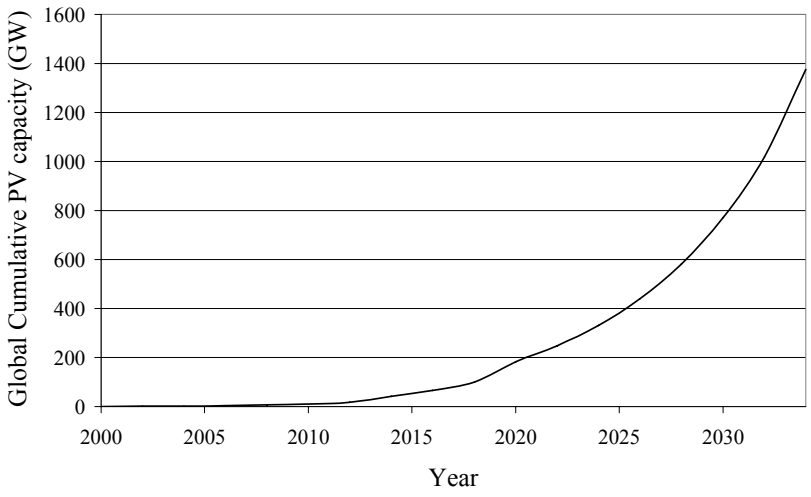

Figure 1-1 Expected global cumulative PV capacity based on EPIA data [3]

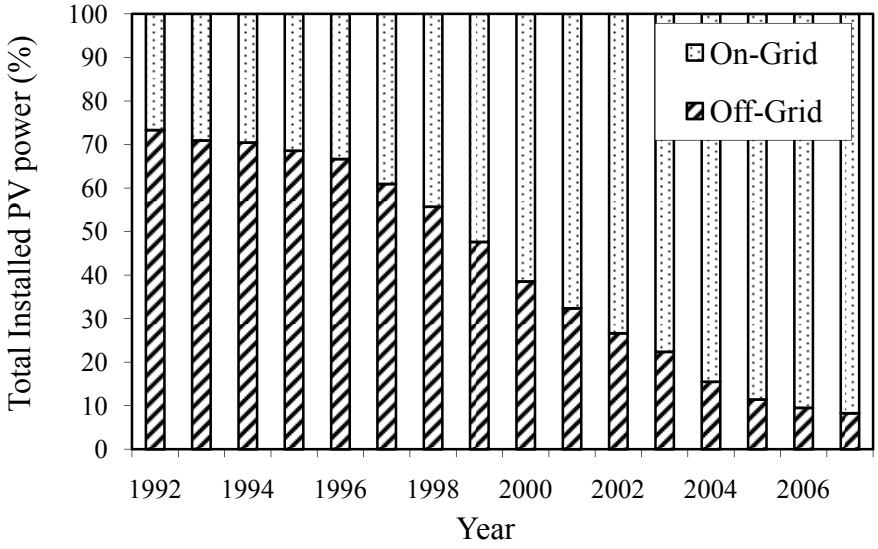

Figure 1-2 Percentages of on-grid and off-grid PV power in the IEA reporting countries [4]

Despite the increasing use of PV systems, these systems still face a major obstacle due to their high capital cost, which is reflected in the cost per KWh of the energy produced by them. This obstacle can be overcome by utilizing the recent technology in developing low cost PV cells and by providing incentives to customers that tend to install these systems.

Another major issue that faces the widespread of PV systems is that the increasing installation of grid-connected PV systems, especially large systems in the order of megawatts, might lead to some operational problems in the electric network. This issue, which is the main focus of the research presented in this thesis, can be tackled by accurately evaluating the impacts of installing PV systems on the performance of the grid and finding solutions that can reduce the operational problems that might arise due to their installation.

Table 1-1 Sample PV projects worldwide

	Project	**Size**	**Location**
1	facades of a skyscraper	14 kW_P[1]	New York
2	subway terminal station at Coney Island	210 kW_P	
3	500 PV houses and some public facilities	1 MW_P	Netherlands
4	Mont-Cenis Academy	1 MW_P	Germany
5	Springerville PV power plant	4.59 MW_P	USA
6	Bavaria solar park	6.3 MW_P	Germany
7	Monte Alto PV power plant	9.5 MW_P	Spain
8	Serpa solar power plant	11 MW_P	Portugal
9*	Sarnia PV power plant	4×10 MW_P	Canada
10**	Nevada PV power plant	18 MW_P	USA
11**	Waldpolenz solar park	40 MW_P	Germany
12**	Girrasol solar power plant	62 MW_P	Portugal
13**	Solar power station in Victoria	154 MW_P	Australia

[1] W_P is the peak power that can be produced by the PV system.
* To be extended to 60 MW
** Project under construction

1.3 Research Objectives

The goal of the research presented in this thesis is to help increasing the penetration level of PV systems in the electric network. This goal can be achieved by accurately evaluating the performance of the PV system without overestimating or underestimating its impacts on the electric network. Upon performing this analysis, the impacts of the fluctuating output power of PV systems should be considered. The main reason behind this consideration is the intermittent nature of the output generated from these systems. Another important factor that can help increasing the penetration level of PV systems is to investigate the suitability of different methods that can improve the performance of the PV system and mitigate its negative impacts, especially due to power fluctuations. Accordingly, the main objectives of the research can be summarized as follows:

1. To develop a method that considers the fluctuations in the PV output power when studying the impacts of installing large PV systems on the performance of distribution networks prior to installing these systems. The method used in such a study can provide accurate evaluation of the possible impacts by covering the following aspects:
    - Estimating the profile of the output power of the PV system using long historical time series data of irradiance and temperature. Hence, the estimated profile will retain information about the fluctuations present in the output power of the system, and thus, can be utilized in the chronological simulations.
    - Considering the actual data of the electric network in the analysis in order to provide realistic results about the performance of the network.
    - Utilizing the long historical time series data of the output power generated from the PV system efficiently by reducing its size while retaining the useful information it contains.
    - Applying statistical analysis in order to identify the periods during which there is high probability of undesirable performance of the electric network. Hence,

proper operational plans can be prepared to solve the expected problems that might arise in the future.

2. To study the possible solutions that can be adopted to reduce the negative impacts of installing a large PV system on the performance of the electric network, especially due to power fluctuations. These solutions include:
   - Selecting the appropriate penetration level of the PV system for a specific distribution network.
   - Selecting the suitable location for installing the system, from a number of candidate locations.
   - Studying the economical value of using energy storage devices, such as batteries, to control the power injected into the grid.
   - Studying the effectiveness of operating the PV system below the maximum power point and its economical impact

1.4 Thesis Outline

To achieve the aforementioned objectives and facilitate the presentation of the results obtained in this research, the thesis is organized as follows:

1- Chapter 2 provides a general overview of grid-connected PV systems. This chapter looks at different components of a grid-connected PV system from the solar irradiance received by the PV arrays to the AC power injected into the grid, and makes a review of the recent achievements and current research activities in the field.

2- Chapter 3 discusses in detail the positive and negative impacts of PV systems on the performance of electric networks. This chapter also presents the methods that can be used for this purpose and discusses their suitability to accurately assess the impacts of power fluctuations generated from PV systems.

3- Based on the review presented in Chapter 3, Chapter 4 introduces the layout of a new method that can be used to investigate the impacts of PV systems on the operation of the electric network, especially due to power fluctuations. The details of each stage of the new method are also presented in this chapter.

4- In Chapter 5, the new method is applied to investigate the performance of a distribution feeder in the presence of a large PV system. Hence, the ability of the method to identify the impacts of power fluctuations and estimate the annual energy loss in the feeder are highlighted. Moreover, the results obtained from the new method are compared with other methods that can be used for the same purpose.

5- After analyzing the impacts of power fluctuations generated from large PV systems, Chapter 6 presents methods that can be used to reduce the impacts of these fluctuations. The main emphasis in this chapter is to quantify the effect of implementing these methods on the economical benefits that the PV system owner gains from selling the PV energy to the grid. To achieve this target, an LP optimization problem is formulated and solved to estimate the maximum profit gained by the system owner. Moreover, the effect of varying different parameters of the problem is investigated through a sensitivity analysis.

6- In the last chapter, Chapter 7, the summary and conclusion of the thesis are presented. Moreover, this chapter outlines the contributions of the presented research.

Chapter 2

Grid-Connected PV Systems: An Overview

2.1 General

This chapter provides an overview of different components of a photovoltaic system and attempts to highlight recent research activities in the field of photovoltaics. More details about the literature relevant to the research presented in this thesis are introduced in later chapters whenever suitable.

2.2 Components of Grid-Connected PV Systems

The building blocks of a grid-connected photovoltaic system are shown in Figure 2-1. The system is mainly composed of a matrix of PV arrays, which converts the sunlight to DC power, and a power conditioning unit that converts the DC power to AC power. The generated AC power is injected into the grid and/or utilized by the local loads. In some cases, storage devices are used to improve the availability of the power generated by the PV system. In the following sub-sections, more details about different components of the PV system are presented and the recent related research activities are discussed.

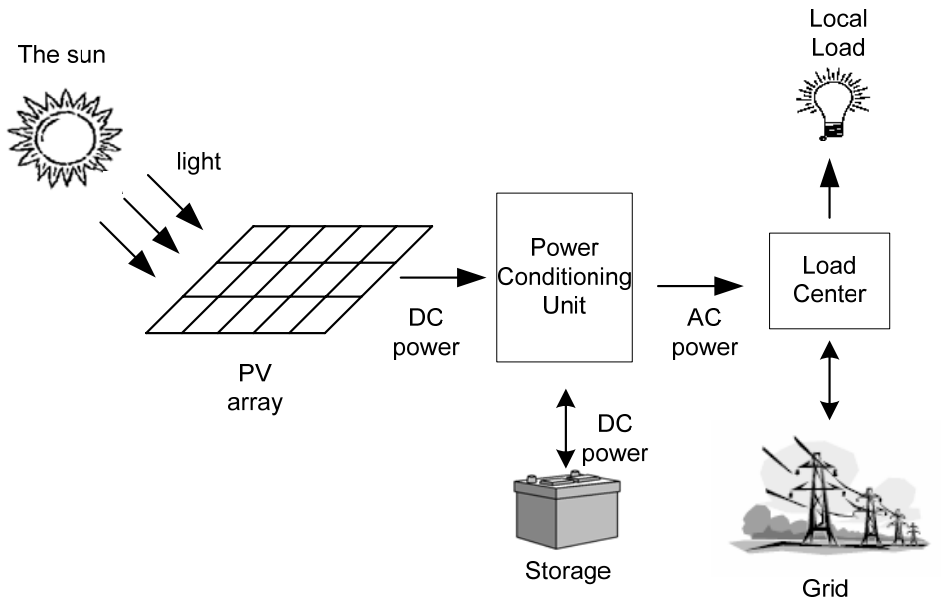

Figure 2-1 Main components of grid-connected photovoltaic systems

2.2.1 The light from the sun

Irradiance or insolation is the instantaneous solar power received on a unit surface area and is normally given in W/m^2. The *global irradiance*, G_g, that reaches a horizontal surface on the earth is the sum of two components [5]: 1) *direct (beam) irradiance,* G_b, that directly reaches the horizontal surface without being scattered by the atmosphere and; 2) *diffuse irradiance,* G_d, that reaches the horizontal surface after being scattered by clouds. Weather stations usually measure the global horizontal irradiance by a Pyranometer placed horizontally at the required location. On the other hand, a Pyrheliometer is used to measure the direct normal irradiance, which is the irradiance received by a surface that is perpendicular to the sun rays. Accordingly, the direct irradiance on the horizontal surface can be calculated. To measure the diffuse irradiance, a shading ball or ring can be used to permanently shade the Pyranometer. Figure 2-2 shows pictures of the three devices used to measure different irradiance components [6].

Figure 2-2 Pyranometer (left top), Two-axis tracked Pyrheliometer (left bottom), Pyranometer with shading ball (right)

PV arrays are usually tilted to maximize the energy production of the system by maximizing the direct irradiance that can be received. Usually the optimum tilt angle with respect to the horizontal surface of the earth is calculated for each specific site; however, it can be roughly set within ± 15^{o} of the site latitude [7]. Thus, the irradiance components received by the tilted surface of the PV array are different from those provided by the weather stations. Accordingly, different models must be used to estimate the different irradiance components on the surface of the PV array from those provided by the weather stations. The accuracy of these models is mainly dependant on the location under study. Further discussions about calculating the different components of irradiance on the surface of the PV array are presented in Chapter 4.

Currently, one of the main research activities in this area focuses on analyzing the short-term fluctuations of irradiance due to passage of clouds. Some of these studies use frequency domain analysis to investigate the smoothing effect of extended area of the PV system on the fluctuation of irradiance [8]. Other studies use frequency domain analysis to analyze the amplitude, and persistence of these fluctuations [9].

Another research activity related to this field focuses on developing models for the different irradiance components at a certain location by using either cloud observations obtained from weather stations [10] or images obtained from satellites [11][12]. These models are important for predicting the output energy produced from PV systems installed at these locations. Short-term prediction of solar irradiance from historical time-series data is very important for short-term planning related to the operation of electric networks in the presence of PV systems, especially in the case of large systems. Different methods, such as ARMA models and neural networks have been used for this purpose [13]-[15]. However, the research in this field still needs more work to become as mature and well-established as wind speed prediction. In fact, predicting the solar irradiance is a complicated task as it is affected by many factors such as types of clouds, cloud heights, wind speed, and wind direction.

2.2.2 PV arrays: technology and modeling

The first silicon solar cell with an efficiency of 6% was developed at Bell Telephone Laboratories in 1954 by Chapin et al. [16]. Nowadays, an efficiency of 18% can be reached and different types of materials are used in manufacturing these cells [17]. However, the most widely used cells are polycrystal silicon cells (54.5% of the world's market share) and single crystal silicon cells (29.36% of the world's market share) [17]. Normally, the electric characteristics of a PV cell are displayed as a relation between the cell voltage and current, and a relation between the cell voltage and power. Accordingly, several electric quantities that are important to the operation of the PV system are identified. These electric quantities include: the cell voltage under open circuit conditions, V_{OC}, the cell current under short circuit conditions, I_{SC}, and the cell voltage, current and power at the maximum power point, V_{MPP}, I_{MPP}, and, P_{MPP}, respectively. Figure 2-3 and Figure 2-4 display the electric characteristics of a common PV cell.

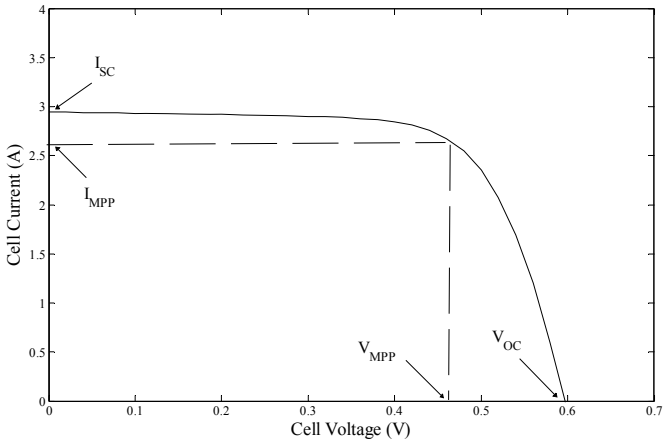

Figure 2-3 I-V characteristics of a single PV cell

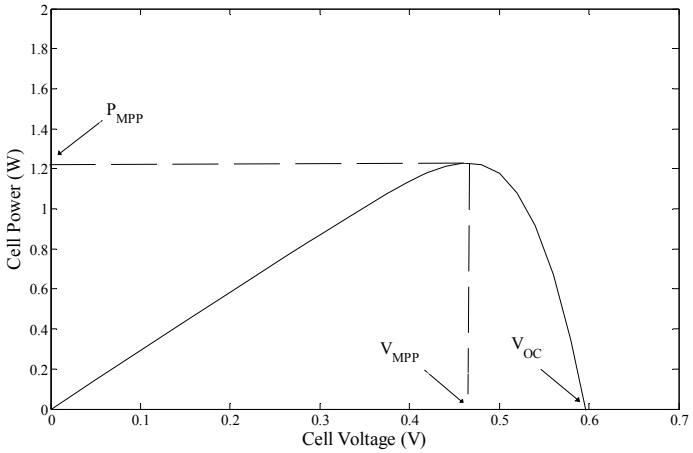

Figure 2-4 P-V characteristics of a single PV cell

The electric characteristics of the PV cell depend mainly on the irradiance received by the cell and the cell temperature. Figure 2-5 displays the electrical characteristics of the cell at different levels of the irradiance and constant temperature. It is clear that the change in irradiance has a strong effect on the short-circuit current and output power of the cell, but negligible effect on the open-circuit voltage. On the other hand, Figure 2-6 shows that the change in temperature at constant irradiance has a strong effect on the open-circuit voltage and output power of the cell, but negligible effect on the short-circuit current.

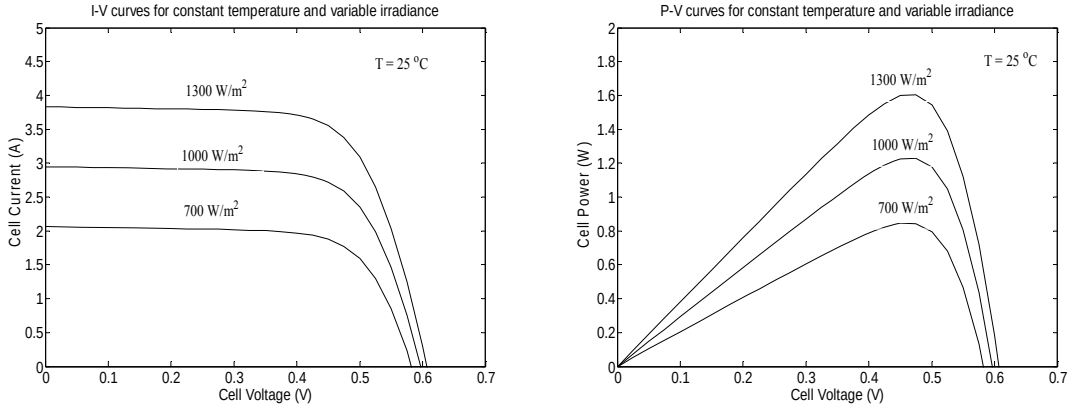

Figure 2-5 Characteristics of the PV cell at constant temperature and variable irradiance

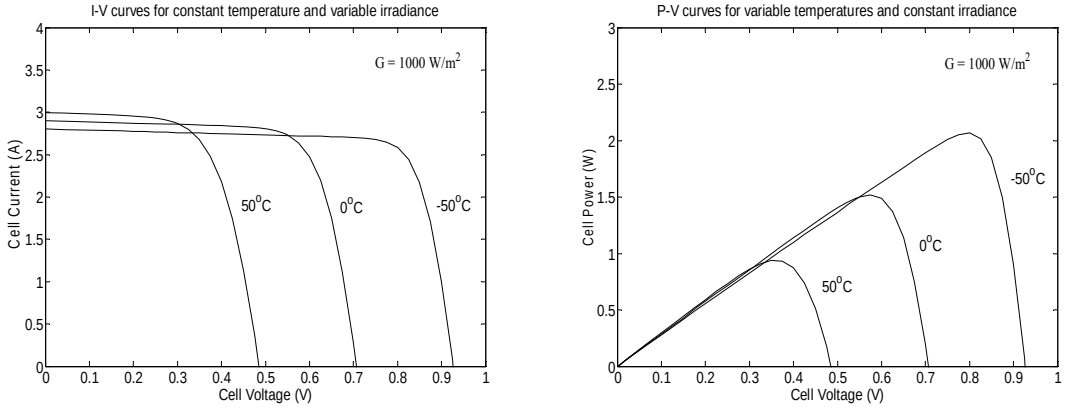

Figure 2-6 Characteristics of the PV cell at variable temperature and constant irradiance

Usually solar cells are connected in series to form a solar module and modules are then connected in series to form a string. Finally, the strings are connected in parallel to form a PV array. The number of modules in each string is specified according to the required voltage level of the array. On the other hand, the number of strings is specified according to the required current rating of the array. Most PV arrays have a power diode, called *bypass* diode, connected in parallel with each individual module or a number of modules. The function of this diode is to conduct the current when one or more of these modules

are damaged or shaded. Another diode, called *blocking* diode, is usually connected in series with each string to prevent reverse current flow and protect the modules. The layout of a PV array is illustrated in Figure 2-7.

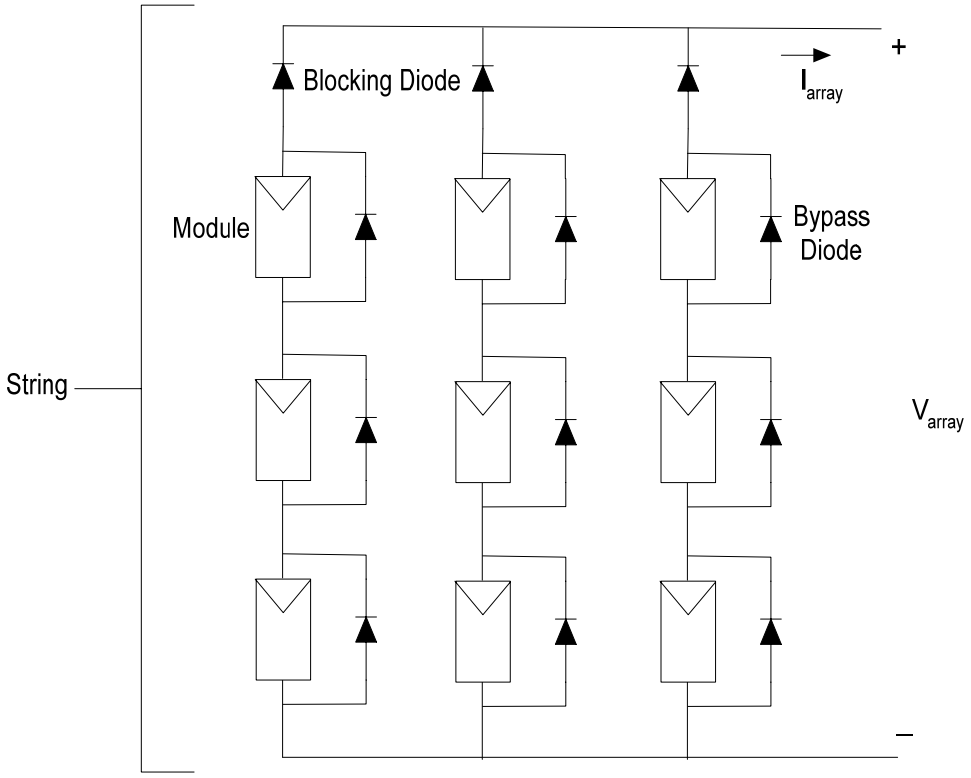

Figure 2-7 Layout of a PV array

The Solar array is the most expensive component in the PV system. The average cost of PV modules is $4.00–5.00/W [18][19]; thus, most of the research activities performed in this area are concerned with manufacturing low-cost solar cells with acceptable efficiencies [20][21]. The advances in this field of research will have a great impact on the large-scale use of PV systems.

Most studies related to the performance of PV systems require the use of a model to convert the irradiance received by the PV array and ambient temperature into the corresponding maximum DC power output of the PV array, P_{MPP}. The models recorded in the literature vary in accuracy and complexity, and thus, appropriateness for different studies. The single-diode model shown in Figure 2-8 is one of the most popular physical models used to represent the electric characteristics of a single PV cell [22]-[24]. The model consists of: 1) a current source, I_{ph}, representing the light-induced current generated in the cell due to the separation and drift of the electron-hole pairs produced by incident photons from the sun, 2) a shunt diode representing the p-n junction of the PV cell, 3) a shunt resistance, R_{sh}, accounting for the leakage currents due to the impurities of the p-n junction (the value of this resistance should be made as high as possible), and 4) a series resistance, R_s, representing all the distributed ohmic resistances in the semiconductor and the resistances of the metallic contacts (ideally, the value of this resistance should be zero).

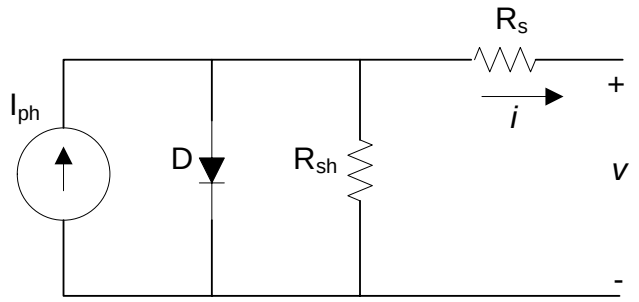

Figure 2-8 Single-diode model of a PV cell

The accuracy of the single-diode model can be further improved by replacing the single diode with two diodes connected in parallel. The first diode represents the diffusion current in the quasi-neutral region of the junction and has an ideality factor of 1. The second diode represents the generation-recombination in the space-charge region of the junction and has an ideality factor of 2 [20]. The main drawback of this model is the

increased complexity of the relationship between the output voltage and current of the cell due to the existence of two diode equations.

The identification of the parameters of the single-diode model from the data sheets of PV cells and the effect of irradiance and temperature on these parameters has been the focus of several studies [25]-[28]. Other studies propose the use of new models for the PV cells that can better represent the cell characteristics. Some of these models enhance the performance of the single-diode model by including the detailed physical processes that occur in the PV cell [29]. Other models use soft computing techniques to model the performance of PV cells under different operating conditions by training the PV model using patterns of the I-V curves at specific operating conditions [30]. Recently, a mathematical model using polynomials has been proposed to represent the performance of the PV cell [31]. This model is useful for the case when real-time identification of the maximum power point is required. However, the model has no physical meaning and its accuracy depends mainly on the available measured data obtained from the cell.

In general, physical PV models provide cell-level information, and thus, are useful for studying the details of the PV system, such as maximum power tracking algorithms and impacts of partial shading. However, these models are not suitable for studying the performance of the electric network in the presence of PV systems. This is mainly because these studies require calculating the power generated from the PV system at different weather conditions over extended periods of time. Thus, simpler models are usually preferred as the amount of calculations is highly reduced. For example, the single-diode model can be simplified by assuming that the shunt resistance is infinitely large or by removing both the shunt and series resistances [27]. Accordingly, the voltage and current are decoupled in the main equation of the model. Other simplified models directly relate the irradiance and temperature at any instant with the maximum power that can be generated from the PV system [32]-[34]. These models are usually used when the performance of the electric network is to be assessed in the presence of a PV system.

However, before using such models in the analysis, they should be validated against one of the comprehensive physical models to examine their accuracies.

2.2.3 Power Conditioning Units

Power conditioning units (PCUs) are used to control the DC power produced from the PV arrays and to convert this power to high-quality AC power before injecting it into the electric grid. PV systems can be divided, according to the number of power processing stages, into single-stage and two-stage systems. In single-stage systems, an inverter is used to perform all the required control tasks. But, in the two-stage system, a DC-DC converter precedes the inverter and the control tasks are divided among the two converters. Two-stage systems provide higher flexibility in control as compared to single-stage systems, but at the expense of additional cost and reduction in the reliability of the system [35]. During the last decade, a large number of inverter and DC-DC converter topologies for PV systems were proposed [35]-[39], almost saturating the research in this area.

In general, PCUs have to perform the following tasks:

1. *Maximum power point tracking (MPPT)*

   One of the main tasks of PCUs is to control the output voltage or current of the PV array to generate maximum possible power at a certain irradiance and temperature. There are many techniques that can be used for this purpose [40]-[42] with the Perturb-and-Observe and Incremental Conductance techniques being the most popular ones. A recent study [43] presented a qualitative comparison between 19 different MPPT techniques to serve as a guideline for choosing a suitable technique.

   Partial shading of PV arrays is considered one of the main challenges that face MPPT techniques. In this case, there might exist multiple local maxima, but only one global maximum power point, as illustrated in Figure 2-9. In this case, the task of the PCU is to identify and operate at the global MPP. The research in this field is active

and several studies have focused on developing new MPPT techniques and PCU topologies that can perform this task [44]-[46].

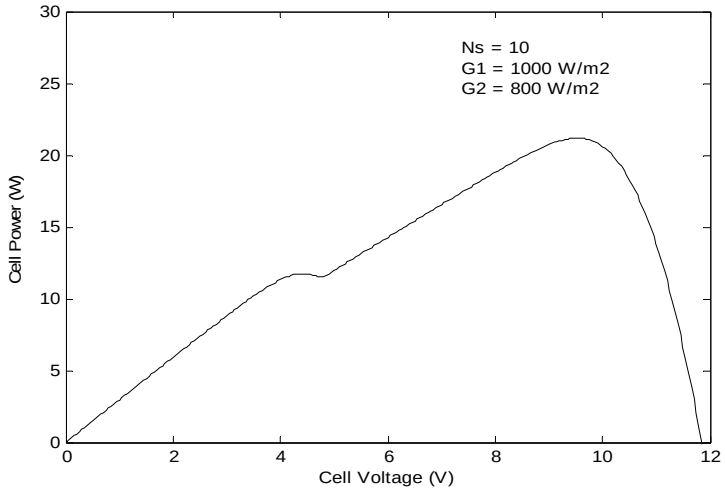

Figure 2-9 P-V curve for two series modules in case of partial shading

2. Control of the injected current

PCUs should control the sinusoidal current injected into the grid to have the same frequency as the grid and a phase shift with the voltage at the point of connection within the permissible limits. Moreover, the harmonic contents of the current should be within the limits specified in the standards. The research in this field is mainly concerned with applying advanced control techniques to control the quality of injected power and the power factor at the grid interface [47]-[49].

3. Islanding detection and protection

Islanding is defined as a condition in which a portion of the utility system containing both loads and distributed resources remains energized while isolated from the rest of the utility system [50]. Most of the standards require that PCUs of PV systems should cease injection of power into the grid under specific abnormal operating conditions of the grid including those leading to islanding [50][51]. Islanding

detection methods can be classified into three categories [52]: 1) Communication-based methods that depend on transmitting signals between the PV system and the grid to identify an islanding condition, 2) Passive methods that depend on monitoring a certain parameter and comparing it with a threshold value, and 3) Active methods that depend on imposing an abnormal condition on the grid such as injecting harmonic current with a specific order at the point of connection with the grid. Most of the recent studies have focused on assessing and comparing different islanding techniques as well as developing new methods with minimized non-detection zones [53]-[56].

4. *Voltage amplification*

Usually, the voltage level of PV systems requires to be boosted to match the grid voltage and to decrease the power losses. This task can be performed using step-up DC-DC converters or multilevel inverters. Three-level inverters can be used for this purpose as they provide a good tradeoff between performance and cost in high-voltage and high-power systems [57].

5. *Additional functions*

The control of PCUs can be designed to perform additional tasks such as power factor correction [58], harmonics filtering [59], reactive power control [60], and operating with an energy storage device and/or a dispatchable energy source such as diesel generator as an uninterruptible power supply [61].

2.2.4 Energy Storage Devices

The use of energy storage devices with PV systems is currently receiving a lot of attention, especially due to the fact that the power generated from these systems is intermittent. The installation of storage devices can enhance the performance of PV systems by bridging their power fluctuations, shifting the time of their peak generation, supplying critical loads during power outages, and providing reactive power support.

There are a variety of storage devices such as batteries, super-capacitors, super-inductors, flywheels, and water pumping. These devices vary in their characteristics, method of operation, and accordingly, the tasks that they can perform. Thus, choosing a storage device that can perform the required function efficiently is a preliminary step. Moreover, due to the fact that the majority of storage devices are expensive, it is essential to study the economical value of using these devices. More details about storage devices are provided in Chapter 6 and Appendix E.

2.3 Connection Topologies of PV Systems

PV systems have different topologies according to the connection of the PV modules with the PCU. Some of the common topologies are shown in Figure 2-10 and a comparison between these topologies is given in Table 2-1.

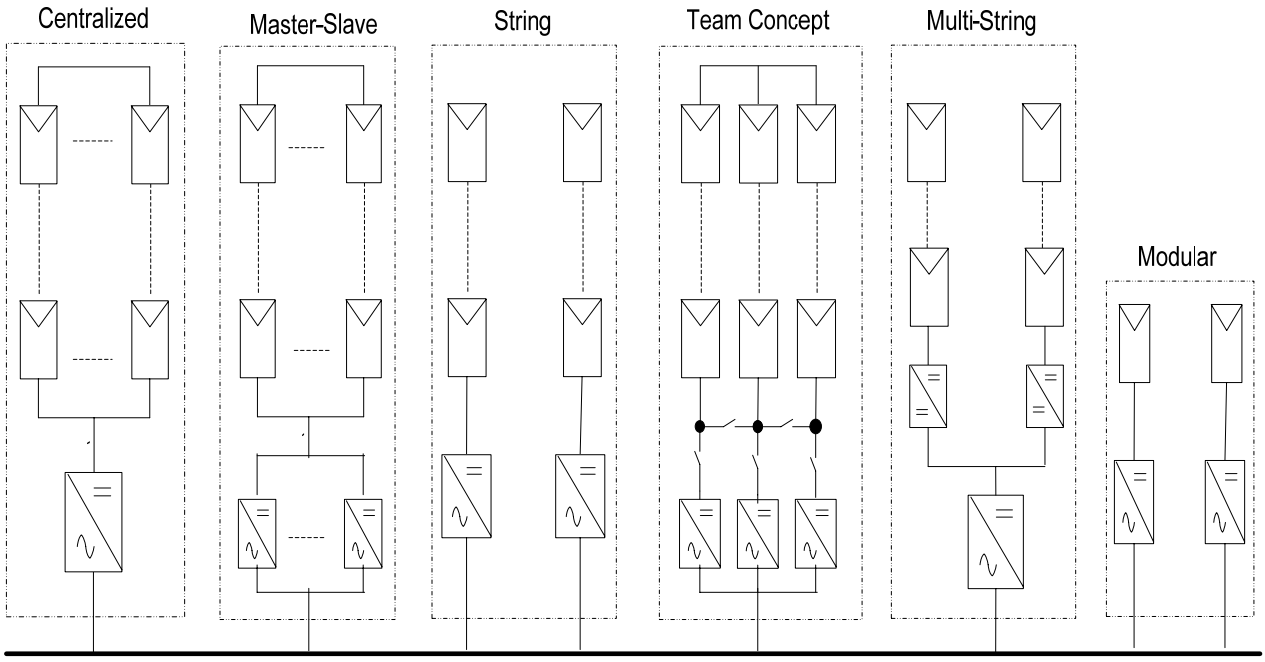

Figure 2-10 Connection topologies of PV systems

A. Centralized topology [35][62][63]

This is one of the well-established topologies. It is usually used for large PV systems with high power output of up to several megawatts. In this topology, a single inverter is connected to the PV array. The main advantage of the centralized topology is its low cost as compared to other topologies as well as the ease of maintenance of the inverter. However, this topology has low reliability as the failure of the inverter will stop the PV system from operating. Moreover, there is significant power loss in the cases of mismatch between the modules and partial shading, due to the use of one inverter for tracking the maximum power point.

B. Master-slave topology [63][64]

This topology aims to improve the reliability of the centralized topology. In this case, a number of parallel inverters are connected to the array and the number of operating inverters is chosen such that if one inverter fails, the other inverters can deliver the whole PV power. The main advantage of this topology is the increase in the reliability of the system. Moreover, the inverters can be designed to operate according to the irradiance level, where for low irradiance level some of the inverters are shut down. This technique of operation extends the lifetime of inverters and overall operating efficiency. However, the cost of this topology is higher than that of the centralized topology and the power loss due to module mismatch and partial shading is still a problem with this topology.

C. String topology [35][62][63][65]

In the string topology, each string is connected to one inverter; hence, the reliability of the system is enhanced. Moreover, the losses due to partial shading are reduced because each string can operate at its own maximum power point. The string topology increases the flexibility in the design of the PV system as new strings can be easily added to the system to increase its power rating. Usually, each string can have a power rating of up to

2-3 kW. The main disadvantage of this topology is the increased cost due to the increase in the number of inverters.

D. Team Concept topology [66]

This topology is used for large PV systems; it combines the string technology with the master-slave concept. At low irradiance levels, the complete PV array is connected to one inverter only. As the irradiance level increases, the PV array is divided into smaller string units until every string inverter operates at close to its rated power. In this mode, every string operates independently with its own MPP tracking controller.

E. Multi-String topology [35][63][65]

In this topology, every string is connected to a DC-DC converter for tracking the maximum power point and voltage amplification. All the DC-DC converters are then connected to a single inverter via a DC bus. This topology combines the advantages of string and centralized topologies as it increases the energy output due to separate tracking of the MPP while using a central inverter for reduced cost. However, the reliability of the system decreases as compared to string topology and the losses due to the DC/DC converters are added to the losses of the system.

F. Modular topology [35][63][65]

This is the most recent topology. It is also referred to as "AC modules", because an inverter is embedded in each module. It has many advantages such as reduction of losses due to partial shading, better monitoring for module failure, and flexibility of array design. However, this topology is suitable only for low power applications (up to 500W) and its cost is relatively high. Moreover, the lifetime of the inverter is reduced because it is installed in the open air with the PV module, thus increasing its thermal stress.

Table 2-1 Comparison between different connection topologies of PV systems

Topology	Advantages	Disadvantages	Power Rating
Centralized	1- Easy to monitor 2- Easy to maintain 3- low cost due to central inverter	1- DC losses in high voltage DC cables 2- Power loss due to centralized MPPT, string diodes and mismatch in PV modules 3- Low reliability 4- Not flexible in design	up to several megawatts
Master-Slave	1- Higher reliability as compared to centralized topology 2- Improved efficiency for the operating inverters 3- Extended lifetime of inverters	1- DC losses in high voltage DC cables 2- Power loss due to centralized MPPT, string diodes and mismatch in PV modules 3- High cost due to use of multiple inverters 4- not flexible in design	up to several megawatts
String	1- Reduction in energy loss that result from partial shading 2- Losses in string diodes are eliminated 3- Good reliability 4- Flexible in the design	1- Higher cost as compared to centralized 2- Used for low power ratings	3-5 kW / string
Team Concept	1- High efficiency due to individual MPPT and increase in the inverters efficiency 2- Higher reliability as compared to centralized topology	1- Losses due to mismatch between PV modules 2- High cost due to the use of several inverters	up to several megawatts
Multi-String	1- Reduction in energy loss that result from partial shading 2- Losses in string diodes are eliminated 3- MPPT and current control are separated 4- Voltage amplification can be achieved by the DC-DC converter	1- All strings are connected to a single inverter thus the reliability of the system decreases 2- Additional losses inside the DC/DC converter 3- The cost is higher as compared to centralized topology	5 kW
AC modules	1- No losses due to partial shading 2- No mismatch losses between modules 3- Easy in failure detection of the modules 4- Flexible & expandable in design	1- High cost 2- Replacement of inverter in case of faults is not easy 3- Reduced lifetime of the power electronic components due to Additional thermal stress	up to 500 W

2.4 Impacts of PV Systems on the Grid

Grid-connected PV systems are usually installed to enhance the performance of the electric network by reducing the power losses and improving the voltage profile of the network. However, this is not always the case as these systems might impose several negative impacts on the network, especially if their penetration level is high. Such negative impacts include power and voltage fluctuation problems, harmonic distortion, malfunctioning of protective devices and overloading and underloading of feeders.

Studying the possible impacts of PV systems on the electric network is currently becoming an important issue and is receiving a lot of attention from both researchers and electric utilities. The main reason for the importance of this issue is that accurate evaluation of these impacts, as well as providing feasible solutions for the operational problems that might arise due to installing PV systems, is considered a major contribution towards facilitating the widespread use of these systems.

Due to their importance and relevance to the research presented in this document, the potential impacts of grid-connected PV systems on the electric network, as well as the appropriate methods that can be used for evaluating these impacts are discussed in detail in Chapter 3.

2.5 Summary and Conclusions

This chapter presented the main components of grid-connected PV systems and discussed the recent research activities regarding these components. Starting with the irradiance, weather stations usually measure the global irradiance on a horizontal surface, and thus, models are required to estimate the irradiance on the tilted surface of the PV system. The accuracy of any of these models is usually dependant on the location where the PV system is being installed, thus, it is important to choose a suitable model for the case under consideration. One of the main activities in this area is the development of

irradiance models suitable for specific locations. The fluctuations in irradiance due to passage of clouds also received a lot of attention from researchers, where most of the work done in this field relied on the frequency domain analysis. One field that still requires more attention is the prediction of irradiance, which is a complicated task as compared to the prediction of wind speed. This is mainly because of the variety of factors that affect the accuracy of prediction including the wind speed and direction and type, height and thickness of clouds.

Modeling of the PV cells is one of the mature areas in the field. There are a variety of models available in the literature and can be divided into two main categories; detailed and simplified models. Detailed models attempt to represent the physics of the PV cell and are usually suitable for studies that require the detailed cell information such as implementation of maximum power techniques and analysis of the effect of change in irradiance and temperature on the performance of the PV cell. On the other hand, simplified models usually provide a direct estimate of the maximum power generated from the PV cell at certain operating conditions. Thus, simplified models are suitable for system studies that try to identify the impacts of PV systems on the electric network.

In the past few years, developing new topologies for power conditioning units and applying new control techniques were the focus of many studies, almost saturating this field of research. Also, the application of new maximum power point tracking algorithms received a lot of attention. However, most of these algorithms fail to operate properly in the case of partial shadings, which is the case where parts of the PV array are shaded by clouds or nearby buildings.

The use of storage devices with PV systems is currently receiving a lot of attention. These devices can be used to bridge fluctuations in the output power of PV systems, shift the peak generation of the system to match the load peaks, and provide reactive power support. One of the main challenges that still face the use of storage devices is the high cost associated with their installation. Thus, studying the economical aspect of installing these devices is of great importance.

Chapter 3

Impacts of Grid-Connected PV Systems on the Electric Network

3.1 General

Photovoltaic systems were first used as stand-alone systems to provide electricity to rural areas where no other sources of energy were present. The advances in the technology and the concerns about global warming are encouraging both utilities and customers to expand the use of grid-connected PV systems. However, the intermittent nature of the output power of these systems might impose some challenges on the operation of the electric network. The aim of this chapter is to explore the pros and cons of installing grid-connected photovoltaic systems and to present some of the methods that can be used to study the impacts of these systems on the electric network [67].

3.2 Definitions

This section presents the definitions of some frequently-used terms related to PV systems.

- ***Availability*** of a PV plant is the ratio of the actual number of operating hours of the PV plant to the number of hours that the plant can potentially operate.

$$Availability = \frac{\text{(number of operating hours of the PV plant)}}{\text{(number of hours with enough insolation for operation)}}$$

- ***Capacity factor (CF)*** of a PV system is ratio of the expected energy production over a certain period (usually one year) to the product of the rated output power of the system and the total number of hours of the same period.

$$CF = \frac{\text{(energy produced from the PVsystem per year)}}{\text{(rated output power of the PVsystem) x (hours per year)}}$$

- ***Penetration level*** is the ratio of the installed PV power to the generation capacity of the utility system to which the PV system is connected.

$$\text{Penetration level} = \frac{\text{(rated output PV power)}}{\text{(generation capacity of the utility system)}}$$

  It should be noted that some definitions of the penetration level relate the output power of the PV system to the loading conditions of the feeder instead of the generated power of the utility.
- ***Short-term fluctuation*** is the sub-hourly fluctuation of the irradiance or the output power of the PV system.
- ***Suitability of a PV system*** is the condition where the peak generation of a PV system matches the peak loading of the network without causing any operational problems.

3.3 Classification of PV Systems

According to the IEEE standard 929-2000 [50], PV systems are divided into three categories: 1) small systems rated at 10 kW or less, 2) intermediate systems, rated between 10 kW and 500 kW, and 3) large systems, rated above 500 kW. However, these ranges are likely to be modified in the near future due to the wide range of power ratings of large systems recently installed or planned to be installed as illustrated in Table 1-1.

The study presented in this research is mainly concerned with large PV systems that range from few megawatts up to a few tens of megawatts.

3.4 Benefits of Grid-Connected PV Systems

Global warming, environmental pollution, and possible scarcity of fossil fuel reserves are some of the main driving forces behind the urge for installing grid-connected PV systems. Moreover, utilities and customers can benefit from installing these systems. The main gain for customers is to take advantage of the incentives provided by the governments upon installing PV systems. For example, the Ontario Power Authority (OPA) has offered to pay ¢42/kWh for the power generated from PV systems under Ontario's Standard Offer Program that was launched in November 2006. In February 2009, the Green Energy Act was introduced and the OPA proposed a new program, the Feed-In Tariff Program, which suggested providing customers with new incentive prices for the kWh generated from PV systems. These prices are summarized in Table 3-1 [68].

Table 3-1 Proposed prices for PV projects in Ontario

Type	**Proposed size tranches**	**Proposed Contract Price ¢/kWh**
Any type	$P \leq 10$ kW	80.2
Rooftop	10 kW $< P \leq 250$ kW	71.3
Rooftop	250 kW $< P \leq 500$ kW	63.5
Rooftop	500 kW $< P$	53.9
Ground Mounted	$P \leq 10$ MW	44.3

For utilities, the gains of installing PV systems are mainly operational benefits, especially if the PV system is installed at the customer side on rural feeders. For example, PV systems can be used to decrease the feeder losses [69][70], improve the voltage profile of the feeder, and reduce the lifetime operation and maintenance costs of transformer load tap changers (LTCs) [71]. Moreover, if the peak output of the PV system matches the peak loading of the feeder, then the loading of some transformers present in the network can be reduced during peak load periods [72].

In order for all the aforementioned benefits to become effective, a number of conditions must be satisfied, including:

1) Strategic placement of the PV system,

2) Proper sizing of the PV system, and

3) Suitability of the output power profile of the PV system.

If one or more of these factors are not satisfied, then the benefits might turn into adverse impacts on the performance of the feeder, as will be discussed in the next section.

3.5 Potential Problems Associated with Grid-connected PV Systems

Despite all the benefits introduced by PV systems to electric utilities, these systems might lead to some operational problems. One of the main factors that lead to such problems is the fluctuations of the output power of PV systems due to the variations in the solar irradiance caused by the movement of clouds. Such fluctuations lead to several operational problems and make the output power forecast of PV systems a hard task. In addition, the high cost of these systems limits the possible solutions that can be adopted by electric utilities to reduce the severity of the operational problems that might arise due to these fluctuations.

The negative impacts of Grid-connected PV systems on the network operation did not receive much attention until lately, after the noticeable increase in installation of these systems. The work done in this area can be classified under three main categories: 1) impacts on the generation side, 2) impacts on the transmission and sub-transmission networks, and 3) impacts on the distribution networks. However, before discussing the possible negative impacts of installing PV systems, it is important to present an overview of the source of power fluctuations in these systems and discuss the data required for analyzing the impact of these fluctuations.

3.5.1 Fluctuation of output power of PV systems

Fluctuation of the solar irradiance due to passage of clouds over a PV array is the main reason behind the fluctuation of the output power of PV systems. There are 10 reported cloud patterns, with cumulus clouds (puffy clouds looking like large cotton balls) and squall lines (a solid line of black clouds) causing the largest variations in the output power of PV systems [73]. Squall lines can cause the output power of a PV system to fall to zero, and thus, they lead to the worst-case scenario for the operation of the system. However, squall lines are predictable, and thus, the periods of time during which the PV system will be out of service can be predicted [73]. On the other hand, cumulus clouds result in lower loss of the PV power, but they cause the output of the PV system to fluctuate more frequently as the irradiance fluctuates due to the passage of such clouds [73]. The time period of fluctuations can range from few minutes to hours depending on the wind speed, the type and size of passing clouds, and the area covered by and topology of the PV system.

The most severe fluctuations in the output power of PV systems usually occur at maximum irradiance level around noon. This period usually coincides with the off-peak loading period of the electric network, and thus, the operating penetration level of the PV system is greatest. The severity of PV power fluctuations on the electric network is governed by several factors, such as:

1. Type of clouds,
2. Penetration level,
3. Size of PV system,
4. Location of the PV system,
5. Topology of the PV system, and
6. Topology of the electric network.

3.5.2 Irradiance data required for studying the impact of PV systems

The time resolution of the irradiance data, required for studying the fluctuations of the output power of PV systems, should be match for the main goal of the study as it plays an important role in the accuracy of the results.

In general, the solar irradiance can be separated into two components [74]: 1) deterministic component defined by the daily, monthly and yearly climate at a certain location, and 2) stochastic component that comprises fluctuations around the deterministic component and is defined by the daily weather. In cases where the expected output energy of a PV system is to be estimated over a period of time, either the deterministic component of the irradiance [74] or the hourly irradiance data can be used [75][76]. On the other hand, if it is required to study the performance of PV systems and their impacts on the electric network, then the time resolution of the irradiance data should be high enough to include the short-term or sub-hourly fluctuations of the irradiance (fluctuations within one hour) [77]. Moreover, irradiance data with high time resolution (e.g., 10-min. resolution) can lead to better prediction accuracy because the auto-correlation coefficients will have higher positive values as compared to those obtained for the data with 1-hour time resolution [78].

In the following subsections, the possible negative impacts of the PV systems are discussed. A summary of these impacts is presented in Figure 3-1.

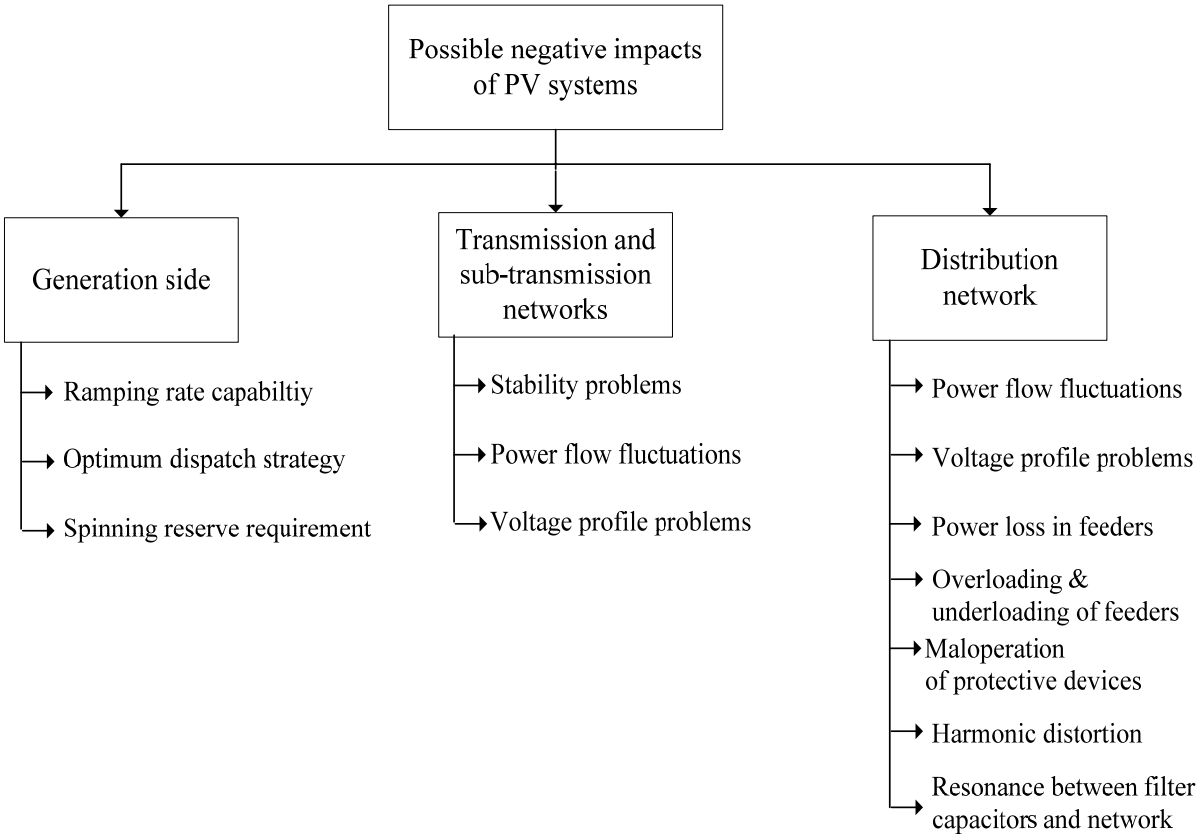

Figure 3-1 Impacts of PV systems on the electric grid

3.5.3 Impact of PV systems on the generation side

Severe fluctuations in the output power of large PV systems might affect the generation in electric utilities. This is mainly due to the fact that the utilities have to follow these fluctuations in order to compensate for any rise and fall in the generation of PV systems. Hence, the generating units that are scheduled to operate during the generation period of PV systems should have ramping rate capabilities that are suitable for the fluctuations of these systems. Moreover, the power fluctuations from the PV system make it difficult to predict the output power of these systems, and thus, to consider them when scheduling

the generating units in the network. Most of the studies performed in this area have addressed this problem and tried to provide some operational solutions that can be adopted by utilities. For example, the studies presented in [79] and [80] discuss the impact of installing large, centralized PV plants on the operation of thermal generation units. The aim of these studies is to identify the penetration level of PV systems that will not lead to generation control problems due to passing clouds. Both studies conclude that the ramping rate capability of the utility is the main factor that controls the penetration level of PV systems. However, the analysis performed in both studies considered the worst-case scenarios only, without providing any details about the frequency and periods during which these scenarios might occur. The work in [81] introduces some factors that can affect the economical and operational values of PV systems for large-scale applications. Some of these factors are the generation mix, maintenance schedules, ramping rates, fuel costs, spinning reserve requirements, PV power fluctuations, and geographical diversification of PV systems. The study suggested some solutions that can be applied to the cases where the severity of the changes in the output power of the PV system is beyond the ramping capacity of the system. These solutions include: 1) increasing scheduled tie-line power, 2) bringing more generating units online to increase the overall ramping capacity of the system, and 3) decreasing the output power of the PV system. A rule-based dispatch algorithm was presented in [82] to take into account the problems that might arise due to the fluctuations in the power of a 100-MW PV plant during the dispatch period. The method is based on predicting the solar irradiance every 10 minutes and assumes that not all the PV power is injected into the grid. A set of rules are proposed to provide operational plans based on the power production of the PV system. These rules depend mainly on the time of the year and the type of electric utility under study.

In general, the generation side of an electric utility can be affected by the PV system if the penetration level of the PV system is comparable to the size of the generating units. However, systems with such large sizes are not expected to be widely installed in the near

future due to the high cost of PV systems. Thus, studying the impacts on the generation side does not seem to be crucial at the time being.

3.5.4 Impact on transmission and sub-transmission networks

PV systems might cause problems in the transmission and sub-transmission networks if their sizes are large enough to affect these networks. The problems arise mainly due to power fluctuations of these systems which might lead to: 1) power swings in lines, 2) power reversal, 3) over and under loading in some lines, and 4) unacceptable voltage fluctuations in some cases [83]. The effect of large PV systems on the voltage levels and the stability of transmission systems after fault conditions was studied in [84]. The results show that replacement of conventional generating units with large PV units affects the voltage levels of the system during normal operating conditions. During fault conditions, rotors of some of the conventional generators present in the network might swing at higher magnitudes due to the existence of PV units. Moreover, at very high penetration levels of PV systems, voltage collapse might occur. In these studies, the sizes of the PV systems required to cause the aforementioned problems were assumed to range from 700 MW to 1500 MW. According to the current market prices of PV systems, such sizes are not expected to be installed soon. Hence, studying the impact of PV systems on transmission and sub-transmission networks does not seem to be important for electric utilities at the time being.

3.5.5 Impact on distribution networks

The impacts of PV systems on the performance of distribution networks are currently one of the main issues for electric utilities. This is because the size and location of the installed PV systems mainly influence these networks. The operational problems introduced by PV systems are similar to those imposed by distributed generators that produce constant active power, such as diesel generators and fuel cells. These problems arise mainly due to the installation of generators at the customer side in a feeder designed

for unidirectional power flow. They include malfunctioning of protective relays, voltage regulation problems, reverse power flow, as well as overloading or underloading of some feeders. Other problems arise due to the use of interfacing electronics that lead to harmonic distortion and parallel and series resonances if a large number of inverters are installed in a certain area. Moreover, the fluctuation of the output power of PV systems adds to the problems faced by the system operator and can deteriorate the power quality of the network.

The impact of small PV systems installed on rooftops of houses has received the attention of many researchers during the last few years. This is mainly due to the increase in installation of these systems due to the incentives provided by governments to residential customers. Typical ratings for PV systems installed on rooftops of houses range from 1 to50 kW.

The issue of harmonic distortion introduced by the power conditioning units used in small PV systems was the main focus of the studies presented in [85]-[87]. All case studies showed harmonic distortions far below the limits specified in the standards. This is mainly because of the great advances made in the inverters technology. However, the filter capacitors of the interfacing inverters might lead to resonances with the electric network if a large number of PV systems are installed in a certain neighborhood [88][89].

The impact of installing small PV systems on the voltage profile of different distribution network topologies was studied in [32]. The results showed that the acceptable voltage limits were exceeded for all networks when the size of each PV system was 200% of the load of the household. The study assumed that PV systems were installed at every node in the network, which might not be a realistic assumption. The results of a real case study presented in [90] indicate the presence of slow transients in the voltage of a medium-voltage distribution feeder corresponding to the frequency of fluctuations of the output power of small PV systems installed on rooftops. Moreover, it was concluded that the presence of PV systems in the network might reduce the lifetime of transformer tap changers due to the increase in their operation. Other studies analyzed the impact of small

PV systems on the voltage profile of a low-voltage grid [91]-[93]. However, these studies did not consider the fluctuations of the irradiance in the analysis.

In general, small PV systems installed on rooftops and facades of buildings might not impose serious problems on the distribution network. This is mainly because the size of these systems requires high concentration in a small area in order to be able to affect the performance of the network. Such situation is not likely to happen often, as the current trends show that small PV systems are usually dispersed over a large area. Such dispersion reduces the impact of fluctuations as the combined irradiance profile over the complete area is more smooth than that over the individual systems [8].

Only few studies have focused on the impact of large centralized PV systems on the distribution network. For example, the study in [94] illustrates that the improper choice of the location of large PV systems can affect the security of the network. Such problem becomes more severe if the generation of the PV system matches the peak loading of the electric network as this might increase the loading of some lines that are already heavily loaded. Thus, to check the network security, the study considered the scenario when the maximum output power of the PV system matched the peak loading conditions of the network. However, the overall performance of the network, including voltage profiles and power losses, was not evaluated because no other scenarios were included. Moreover, no information was provided about how often or when the case of peak matching might occur.

In [95], the impact of installing a 5-MW PV system on the voltage regulation and overcurrent protection of a real distribution feeder was studied. The study shows that the PV system might cause the voltage to reach unacceptable levels during certain periods. On the other hand, the overcurrent protection was not affected by the operation of the PV system, as the inverter of the system seized to operate as soon as the fault was detected. The main advantage of this study is in the fact that a real case is analyzed where the corrective devices used for voltage regulation (transformer LTCs and shunt capacitors) and protection were included. However, the conclusions drawn are based on simulating

the output power of the PV system over a five-day period only and with time resolution of 1 hour. Hence, the variations of loading during different seasons and the sub-hourly fluctuations are not considered in this study, even though they are essential for proper evaluation of the performance of the network.

The impact of increasing the penetration level of PV systems on the network losses was analyzed in [96]. However, the analysis did not investigate the impacts of the PV system on other performance parameters such as the voltage profile of the network and power flow in the lines. To perform such a study, the power fluctuations of the PV output power should be simulated accurately. Thus, the hourly irradiance data used in the analysis of [96] might not be appropriate for this case.

From the above discussions, it can be concluded that large centralized PV systems should be the main concern when studying the impacts of PV systems on the performance of distribution networks. Upon studying these impacts, it is important to consider the fluctuations in the output power of the PV system as it constitutes an inherent characteristic for these systems. Moreover, to obtain accurate results, it is important to examine the performance of the network for an extended period of time in order to consider different possible patterns of generations from the PV system and loading conditions of the feeder under study. To consider these aspects, it is essential to use a method that can manipulate the available data efficiently to be able to provide realistic evaluations about the performance of the network.

3.6 Methods used for studying the impact of PV systems on the electric network

The performance of the electric network in the presence of PV systems can be assessed using three main approaches [97]. First, deterministic approach that considers a certain generation for PV systems and certain loading conditions of the network. Some of the methods that use this approach assume a constant output power for the PV system using

capacity factors specified for each location. Other methods simulate the expected worst-case scenarios that might occur in certain situations. Such scenarios include high PV output power at low loading conditions of the network, high PV output power at high loading conditions and low PV output power at high loading conditions. Second, probabilistic approach that treats the output of the PV system and the load as random variables modeled by appropriate probability density functions. Finally, the approach based on chronological simulations utilizes time series data of the irradiance to calculate the actual profile of the output power of the PV system. This power is used with the actual load profile of the network in simulating the performance of the electric network.

Each of the three approaches has its advantages and disadvantages. Methods based on the deterministic approach are simple and straightforward and can be used to provide an overview of the expected performance of the system under specific operating conditions. For example, methods that use capacity factors can provide a good estimate of the energy production of the PV system during a certain period of time. On the other hand, methods that simulate specific scenarios can provide an estimate about the performance of the network for these scenarios. However, the results obtained from applying these methods do not provide any information about the actual performance of the network and hence no general conclusions can be obtained. Methods based on the probabilistic approach have the advantage of providing realistic information about the performance of the system if the random variables are modeled correctly. However, the major drawback of these methods is the loss of the temporal information that can be obtained using actual data. This, in return, will limit the solutions to any possible problems arising from the integration of PV systems. Methods based on chronological simulations avoid this problem and provide accurate results about the impact of the PV system on the electric network. The main drawback of these methods is the requirement of extensive time series data of the irradiance and loads, which are usually not available [97]. However, the increasing use of PV systems will force utilities to consider such data. Another drawback of these methods is that the number of simulations required in evaluating the performance

of the network will be very large if the time resolution of the data is high and the period of study is long, which makes their implementation impractical for utility studies [98]. Moreover, the huge amount of data obtained from these simulations can make the evaluation of the performance of the network a difficult task.

The deterministic approach was used in [94] to evaluate the system security in the presence of centralized PV systems. In [97] and [99], the probabilistic approach was adopted to evaluate the impact of small PV systems on the electric network. However, as mentioned before, both approaches fail to provide accurate information about the performance of the network, especially due to the power fluctuations of the PV system.

A general layout of a model based on chronological simulations was presented in [100] allowing the interactions between the installed PV plants and the network operation to be studied in detail. The proposed model can be used in studying the system security, generation scheduling, economic dispatch and unit commitment in the presence of PV systems. The model has the advantage of being comprehensive and detailed. However, this model is concerned mainly with power system operation rather than investigating the impacts on the distribution network. Moreover, the method cannot be applied for large data sets due to high analysis cost. Hence, no statistical conclusions about the power fluctuations, such as frequency and periods of occurrence, can be obtained.

In [101], wavelet analysis was applied to the time series of the solar irradiance and the node voltages. Based on the analysis, two indices were proposed to determine the power content of the fluctuations produced by PV systems. The indices have the advantage of being able to provide information about the persistence and severity of the short-term fluctuations. However, the method applied in this study has to be modified in order to be able to deal with extended time series data. Another drawback of the analysis performed in this study is that the loads were assumed constant in the simulations. This assumption might be suitable for the short period of the simulations presented in the study; however, for extended periods, the load profile must be considered.

The sub-hourly irradiance data obtained over one year was used in [98] to calculate the output powers of small PV systems randomly distributed along a distribution feeder. The output powers of the PV systems and the active and reactive powers of the loads were used to calculate the power loss and voltage profile of the distribution feeder. To reduce the number of simulations, the *K-means* algorithm was used to cluster similar data points together. The reduced data points were then used to calculate the annual duration curves of the total power loss of the feeder, the active and reactive powers of the substation, and the voltage at different nodes. The method presented is a powerful tool that can be used to assess the performance of the feeder over an extended period by reducing the number of required simulations. However, the method neglects the temporal information completely; hence, it fails to provide any information about the magnitude and frequency of occurrence of fluctuations in the PV power. Moreover, the periods when high fluctuations are likely to occur, or when the output power of the PV systems is expected to be low, cannot be identified. Hence, no operational plans can be adopted to overcome the possible problems that might arise during these periods.

3.7 Summary and Conclusions

Grid-connected PV systems can provide a number of benefits to electric utilities, such as power loss reduction, improvement in the voltage profile, and reduction in the maintenance and operational costs of the electric network. However, improper choice of the location and size of the PV systems and unsuitability of the output power profile of the PV system to the profile of the electric network can impose operational problems on the network. Moreover, the fluctuations in the output power of these systems add to the complexity of the problem.

Large, centralized PV systems, installed in distribution networks, require more attention at the time being. Detailed studies should be carried out prior to installing these systems to predict the performance of the electric network under different operating conditions of both PV systems and existing loads. Such studies should be carried out over an extended

period of time and should include detailed information about the feeder and the profile of the output power of the PV system to provide accurate conclusions about the performance of the network.

There are a number of methods that can be used to assess the performance of the electric network in the presence of PV systems. The deterministic and probabilistic methods cannot provide any information about the impacts of the power fluctuations generated from the PV system. On the other hand, the approach based on chronological simulations can achieve this task by including the temporal information in the analysis. However, none of the studies presented in the literature that have utilized this approach can be applied using long historical data sets (historical data of the last few years) while preserving the temporal information on the PV power fluctuations. The importance of utilizing long historical data in the analysis is to provide accurate evaluation of the performance of the system by including many possible patterns that can be generated from the PV system. Moreover, the results obtained from long historical data can help in providing the system operator with information about "when" and "how often" unacceptable performance of the network is likely to occur. Hence, suitable operational plans can be prepared to reduce the severity of the problems that might arise due to installing the PV system. Such plans include the choice of the operating power factor, requirement of storage and operation below the maximum power point (MPP).

As a conclusion, it is essential to develop a method that can overcome the aforementioned drawbacks of the existing methods and facilitate the study of different solutions that can reduce the severity of the operational problems in distribution networks resulting from the installation of large PV systems.

Chapter 4

A Clustering-Based Method for Studying the Impacts of Large PV Systems

4.1 General

In the previous chapter, it was concluded that the methods used to evaluate the impacts of large PV systems on the performance of the distribution network require further improvement. This is mainly because most of the existing methods cannot be used in simulating the performance of the electric network over a long period of time while preserving the temporal information.

The main advantage of the methods that are based on chronological simulations is their ability to include the temporal information in the analysis. However, these simulations should be carried out over an extended period of time in order to provide accurate evaluation of the performance of the system by including many possible patterns that can be generated from the PV system. Such a study requires collecting historical time series data of irradiance and temperature at the location where the PV system is to be installed. The time resolution of the collected data should be suitable for taking the short-term fluctuations in the output power of the PV system into account. Moreover, the load profile of the network under study should be utilized in the analysis to allow for

evaluating the impacts of PV power fluctuations at different loading conditions of the network.

In most cases, using the whole data in simulating the performance of the network might not be a practical approach for performing such a study. For example, if the irradiance and temperature data of the past five years are available with time resolution of 10 minutes, then it is required to simulate 262,800 case studies. This number can jump to a couple of million if, for example, the location and size of a PV system is required to be identified. Simulating such an overwhelming number of case studies might make the extraction of useful information from the obtained results a difficult task. Moreover, the analysis of all the possible scenarios might be a time consuming task, especially if the electric network is complicated.

The main purpose of this chapter is to present a new method that can overcome many of the drawbacks of the methods presented in the previous chapter [102]. The general layout of the method is presented in the following section and the application of different stages of the method is discussed in detail in the sections that follow.

4.2 Layout of the Proposed Method

The main idea of the proposed method is to use the huge amount of the available data in an efficient and intelligent manner, while preserving the temporal information. This can be achieved by first dividing the long historical time series of the calculated PV power into segments. The next step is to group together the segments that have close profiles and to choose a representative for each group. The representative segment of the group can then be used to evaluate the performance of the electric network, and thus, provide information about the expected performance of the segments included in the group. The results generated from using the representative segments can either be utilized directly to evaluate the performance of the feeder or can help identify the groups of segments that require further analysis. The general layout of the proposed method is presented in Figure 4-1.

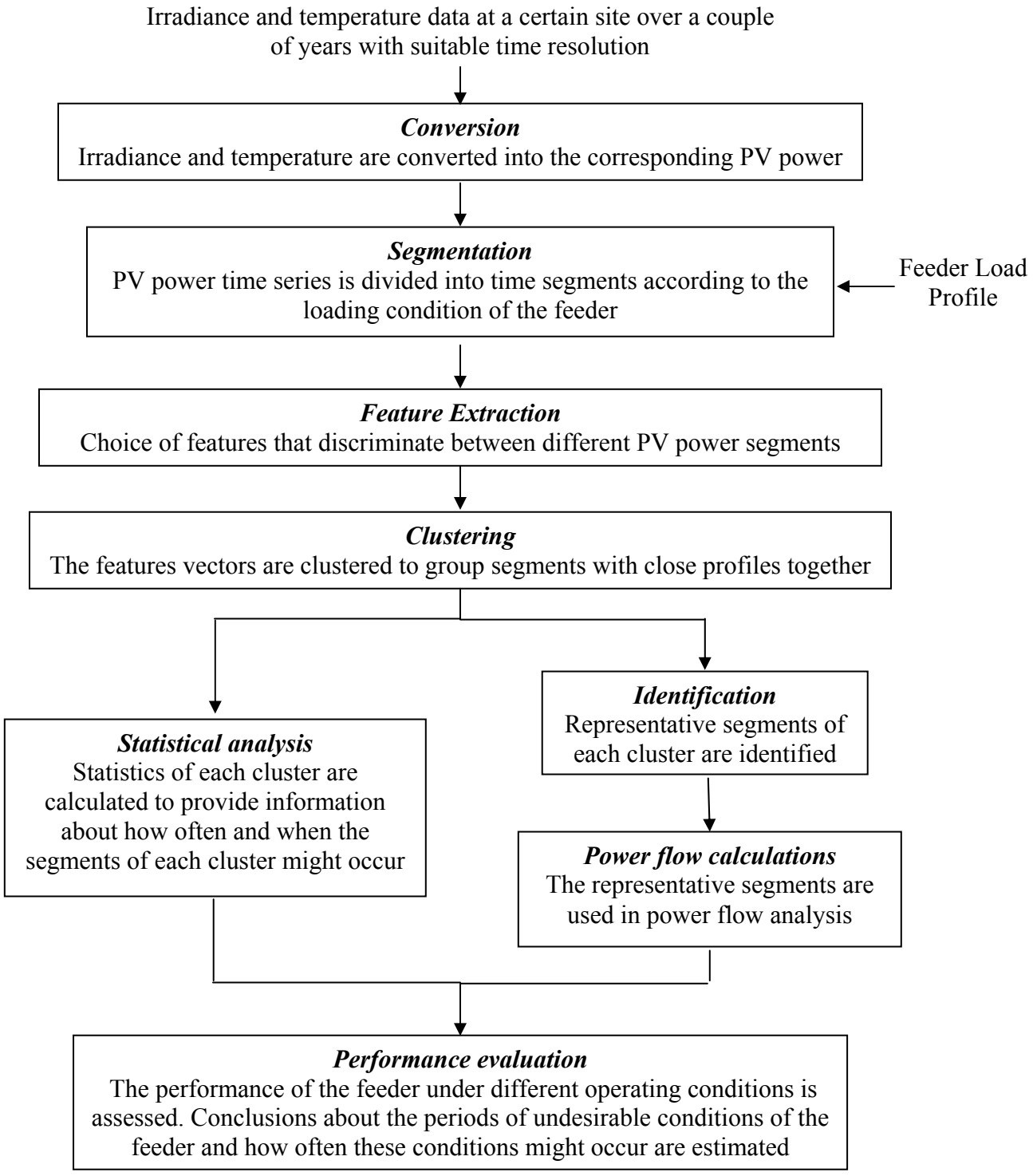

Figure 4-1 Flow chart of the proposed method

In the following the blocks in Figure 4-1 are described.

A. Conversion Stage

Input: Historical irradiance and temperature data

Output: Time series of the output AC power of the PV system

Description: The time series of the irradiance and temperature at a certain site are obtained for the past couple of years with appropriate time resolution. This data can be obtained from the weather stations or satellites. The maximum available DC output power of the solar arrays is then calculated using an appropriate model. Finally, the AC power produced by the PV system is estimated from the calculated DC power using the manufacturer's efficiency curves of a typical power-conditioning unit.

B. Segmentation Stage

Input: Time series of the PV system AC power, loading profile of the feeder

Output: Segmented time series of the power of the PV system

Description: The PV power time series is categorized according to the annual loading conditions of the feeder. Furthermore, each category of the PV power time series is divided into segments, each representing one day. Other types of segments can be identified based on the loading profile of the feeder as will be explained later.

C. Feature Extraction Stage

Input: Segments of PV power for each feeder loading category

Output: Features vectors for each segment

Description: In this stage, it is required to extract the features of the power segments that can be used to group together similar power segments present in each category.

D. Clustering Stage

Input: Features vectors of all segments for each category

Output: Groups of power segments in the same category that have close profiles

Description: In this stage, a clustering algorithm is applied to the features vectors obtained from the previous stage. The vectors present in the same cluster should reflect segments having close profiles of PV power.

E. Identification Stage

Input: Clusters of each category

Output: Representative segments for each cluster

Description: In this stage, each cluster is represented by its cluster representative. These cluster representatives can be used for simulating the performance of the system. The results of these simulations can provide information about the performance of the whole cluster.

F. Power Flow Calculations

Input: Representative PV power segments and the corresponding loading of the feeder

Output: Performance of the electric network

Description: The representative PV power segments for each category are used with the corresponding active and reactive power loading of the feeder in a power flow algorithm. The active and reactive powers flowing in the network, the voltage profiles at different nodes of the network and the power losses in the network are calculated for different operating conditions.

G. Statistical Analysis Stage

Input: Clusters of each category

Output: Statistical results about each cluster

Description: The main objective of this stage is to provide statistical information about the frequency of occurrence of clusters having power segments with important profiles and the time periods during which the members of each cluster are most likely to occur. Examples of these profiles include days with high power output, overcast days and days with high power fluctuations. The generalization of the conclusions obtained from this stage depends on the amount of historical data used in the study. Such conclusions can help the system operator in predicting the performance of the system during similar periods, and thus, preparing to implement appropriate corrective measures.

H. Performance Evaluation Stage

Input: Results from power flow and/or statistical analysis stages

Output: Operational plans for different periods of the year

Description: The aim of the final stage of the method is to evaluate the performance of the electric network if a PV system is to be installed at a certain location. This evaluation leads to identifying suitable decisions and plans to mitigate possible future operational problems such as voltage fluctuations, improper operation of protective relays and over and under loading of feeders.

4.3 Conversion Stage

In the first stage of the proposed method, the output AC power time series of the PV system is estimated from the irradiance and temperature time series provided from the weather station. This stage can be divided into three sub-stages: 1) estimation of the irradiance on the surface of the PV array using the irradiance data available from the weather station, 2) calculation of the DC output power generated from the PV array using a suitable PV model, and 3) calculation of the AC output power generated from the PV

system using the manufacturer's efficiency curve for the PCU. The following sub-sections provide the details of each sub-stage.

4.3.1 Estimation of the irradiance on the surface of the PV array

As mentioned in Chapter 2, weather stations usually measure the global irradiance on a horizontal surface, G_g, and direct normal irradiance, G_{bn}. However, most of the PV arrays are tilted by an angle τ with respect to the horizontal surface, depending on the site of installation, to maximize their production of electric power. Therefore, it is essential to calculate the global irradiance on the tilted PV array.

The global irradiance on a tilted surface, G_{tg}, is composed of three components:

1) *Direct (beam) irradiance,* G_{tb}, which directly reaches the PV array and is considered the most effective part for generating electricity.

2) *Diffuse irradiance,* G_{td}, which reaches the PV array after being scattered by clouds.

3) *Albedo,* G_{tr}, which is the irradiance reflected from the ground and is effective if the PV array is tilted.

Accordingly, the global irradiance on the tilted surface can be calculated by:

$$G_{tg} = G_{tb} + G_{tr} + G_{td} \tag{4-1}$$

To calculate the three irradiance components on the tilted surface, G_{tb}, G_{tr} and G_{td}, from the two irradiance components obtained from the weather stations, G_g and G_{bn}, the following model can be used:

a) Calculation of the direct irradiance on the horizontal surface [103]:

$$G_b = G_{bn} \cos\theta_z \tag{4-2}$$

where

θ_z is the Zenith angle in degrees defined as the angle between the vertical line from the earth and the line to the sun. It is also called the angle of incidence of the sun on a horizontal surface on the earth.

b) Calculation of the direct irradiance on the tilted surface [5]:

$$G_{tb} = G_b \frac{\cos\theta_i}{\sin\alpha_s} \tag{4-3}$$

where

θ_i is the angle of incidence in degrees defined as the angle between the beam irradiance on a tilted surface and the normal to that surface, and

α_s is the solar altitude in degrees defined as the angle between the horizontal and the line to the sun.

c) Calculation of the reflected irradiance or albedo [5]:

$$G_{tr} = 0.5\rho G_g(1-\cos\tau) \tag{4-4}$$

where ρ is a constant which depends on the type of ground surrounding the tilted surface and is called the ground reflectance. The values of the ground reflectance for different types of surroundings are given in Table 4-1 [104].

Table 4-1 Ground reflectance for different ground surroundings

ρ	**Locations**
0.2	Temperate and humid tropical location (most commonly used)
0.5	Dry tropical locations
0.9	Snow covered ground

d) Calculation of the diffuse irradiance on the tilted surface:

Estimating the diffuse irradiance on a tilted surface is considered the most sophisticated part in the calculations. Thus, many models have been proposed to estimate the diffuse irradiance on the tilted surface from that on the horizontal surface. An important factor that affects the model accuracy is the location of the site under study. In [105], a survey about the use of various models for different locations is presented and it is shown that the Klucher model [106] provides small errors when compared with the actual irradiance measured at a location with Latitude 42.42°N and Longitude 73.50°W. This site is close to the location of the weather station from which the data used in this research was obtained. Accordingly, the Klucher model is chosen to calculate the diffuse irradiance on the tilted surface of the PV array. The model can be given by the following equations [106]:

$$G_{td} = \frac{G_d}{2}\left[1+\cos\tau\right]\left[1+F\sin^3\left(\frac{\tau}{2}\right)\right]\left[1+F\cos^2\theta_i\sin^3\theta_z\right] \tag{4-5}$$

$$F = 1-\left(\frac{G_d}{G_g}\right)^2 \tag{4-6}$$

$$G_d = G_g - G_b \tag{4-7}$$

4.3.1.1 Calculation of Solar angles [24][103]

The irradiance model presented in the previous sub-section requires the calculation of several solar angles that are needed in estimating the irradiance components on the surface of the PV array. In this sub-section the calculation of these solar angles is presented. Figure 4-2 displays some of the angles that are used for this purpose.

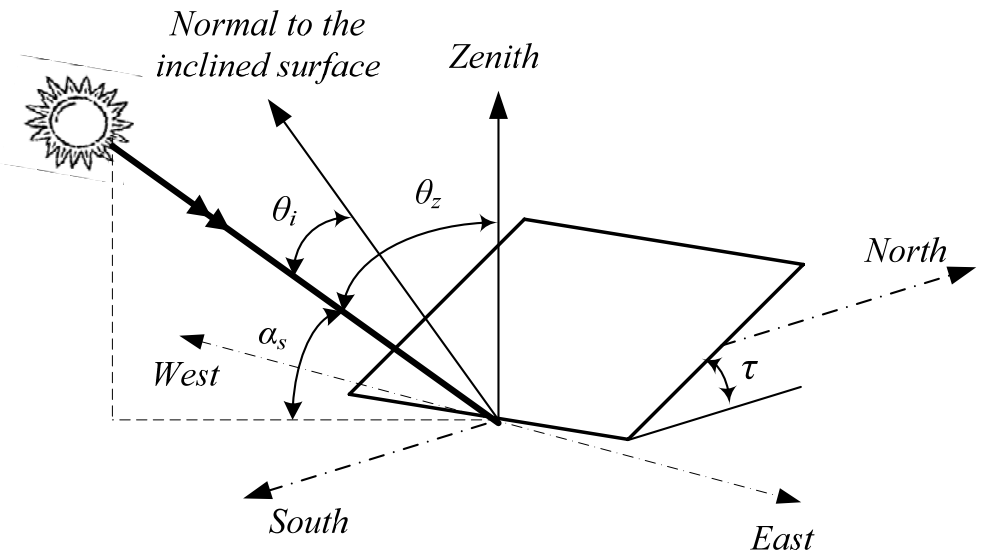

Figure 4-2 Solar angles for an inclined surface

- The Zenith angle, θ_z, is calculated by:

$$\cos\theta_z = \sin\delta_s \sin\varphi_{lat} + \cos\delta_s \cos\varphi_{lat} \cos\omega \tag{4-8}$$

where

δ_s is the Declination angle in degrees and is defined as the angular position of the sun at solar noon with respect to the plane of the equator ($-23.45^o \le \delta_s \le 23.45^o$, $\delta_s>0$ for north),

φ_{lat} is the Latitude in degrees and is defined as the angular location north or south to the equator ($-90^o \le \varphi_{lat} \le 90^o$, $\varphi_{lat} > 0$ for north), and

ω is the Hour angle in degrees and is defined as the angular displacement of the sun east or west of the local meridian due to the rotation of the earth on its axis at 15^o/hour, negative before solar noon and positive after solar noon.

- The Declination angle, δ_s, can be calculated for the Northern hemisphere in terms of an integer representing the day of the year, d, by:

$$\delta_s = 23.45 \sin(360 \frac{284 + d}{365}) \tag{4-9}$$

- To calculate the Hour angle, the Solar time is required to be calculated from the standard time. This can be achieved by using the following equations:

$$\text{SolarTime} = \text{Standard Time} + 4(L_{st} - \varphi_{long}) + Et \tag{4-10}$$

$$\begin{aligned} Et = 229.2\big[&0.000075 + 0.001868 \cos \psi - 0.032077 \sin \psi \\ &-0.014615 \cos(2\psi) - 0.04089 \sin(2\psi)\big] \end{aligned} \tag{4-11}$$

$$\psi = (d - 1) \frac{360}{365} \tag{4-12}$$

where

L_{st} is the standard meridian for the local time zone in degrees (for example, the standard meridian for continental US time zones are: Eastern, 75°W, Central time, 90°W, Mountain, 105°W, Pacific, 120°W), and

φ_{long} is the Longitude in degrees and is defined as the East-West geographic coordinate measurement.

- The angle of incidence, θ_i, which is the angle in degrees between the direct irradiance on a tilted surface and the normal to that surface can be calculated by:

$$\cos \theta_i = \sin \delta_s \sin(\varphi_{lat} - \tau) + \cos \delta_s \cos(\varphi_{lat} - \tau) \cos \omega \tag{4-13}$$

- Finally, the Solar altitude, α_s, is calculated by:

$$\alpha_s = 90^o - \theta_z \tag{4-14}$$

4.3.1.2 Available data

For the purpose of this study, the 10-min. global and direct normal irradiance data measured on a horizontal surface, G_g and G_{bn}, respectively, are obtained for the period 1997 – 2005. The data is obtained from the Solar Radiation Research Laboratory [107] at a location with Latitude of 39.74^{o}N and Longitude of 105.18^{o}W.

4.3.1.3 Quality control of the measured data

Usually, the raw data obtained from weather stations include outliers due to the possible errors in measurements. Thus, the data obtained from the weather station are subjected to quality control rules similar to those used in [105]:

1- Reject all data with a Zenith angle larger than 90^o.

2- Reject all global irradiance data that are greater than the corresponding extraterrestrial irradiance calculated for a horizontal surface.

3- Reject all direct normal irradiance data that are greater than the extraterrestrial irradiance received on a surface normal to the sun rays outside the atmosphere.

Rejecting the outliers, as suggested in [105], can affect the resolution of the data. This can affect the segmentation and feature extraction stages as these stages require data resolution to be the same for all segments. Accordingly, the outliers are adjusted using the following relations:

1- All the measured irradiance components corresponding to a Zenith angle larger than 90^o are set to zero.

2- The measured global irradiance, $G_{g\text{-}m}(i)$, that is greater than the corresponding extraterrestrial irradiance calculated on a horizontal surface outside the earth, $G_o(i)$, is replaced by:

$$G_{g-m}(i) = G_o(i) \text{x} \frac{G_{g-m}(i-1)}{G_o(i-1)} \tag{4-15}$$

where $$G_o(i) = G_{sc}\left(1 + 0.033\cos\frac{360 \times d}{365}\right)\cos\theta_z \tag{4-16}$$

i is the order of the data point and G_{sc} is the solar constant = 1367 W/m^2.

3- The measured direct normal irradiance data at any instant, $G_{bn\text{-}m}(i)$, that is greater than the corresponding extraterrestrial irradiance, $G_{on}(i)$, is replaced by:

$$G_{bn-m}(i) = G_{on}(i) \text{x} \frac{G_{bn-m}(i-1)}{G_{on}(i-1)} \tag{4-17}$$

where

$$G_{on}(i) = G_{sc}\left(1 + 0.033\cos\frac{360 \times d}{365}\right) \tag{4-18}$$

4.3.1.4 Verification of the irradiance model

The irradiance model described by Equations (4-1) to (4-18) is implemented in the MATLAB environment to estimate the global irradiance on the surface of a tilted PV array from the data on the horizontal surface provided by the weather station. The weather station also provided the measured global irradiance on a surface tilted by 40^o for the year 2005. Thus, this data was used to examine the performance of the irradiance model by comparing the measured irradiance data with that calculated from the model. Figure 4-3 to Figure 4-5 display samples of the irradiances on the surface of the PV array calculated and measured for three different days of the year: Days 2, 182 and 337.

To compare the calculated and measured irradiance data for each day of the year, two statistical measures are used:

a) Mean Bias Error (*MBE*) [103]:

The Mean Bias Error indicates the average deviation of the measured daily irradiance values from the calculated values, and is defined by:

$$MBE = \frac{1}{N}\sum_{i=1}^{N} G_{tg}(i) - G_{tg-m}(i) \tag{4-19}$$

where G_{tg} is the calculated irradiance on the tilted surface, $G_{tg\text{-}m}$ is the measured irradiance on the same surface and N is the number of compared irradiance points during the day. The *MBE* has the same units as that of the calculated variable and should be equal to zero for ideal cases.

To obtain a percentage value, the Relative Mean Bias Error (*RMBE*) [105] can be obtained in terms of the measured average daily irradiance, $\overline{G}_{tg-m}$,by:

$$RMBE = \frac{1}{N}\frac{\sum_{i=1}^{N} G_{tg}(i) - G_{tg-m}(i)}{\overline{G}_{tg-m}} \tag{4-20}$$

b) Mean Absolute Error (*MAE*):

The Mean Absolute Error calculates the average of the absolute errors between the calculated and measured irradiance by:

$$MAE = \frac{1}{N}\sum_{i=1}^{N} \left| G_{tg}(i) - G_{tg-m}(i) \right| \tag{4-21}$$

The *MAE* has the same units as that of the calculated variable and is always positive. It should be equal to zero for ideal cases.

To obtain a percentage value, the Relative Mean Absolute Error (*RMAE*) can be obtained in terms of the average daily measured irradiance, $\overline{G}_{tg-m}$, by:

$$RMAE = \frac{1}{N} \frac{\sum_{i=1}^{N} \left| G_{tg}(i) - G_{tg-m}(i) \right|}{\overline{G}_{tg-m}} \tag{4-22}$$

The results displayed in Figure 4-6 to Figure 4-9 indicate that the calculated errors are small for all days of the year except for the Day 337, where the error is relatively large. A possible reason for this large error is an error in one of the measured irradiance components that was recorded on that day. This can be true because the errors calculated for all other days are very small. As a conclusion, the irradiance model can be used to estimate the irradiance components for the data set of the 9 years.

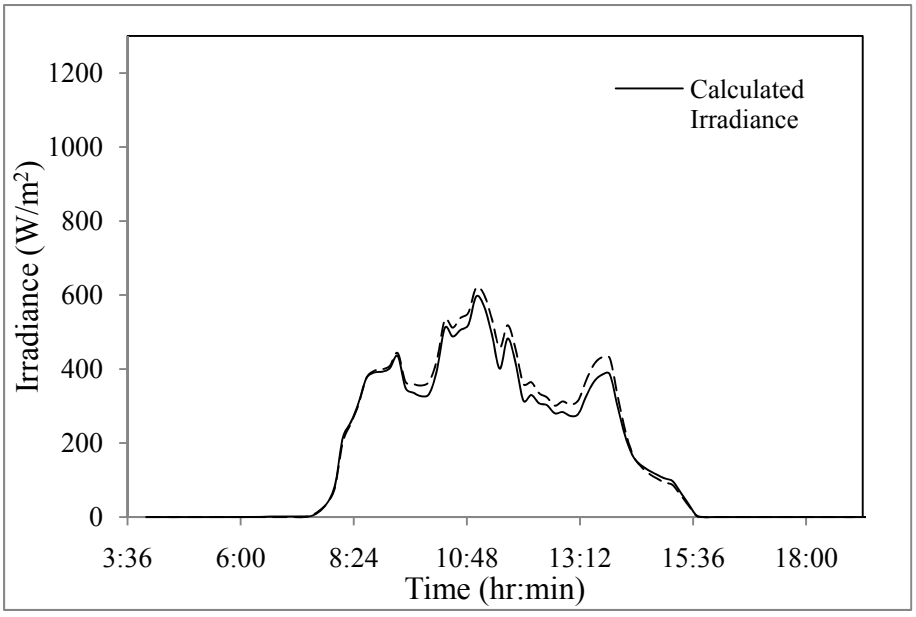

Figure 4-3 Calculated and measured irradiances for Day 2

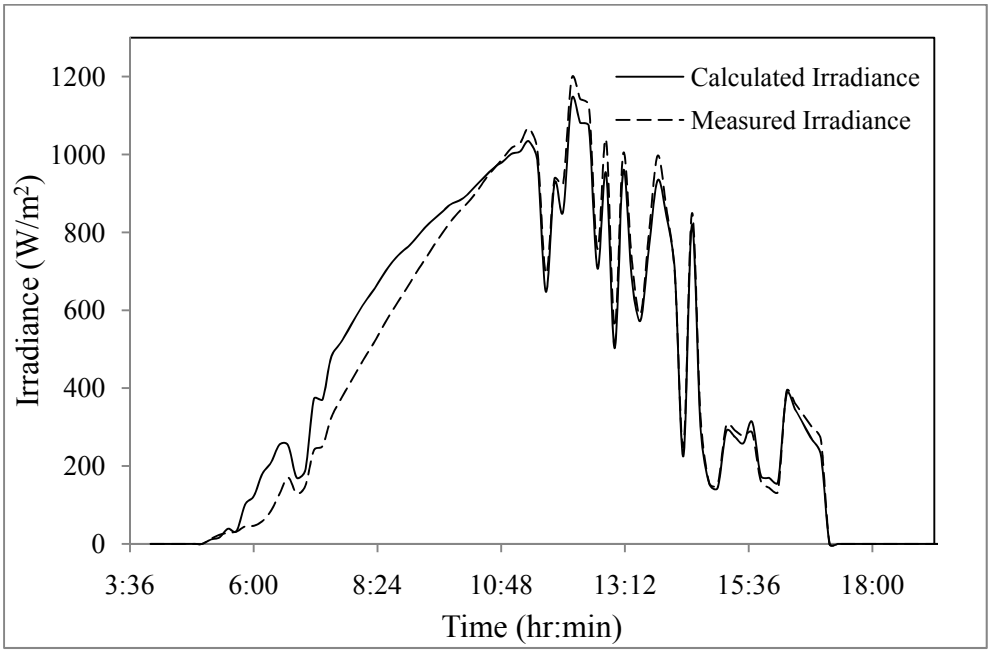

Figure 4-4 Calculated and measured irradiances for Day 182

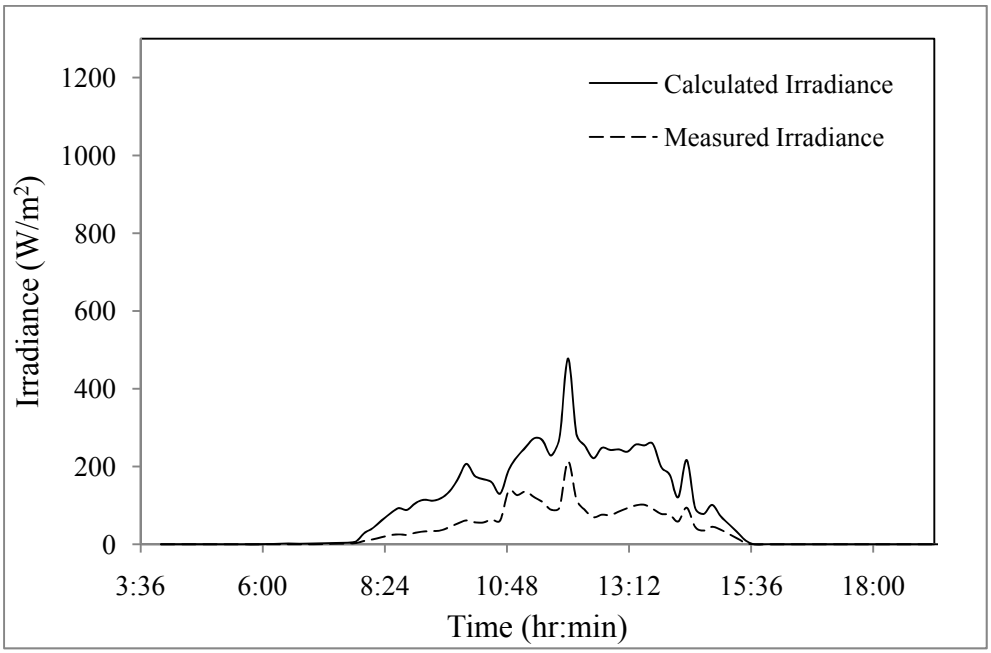

Figure 4-5 Calculated and measured irradiances for Day 337

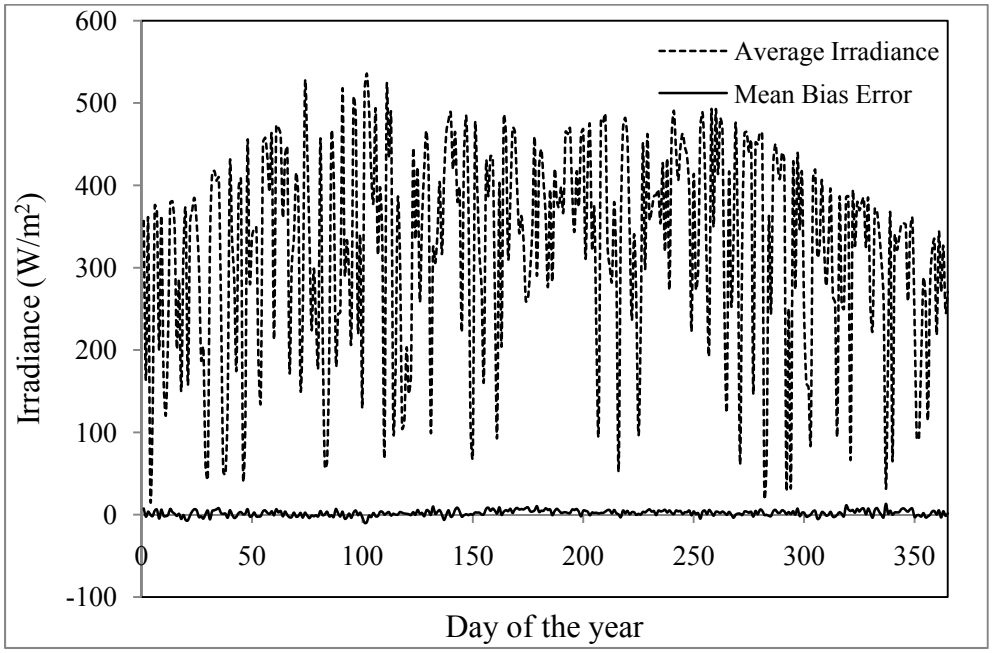

Figure 4-6 Average Irradiance and Mean Bias Error for the year 2005

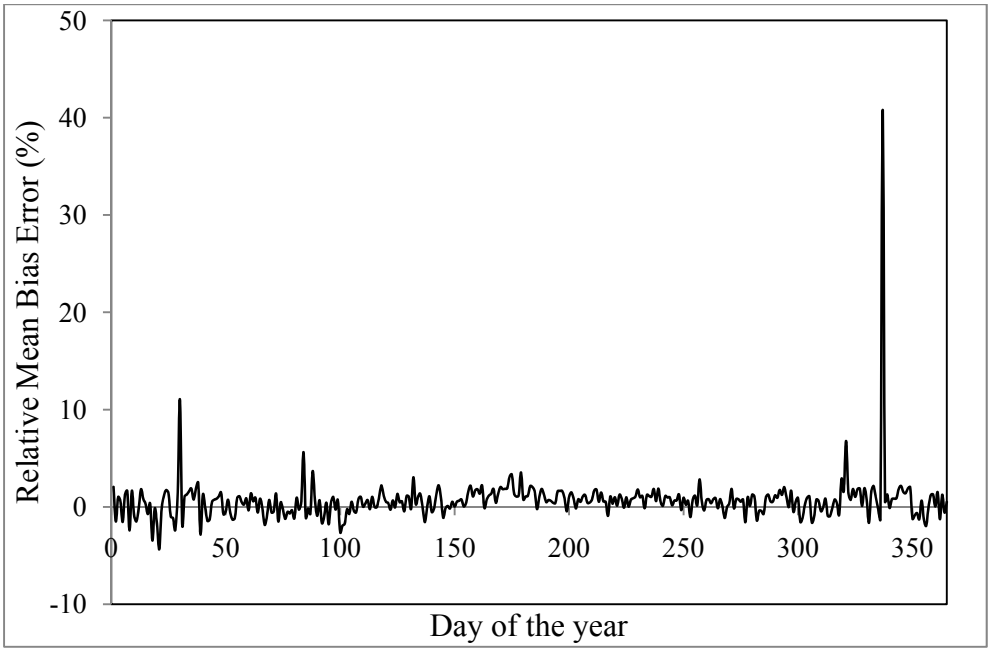

Figure 4-7 Relative Mean Bias Error for the year 2005

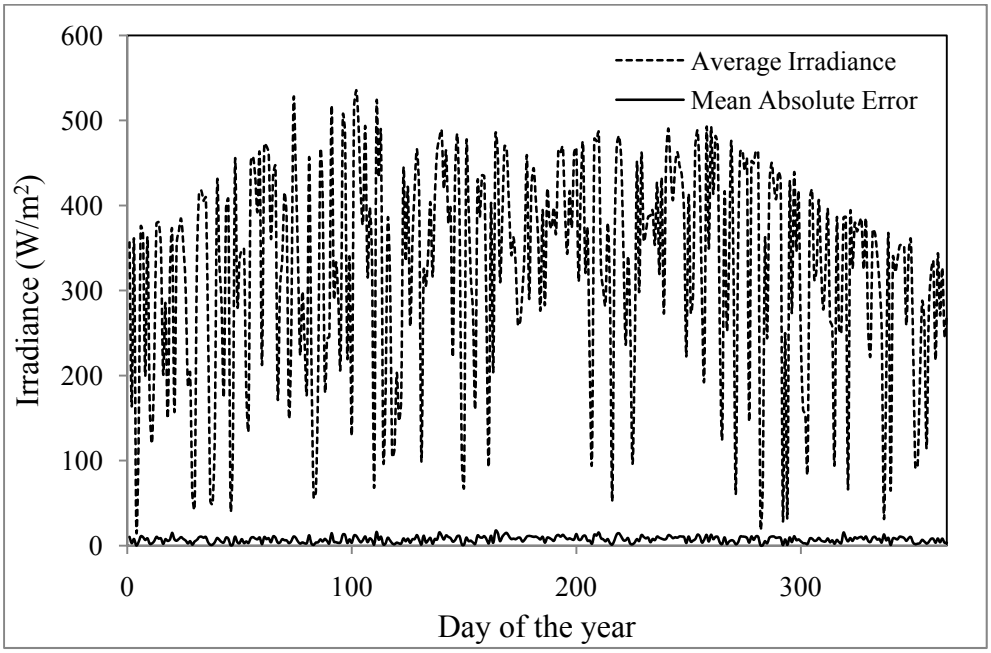

Figure 4-8 Average Irradiance and Mean Absolute Error for the year 2005

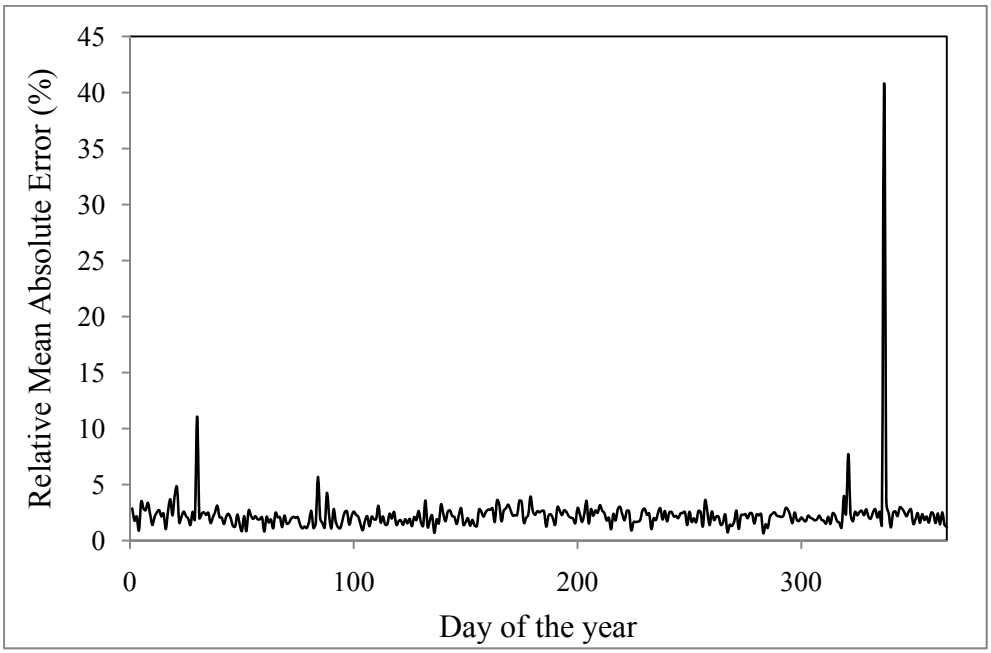

Figure 4-9 Relative Mean Absolute Error for the year 2005

4.3.2 Calculation of the DC output power of the PV array

After estimating the global irradiance on the surface of the PV array, the irradiance and the ambient temperature data are used in a PV model to calculate the maximum output DC power of the PV system, P_{MPP}. As mentioned in Chapter 2, the PV models recorded in the literature vary in their accuracy and complexity. These models can be divided into two main categories: 1) detailed physical models, and 2) simplified mathematical models. Physical models have the advantage of representing the behavior of PV arrays accurately, however, they are complicated [108]. These models are useful for studying the details of the PV system as they provide cell-level information. On the other hand, simplified mathematical models can provide an estimation of the output power of the PV system directly. Thus, these models are more appropriate for studying the impact of PV systems on the electric network [32][109]. The main challenge in using the simplified models is to choose the model that can provide results with good accuracy. In this sub-section, the performance of two of the simplified models is compared with the performance of a detailed physical model. Hence, the model producing more accurate results is used in the conversion stage of the proposed method.

4.3.2.1 Simplified models

The two investigated simplified models estimate the maximum DC output power from the irradiance, G_{tg}, received by a PV array with an area of A_a and a module efficiency of η_m by [32]:

$$P_{MPP} = A_a \times G_{tg} \times \eta_m \tag{4-23}$$

The main difference between the two models is in how to calculate the module efficiency at any irradiance and temperature.

A. Model 1

The module efficiency at any irradiance and temperature can be obtained as a function of the rated efficiency at Standard Test Conditions (*STC*: G_{tg}=1000 W/m^2, T_a = 25 ^{o}C), η_o, by Equation (4-24) [32]:

$$\eta_m = \eta_o \left[1 - 0.0042 \left(\frac{G_{tg}}{18} + T_a - 20 \right) \right] \qquad (4\text{-}24)$$

B. Model 2

A different relation that estimates the module efficiency at any irradiance and temperature can be obtained as a function of the rated efficiency at Standard Test Conditions by [34]:

$$\eta_m = \eta_o \left[1 + K_T \left(T_c - T_{a,STC} \right) \right] \qquad (4\text{-}25)$$

where K_T is the thermal derating coefficient of the PV module in ^{o}C^{-1} and is estimated experimentally to be -0.0037 ^{o}C^{-1}.

The cell temperature, T_c, in ^{o}K is calculated by Equation (4-26) [110]:

$$T_c = T_a + \frac{G_{tg}}{G_{tg,NOCT}} \left(T_{c,NOCT} - T_{a,NOCT} \right) . \left(1 - \frac{\eta_o}{0.9} \right) \qquad (4\text{-}26)$$

where T_a is the ambient temperature in ^{o}K, and

The subscript *NOCT* implies that the parameter is given at nominal operating cell temperature which is the cell or module temperature when the irradiance reaching the module is 800 W/m^2 and the ambient temperature is 20^{o}C.

4.3.2.2 Physical model: PC1D simulator

The physical model used as the basis of comparison is that implemented in the simulator software PC1D [111]. This is a commercial PV cell simulator developed by the University of New South Wales in Australia. The simulator attempts to solve the coupled semiconductor device equations numerically using Newton's method. The input to the model is the detailed parameters of the cell such as the area, thickness, material, doping, and contact resistances. The model also receives the irradiance and temperature in order to calculate the maximum DC power, the open circuit voltage and the short circuit current of the cell.

4.3.2.3 Comparison between the simplified models

To compare between the simplified models, the following steps are taken:

- The first step is to calculate the rated efficiency of the PV cell, η_o. This can be done using the PC1D simulator. The parameters used for the cell are those of a typical low-cost commercial 100-cm^2 silicon solar cell, with series resistance and shunt conductance, provided by the test problem included in the simulator. The base contact resistance is 0.15 Ω and the internal conductance is 0.3 S. The exterior front reflectance of the cell is 10% across the solar spectrum.

  The maximum output DC power of the cell calculated at STC is 1.362W. Accordingly, the rated efficiency of the cell, η_o, is calculated from:

$$\eta_o = \frac{P_{MPP}}{G_{tg} \times A_c} \tag{4-27}$$

  Moreover, the DC powers and efficiencies for a wide range of irradiances and temperatures are also obtained from the simulator and recorded in Table 4-2.

- In the second step, the efficiencies at different irradiances and temperatures are calculated from models 1 and 2 based on the rated efficiency obtained from Step

1. Accordingly, the maximum DC power obtained from each model is estimated from Equation (4-23). The efficiencies and maximum DC powers for both models are recorded in Table 4-2.

- The third step is to calculate the error in estimating the output DC power at each irradiance and temperature for both models. This is useful for identifying the levels of irradiances and temperatures where the performances of the models are not satisfactory. To achieve this task, the Absolute Percentage Error is calculated for each irradiance and temperature using the following relation:

$$APE = \frac{\left| P_{MPP\_PC1D} - P_{MPP\_Model} \right|}{P_{MPP\_PC1D}} \times 100 \tag{4-28}$$

The errors calculated for each of the simplified models are displayed in Table 4-3.

- The last step is to compare the overall performances of the two models by calculating the Mean Absolute Percentage Error, *MAPE*, given by:

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} APE(i) \tag{4-29}$$

where N is the total number of cases examined for each model. Accordingly, the model that produces lower *MAPE* is chosen for the conversion stage of the proposed method.

The results displayed in Table 4-3 shows that Model 2 produces lower *MAPE* as compared to Model 1, and thus, is more accurate. Moreover, the results displayed in this table show that the *APE* of Model 2 is low for almost all cases except for the cases where the irradiance is low (200 W/m^2). In fact, this is considered another advantage of this model as the power generated at low irradiance is also low. Thus, the errors produced in

these cases will not have a significant impact upon studying the performance of the electric network. Accordingly, Model 2 is chosen for use in the Conversion Stage.

Table 4-2 Powers and efficiencies for different models

G_{tg} (W/m^2)	T_c (^{o}C)	T_a (^{o}C)	PC1D		Model 1		Model 2	
			η_m	P_{MPP} (W)	η_m	P_{MPP} (W)	η_m	P_{MPP} (W)
1400	0	-37.13	0.1524	2.1340	0.1244	1.7414	0.1488	2.0832
1400	25	-12.13	0.1347	1.8860	0.1101	1.5412	0.1362	1.9068
1400	50	12.87	0.1185	1.6590	0.0958	1.3410	0.1236	1.7304
1000	0	-26.52	0.1516	1.5160	0.1310	1.3103	0.1488	1.4880
1000	25	-1.52	0.1362	1.3620	0.1167	1.1673	0.1362	1.3620
1000	50	23.48	0.1185	1.1850	0.1024	1.0243	0.1236	1.2360
600	0	-15.91	0.1487	0.8923	0.1377	0.8261	0.1488	0.8928
600	25	9.09	0.1321	0.7924	0.1234	0.7402	0.1362	0.8172
600	50	34.09	0.1166	0.6994	0.1091	0.6544	0.1236	0.7416
200	0	-5.30	0.1192	0.2383	0.1443	0.2886	0.1488	0.2976
200	25	19.70	0.1114	0.2228	0.1300	0.2600	0.1362	0.2724
200	50	44.70	0.0983	0.1965	0.1157	0.2314	0.1236	0.2472

Table 4-3 Errors calculated for the two models

G_{tg} (W/m^2)	T_c (^{o}C)	T_a (^{o}C)	*Model 1*	*Model 2*
			APE (%)	*APE (%)*
1400	0	-37.13	18.3958	2.3815
1400	25	-12.13	18.2810	1.1029
1400	50	12.87	19.1678	4.3051
1000	0	-26.52	13.5674	1.8480
1000	25	-1.52	14.2946	0.0000
1000	50	23.48	13.5614	4.3051
600	0	-15.91	7.4244	0.0550
600	25	9.09	6.5817	3.1297
600	50	34.09	6.4283	6.0350
200	0	-5.30	21.1238	24.8833
200	25	19.70	16.7127	22.2621
200	50	44.70	17.7781	25.8031
MAPE (%)			**14.443**	**8.009**

Usually, the DC power generated from the PV array is affected by several factors, such as, the power loss due to dust on the PV array, the power loss due to module parameter mismatch, the power losses in the DC circuit and the power loss due to DC current ripple and "algorithm error" caused by the switching converter which performs the maximum power point tracking function. These factors can be incorporated in the model by modifying Equation (4-23) to [34]:

$$P_{MPP} = A_a \times G_{tg} \times \eta \tag{4-30}$$

$$\eta = \eta_m \times \eta_{dust} \times \eta_{mismatch} \times \eta_{DCloss} \times \eta_{MPPT} \tag{4-31}$$

where

η_m = module efficiency calculated by Equation (4-25)

η_{dust} = 1 – the fractional power loss due to dust on the PV array

$\eta_{mismatch}$ = 1 – the fractional power loss due to module parameter mismatch

η_{DCloss} = 1 – the fractional power loss in the DC side

η_{MPP} = 1 – the fractional power loss due to DC current ripple and MPPT algorithm error caused by the switching converter

The values of the efficiencies used in this model to estimate the DC power of a 10 MW PV plant are displayed in Table 4-4 [34]. The PV arrays used in this research are assumed to have a total area of 75×10^3 m^2, similar to that of the 10-MW Pocking PV plant in Germany [112].

Table 4-4 Efficiencies used for calculating the DC power of the PV array

η_{dust}	96%
$\eta_{mismatch}$	95%
η_{DCloss}	98%
η_{MPP}	95%

4.3.3 Calculation of the AC output power of the PV system

The last step in the conversion stage is to estimate the AC power generated from the PV system, P_{PV}. The DC power generated from the PV arrays, P_{MPP}, is converted into AC power through an inverter. As mentioned in Section 2.3, a single centralized inverter is usually used with large PV systems, in the order of megawatts, to convert the DC power into AC power. One of the methods that can be used to calculate the converted AC power is to use the manufacturer's efficiency curve for the inverter. This curve relates the input DC power to the inverter, as a percentage of the rated power, to the inverter efficiency, η_{inv}. Accordingly, the AC power, P_{PV}, can be calculated by:

$$P_{PV} = P_{MPP} \times \eta_{inv} \tag{4-32}$$

The manufacturer's efficiency curve used in this research is shown in Figure 4-10 [113].

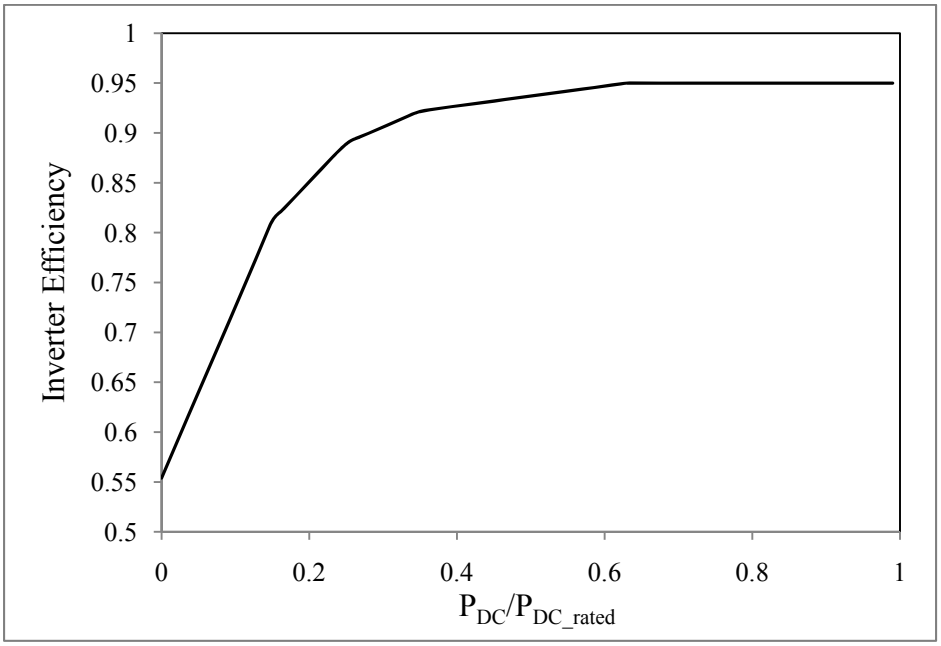

Figure 4-10 Manufacturer's efficiency curve for the inverter

4.4 Segmentation Stage

The output obtained from the Conversion stage is a long time series representing the AC power of the PV system, as shown in Figure 4-11. The main objective of the Segmentation stage is to divide this long time series into short segments of power that can be utilized in the clustering stage. This objective can be achieved in many ways depending on the purpose of the study and the loading profile of the electric network. For example, if the loading of the network is nearly flat during the period when the sun is available, then the daily PV power can be considered as one segment. In this case, 365 power segments will be available for each year. Moreover, if the time resolution of the data is 10 minutes, then each segment would include 144 sample points, as shown in Figure 4-12. In the cases where the load profile of the network has one or more peaks and valleys, dividing the PV power into segments that correspond to the periods of these peaks and valleys can be more useful. For example, the loading of a certain feeder might increase significantly from 8:00 am to 2:00 pm due to the operation of a manufacturing facility installed at that feeder. From 2:00 pm to 8:00 pm, the loading of the feeder decreases. In this case, the long PV time series can be divided into two categories, the first includes the segments of power from 8:00 am to 2:00 pm and the second contains segments of power from 2:00 pm tp 8:00 pm. Accordingly, one year will have 2 categories, each containing 365 segments to be clustered independently. If the time resolution of the data is 10 minutes, then each segment will include 36 sample points.

In this research, the segments were chosen by dividing each year into three categories that represent different seasonal loadings, i.e., winter, summer and spring/fall. Accordingly, for each category, the daily power generated from the PV system is considered as one segment. The main reason behind this choice is that the rural distribution feeder used in this research has three different smooth loading patterns that correspond to the aforementioned seasons. The feeder and its loading profile will be introduced in the next chapter.

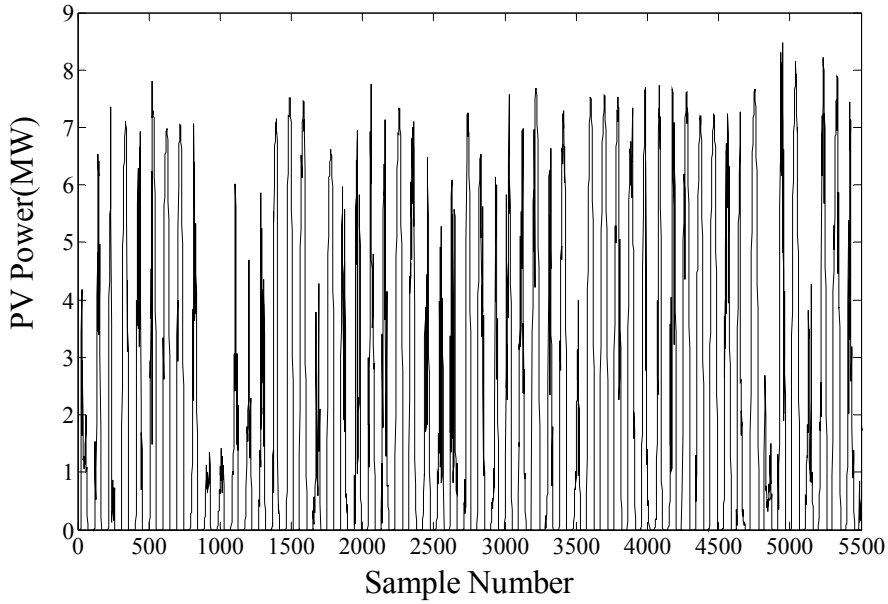

Figure 4-11 Long time series representing the AC power of the PV system

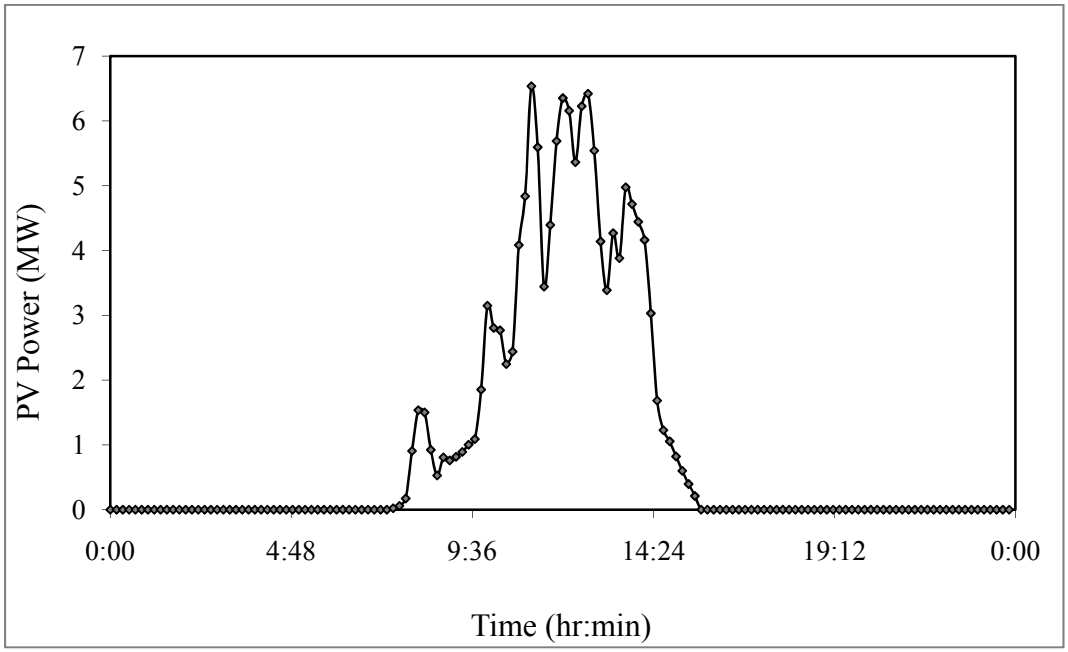

Figure 4-12 A Segment representing the power of one day

4.5 Feature Extraction Stage

After dividing the long time series of the PV power into segments, these segments are passed to the Feature Extraction stage. The main objective of this stage is to identify features that can provide information about the profiles of the PV power segments to facilitate grouping segments having similar profiles together. There are a variety of techniques that can be used for this purpose. However, this research will focus on using two sets of features extracted using two different techniques.

First features set

One possibility is to consider the power values corresponding to each time step of the PV power segment as the features for this segment. Accordingly, each segment has 144 sample points or features as explained in the previous section. This number of features can be reduced to 96 if only the period from 4:00 am till 8:00 pm is considered, the period during which the sun is available. The reason behind eliminating the rest of the 144 features from the feature data set is that the values of these features are equal to zero in all the PV power segments. Thus, the information contained in these features is redundant.

Second features set

Another set of features can be extracted using the Principal Component Analysis (PCA). PCA, also called Karhunen-Loève transform, is a technique that is usually used for reducing the dimensionality of large data sets while retaining, as much as possible, the variation present in the data [114]. The technique is based on orthogonal linear transformation that transforms the correlated original data to a new coordinate system, or a new feature space, where the components of the new system are uncorrelated. The first coordinate in the new system coincides with the direction of the greatest variance of the original data; this coordinate is defined as the first principal component. The second coordinate lies in the direction of the second greatest variation of the data, and so on.

Thus, for dimensionality reduction, it is possible to neglect the higher principal components that retain the least variance of the data [114][115]. Moreover, the PCA can be used in identifying patterns in data, and expressing the data in such a way that their similarities and differences are highlighted [116].

PCA is widely applied to different fields of science such as spatial and temporal variation in atmospheric science, stock market prices, properties of chemical compounds and anatomical measurements [114]. In the field of electrical engineering, many studies utilize PCA for feature extraction. For example, in [117], it was required to group similar partial discharge signals obtained from a power transformer together. The method adopted for feature extraction and dimensionality reduction in this study was PCA. In [118], PCA was used to group generators having similar features together; such features included bus angles and generators' speeds. Furthermore, PCA was one of the methods examined in [119] to identify customers having similar electric load patterns.

In the proposed research, the main goal of using PCA is to extract the common modes of variation present in the PV power segments, to be able to group similar segments together. To achieve this goal, the principal components (PCs) that represent all the segments are calculated. As a result, each segment is related to the calculated PCs by scalar coefficients called the transformation coefficients. The number of transformation coefficients for every segment is equal to the number of calculated PCs. The next step is to choose the PCs that retain most of the variation in the data. Let us assume that the number of these components is L. Thus, every segment has a corresponding vector that contains L transformation coefficients. If the transformation coefficients of two segments are close, then these segments have close profiles. Hence, the vectors containing the transformation coefficients can be considered as features vectors that can be clustered to group together segments having close patterns. The steps of applying PCA are summarized in Appendix A [116].

To choose the set of features that can generate better groups of segments, a clustering algorithm should be used. Accordingly, the comparison between the two features sets will be presented later after presenting the clustering algorithms that are used in this research.

4.6 Clustering Stage

The main objective of the clustering stage is to group together PV power segments having close features in the same cluster. The formed clusters are then passed to the identification stage where a representative segment, called the cluster representative, is chosen to represent the corresponding cluster. These representatives can be used to evaluate the performance of the electric network due to the installation of the PV system. In this case, all the cluster representatives, forming the reduced data set, are used in the power flow analysis and the monitored electric quantities are recorded for each representative. The representatives that have the potential to violate the acceptable limits of the monitored electric quantities are then identified along with their corresponding clusters. Accordingly, the segments in the identified clusters can be used in detailed analysis in order to accurately assess the performance of the feeder.

There are a number of clustering algorithms that can be used to group together the features of the PV power segments extracted from the previous stage. Each algorithm produces different clusters, and thus, different cluster representatives, i.e., different reduced data sets. This means that the results obtained from the power flow analysis using the reduced data set are highly dependent on the applied clustering algorithm. Therefore, it is important to apply more than one clustering algorithm on the data set under study in order to choose the most suitable algorithm for this data set.

In general, Clustering algorithms can be divided into two main categories: partitional and hierarchal [120]. In partitional algorithms, all the clusters are determined at once. An example for this category is the K-means algorithm. In this algorithm, the number of clusters, n_C, should be specified in advance. Hence, the algorithm starts by choosing n_C

random centroids for the data, which is called the initialization process. Each data point is assigned to the closest cluster centroid and then the centroids are updated. The assigning and updating processes are repeated until there is no change in the cluster centroids. This corresponds to the minimum Sum of Squared Errors (SSE) between the data points in each cluster and the cluster centroid. The main advantage of the K-means is its simplicity. However, the algorithm is sensitive to the initialization process, as the random choice of initial centroids can lead to the formation of poor clusters. This issue can be overcome by repeating the clustering several times and choosing the clusters that produce the least SSE [121].

Hierarchical algorithms are based on two approaches: 1) agglomerative, which starts with every data point as a single cluster and then merges the closest pairs of clusters, and 2) divisive, which starts with all the data points in one cluster and then splits the cluster until every data point is put in one cluster [122]. Hierarchical clustering yields a dendrogram that represents the clusters and sub-clusters in the order by which they were grouped or divided, as shown in Figure 4-13. The agglomerative approach is the most commonly used and requires calculating the proximity between clusters to decide which clusters should be grouped together. Three main methods can be used to calculate the proximity between clusters: 1) Single Linkage, where the distance between two clusters is the minimum distance between all data points in both clusters (one data point from each cluster); 2) Complete Linkage, where the maximum distance between data points in both clusters is considered; and 3) Average Linkage, where the average distances between all pairs of data points in both clusters are considered. Another hierarchical clustering algorithm is the Ward method which uses the SSE as the proximity measure.

The three clustering algorithms compared in this research are: 1) the Average Linkage hierarchical algorithm, 2) the K-means algorithm, and 3) a hybrid algorithm that uses the Ward method to generate clusters and hence use the centroids of these clusters to initialize the K-means algorithm.

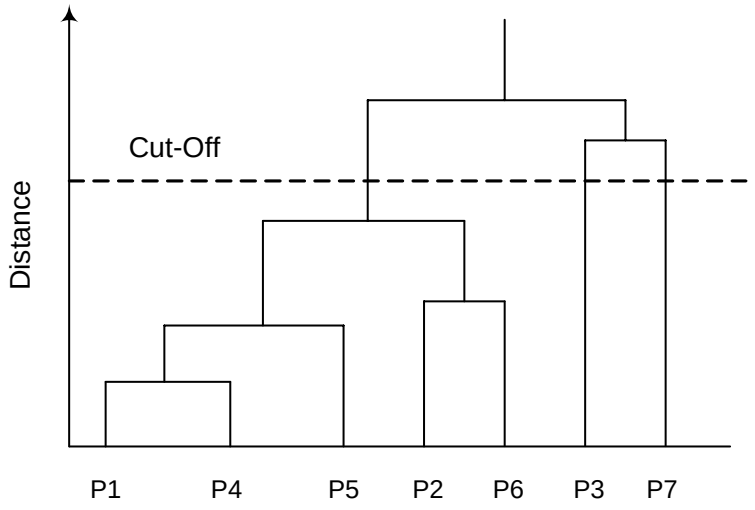

Figure 4-13 Dendrogram obtained from hierarchical clustering

To perform clustering, a dissimilarity measure (or a similarity measure) is required to measure the distances between data points. In this thesis, the Euclidian metric, which is the most commonly used dissimilarity measure, is used. This metric measures the distance between the i^{th} segment and the k^{th} segment, *Dist(i,k)* and is given by [122]:

$$Dist(i,k) = \left[\sum_{j=1}^{n_f} \left(f_{ij} - f_{kj} \right)^2 \right]^{\frac{1}{2}} \tag{4-33}$$

where f_{ij} is the j^{th} feature of the i^{th} segment, f_{kj} is the j^{th} feature of the k^{th} segment, and n_f is the number of features.

To choose a suitable clustering algorithm, the three algorithms should be compared using cluster validity indices. These indices, as well as the comparison between the clustering algorithms, will be presented in the next chapter.

4.7 Identification Stage

AS mentioned before, the main goal of this stage is to choose a representative segment for each cluster that can represent all the other segments in the cluster in the power flow

analysis. Accordingly, the clusters whose representatives have the potential to negatively affect the acceptable performance of the electric network can be identified. In that case, more thorough analysis can be carried out using the segments of these clusters.

One of the most popular cluster representatives is the centroid of the cluster which is a segment formed by the calculated mean of features of all segments included in the cluster. Another popular cluster representative is the medoid which is a segment that is originally included in the cluster and is the closest to the calculated centroid. In some cases when the cluster has an elongated shape, the boundary segments of the cluster can be used as the cluster representatives [120]. However, this is not a very common representative as compared to the centroid and the medoid.

In this research, the centroids and medoids, two sets of cluster representatives, are compared to choose the most suitable representative for the data set in hand. This will be presented in the next chapter after choosing the suitable feature extraction technique and clustering algorithm.

4.8 Summary and Conclusions

This chapter introduced the layout of a new method that can be used in studying the impacts of large grid-connected PV systems on the performance of the electric network. One of the advantages of the new method is its ability to use the long historical time series data in an efficient and intelligent way in order to preserve the temporal information contained in the data. The main idea of the proposed method is to divide the long PV power time series into segments that can be grouped together according to their profiles, such that the segments with close profiles are grouped in one cluster. The next step is to choose a representative segment for each cluster and to use these representatives in the power flow simulations. Accordingly, the representatives and their respective clusters that have the potential to violate the acceptable performance of the electric

network can be identified. In the cases when an undesirable performance of the network is observed, a thorough analysis can be carried out by using all the segments of the group.

This chapter also introduced the details of several stages of the proposed method. In the first stage, the models that are used to convert the temperature and irradiance obtained from the weather stations into the AC power generated from the PV system were presented. Moreover, the accuracy of these models was examined in order to choose the most suitable models to be used in this stage.

In the second stage, the different methods that can be used to segment the PV power time series into segments were explained. These methods depend mainly on the loading profile of the electric network, which is considered an advantage for the proposed method.

In the third stage, the feature extraction stage, the use of two different sets of features was proposed. The first set of features is the sampled data of each segment while the second set is the conversion coefficients obtained from applying the Principal Component Analysis. The reason for choosing these sets of features is their ability to reflect the profile of the power segments, and accordingly, facilitate their grouping according to their profile in the clustering stage.

Three different clustering algorithms that can be applied in the proposed method were introduced in the fourth stage. Each clustering algorithm uses a different approach in grouping the data points together. Thus, different groups or clusters will be formed from the different clustering algorithms when applied on the same data set. This requires the use of an index to compare the performances of the clustering algorithms, and hence, choose a suitable algorithm.

In general, different feature extraction techniques and different clustering algorithms, as well as the use of different cluster representatives, generate different results obtained from the power flow analysis. Thus, one of the most important steps in the proposed method is to compare different techniques used in these stages to be able to choose the technique that is most suitable for the data set in hand.

In the following chapter, the comparison between different techniques used in the last three stages of the proposed method is presented. Moreover, the application of the proposed method to investigate the impacts of a large grid-connected PV system on the performance of the electric network is illustrated.

Chapter 5

Analyzing the Impacts of Large Grid-Connected PV Systems

5.1 General

In the previous chapter, the layout and details of different stages of the clustering-based method were introduced. As explained before, some stages of the proposed method require comparison between techniques and methods that can be used in these stages. The main objective of this comparison is to choose the techniques that are most suitable for the data set used in the analysis. Thus, the first goal of this chapter is to introduce the indices that can be used to compare these techniques with one another, especially those used in the Clustering stage. The performance of these indices is evaluated and critiqued, and accordingly, a new index that is more suitable for the analysis of the electric network is proposed.

The second goal of this chapter is to highlight the merits of applying the proposed method by using it to investigate the performance of a rural distribution feeder in the presence of a large grid-connected PV system. Moreover, the results obtained from the proposed method are compared with those obtained from other methods that are usually used for the same type of studies.

5.2 Comparing Between Techniques Used in Different Stages of the Proposed Method

In the Clustering stage, the objective of the comparison is to choose an algorithm that generates more homogeneous clusters containing segments with similar profiles. Usually, cluster validity indices are used to compare different clustering algorithms with one another[117]. In the following sub-section, the cluster validity indices that are used in this research are presented.

5.2.1 Cluster Validity Indices

To evaluate the clustering algorithms, three types of indices can be used [123]: 1) external or supervised validity indices, 2) internal or unsupervised validity indices, and 3) relative validity indices. External validity indices require the knowledge of external information about the data such as the data labels, which is usually not available for the type of study carried in this research. On the other hand, internal validity indices usually measure the quality of clustering by measuring the cohesion within each cluster and the separation between different clusters. Hence, the clustering algorithm that produces tighter and better separated clusters is considered suitable for the data set under study. The relative validity indices use either the external or internal indices to compare different clustering algorithms with one another. The clustering algorithms considered in this research are compared using the relative validity indices that utilize the following internal validity indices:

a- Silhouette Index, SI [124]:

This index calculates the average distance, Aa_i, between the i^{th} segment and all other segments in the same cluster. It also calculates the average distance, Ba_i, between the i^{th} segment and all other segments not included in the same cluster. Hence, the Silhouette coefficient of the i^{th} segment, Sil_i, can be obtained by:

$$Sil_i = \frac{Ba_i - Aa_i}{\max(Aa_i, Ba_i)} \tag{5-1}$$

Accordingly, the Silhouette index, *SI*, can be found by calculating the average of the individual Silhouette coefficients. The value of this index can vary between -1 and 1, where a value closer to 1 means a better clustering.

b- Davis-Bouldin Index, DBI [125]:

This index is a function of the ratio of the sum of within-cluster scatter to between-cluster separation. The index calculates the average distance, Ca_i, between all segments in cluster i and their centroid. It also calculates the distance, $CC_{i,j}$, between the centroids of clusters i and j. Hence, the Davis-Bouldin index, *DBI*, can be calculated for n_C clusters by:

$$DBI = \frac{1}{n_C} \sum_{i=1}^{n_C} \max_{j, i \neq j} \frac{Ca_i + Ca_j}{CC_{i,j}} \tag{5-2}$$

A small *DBI* indicates that the clusters are compact and far from each other. Consequently, this index will have a small value for a good clustering algorithm.

c- Partition Index, SC [126]:

This index was originally proposed for evaluating clusters formed using fuzzy clustering algorithms. However, the index can be applied for crisp clustering algorithms if the fuzzy membership, μ_{ik}, is limited to the values 0 and 1. The index is the summation of the ratios of the compactness of each cluster to its separation from other clusters; thus, a lower value for the index indicates better clustering. The index is given by:

$$SC = \sum_{i=1}^{n_C} \frac{\sum_{k=1}^{n_s} \mu_{ik}^2 \left\| s_k - C_i \right\|^2}{n_i \sum_{t=1}^{n_C} \left\| C_i - C_t \right\|^2} \tag{5-3}$$

where s_k is segment number k, C_i is the centroid of cluster i, n_C is the total number of clusters, n_s is the total number of segments, and μ_{ik} is the fuzzy membership. For crisp clustering, $\mu_{ik} = 1$ if segment k belongs to cluster i, and 0 otherwise.

5.2.2 Comparison of the clustering algorithms using the internal validity indices

The three clustering algorithms presented in Section 4.6 are compared using the validity indices presented in sub-section 5.2.1. To compare the three algorithms properly, the dendrograms generated from the hierarchical algorithms are cut off to produce the same number of clusters specified for the *K-means* algorithm. For each algorithm, the number of clusters is increased in steps from 2 clusters up to 100 clusters. For each number of clusters, the three validity indices, *SI*, *DBI*, and *SC*, are calculated. This procedure is applied to the 276 PV power segments of the summer season of the 3-year data set obtained for the years 2005 to 2007. The indices calculated for the three clustering algorithms are displayed in Figure 5-1 to Figure 5-3 and the clusters obtained from the three algorithms for 25 clusters are displayed in Figure 5-4 to Figure 5-6.

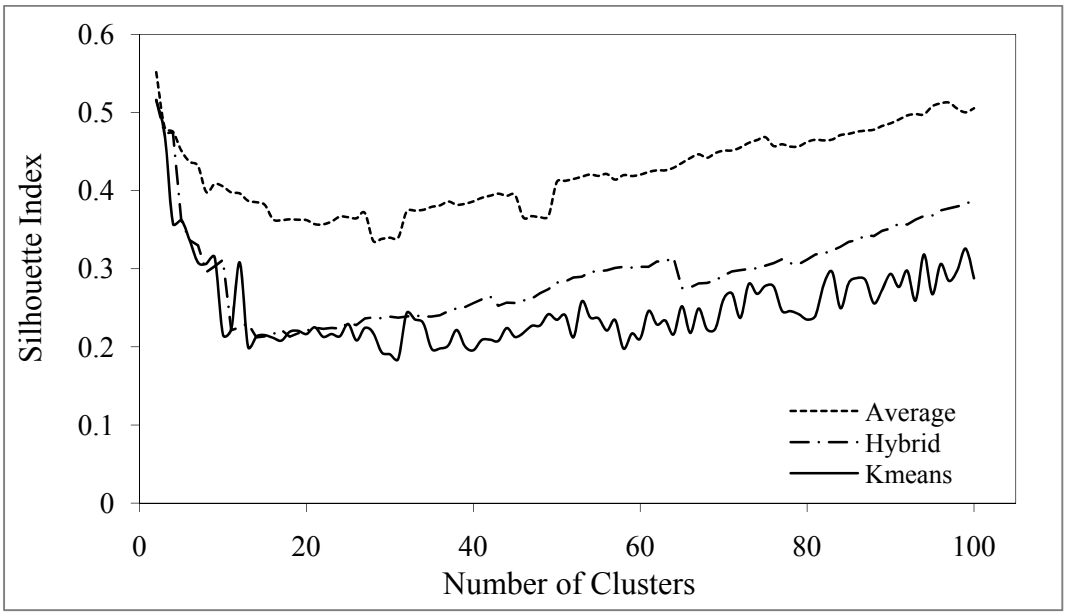

Figure 5-1 Silhouette index for the three clustering algorithms

Figure 5-2 Davis-Bouldin index for the three clustering algorithms

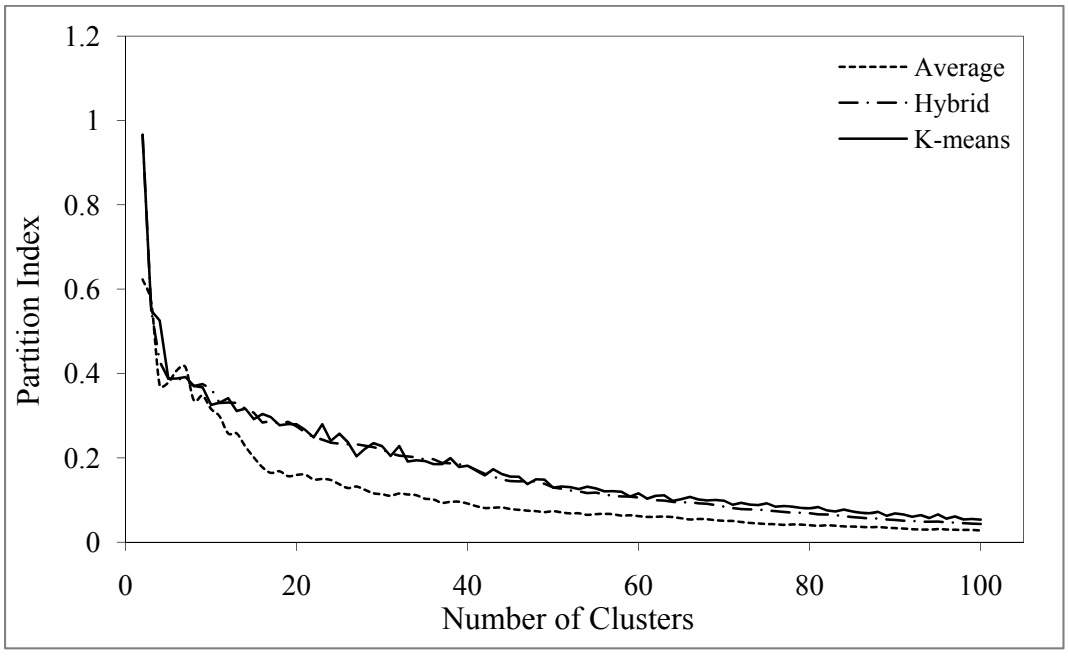

Figure 5-3 Partition index for the three clustering algorithms

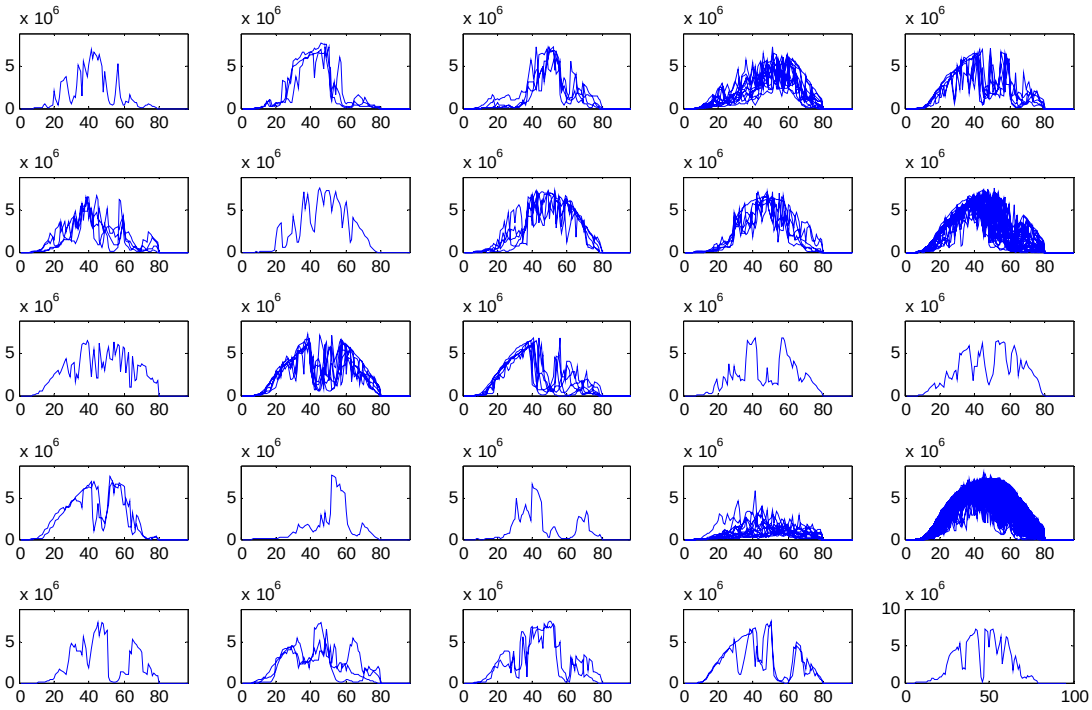

Figure 5-4 Clusters obtained using the Average Linkage hierarchical algorithm

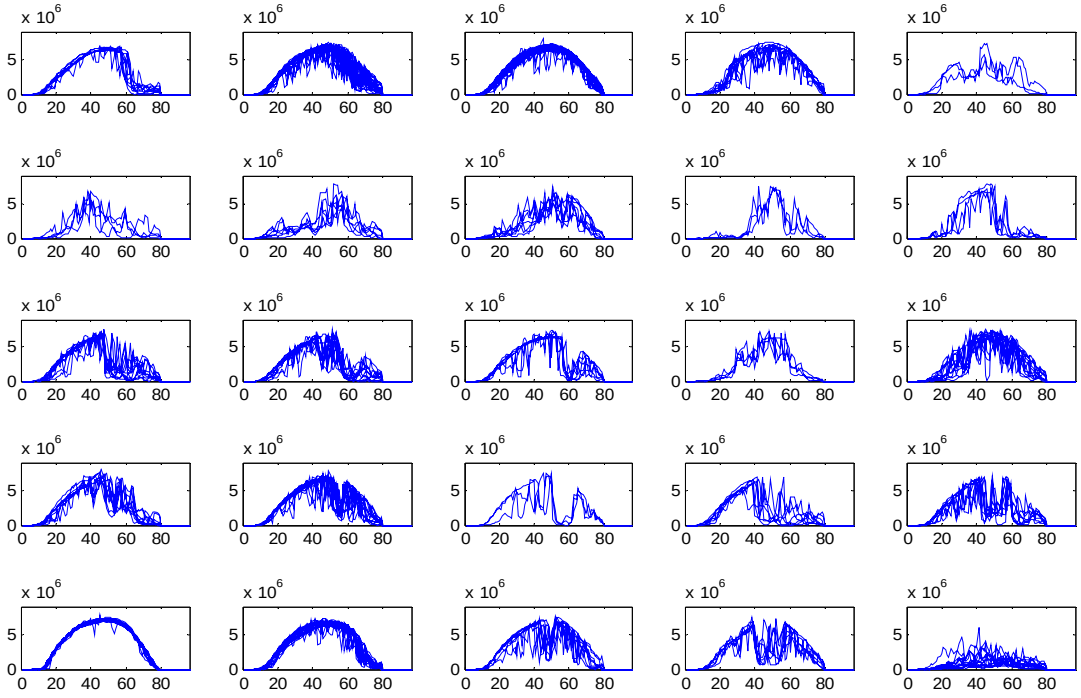

Figure 5-5 Clusters obtained using the Hybrid algorithm

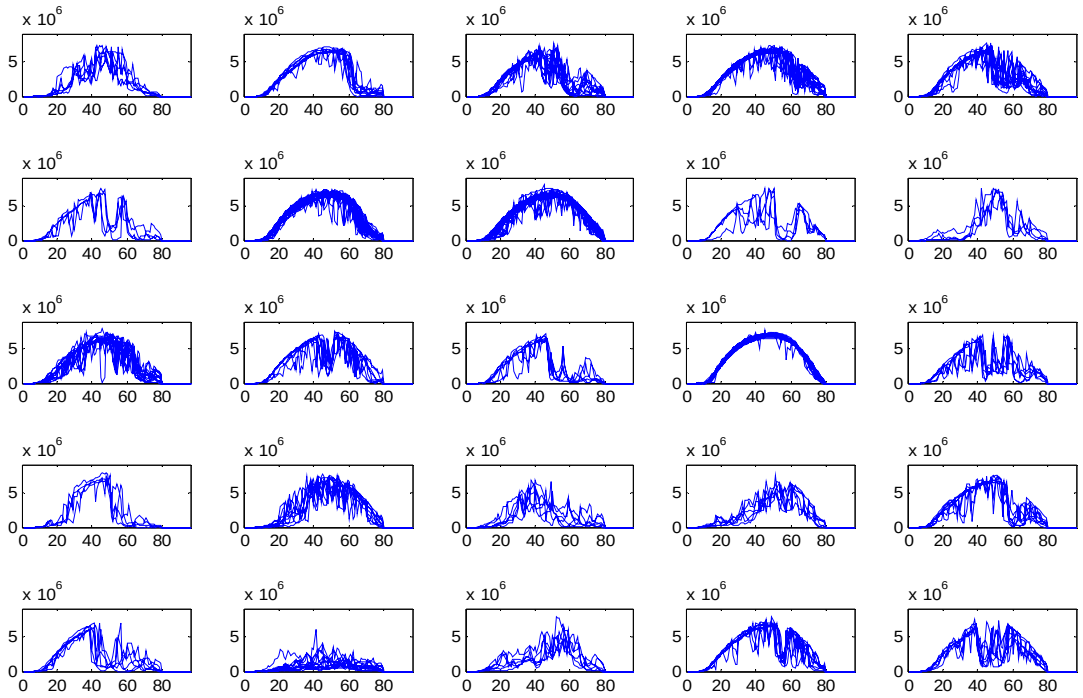

Figure 5-6 Clusters obtained using the K-means algorithm

The indices displayed in Figure 5-1 to Figure 5-3 indicate that for all cluster numbers, the Average Linkage hierarchical clustering algorithm generates clusters that are more compact and separated than those generated by the other two algorithms. However, by examining the clusters formed by applying this algorithm, Figure 5-4, it can be observed that there are a number of clusters each containing a single segment (singletons) and other clusters containing segments that do not have close profiles, e.g., cluster number 10 (Row 2, Column 5) and cluster number 20 (Row 4, Column 5). On the other hand, Figure 5-5 shows that the Hybrid algorithm produces clusters that are more homogeneous than those obtained from the Average Linkage hierarchical algorithm. For example, segments that represent sunny days, i.e., small fluctuations, are grouped together in cluster number 21 (Row 5, Column 1), and the segments that represent overcast days, i.e., low output power from the PV system, are grouped together in cluster number 25 (Row 5, Column

5). The same observation can be stated for the clusters obtained from applying the K-means algorithm, displayed in Figure 5-6.

The contradiction between the results obtained from the validity indices and those observed from inspecting the formed clusters can be explained by examining the formulas that calculate these indices, Equation (5-1) to Equation (5-3). One of the aspects examined by these indices is the compactness of each cluster, which is mainly evaluated by calculating the distance between the segments or the distance between the segments and the centroid of the same cluster. Accordingly, for the clusters containing only one segment (singletons), the compactness is zero, thus, producing better values for these indices. In other words, the examined indices tend to prefer the clustering algorithm that produces more clusters containing single segments, as in the case of the Average Linkage hierarchical algorithm. This preference does not serve the main purpose of the proposed method, which is to group power segments that have similar profiles in the same cluster. Moreover, these indices cannot provide any information about the goodness of using a certain group of cluster representatives in representing the segments present in the cluster when performing power flow analysis. As a conclusion, it is essential to develop a new index that is suitable for the main purpose of the proposed method.

5.2.3 The proposed validity index for utility studies

In this sub-section, a new index that can be used to evaluate the clustering algorithms for the purpose of electric network analysis is introduced [127]. However, before introducing the new index, it is important to recall the main purpose of applying the clustering algorithms in the proposed method. As mentioned before, the idea is to group the PV power segments having similar profiles together. This means that the features of the segments in the same cluster should be close to each other. In other words, the values of the power at each time step should be close for all segments contained in the same cluster. The next step that follows the grouping of power segments is to identify cluster representatives that can replace the segments included in the clusters when performing

power flow analysis. Therefore, evaluation of clustering algorithms should achieve two main goals: 1) to ensure that the segments in each cluster have close features as compared to the segments that are not included in the cluster, and 2) to evaluate how accurate the cluster representatives can represent the segments in the corresponding clusters.

To achieve these goals, the proposed index compares two data sets for each clustering algorithm. The first data set is the original set formed from all the available power segments. The second data set is a virtual set formed from the cluster representatives each replicated according to the number of segments contained in the original cluster. The steps for calculating the proposed index are:

1. For the *n* segments included in the cluster, calculate the cluster representative, e.g., the centroid.
2. Form a virtual cluster containing *n* replicas of the cluster representative.
3. Specify a number of intervals that cover the whole range of the power generated from the PV system. For example, for a 10-MW PV system 10 intervals can be specified.
4. Using the specified intervals, calculate the histograms of the PV power, $P_{PV\text{-}freq}$ and $P^{*}_{PV\text{-}freq}$, for both the original and virtual clusters, respectively. These histograms display the frequency of occurrence of the power in each of the specified ranges as calculated from the original and virtual clusters.
5. Compare the histogram calculated from the original clusters with that obtained from the virtual clusters using the Mean Absolute Percentage Error, *MAPE*:

$$MAPE = \frac{1}{N_i}\sum_{i=1}^{N_i}\frac{\left|P_{PV-freq}(i) - P^{*}_{PV-freq}(i)\right|}{P_{PV-freq}(i)} \tag{5-4}$$

where

N_i = number of intervals

$P_{PV\text{-}freq}$ = frequency of occurrence of the PV power in the i^{th} interval calculated from the original clusters

$P^*_{PV\text{-}freq}$ = frequency of occurrence of the PV power in the i^{th} interval calculated from the virtual clusters

Thus, to make a comparison among the clustering algorithms, the algorithm that produces a lower *MAPE* value indicates that the clusters generated from this algorithm are more homogeneous and the calculated cluster representatives are more capable of representing the original PV power segments. Accordingly, this algorithm is considered more suitable for application with the data set under study.

5.2.4 Comparing the clustering algorithms using the proposed index

The proposed index is used to compare the three clustering algorithms when applied on the summer seasons of the 3-year data set used in Sub-section 5.2.2. Figure 5-7 shows that both the Hybrid and K-means algorithms almost always produce lower values for the calculated *MAPE* as compared to the Average Linkage hierarchical algorithm, which agrees with the results displayed in Figure 5-4 to Figure 5-6. Accordingly, it can be concluded that the clusters generated from the two algorithms are more homogeneous and the cluster representatives calculated from these algorithms can better replace the full data set in network analysis. The results displayed in Figure 5-7 also show that the K-means algorithm is sensitive to the initialization process, even when repeating the algorithm 50 times, as done in the current study. This sensitivity is seen from the fluctuations in the values of the *MAPE* calculated for the K-means algorithm. On the other hand, the *MAPE* calculated for the Hybrid clustering algorithm does not have fluctuations and is close to the *MAPE* calculated for the K-means algorithm.

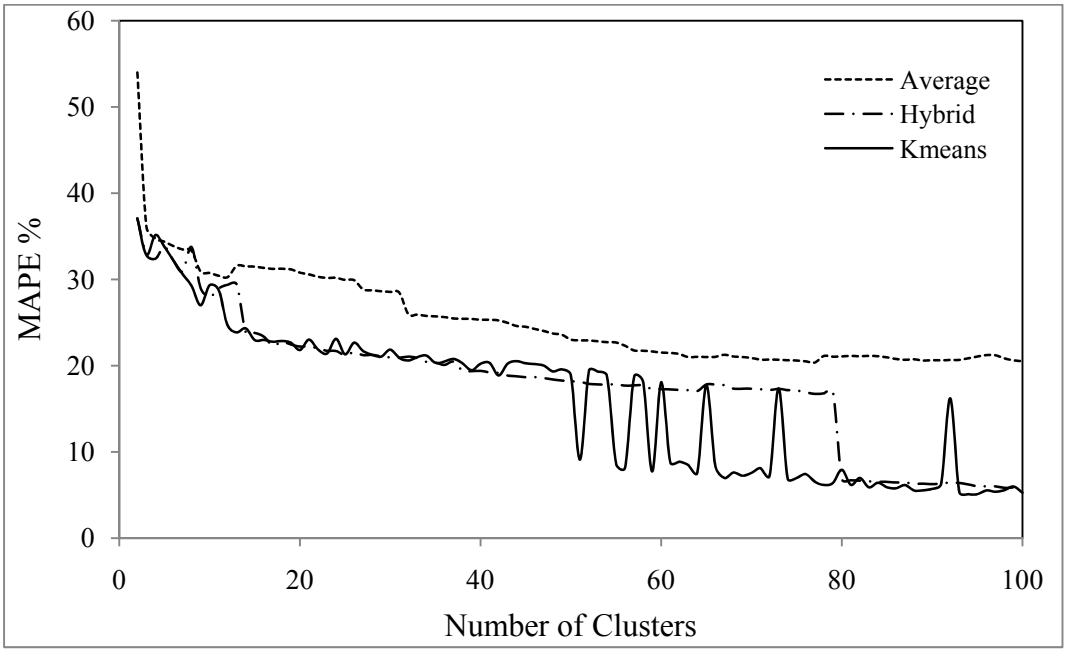

Figure 5-7 MAPE index for the summer season of the 3-year data set

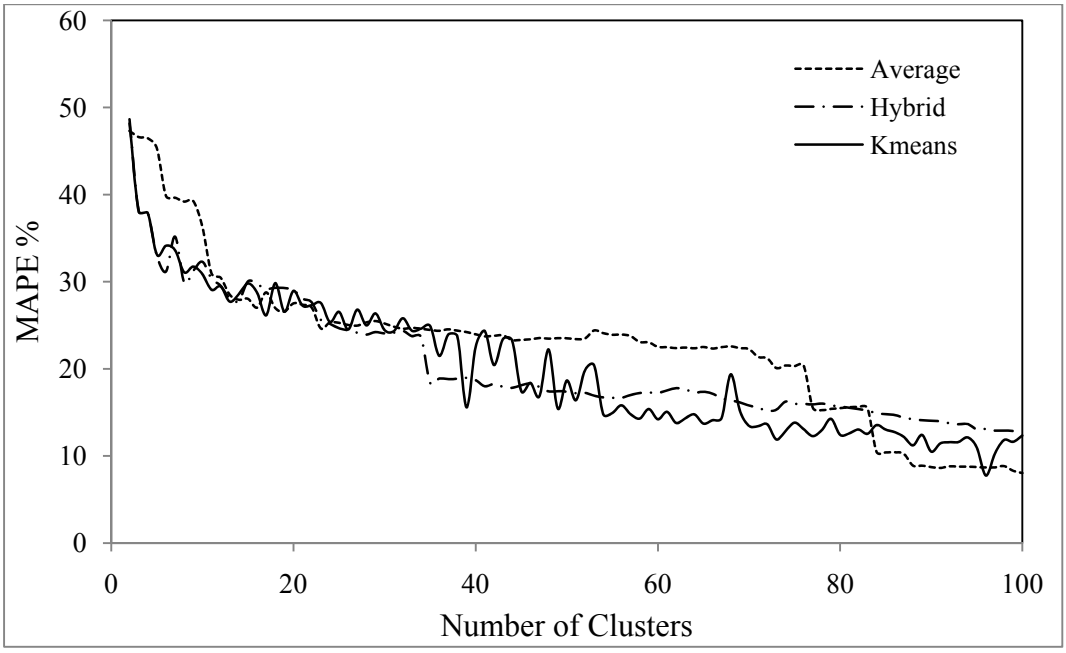

Figure 5-8 MAPE index for the winter season of the 3-year data set

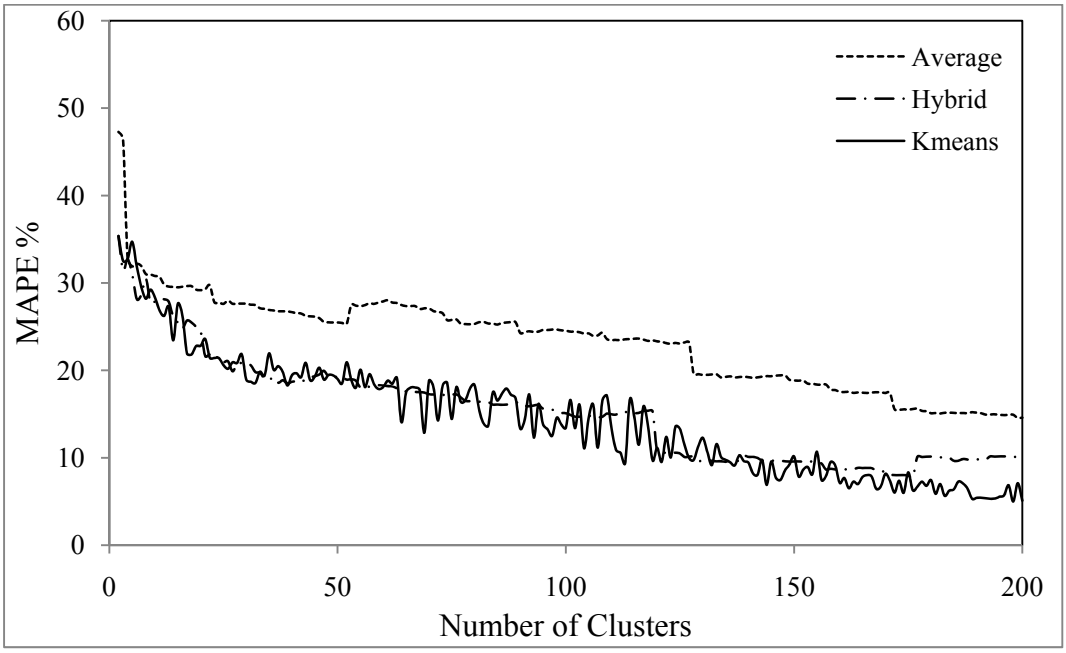

Figure 5-9 MAPE index for the spring/fall season of the 3-year data set

Moreover, the Hybrid algorithm is much less computationally expensive because the algorithm does not require to be repeated a number of times as in the case of the K-means algorithm. The same results can be observed from Figure 5-8 and Figure 5-9 obtained upon applying the three clustering algorithms on the 270 power segments of the winter season and the 549 power segments of the spring/fall season. Accordingly, the Hybrid algorithm is favored for clustering these data sets. The comparison between the three clustering algorithms using the proposed index when applied on the three seasons of the year 1999 (1-year data set) and the years 2000 to 2004 (5-year data set) are displayed in Appendix B. All the results agree with the conclusions stated in this section.

5.2.5 Comparing the feature extraction techniques

The feature extraction techniques used in the Feature Extraction stage can be compared using the index proposed in this chapter. To achieve this task, the two sets of features presented in Section 4.5 are used with one of the clustering algorithms. The first set of

features is the 96 features representing the data points of each power segment. The second set is the transformation coefficients relating each segment to the principle components obtained from the complete data se. In this case, the PCA is applied on the 3-year data set and the principle components that retain 90% variation of the data are calculated from Equation (A-2) in Appendix A. Thus, the 96 features of the segments of the three seasons are reduced to 8 features for the spring/fall season, 13 features for the summer season, and 6 features for the winter season. The Hybrid clustering algorithm is applied using the two sets of features, and the results displayed in Figure 5-10 to Figure 5-12 show that the values of the *MAPE* for the two feature extraction techniques are almost always close. This indicates that a reduced number of features obtained by PCA can be used in the analysis without sacrificing the accuracy. Reducing the dimensions of data reduces the computational effort, which is a great advantage, especially in the case of very large data sets.

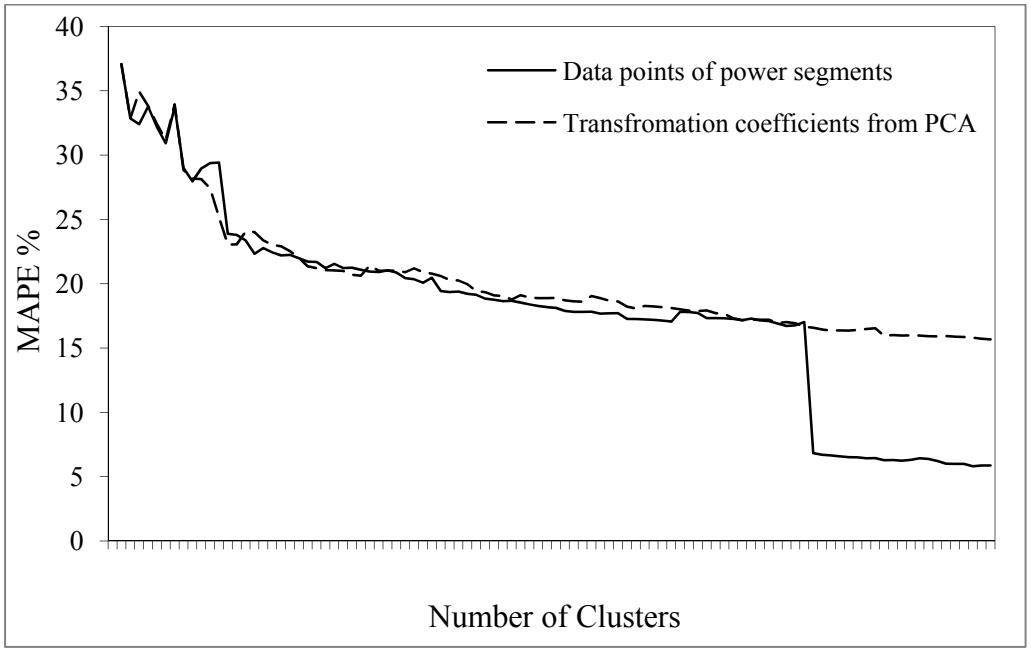

Figure 5-10 Comparison between the two sets of features for the summer season of the 3-year data set

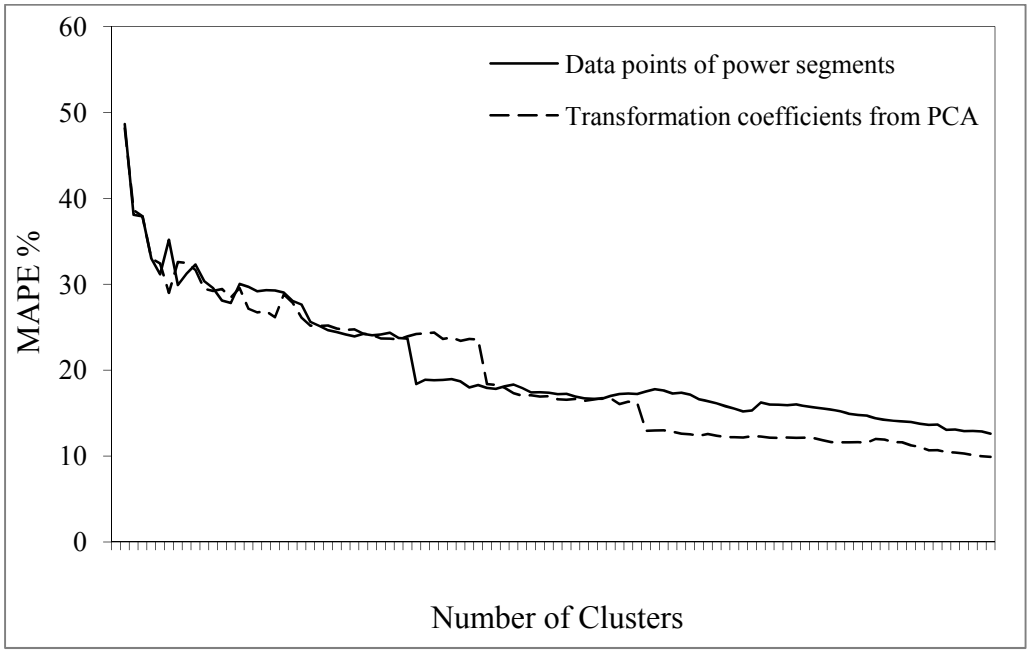

Figure 5-11 Comparison between the two sets of features for the winter season of the 3-year data set

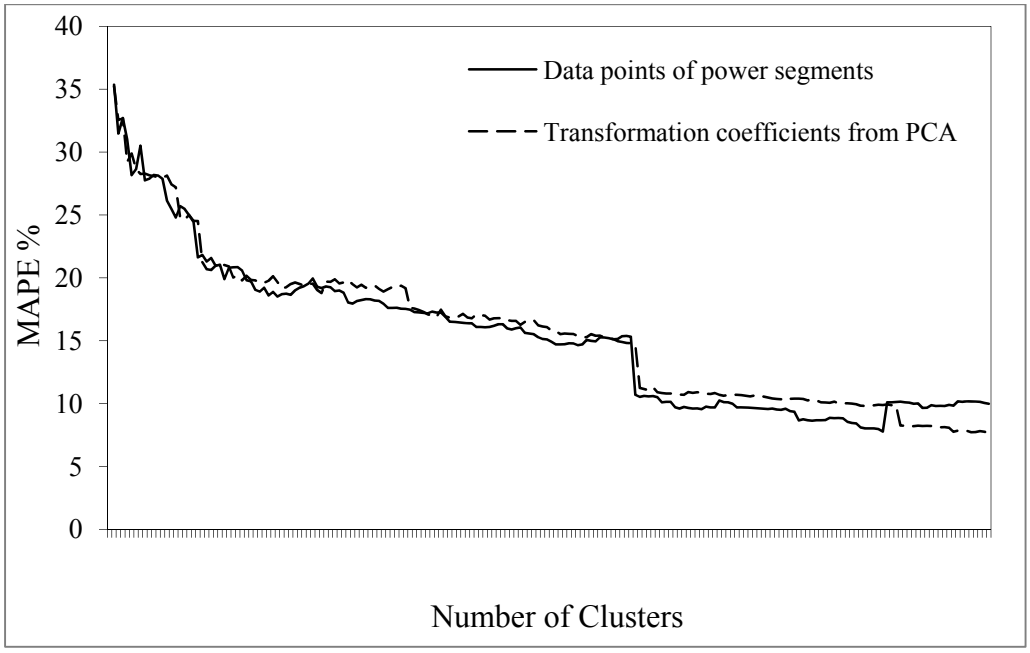

Figure 5-12 Comparison between the two sets of features for the spring/fall season of the 3-year data set

5.2.6 Comparing the cluster representatives

As mentioned before, one of the main advantages of the proposed index is its ability to provide information about using a certain group of cluster representatives in representing the full data set. This information facilitates the comparison between different groups of cluster representatives to choose the most suitable for the data set they represent. In this sub-section, two groups of cluster representatives are compared together to choose the most suitable one for different seasons of the 3-year data set. The first group of representatives is the one obtained by calculating the centroid for each cluster, while the second group is that obtained by identifying the medoid for each cluster. The main difference between the two groups is that the centroid is a virtual segment calculated from all other segments, while the medoid is a segment that already exists in the cluster and is the closest to the centroid.

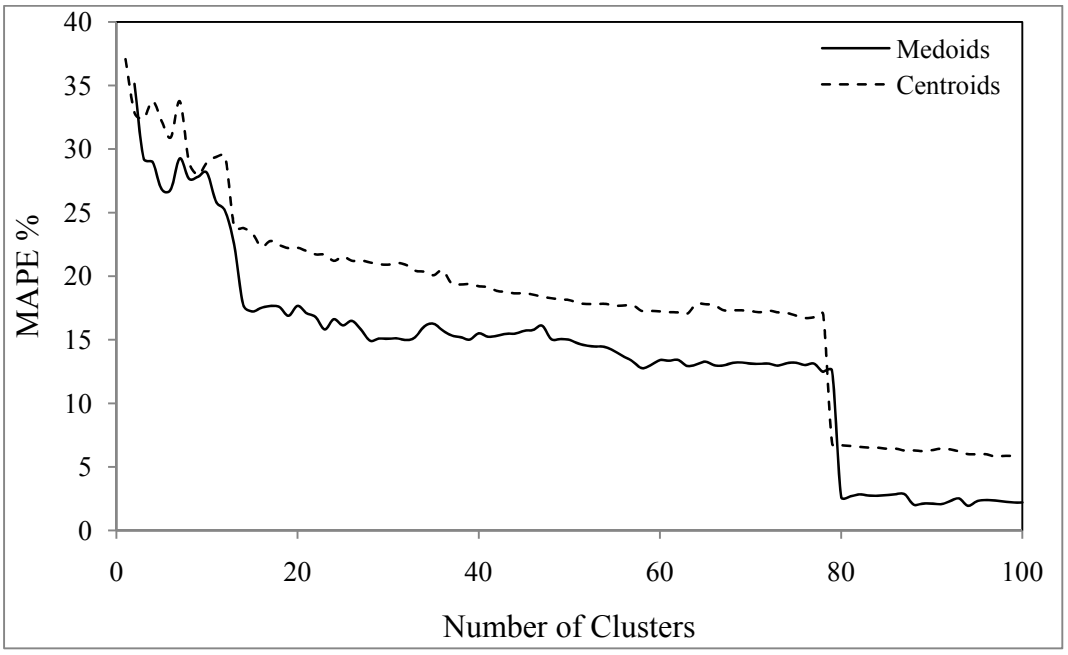

Figure 5-13 Comparison between the two groups of cluster representatives for the summer season of the 3-year data set

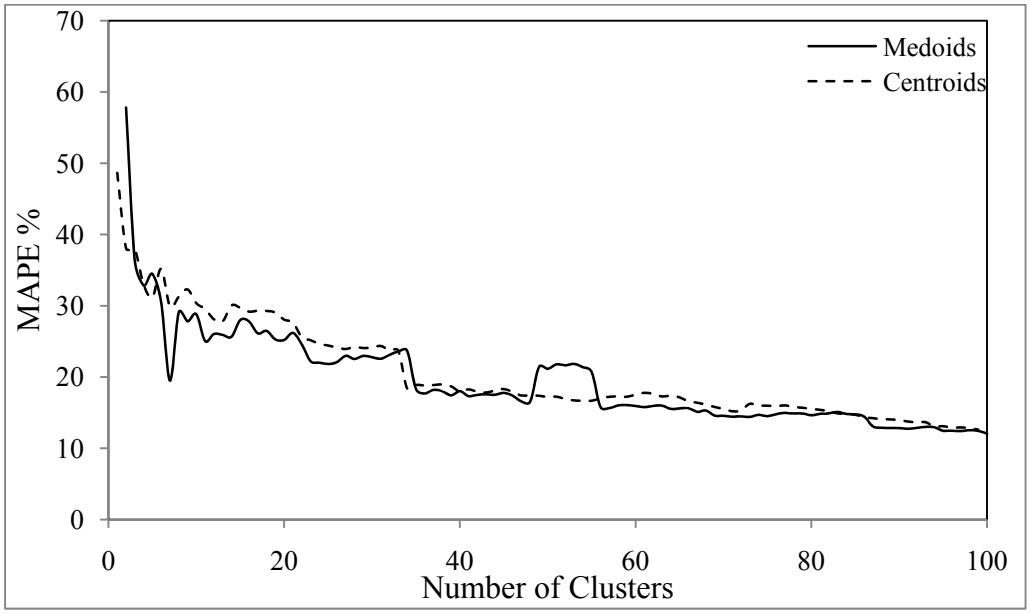

Figure 5-14 Comparison between the two groups of cluster representatives for the winter season of the 3-year data set

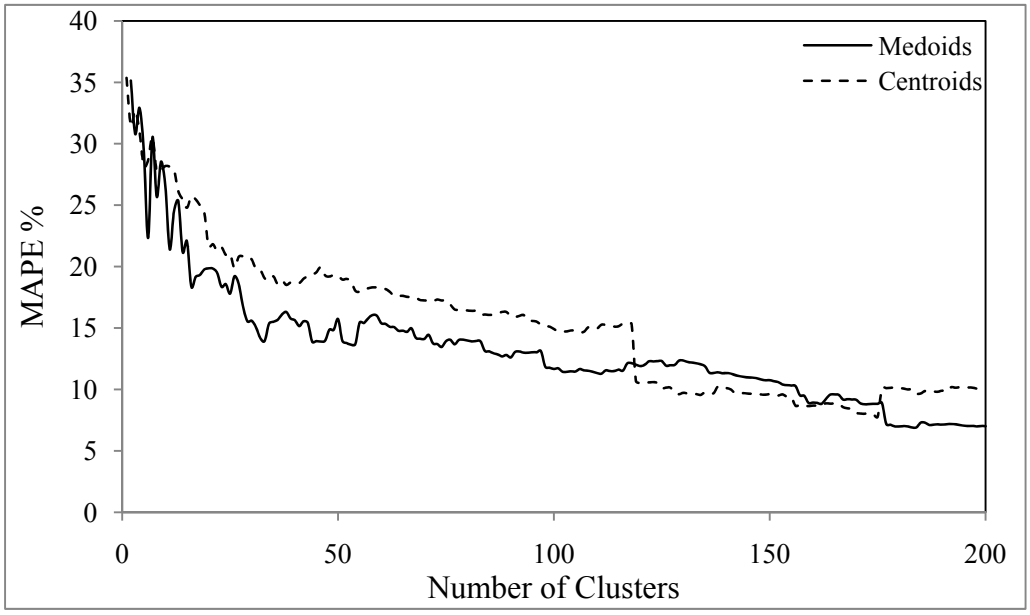

Figure 5-15 Comparison between the two groups of cluster representatives for the spring/fall season of the 3-year data set

The results for the summer and spring/fall seasons of the 3-year data set, displayed in Figure 5-13 and Figure 5-15, show that the medoids can better represent the segments of these data sets. This is because the values of the *MAPE* obtained when using the medoids are lower than those obtained when using the centroids for most of the cluster numbers. On the other hand, Figure 5-14 shows that the values of the *MAPE* obtained for medoids and centroids are close for the winter season of the 3-year data set. Thus, the medoids are chosen to represent the segments in the clusters of the data sets for all three seasons.

5.3 Application of the Clustering-Based Method on a Rural Distribution Feeder

This section is dedicated to the final stage of the proposed method, which is evaluating the performance of the electric network in the presence of the PV system. The main goal of this analysis is to highlight the benefits of the method and demonstrate its ability to overcome the drawbacks of other methods that can be used for the same purpose. The analysis presented in this section covers the following tasks: investigating the impacts of power fluctuations on the performance of the network, estimating the annual energy loss of the feeder, and indentifying a suitable size and site for the PV system.

5.3.1 The rural distribution feeder under study

A rural distribution feeder provided by a local utility in the province of Ontario, Canada, is used to demonstrate the application of the proposed method. The single line diagram of the feeder is shown in Figure 5-16.The feeder is connected to a 27.7-kV substation and the peak loading level of the feeder is 15.213 MW and 5.528 MVAR, which is assumed to be equally distributed among the three phases. The seasonal loading of the feeder is displayed in Figure 5-17, where it is assumed that the active and reactive powers of the loads follow the same pattern. The complete data for the feeder is displayed in Appendix C.

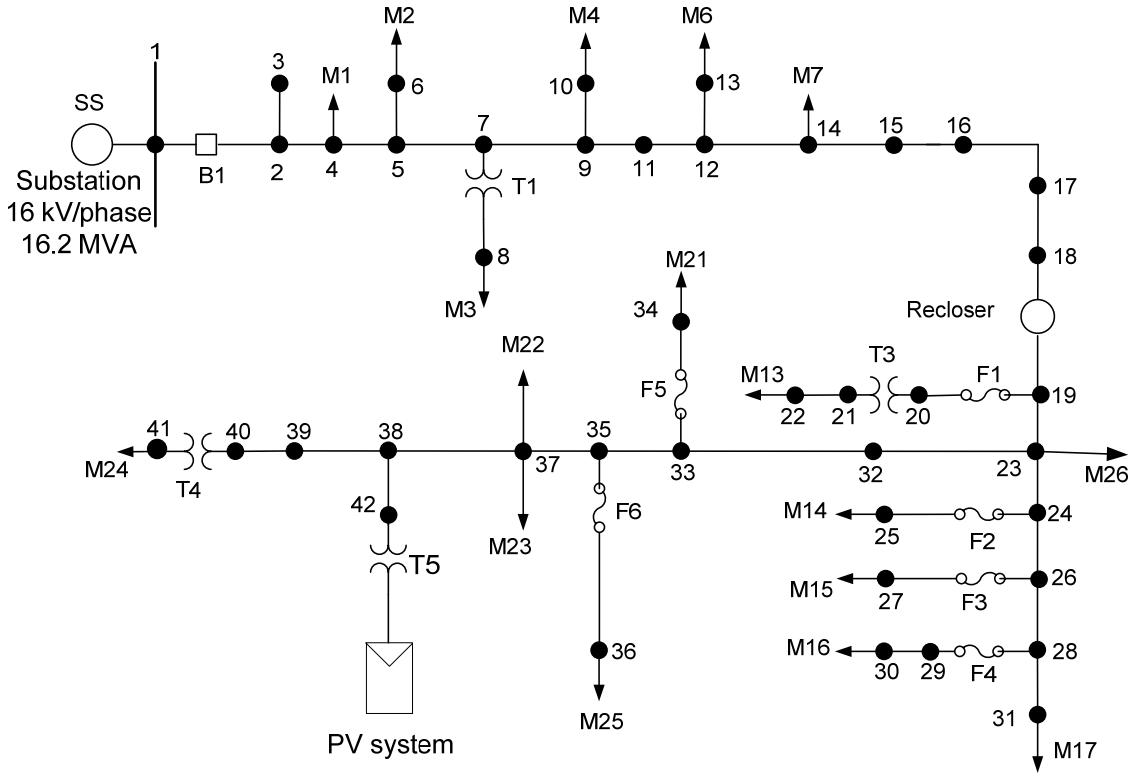

Figure 5-16 Single-line diagram of the distribution feeder under study

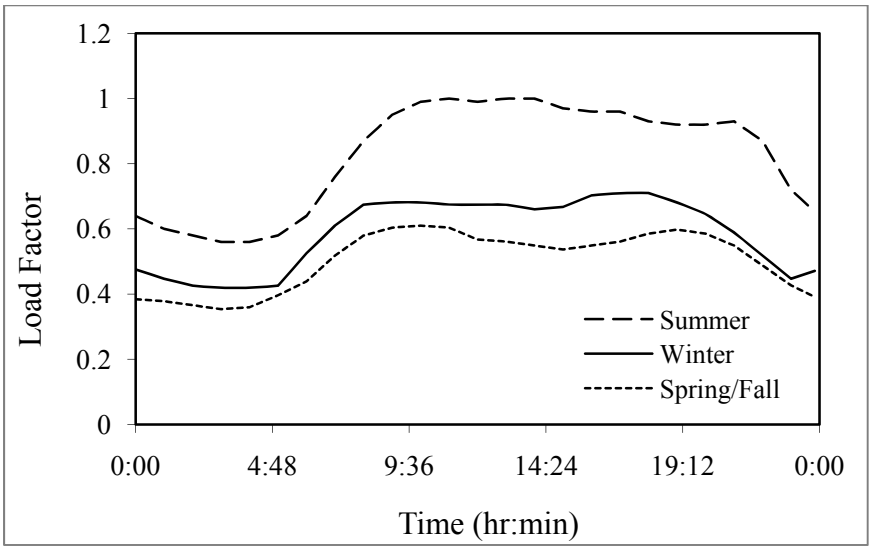

Figure 5-17 Seasonal Loading of the feeder under study

To investigate the impacts of the PV system using the proposed method, the following steps are implemented. First, the irradiance and temperature data of the three seasons of the 3-year data set and the seasonal loading of the feeder are used as inputs to the method. Accordingly, the AC output power time series of a 10-MW PV plant connected at Node 38 of the distribution feeder is estimated for each season using the Conversion stage illustrated in Section 4.3. In the second step, the estimated time series of AC power is divided into segments, taking the PV power of each day as one segment. Also, the data points of each segment are considered as the features of the segment that are required to be clustered. Finally, the Hybrid clustering algorithm is applied to group together the segments that have similar profiles, and hence, the cluster representatives are identified. The identified cluster representatives are considered as the reduced data set that can replace all the PV power segments, i.e., the full data set, in the power flow analysis.

5.3.2 Choice of number of clusters for power flow analysis

To perform the power flow analysis using the reduced data set, it is essential to choose a suitable number of representatives, i.e., a suitable number of clusters. In general, the suitable number of clusters can be identified using the cluster validity indices where the number of clusters corresponding to the best index value is chosen [121] (the best value being the highest or lowest value depending on the index used). This strategy is useful if the data set contains natural clusters; however, in cases where there are no natural clusters and the main target is to group similar data points, as in the case under study, choosing higher number of clusters can produce better results. This can be seen from the overall trend of the internal validity indices as well as the proposed index. On the other hand, choosing a lower number of clusters reduces the volume of simulations and analysis to be performed in the network analysis. Thus, in such cases, the choice of a suitable number of clusters is usually governed by the purpose of the study and the experience of the system operator. However, it is possible to provide a guideline for choosing the suitable number of clusters by noting that in most cases the power generated

from the large PV system is the main source of fluctuations in the feeder under study. This is because severe fluctuations usually occur during periods of high irradiance, i.e., around noon [73]. During these periods, the loading profiles of most of the distribution feeders are almost flat as shown in Figure 5-17. Thus, the deviation of the histograms of the PV power segments calculated from original clusters and the virtual clusters (cluster representatives repeated according to the number of segments in their respective clusters) be used as an indication of the difference between the histograms of the active powers flowing in different sections of the feeder obtained from the full and reduced data sets in the power flow analysis. The deviation of all other studied electric quantities is expected to be lower than that of the active power due to the fact that the fluctuations are mainly in the active power generated from the PV system. In other words, the *MAPE* calculated from the reduced and full PV power segments obtained for a specific number of clusters can be used as a guideline to estimate the values of *MAPE* calculated for the different electric quantities obtained from the power flow analysis using the same number of clusters.

To illustrate this concept, the full and reduced PV power segments contained in the three seasons of the 3-year data set are used in power flow analysis. The analysis is repeated for two reduced data sets, where the first set contains 11% of the full data set (30 clusters for the summer and winter seasons and 60 clusters for the spring/fall season), while the second set includes 22% of the full data set (60 clusters for the summer and winter seasons and 120 clusters for the spring/fall season). The power flow analysis is performed using the full and reduced data sets in MATLAB environment using a code obtained from [128]. Hence, the electric quantities are obtained for different data sets and compared using the *MAPE* index. The electric quantities that are considered in this sub-section are the histograms displaying the frequency of occurrence of the active and reactive powers, $P_{X\text{-}Y\_freq}$ and $Q_{X\text{-}Y\_freq}$, average powers, $P_{X\text{-}Y\_av}$ and $Q_{X\text{-}Y\_av}$, and maximum powers, $P_{X\text{-}Y\_max}$ and $Q_{X\text{-}Y\_max}$, flowing in branches 1-2, 19-23 and 37-38. Moreover, the histograms for the frequency of occurrence of the voltages, V_{Z_freq}, average voltages, V_{Z_av}, and maximum

voltages, V_{Z_max}, at nodes 4, 23, 38 and 41 are included in the results. Hence, the *MAPE* for different electric quantities are calculated and compared with the values of the *MAPE* obtained from the comparison of the histograms calculated from the PV power segments. Table 5-1 and Table 5-2 display the *MAPE* values obtained using the 11% data set in the power flow analysis and in calculating the histograms of the PV power segments. The values obtained for the 22% data set are displayed in Appendix D.

Table 5-1 *MAPE* calculated for the active and reactive powers of the 11% data set

Season	**Spring/Fall**	**Summer**	**Winter**
Number of clusters: 11% of the full data set			
MAPE in P_{PV_freq}	***15.40***	***15.08***	***22.79***
MAPE in $P_{1\text{-}2\_freq}$	12.78	22.40	15.11
MAPE in $P_{19\text{-}23\_freq}$	17.64	16.00	18.91
MAPE in $P_{37\text{-}38\_freq}$	16.99	16.60	14.57
MAPE in $Q_{1\text{-}2\_freq}$	8.59	3.68	9.87
MAPE in $Q_{19\text{-}23\_freq}$	10.14	8.76	9.13
MAPE in $Q_{37\text{-}38\_freq}$	16.80	10.33	10.40
MAPE in $P_{1\text{-}2\_av}$	0.09	0.51	0.10
MAPE in $P_{19\text{-}23\_av}$	-2.88	5.75	3.77
MAPE in $P_{37\text{-}38\_av}$	-0.46	-12.72	-1.02
MAPE in $Q_{1\text{-}2\_av}$	0.09	0.11	0.07
MAPE in $Q_{19\text{-}23\_av}$	0.15	0.10	0.18
MAPE in $Q_{37\text{-}38\_av}$	0.14	0.17	0.19
MAPE in $P_{1\text{-}2\_max}$	1.51	1.31	0
MAPE in $P_{19\text{-}23\_max}$	1.68	4.04	0
MAPE in $P_{37\text{-}38\_max}$	1.68	4.05	0
MAPE in $Q_{1\text{-}2\_max}$	3.27	0.69	4.58
MAPE in $Q_{19\text{-}23\_max}$	3.85	5.23	8.78
MAPE in $Q_{37\text{-}38\_max}$	3.86	5.60	8.97

Table 5-2 *MAPE* calculated for the voltages of the 11% data set

Season	Spring/Fall	Summer	Winter
Number of clusters: 11% of the full data set			
MAPE in P_{PV_freq}	***15.40***	***15.08***	***22.79***
MAPE in V_{4_freq}	0	0	0
MAPE in V_{23_freq}	0	0.67	0
MAPE in V_{38_freq}	0	0.88	0.26
MAPE in V_{41_freq}	0.11	15.06	0.02
MAPE in V_{4_av}	0	0.01	0
MAPE in V_{23_av}	0	0.02	0
MAPE in V_{38_av}	0.01	0.04	0
MAPE in V_{41_av}	0.01	0.04	0
MAPE in V_{4_max}	0	0	0
MAPE in V_{23_max}	0.02	0	0.02
MAPE in V_{38_max}	0.07	0.16	0.16
MAPE in V_{41_max}	0.08	0	0.16

The results show that most of the values of the *MAPE* obtained from the power flow analysis are lower than those obtained from the histograms of the PV power segments. This is especially true for the average and maximum values of all monitored quantities. In few cases, the values of *MAPE* calculated from the histograms displaying the frequency of occurrence of powers flowing in some branches are higher than those obtained from the PV power segments. These values are highlighted by shading their cells in the aforementioned tables. The main reason behind this difference is the fact that the loading of the feeder is not constant over the whole period of operation of the PV system. Accordingly, this can affect the values of the *MAPE* calculated for the histograms of the powers flowing in the branches. For example, if the actual power of the PV system (obtained from the full data set) is 10 MW, while that of the cluster representative (obtained from the reduced data set) is 8 MW, for two different loading conditions of the

feeder, say 0.5 loading factor and 0.9 loading factor, the values of the *MAPE* calculated from the histograms of the frequency of occurrence of powers flowing in some branches might be different. However, as the results displayed in the table indicate, these differences are not high and, thus, the *MAPE* calculated for the PV power segments can be used as a guideline for the maximum *MAPE* obtained using the reduced data set in the power flow analysis.

5.3.3 Studying the impacts of power fluctuations

The power fluctuations generated from PV systems can negatively impact the performance of electric networks, as they can lead to fluctuations in node voltages, power swings in feeders, and can reduce the lifetime of voltage regulators. Thus, it is important to estimate the impacts of these fluctuations prior to installing the PV system.

As mentioned in Section 3.6, there are three main methods that can be used for investigating the impacts of PV systems on the performance of the network. However, the methods based on the probabilistic approach cannot be used to investigate the impacts of power fluctuations because they do not consider the temporal information in the analysis. The deterministic approach based on using the capacity factors of the PV system at a certain location also fails in providing any information about the power fluctuations. This is because of the assumption that the PV system generates constant power corresponding to the capacity factor. On the other hand, the deterministic approach based on using specific scenarios in the analysis can provide some information about the impacts of worst case fluctuations. However, this method cannot provide any information about other possible scenarios or the probability of occurrence of these worst case scenarios. Moreover, using only these scenarios in predicting the impacts of the power fluctuations on the performance of the network can lead to an overestimation of these impacts. The only methods that can provide detailed information about the impacts of power fluctuations are those based on chronological simulations, such as the clustering-based method proposed in this research.

In this subsection, three techniques that implement the available data in the chronological simulations are applied and compared. The first technique is the one based on using the cluster representatives obtained from applying the proposed method on the 10-min. data of the winter season of the 3-year data set [129]. In the second technique, the data set is obtained by reducing the resolution of the 10-min. data to 1 hour. This is achieved by recording the power every 1 hour rather than 10 minutes, which is a commonly-used technique. The third technique reduces the resolution of the data to 1 hour by averaging the data points of 10-min. data set.

Figure 5-18 shows the 20 clusters formed using the clustering-based method. This number constitutes approximately 7.4% of the full data set (270 segments). Figure 5-19 displays the medoid, the cluster representative, of cluster number 12 (Row 2, Column 4). As the definition implies, a medoid is also a power segment that belongs to the full data set. Thus, the second and third techniques used to reduce the data can be applied on this medoid. The resultant power segments are displayed in Figure 5-20 and Figure 5-21, respectively. It is clear from both figures that the 1-hr power segment is smoother than the 10-min. segment. Accordingly, these segments do not provide any information about the sub-hourly fluctuations, and thus, cannot provide accurate information on the performance of the network in the presence of these fluctuations.

Figure 5-22 to Figure 5-25 show samples of the results obtained from the power flow analysis using the PV power segments displayed in Figure 5-19 to Figure 5-21. Based on the results displayed in these figures, the following can be stated:

1- Observing the solid line in Figure 5-22 (10-min. cluster representative) that displays the active power flowing through the section connecting nodes 19 and 23, it can be seen that the active power flows from Node 19 to Node 23 until 8:10 am when the generation of the PV system starts to increase above 2.56 MW. At this instant, and until noon, the direction of power flow is reversed where the PV system feeds most of the loads in the feeder. During the period, 12:00 pm to 3:00 pm, the power flowing in this section fluctuates due to the passage of clouds over

the PV array. After 3:00 pm, and until almost sunset, the power flows back in its original direction (from Node 19 to 23). Now, observing the results obtained from the dotted lines (1-hr and averaged 1-hr), it is clear that much of this detailed information is lost. Moreover, the severe fluctuations in the power flowing in this section of the feeder are not present in the results.

2- Figure 5-23 shows that the reactive power flowing in this section of the feeder is not highly affected by the fluctuations of the PV power as compared to the active power flowing in the same section. The profile of the reactive power in this case is similar to the seasonal loading profile for the winter season. The only exception is the period from 12:00 pm to 3:00 pm, where the PV power fluctuates due to the passage of clouds.

3- Figure 5-24 shows the power loss in the same section of the feeder. The power loss increases as the injected PV power increases. The maximum power loss occurs at around 10:30 am corresponding to the maximum generation the PV system.

4- Finally, the voltage profile at Node 41, the furthest node in the feeder, is shown in Figure 5-25. The fluctuations in the voltage of this node are mainly due to the fluctuations in the PV power. Moreover, the voltage at this node can reach a value less than 0.95 p.u., which is not acceptable for some networks. On the other hand, the results obtained from the PV power segments with 1-hr resolution indicate that the voltage at this node is smooth, which is not true.

Based on the above discussions, it can be concluded that the 1-hr resolution data cannot provide accurate information about the impacts of PV power fluctuations. On the other hand, using the cluster representatives can achieve this task. Moreover, whenever a cluster representative indicates the possibility of occurrence of an undesirable behavior of the feeder, more detailed analysis can be performed by considering all the segments present in this cluster in the analysis.

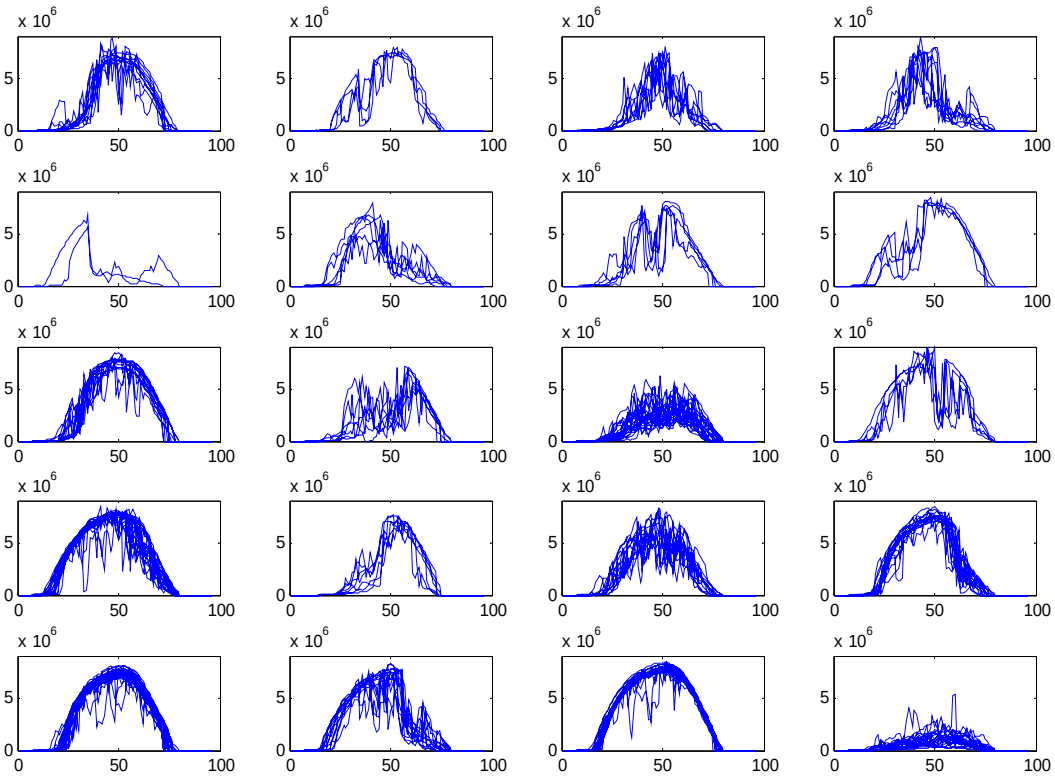

Figure 5-18 Clusters obtained for the winter season of the 3-year data set

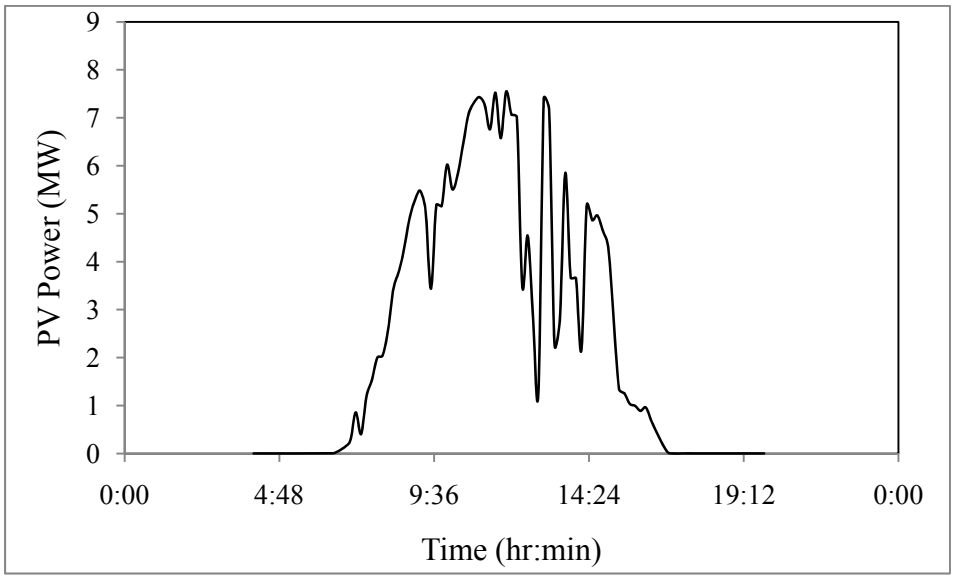

Figure 5-19 Cluster representative for cluster number 12

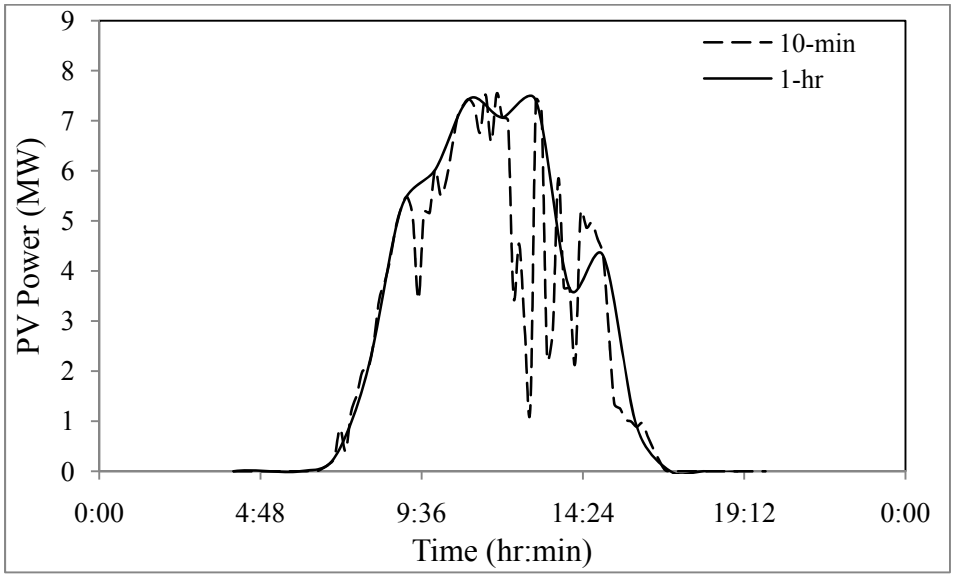

Figure 5-20 Comparison between fluctuations in the 10-min. and 1-hr segments

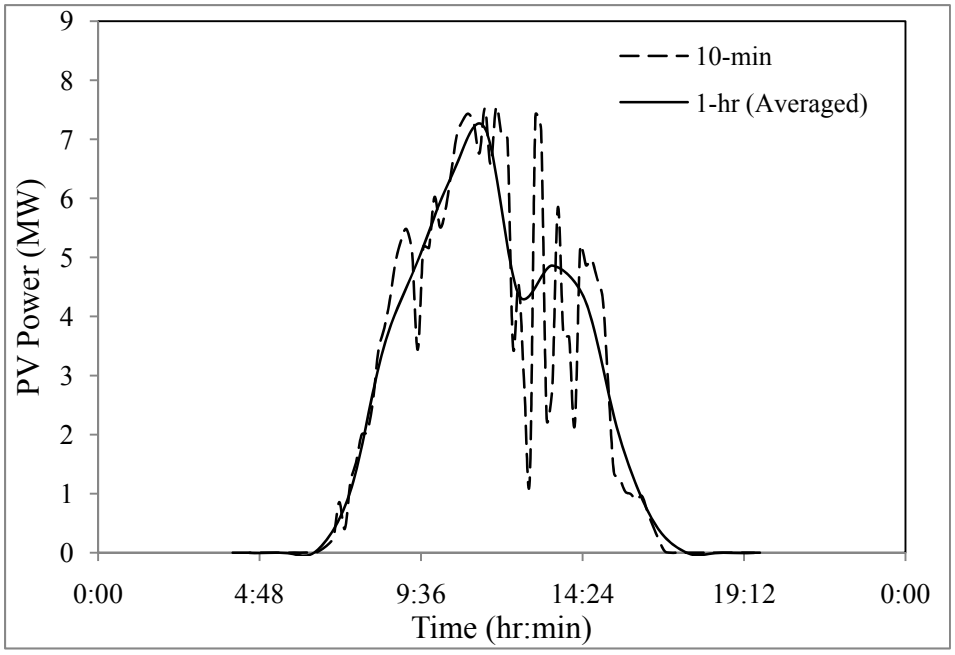

Figure 5-21 Comparison between fluctuations in the 10-min. and averaged 1-hr segments

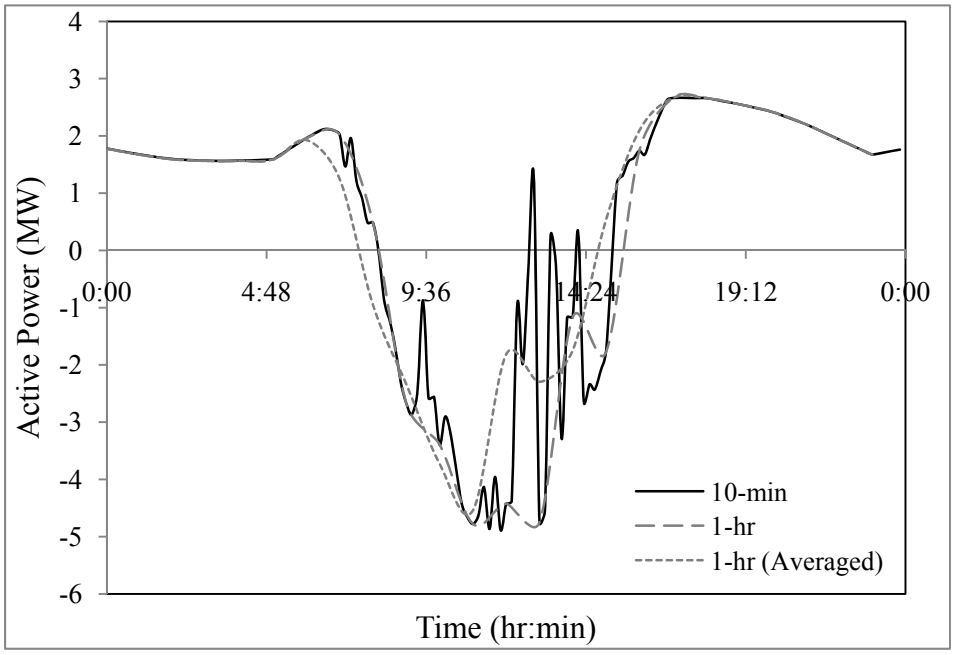

Figure 5-22 Active Power flowing in the section connecting nodes 19 and 23

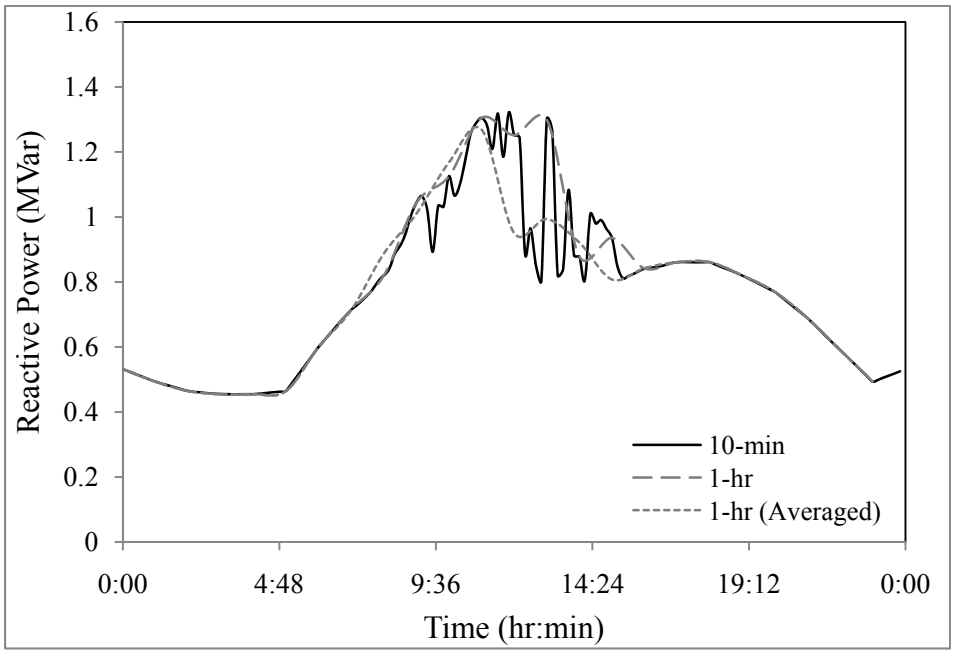

Figure 5-23 Reactive power flowing in the section connecting nodes 19 and 23

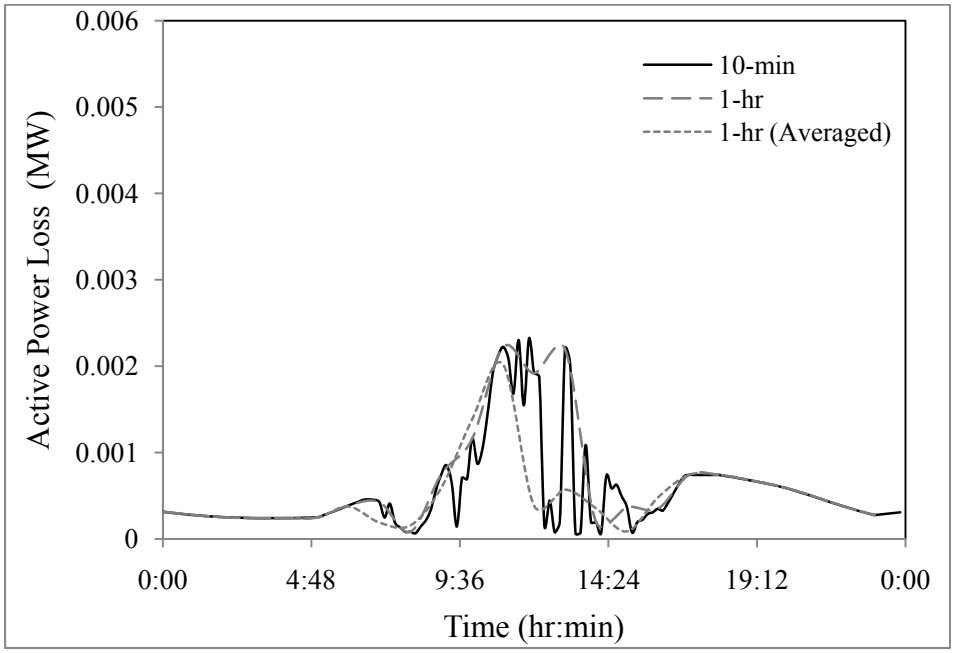

Figure 5-24 Power loss in the section connecting nodes 19 and 23

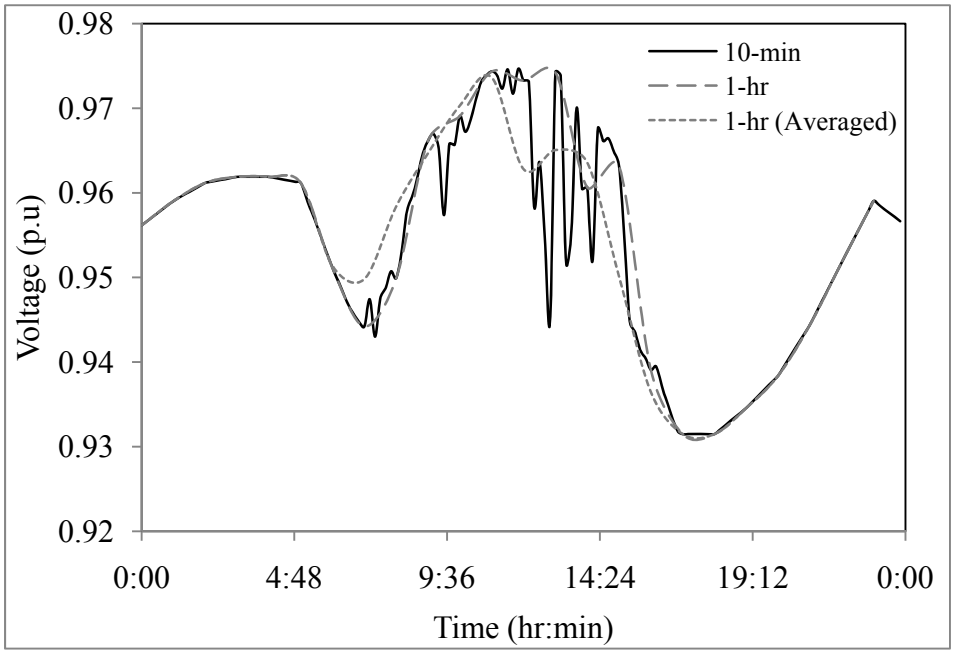

Figure 5-25 Voltage profile of Node 41

5.3.4 Statistical analysis

One of the advantages of the proposed method is its ability to provide information about how often and when the specific clusters of interest might occur. For example, Figure 5-26 displays the probability of occurrence for the 20 clusters obtained for the winter season of the 3-year data set, displayed in Figure 5-18. This figure shows that cluster number 17 (Row 5, Column 1) has the highest probability of 16.3%. This cluster contains segments that represent high generation of power from the PV system. On the other hand, the probability of occurrence of cluster 14 (Row 4, Column 2), which includes segments corresponding to days with cloudy mornings and sunny afternoons, is 2.59%. Moreover, the majority of segments present in this cluster occur in January.

It is worth to mention that one of the important aspects that affect the accuracy of the results obtained from the statistical analysis is the amount of historical data used in the study. In fact, using long historical data (e.g., data of the past 10-20 years) in the analysis can help generalizing the obtained conclusions. These conclusions can help the system operator expect the performance of the network during specific periods of time, and thus, corrective measures can be applied if necessary.

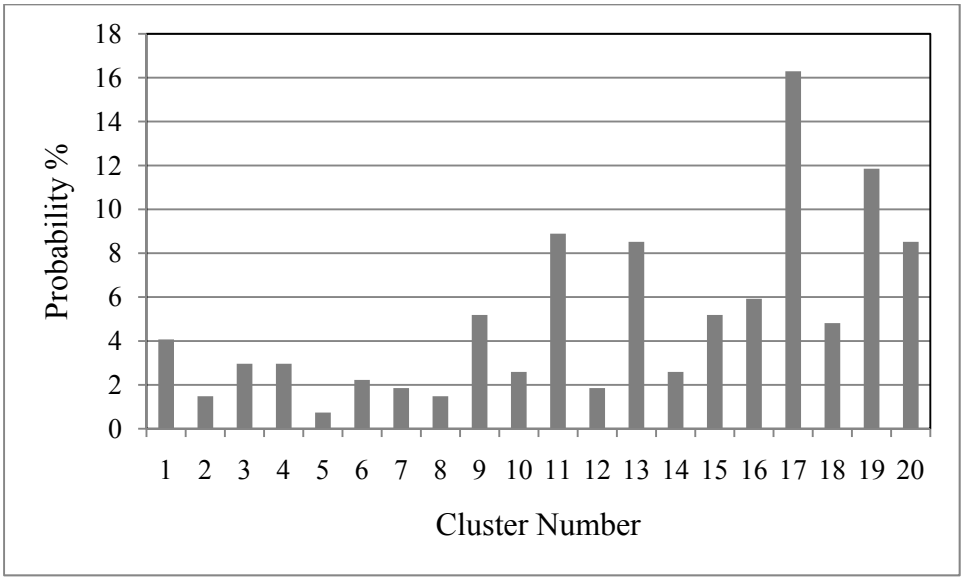

Figure 5-26 Probability of occurrence of the 20 clusters of Figure 5-18

5.3.5 Estimating the annual energy loss of the feeder

The change in the annual energy loss due to the installation of the PV system is one of the important quantities that should be estimated prior to installing the PV system. Moreover, the annual energy loss is used as one of the criteria for choosing the suitable site and size of a PV system. Thus, it is important to accurately estimate this important quantity.

In this section, the estimated annual energy losses of the feeder obtained by applying some of the commonly used methods are compared with that obtained from the clustering-based method. The five methods that are applied and compared are: 1) the full 3-year data set of the power generated from the PV system with 10-minute time resolution, which is considered as the reference case, 2) the average PV power obtained using the capacity factors for each season, calculated from Equation (5-5), 3) the 3-year data of the PV power with 10-min. time resolution reduced to obtain a 1-hr. time resolution, 4) the 3-year data of the PV power with 10-minute time resolution averaged to obtain a 1-hr. time resolution, and 5) the cluster representatives of the 3-year data set with 10-min. time resolution (7.4% of the full data set).

$$CF = \frac{\text{Energy generated by the PVsystem in time } t}{\text{Rated output power of the PVsystem } \times t} \tag{5-5}$$

Table 5-3 shows the seasonal capacity factors calculated from the estimated energy production of the 10-MW PV system obtained using the 10-min. data of irradiance and temperature over the past 9 years. On the other hand, the average annual energy loss obtained from each of the five methods is displayed in Table 4 displays, and the error for each method is calculated from Equation (5-6).

Table 5-3 Seasonal capacity factors

Season	Spring/fall	Summer	Winter
Capacity Factor	0.1473	0.1591	0.1423

$$\text{Error (\%)} = \frac{E_{loss} - E^{*}_{loss}}{E_{loss}} \times 100 \qquad (5\text{-}6)$$

where E_{loss} is the annual energy loss estimated from the full data set and E^{*}_{loss} is the annual energy loss estimated from any of the other four methods.

Table 5-4 Average annual energy loss calculated from different methods

Method	Spring/Fall Energy loss (MWh)	Summer Energy loss (MWh)	Winter Energy loss (MWh)	Annual Energy loss (MWh)	Error (%)
Full data, 10-min. resolution	393.2	483.8	266.2	1143.2	0
Capacity Factor	417.2	609.4	292.5	1319.1	- 15.4
Data with 1-hr. resolution	385.5	472.2	261.7	1119.4	2.1
Data with 1-hr. resolution (Averaged)	274.3	234.5	170.6	679.4	40.6
Cluster Representatives	390.0	477.1	264.4	1131.6	1.0

The results displayed in Table 5-4 show that the use of capacity factors cannot provide accurate estimation for the annual energy loss. In the case analyzed in this sub-section, the capacity factor overestimates the annual energy loss of the feeder. On the other hand, the averaged 1-hr. resolution data highly underestimates the annual energy loss. This underestimation is reduced significantly when the 1-hr resolution data is used. Using the cluster representatives to estimate the energy loss of the feeder produces the most accurate results, as it generates the least error (1%). Moreover, the amount of data used in

the case of cluster representatives (7.4% of the full data set) is less than half the data used in the case of 1-hr resolution data (16.67 % of the full data set).

As a conclusion, using the cluster representatives only to estimate the annual energy loss of the feeder is accurate and reliable.

5.3.6 Sizing and siting the PV system

Choosing a suitable size and location for installing a large PV system is one of the important decisions that has be considered prior to installing the PV system. This choice is usually governed by a number of factors, such as reducing the annual energy losses of the feeder and preserving the voltage limits at different nodes. Another important factor that governs this choice is the land availability. This is due to the fact that large PV systems require significant areas for installing the PV arrays which makes the choice of the location limited to some specific locations. In most cases, the choice of a suitable location and size for the PV system is identified by performing the power flow analysis using the available historical time series data for different sizes and sites of the PV system. However, this analysis is time consuming and computationally expensive, especially if the historical time series data is long and has a high time resolution. For example, if it is required to choose between 4 locations and 5 sizes of the PV system using the full data set of the past three years, then, 3,153,600 power flow case studies should be simulated. To overcome this problem, the cluster representatives obtained from the clustering-based method can be used to estimate the energy loss of the feeder, and thus reduce significantly the required analysis while preserving the accuracy. To check the voltage limits at different nodes, the worst case scenario can be used. One of the important scenarios that must be used in this case is the one corresponding to maximum generation of the PV system under minimum loading conditions.

For the analysis presented in this sub-section, it is assumed that three different locations are available for installing the PV system, and thus, it is required to choose the location and size that can lead to a maximum reduction in the annual energy loss of the feeder.

The first available location is Node 38, near the end of the feeder and close to Node 41 to which a large load is connected. The second location is Node 31, which is located at the other far end of the feeder where a number of medium loads are connected. The third location is Node 23, in the middle of the feeder. The annual energy loss of the feeder was calculated using the cluster representatives of the 3-year data set for each of the three locations while the size of the PV system is varied between 0 and 12 MW.

Figure 5-27 shows the annual energy loss for different sizes of the PV system and different connection nodes. The results indicate that a 10-MW PV system connected at Node 23 results in minimum annual energy loss in the feeder. This will result in a 20% reduction in the annual energy loss of the feeder as compared to the case where no PV system is installed. The results also indicate that any further increase in the size of the PV system will increase the energy loss of the feeder. For this size and location of the system, the voltages at different nodes obtained for the case of maximum generation of the PV system and minimum loading conditions are within the permissible limits (0.9 – 1.1 p.u.). As a result, the 10-MW PV system can be connected at this node.

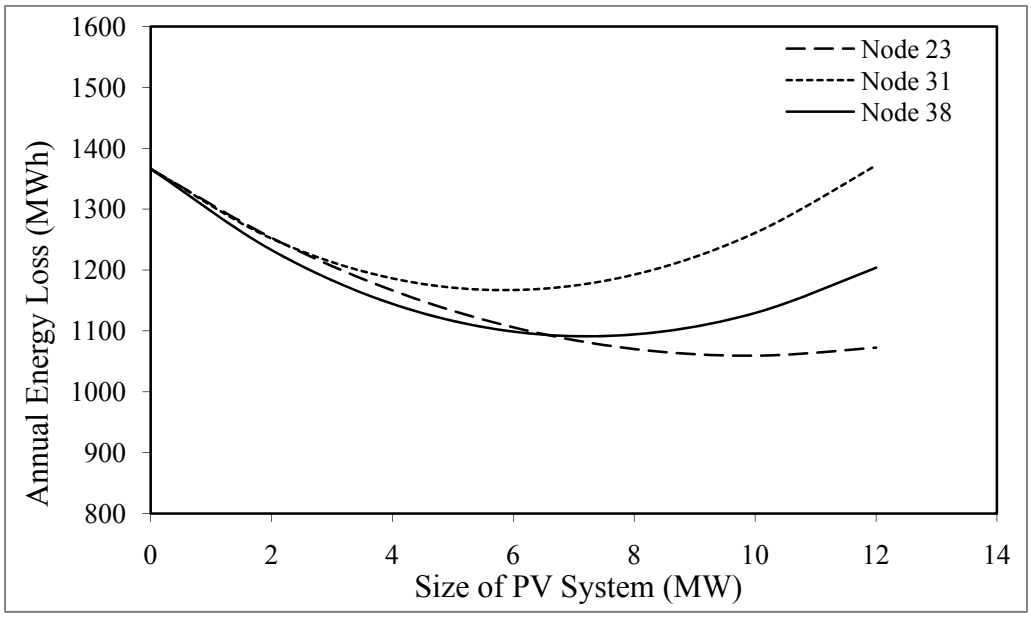

Figure 5-27 Sizing and siting of a large PV system

Table 5-5 Average annual energy loss calculated for different sizes and locations

Size (MW)	Spring/Fall Energy loss (MWhr)	Summer Energy loss (MWhr)	Winter Energy loss (MWhr)	Annual Energy loss (MWhr)
Node 23				
0	435.5913	627.3651	303.0686	1366.025
2	392.5527	582.6996	277.6878	1252.9401
4	361.9724	545.509	259.0464	1166.5278
6	343.4699	515.4919	246.9124	1105.8742
8	336.7612	492.4138	241.1163	1070.2913
10	341.649	476.0979	241.544	1059.2909
12	358.0168	466.4192	248.133	1072.569
Node 31				
0	435.5913	627.3651	303.0686	1366.025
2	393.0498	580.85	277.7097	1251.6095
4	373.4773	547.9512	264.7245	1186.153
6	375.7451	527.8683	263.446	1167.0594
8	399.0259	520.0036	273.399	1192.4285
10	442.7455	523.9247	294.2673	1260.9375
12	506.5531	539.3405	325.8749	1371.7685
Node 38				
0	435.5913	627.3651	303.0686	1366.025
2	385.6112	573.6426	273.3814	1232.6352
4	356.6024	532.6617	255.0507	1144.3148
6	347.575	503.6496	247.4793	1098.7039
8	357.8182	486.0348	250.2501	1094.1031
10	386.8584	479.4083	263.0967	1129.3634
12	434.4343	483.5011	285.8881	1203.8235

5.4 Merits of the Clustering-Based Method

The previous sub-sections illustrated how the clustering-based method can be applied to analyze the impacts of large PV systems on the performance of electric networks. To finalize this part, it is important to highlight the merits of applying this method. These can be summarized in the following points:

1- Usually, when studying the system performance, dealing with overwhelming number of simulated scenarios makes it difficult to benefit from the obtained results. Thus, the proposed method aims to use the available data in an intelligent and efficient manner in order to help the system operator extract useful and comprehensive information from the available data.

2- The performance of the electric network due to the installation of a large PV system can be evaluated with high accuracy and with a highly reduced number of simulations using the proposed method. To achieve this task, the representative segments are used in the power flow analysis and the monitored electric quantities are recorded for each representative. The representatives that have the potential to violate the acceptable limits of the monitored electric quantities are then identified along with their corresponding clusters. Accordingly, the segments in the identified clusters can be used in detailed analysis in order to accurately assess the performance of the feeder.

3- In cases when there are more than one loading profile for the electric network, such as profiles for weekdays, weekends, and each month, using the cluster representatives in the power flow simulations facilitates studying the impacts of possible combinations of different generating profiles of the PV system with different loading profiles of the electric network. For example, if 2 loading profiles are available for each of the 12 months of the year, this means that 24 load profiles are available. Thus, if 50 representative PV power segments are chosen to represent the power generated from the PV system, a total of 24×50

profile combinations should be simulated. On the other hand, using the complete data set cannot be used for this type of analysis, as it is almost impossible to simulate all the possible cases.

4- The proposed method serves in grouping together segments having similar profiles to facilitate the discovery of important profiles that can impact the performance of the electric network. Examples of these patterns are those with high power output as in the case of sunny days, those with low power output as in the case of overcast days and those with severe power fluctuations.

5- Clustering of PV power segments facilitates performing statistical analysis to estimate the probability of occurrence of segments that might impact the performance of the electric network. This analysis requires the use of large data sets, such as data of 10 years, in order to obtain accurate statistics. Moreover, this analysis can help in predicting the seasons or months in which the patterns that might impact the performance of the electric network are most likely to occur. Thus, suitable operational plans can be adopted and prepared.

6- In the cases where power flow simulations have to be repeated a large number of times, reduction of data becomes essential. An example of such cases is when it is required to choose a suitable location and size of the PV system that result in reducing the annual energy loss in the feeder. In this case, the number of power flow case studies that should be simulated can be in the order of millions. The load of performing this overwhelming number of simulations, which is a tedious task, can be highly reduced if the cluster representatives obtained from the proposed method are used. For example, if the cluster representatives constitute 10% of the full data set, the number of simulations, and thus, the simulations time, will be reduced by 90%.

5.5 Summary and Conclusions

This chapter presented some of the validity indices that can be used for evaluating different clustering algorithms. The investigation of the results obtained from these indices showed their tendency to prefer the clustering algorithm that generates clusters with single segments. However, this preference does not comply with the main purpose of the clustering process. Thus, a new index that is more suitable for utility analysis was proposed and examined. The proposed index can successfully identify the clustering algorithm that generates more homogenous clusters and can provide information about using a certain group of cluster representatives to represent the power segments contained in the clusters. Moreover, the proposed index provides guidelines on choosing the suitable number of clusters. This is especially true in the case when the severe power fluctuations of the PV system occur when the loading profile of the feeder is flat. In this case, the PV system is the main source of fluctuations in the feeder, and thus, the *MAPE* calculated from the PV power segments can provide estimation for the highest values of the *MAPE* calculated from the power flow analysis.

The use of the proposed index indicated that the Hybrid clustering algorithm is more favorable as compared to the Average Linkage hierarchical algorithm. Moreover, the index showed that there is no obvious difference between using the data points of the power segments as the features of these segments and using the transformation coefficients obtained from PCA. However, the index suggested that the use of medoids as cluster representatives is more favorable than using the centroids.

In the second part of this chapter, the clustering-based method was used to investigate the impacts of a large PV system on the performance of a rural distribution feeder. One of the investigated impacts was the effect of power fluctuations generated from the PV system on the different electric quantities of the feeder. The discussions presented in this part showed that using the probabilistic or deterministic approaches cannot provide information about the power fluctuations. In addition, the results presented in this part

indicated that the 1-hr. resolution data cannot provide accurate information about the sub-hourly fluctuations. On the other hand, the cluster representatives can provide detailed information about these fluctuations. Moreover, in the cases when a cluster representative leads to undesirable performance of the feeder, the segments of the corresponding segments can be used in a more thorough analysis. This approach can provide accurate results while utilizing a reduced number of simulations.

The analysis presented in this chapter also showed that the capacity factor and the averaged 1-hr. data fail to estimate the annual energy loss of the feeder accurately. On the other hand, the cluster representatives estimate the energy loss with a very high accuracy and with a highly reduced number of clusters. This conclusion is very useful, especially for cases when the size and location of the PV system is required to be identified.

The results and discussions presented in this chapter illustrated the usefulness of applying the clustering-based method to identify the impacts of large PV systems on the performance of the electric network, especially due to power fluctuations. In the following chapter, some of the methods that can be used to reduce the impacts of these power fluctuations are investigated and analyzed.

Chapter 6

Investigation of Methods for Reducing Power Fluctuations of Large PV Systems

6.1 General

Chapter 3 highlighted some of the impacts of power fluctuations generated from PV systems on the operation of the electric network. To name a few, power fluctuations from large PV systems can affect the scheduling of utility generating units [81][82], can lead to unstable operation of the electric network prior to the fault conditions [84], high power swings in the feeders [83], and unacceptable voltage fluctuations at certain nodes in the electric network [83][101]. Moreover, the random fluctuations of the power output generated from PV systems do not allow them to be considered in the scheduling process of electricity generation. Chapter 4 and Chapter 5 introduced the layout and details of a new method that can help in quantifying the impacts of power fluctuations, and evaluating the performance of the electric network. In the present chapter, the main objective is to investigate some methods that can be used to reduce the fluctuations in the power generated from a large customer-owned PV system [130]. This chapter focuses on three methods: 1) the use of battery storage systems, 2) the use of dump loads, and 3) curtailment of the generated power by operating the power conditioning unit (PCU) of the PV system below the maximum power point (MPP). The emphasis in the analysis

presented in this chapter is on investigating the impacts of implementing these methods on the economical benefits gained by the PV system owner. To estimate the maximum profit gained by the system owner, an LP optimization problem is formulated and solved. Moreover, the effect of varying different parameters of the problem is investigated through sensitivity analysis.

6.2 Energy Storage Systems

The use of energy storage devices with PV systems is currently receiving a lot of attention due to the fact that the power generated from these systems is intermittent. The installation of storage devices can enhance the performance of PV systems in one or more of the following ways:

1- *Bridging power fluctuations*: Smoothing the output power of the PV system can be achieved by storing energy during high generation periods and delivering energy when the output of the system falls due to passage of clouds. In this case, the storage device should have a response time that matches the speed of the fluctuations and should be able to deliver energy over the range of few minutes to few hours.

2- *Shifting the time of peak generation*: This is an important task whenever the peak loading of the feeder is not around noon, which corresponds to the peak generation period of PV systems. In this case, the storage device stores energy during the peak generation period and injects this energy during the peak loading period. Thus, the storage device should be able to operate for up to a few hours.

3- *Feeding critical loads in case of outages*: In this case, the storage device along with the PV system can be operated during electricity outages to feed critical loads in the feeder. This has to be done after coordinating with the local utility so as not to violate the islanding protection regulations.

There are a number of technologies that can be used for storing and delivering electric energy. These technologies vary in performance and suitability for use with PV systems. For example, Pumped Hydro Storage systems require significant land area with suitable topography for placing the reservoirs at different altitudes. Thus, they might not be suitable for use with large PV systems installed on large flat land areas. Also, Superconducting Magnetic Energy Storage systems have very low energy densities, and thus, are not suitable for bridging power fluctuations in the order of few minutes up to few hours, as in the case of PV systems. These systems are rather used for pulsed-power and system-stability applications [131]. Ultracapacitors also have low energy densities [132], and thus, cannot be used alone to with PV power systems. However, they can be used with batteries to achieve high power and energy densities and extend the lifetime of batteries [133]. Conventional low-speed flywheels have high power and low energy densities while expensive advanced high-speed flywheels have higher power and energy densities [134]. Flywheels can be used with PV systems but are more suitable for wind turbines due to electromechanical nature of these devices.

Batteries store energy through an electrochemical process, and thus, have quick response in both charge and discharge operations [135]. Moreover, batteries have medium power and energy densities, and thus, they are one of the most suitable technologies that can be used to bridge the power fluctuations generated from PV systems. During the past few years, a number of studies focused on analyzing the different aspects of using battery storage systems with PV systems, especially stand-alone PV systems [136]-[138]. Other studies focused on using BS systems with small grid-connected PV systems to supply power to critical loads during power outages of the grid [139][140]. Recently, some studies investigated the use of BS systems to shift the peak generation of the PV system to the periods of peak loading conditions [141]-[143].

There are a number of battery technologies that can be used for storing electric energy, with Lead-Acid (LA) batteries being the benchmark for storage batteries. This is mainly

due to their low cost, maturity of technology and availability in the market [144]. However, these batteries have short lifetime and require frequent maintenance [131]. Large LA plants can operate from minutes to several hours; for example, a 10-MW LA plant in California can deliver its rated power for 4 hours [131].

Lithium-Ion (Li-Ion) batteries have energy densities that are three times greater than those of LA batteries [145]. They have long cycle life and very high efficiency close to 95% [146]. Their main disadvantages are the safety issues associated with their operation and their high cost, restricting their use to small-scale applications [147][148] such as laptop computers and hybrid vehicles.

The new technology Sodium-Sulfur (NaS) batteries have high power and energy densities and are suitable for large-scale applications. These batteries do not suffer from self-discharge and their efficiency is around 90% [132]. However, they operate at high temperatures, 300 – 400 ^{o}C, and thus, require additional heating [144]. In Japan, a 6-MW/48-MWh NaS battery plant is used for peak load shaving [148]. Another 34-MW/245-MWh NaS battery storage unit was installed in 2008 for the 51-MW Futamata wind farm in Japan [146]. NaS batteries are one of the most suitable technologies for energy storage with large PV systems.

Flow batteries are also suitable for use with large PV systems. They are charged and discharged by a reversible chemical reaction between the two liquid electrolytes of the battery. These electrolytes are stored separately in storage tanks and are pumped into the reactor during operation. This provides the battery with flexibility in operation as the power and energy can be specified separately. The power is determined by the power electronic devices that operate the battery while the energy is specified by the size of the storage tanks [148]. Vanadium Redox Batteries (VRB) are flow batteries with efficiencies in the order of 75% and their cycle lifetime can reach 14,000 cycles [131]. In Japan, a 4-MW/6-MWh VRB-based storage system is currently operating with a large wind farm [147].

In general, one of the main issues with using storage devices for large PV systems is the increase in the cost of already expensive PV systems. This cost can reach up to $1.5 million/MW [149]. Thus, performing an economical analysis is a crucial step before deciding to use any storage device. Moreover, the choice of the suitable storage technology to be integrated with large PV systems is of great importance. Appendix E provides a comparison between different storage technologies.

6.3 Smoothing Power Fluctuations Using a Battery Storage System

The previous section and the information presented in Appendix E suggest that battery storage (BS) systems are strong candidates for smoothing the output power of PV systems. In this section, it is assumed that the utility requests the system owner to limit the power fluctuations of a 10-MW PV system within a certain range during a specific period of time (e.g., ±250kW/10min., between 9:00 am and 3:00 pm). Accordingly, the system owner considers installing a BS system, as shown in Figure 6-1, to comply with the power fluctuations limit. In this case, it is important for the system owner to choose the proper power and energy ratings for the battery and estimate the profit gained from selling the energy of the PV/BS system over the lifetime of the BS system. This information can be obtained by solving the Linear Programming optimization problem defined in the following sub-section. The input to the optimization problem is the historical time series of the power generated from the PV system alone, P_{PV}, calculated over a number of years. This power can be calculated from the ambient temperature and the irradiance data over the tilted PV array using the conversion model presented in Chapter 4.

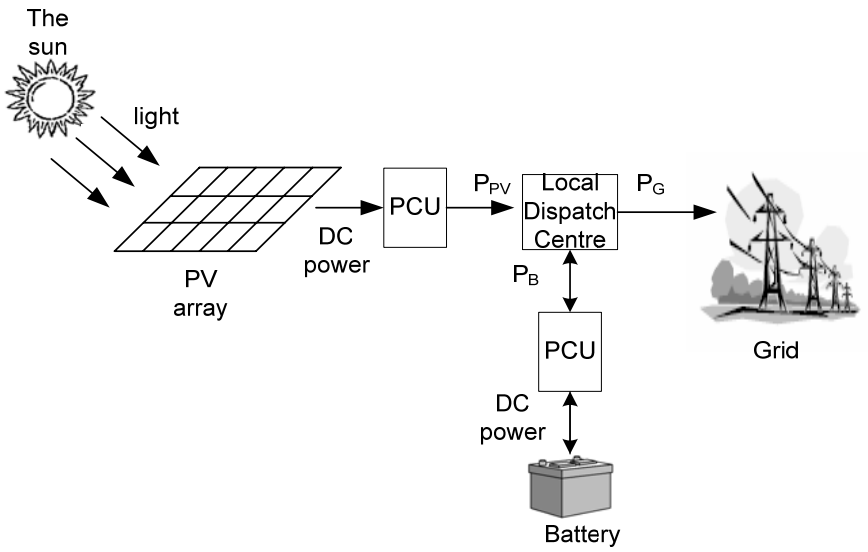

Figure 6-1 A grid-connected photovoltaic/battery system

6.3.1 Formulation of the optimization problem

The optimization problem is defined by Equations (6-1) to (6-13). The objective function, Equation (1), calculates the profit gained from the PV/BS system, $R_{PV/BS}$. The first term in this equation calculates the revenues gained from selling the energy generated from the system to the grid, where it is assumed that the average discount rate over the period of study is d_r. On the other hand, the second and third terms count for the net costs related to the battery power and energy ratings, respectively. The net costs of the battery are calculated from Equations (6-2) and (6-3), where the first terms in these equations estimate the present worth of the capital cost related to the battery power and energy ratings, respectively [150]. These terms consider the fact that the battery might have to be changed a number of times during the period of study. For example, if the battery lifetime is 2 years, and the study period is 6 years, the battery has to be replaced 2 times after the first installation. The second terms in both equations account for the present worth of the annual operating and maintenance costs while the third terms consider the present worth of the salvage value at the end of life of the battery [150].

The problem is subject to several technical constraints. The first constraint, Equation (6-4), ensures that the power balance of the system is maintained, where it is assumed that the battery can be energized from the PV system only. The second constraint, Equation (6-5), ensures that the change in the power output of the PV/BS system during each time step remains within the upper and lower bounds specified by the utility. It is worth mentioning that this constraint also facilitates the prediction of the generation of the system during a short period of time. For example, for a data set with 10 minutes time resolution, if the generation is 5 MW at a certain instant, and the values of x and y are specified to be 100kW/10min., then, the power generated from the system after 1 hour is expected to be between 4.4 and 5.6 MW. Equation (6-6) displays the constraint that ensures the energy balance of the battery is maintained at any instant during the day. On the other hand, Equation (6-7) ensures that the energy of the battery at the end of each day is available for the beginning of the next day in the same year. Similarly, Equation (6-8) ensures that the energy of the battery at the end of the last day of the year is available for the beginning of the first day in the next year. Equation (6-9) assumes that the battery is fully charged at the beginning of the period of study. Equation (6-10) ensures that the battery power during charge or discharge does not exceed the power rating of the battery. To limit the depth of discharge of the battery to the recommended level, Equation (6-11) ensures that the battery energy at any instant is not less than the minimum allowable state of charge. This equation also does not permit the energy inside the battery to exceed the battery rating during the charging period. In general, the contract between the utility and the system owner specifies the power rating of the PV system installed at a certain location. Thus, the power generated from the PV/BS system should not exceed the ratings specified in the contract as shown in Equation (6-12). The last constraint in this optimization problem is the non-negativity constraint, where the power and energy ratings of the battery, the energy of the battery at any instant and the power injected into the grid at any instant are all positive. To solve this optimization problem, all thirteen equations are modeled in GAMS and solved using the MOSEK solver [151].

Objective function:

$$R_{PV/BS} = \max\left\{\sum_{k=1}^{T}\left[\left(\sum_{j=1}^{N}\sum_{i=1}^{M}\alpha\times\frac{\Delta t}{60}\times P_G\left(i,j,k\right)\right)\times\frac{1}{\left(1+d_r\right)^k}\right]-\beta\frac{1}{\eta_{PB}}P_{B\max}-\gamma\frac{1}{\eta_{ch}}E_{B\max}\right\} \quad (6\text{-}1)$$

$$\beta=\sum_{i=1}^{N_B}C_P\frac{1}{\left(1+d_r\right)^{(i-1)T_B}}+O_P\frac{\left(1+d_r\right)^T-1}{d_r\left(1+d_r\right)^T}-\sum_{i=1}^{N_B}S_P\frac{1}{\left(1+d_r\right)^{iT_B}} \quad (6\text{-}2)$$

$$\gamma=\sum_{i=1}^{N_B}C_E\frac{1}{\left(1+d_r\right)^{(i-1)T_B}}+O_E\frac{\left(1+d_r\right)^T-1}{d_r\left(1+d_r\right)^T}-\sum_{i=1}^{N_B}S_E\frac{1}{\left(1+d_r\right)^{iT_B}} \quad (6\text{-}3)$$

Subject to:

$$P_G\left(i,j,k\right)=P_{PV}\left(i,j,k\right)+P_B\left(i,j,k\right) \qquad \forall i,j,k \quad (6\text{-}4)$$

$$P_G\left(i-1,j,k\right)-x\le P_G\left(i,j,k\right)\le P_G\left(i-1,j,k\right)+y \qquad \forall i>1,j,k \quad (6\text{-}5)$$

$$E_B\left(i,j,k\right)=E_B\left(i-1,j,k\right)-\frac{\Delta t}{60}P_B\left(i-1,j,k\right) \qquad \forall i>1,j,k \quad (6\text{-}6)$$

$$E_B\left(1,j,k\right)=E_B\left(M,j-1,k\right) \qquad \forall j>1,k \quad (6\text{-}7)$$

$$E_B\left(1,1,k\right)=E_B\left(M,N,k-1\right) \qquad \forall k>1 \quad (6\text{-}8)$$

$$E_B\left(1,1,1\right)=E_{B\max} \quad (6\text{-}9)$$

$$-P_{B\max}\le P_B\left(i,j,k\right)\le P_{B\max} \qquad \forall i,j,k \quad (6\text{-}10)$$

$$SOC_{\min}E_{B\max}\le E_B\left(i,j,k\right)\le E_{B\max} \qquad \forall i,j,k \quad (6\text{-}11)$$

$$0\le P_G\left(i,j,k\right)\le P_{G\max} \qquad \forall i,j,k \quad (6\text{-}12)$$

$$P_{B\max},E_{B\max},E_B,P_G\ge 0 \quad (6\text{-}13)$$

The parameters in Equations (6-1) to (6-13) are defined as follows:

C_E = Capital cost related to the battery energy capacity ($/kWh)

C_P = Capital cost related to the battery power capacity ($/kW)

d_r = Average annual market discount rate

E_B = Energy discharged from the battery (kWh)

$E_{B\,max}$ = Maximum energy storage capacity of the battery (kWh)

M = Number of data points in each power segment (corresponding to one day)

N = Number of PV power segments in each year

N_B = Number of times the battery will be replaced during period T ($N_B = {}^{T}/_{T_B}$)

O_E = Annual operating cost related to the battery energy storage capacity ($/kWh)

O_P = Annual operating cost related to the battery power capability ($/kW)

P_B = Power generated from the battery (kW), positive: discharging and negative: charging

$P_{B\,max}$ = Maximum power capability of the battery (kW)

P_G = Power injected into the grid (kW)

P_{Gmax} = Maximum power injected into the grid (kW)

P_{PV} = Power generated from the PV panels (kW)

$R_{PV/BS}$ = Profit gained from the PV/BS system over period T ($)

S_E = Salvage value related to the battery energy storage capacity ($/kWh)

S_P = Salvage value related to the battery power capability ($/kW)

SOC_{min} = Minimum allowable state of charge

T = Number of years studied

T_B = Lifetime of the battery (years)

x = Lower limit for power fluctuations (kW)

y = Upper limit for power fluctuations (kW)

α = Price of energy sold from the PV system ($/kWh)

β = Present worth related to the battery power capability ($/kW)

γ = Present worth corresponding to the battery energy storage capacity ($/kWh)

Δt = Minute resolution of data points (minute)

η_{PB} = Power efficiency of the battery charger

η_{ch} = Energy efficiency of the battery

The output power of the PV system, P_{PV}, in the optimization problem is obtained from the historical irradiance and temperature at each time step. The decision variables in the problem are E_{Bmax}, P_{Bmax}, E_B, P_B, P_G, $R_{PV/BS}$. Thus, the problem is solved to find the ratings of the battery, the power and energy of the battery at each time step, and the power injected into the grid at each time step that satisfy the fluctuation limits. Moreover, the problem is also used to choose the battery type by comparing the profit gained from selling the energy generated from the system with different battery types. The chosen type of battery is further examined by changing the different parameters of the battery and investigating the corresponding changes in the profit.

6.3.2 Choice of battery types

The first battery type considered in this study is the traditional LA battery and the second is the new technology NaS battery. The main differences between the two types of battery are: 1) the NaS battery has an operating lifetime that is almost three times that of the LA battery [132], 2) the operating cost of the NaS battery is much lower than that of the LA battery [152], and 3) the capital cost of the LA battery is lower than that of the NaS battery [146]. In this study, the lifetime of the LA battery is assumed to be 2 years [153], and thus, that of the NaS battery is assumed to be 6 years. For both batteries, the

charging efficiencies and the capital costs, which constitute the main portion of the batteries cost, are obtained from [146]. The operating cost of the LA battery is assumed to be 10% of the capital cost, and ten times that of the NaS battery. The salvage values for both batteries are assumed to be 1% of the capital costs. The parameters used in the optimization problem are displayed in Table 6-1.

Table 6-1 Parameters used in the optimization problem

	LA Battery	***NaS Battery***
T_B (years)	2	6
η_{PB}	0.85	0.85
η_{ch}	0.75	0.85
C_P (\$/kW)	300	1000
C_E (\$/ kWh)	150	170
O_P (\$/ kW)	30	3
O_E (\$/ kWh)	15	1.5
S_P (\$/kW)	3	10
S_E (\$/kWh)	1.5	1.7
SOC_{min}	0.2	0.2
T (years)	6	
α (\$/kWh)	0.42	
d_r	10%	

The profit, R_{PV}, of the electricity produced by the PV power to the grid without smoothing the power fluctuations, i.e., without using BS system, is considered as the base case, and is calculated by:

$$R_{PV} = \sum_{k=1}^{T}\left[\left(\sum_{j=1}^{N}\sum_{i=1}^{M}\alpha\times\frac{\Delta t}{60}\times P_G\left(i,j,k\right)\right)\times\frac{1}{\left(1+d_r\right)^k}\right] \tag{6-14}$$

Thus, the profit gained in the base case during the 6-year period of study is \$23,666,808.

For each battery type, and for all the other case studies, the percentage change in profit is calculated by:

$$\Delta R\,\% = \frac{R_{PV/BS} - R_{PV}}{R_{PV}} \times 100 \qquad (6\text{-}15)$$

The results obtained from using the two types of battery to limit the power fluctuations to ±250kW/10min. during the period 9:00 am to 3:00 pm are displayed in Table 6-2. The profit calculated in this table is that obtained from selling the energy to the grid during the 6-year period of study.

Table 6-2 Profit obtained from the two types of batteries

Type of Battery	***Profit with battery, $R_{PV/BS}$, ($)***	***Change in profit, $\Delta R\%$***	***Battery power rating (kW)***	***Battery energy rating (kWh)***
LA	16,483,352	-30.4	4153	5698
NaS	18,359,992	- 22.4	3677	6491

The results of Table 6-2 indicate that for both types of battery there is a loss in profit as compared to the case when no battery is installed. However, the loss in profit is less if the expensive NaS battery with the parameters displayed in Table 6-1 is used. The results also show that the required power ratings of the two batteries are not the same, even with the power-related efficiency being the same. This is due to the fact that the optimization problem aims to maximize the profit while limiting the fluctuations to a certain value. Thus, for different parameters of the batteries, different power and energy patterns are generated, and accordingly, different ratings for the batteries are specified. Figure 6-2 illustrates this fact as it shows the power generated from the two PV/BS systems for the same PV power profile. Figure 6-3 and Figure 6-4 show the corresponding power and energy profiles of the batteries indicating different maximum values for the two battery types.

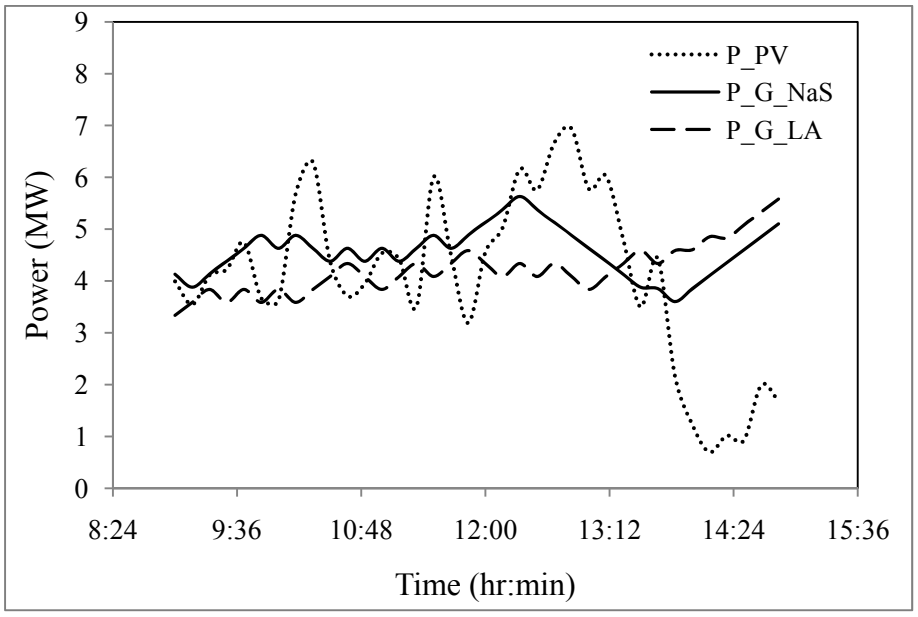

Figure 6-2 Power patterns generated from the PV/BS systems for a cloudy day

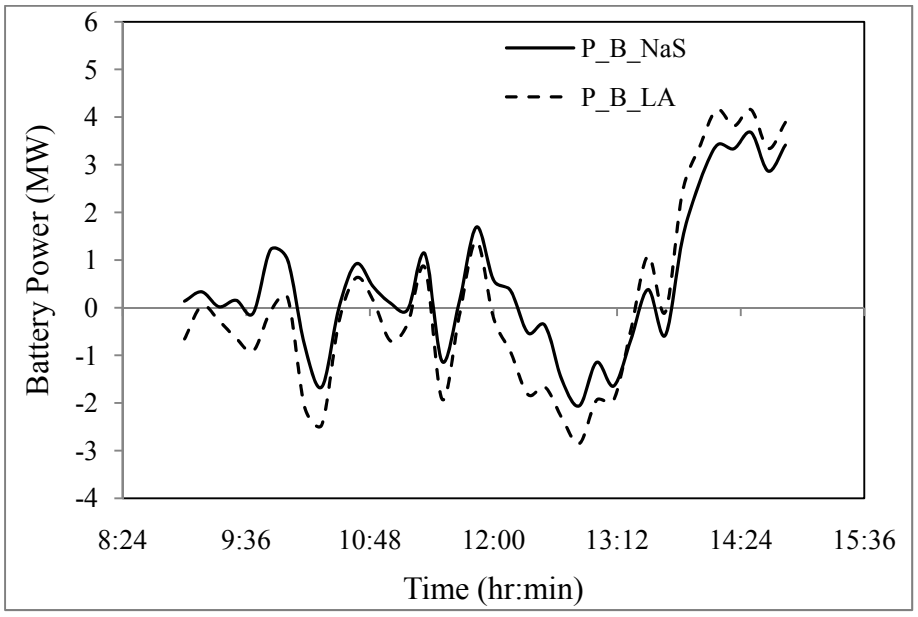

Figure 6-3 Power patterns for the two types of battery

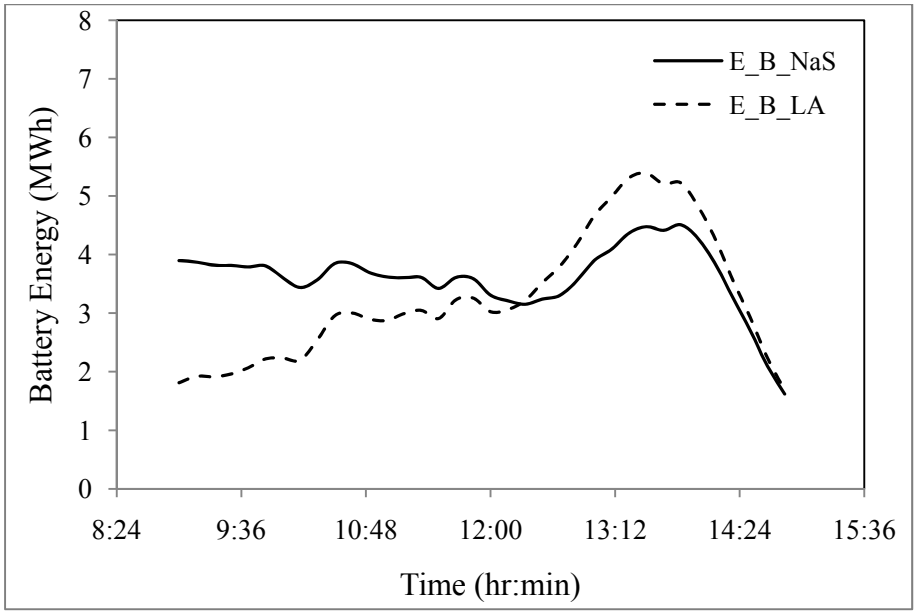

Figure 6-4 Energy patterns for the two types of batteries

6.3.3 Effect of changing the operating period of the battery

In this sub-section, the effect of changing the operating period of the battery is examined. The NaS battery is used to limit the fluctuations within each time step to ±250kW/10min; however, the time period during which fluctuations are limited is varied. The results displayed in Table 6-3 show that the power rating of the battery is constant for all time periods, indicating that the maximum power fluctuations occur around noon. On the other hand, the energy rating of the battery depends on the period of operation because the energy rating of the battery is determined by the power pattern of the battery during the operating period. The results of Table 6-3 also show that the gained profit is almost the same for different operating periods of the battery, with the 2-hour operating period producing slightly higher profit. The main reason behind the independence of the profit from the operating period is that the capital cost of the battery per kW is relatively high compared to its capital cost per kWh. Thus, for a constant power rating of the battery, the profit gained from the system will not be significantly affected.

Table 6-3 Effect of changing the operating period of the battery

Time Period	Profit with battery, $R_{PV/BS}$, ($)	Change in profit, $\Delta R\%$	Battery power rating (kW)	Battery energy rating (kWh)
11:00 am- 1:00 pm	18,802,269	-20.6	3677	3349
10:00 am- 2:00 pm	18,439,918	-22.1	3677	5881
9:00 am- 3:00 pm	18,359,992	-22.4	3677	6491
8:00 am- 4:00 pm	18,508,659	-21.8	3677	5830
7:00 am- 5:00 pm	18,470,353	-22	3677	6033

6.3.4 Effect of changing the power fluctuation limit

To examine the relation between the gained profit and the power fluctuations limit, the NaS battery is used to limit the fluctuations to ±400kW/10min., ±250kW/10min., ±100kW/10min. and 0kW/10min. (constant power). The profits displayed in Table 6-4 for all fluctuation limits are lower than the case when no battery is used. Moreover, the profit decrease as the fluctuation limit decreases, with the lowest profit corresponding to the case when the PV/BS system generates constant power. This is a logical result due to the fact that the power and energy ratings of the battery have to increase as the fluctuation limit decreases to satisfy the constraint of Equation (6-5).

Table 6-4 Effect of changing the power fluctuation limit

Fluctuation Limit (kW/10min.)	Profit with battery, $R_{PV/BS}$, ($)	Change in profit, $\Delta R\%$	Battery power rating (kW)	Battery energy rating (kWh)
±400	19,298,635	-18.5	3377	3535
±250	18,470,353	-22.4	3677	6491
±100	16,991,448	-28.2	3977	11285
0	15,499,218	-34.5	4805	12959

6.3.5 Effect of changing the cost and efficiency of the battery

In this sub-section, the effects of reducing the capital cost and increasing the efficiency of the battery due to the possible technology advances are investigated. Any change in the capital cost of the battery is accompanied by a similar change in the salvage value of the battery.

Table 6-5 shows that reducing the capital cost related to the power and energy ratings of the battery by 50% reduces the loss in profit by approximately 50%. The results also indicate that the reduction in the capital cost per kW has a higher impact on decreasing the loss in profit when compared to the reduction in the capital cost per kWh. This is because the capital cost per kW used in this research constitutes the main portion of the battery cost. On the other hand, the results show that increasing the power and energy efficiencies of the battery does not have a considerable impact on the change in the profit.

It is worth mentioning that, for all cases, the power and energy ratings of the battery are estimated by the optimization problem to be 3,677 kW and 6,491 kWh, respectively. These are the ratings of the battery that can satisfy the ±250kW/10min. power fluctuations constraint. Any increase in the battery ratings will lead to additional loss in the profit as it will lead to additional costs to the system owner.

Table 6-5 Effect of changing the capital cost of the battery

Capital costs of battery, ($)		Salvage values of battery, ($)		Efficiencies of battery		Profit with battery ($)	Change in profit
C_P	C_E	S_P	S_E	η_P	η_E	$R_{PV/BS}$	$\Delta R\%$
1000	170	10	1.7	0.85	0.85	18,359,992	-22.4
1000	85	10	0.85	0.85	0.85	19,005,399	-19.7
500	170	5	1.7	0.85	0.85	20,510,724	-13.3
500	85	5	0.85	0.85	0.85	21,156,131	-10.6
1000	170	10	1.7	0.95	0.95	18,959,853	-19.9

6.4 Smoothing the Power Fluctuations by Installing a Dump Load

Instead of using a BS system, the system owner can reduce the fluctuations in the power generated from the PV system by reducing the power injected into the grid during periods of high fluctuations. One method that can be used for this purpose is to replace the BS system in Figure 6-1by a dump load. The dump load (DL) consists of a resistance and a controller to control the power flow through the load. In some cases, the load is cooled to avoid excessive heating during operation. The main function of the dump load is to absorb the excess power generated from the PV system, and thus, to smooth the power injected into the grid. In the analysis presented in this section, the dump load is assumed to operate during the 6-year period of study without replacement. The capital cost of the dump load including the controller is assumed to be $200/kW while the operating costs and the salvage value of the dump load are considered to be zero. To estimate the power rating of the dump load and the profit obtained from selling the energy generated from this system, the optimization problem is modified to:

Objective function:

$$R_{PV/DL} = \max\left\{\sum_{k=1}^{T}\left[\left(\sum_{j=1}^{N}\sum_{i=1}^{M}\alpha\times\frac{\Delta t}{60}\times P_G\left(i,j,k\right)\right)\times\frac{1}{\left(1+d_r\right)^k}\right]-\delta P_{D\max}\right\} \quad (6\text{-}16)$$

Subject to:

$$P_G\left(i,j,k\right)=P_{PV}\left(i,j,k\right)-P_D\left(i,j,k\right) \qquad \forall i,j,k \quad (6\text{-}17)$$

$$P_G\left(i-1,j,k\right)-x\le P_G\left(i,j,k\right)\le P_G\left(i-1,j,k\right)+y \quad \forall i>1,j,k \quad (6\text{-}18)$$

$$0\le P_D\left(i,j,k\right)\le P_{D\max} \qquad \forall i,j,k \quad (6\text{-}19)$$

$$0\le P_G\left(i,j,k\right)\le P_{G\max} \qquad \forall i,j,k \quad (6\text{-}20)$$

$$P_D, P_G \geq 0 \tag{6-21}$$

The parameters in Equations (6-16) to (6-21) are defined as follows:

d_r = Average annual market discount rate

M = Number of data points in each power segment (corresponding to one day)

N = Number of PV power segments in each year

P_D = Power lost in the dump load (kW)

$P_{D\,max}$ = Maximum power capability of the dump load (kW)

P_G = Power injected into the grid (kW)

P_{Gmax} = Maximum power injected into the grid (kW)

P_{PV} = Power generated from the PV panels (kW)

$R_{PV/DL}$ = Profit gained from the PV/DL system over period T ($)

T = Number of years studied

x = Lower limit for power fluctuations (kW)

y = Upper limit for power fluctuations (kW)

α = Price of energy sold from the PV system ($/kWh)

δ = Present worth of the dump load ($/kW)

Equations (6-16) to (6-21) are solved to estimate the profit gained from the PV/DL system with the same fluctuations limits specified in Sub-section 6.3.4. The power rating of the dump load and the percentage change in profit as compared to the case when no power is dumped are shown in Table 6-6 for different fluctuation limits. The base case for calculating the percentage change in profit is the case when no power is dumped.

Comparing the results displayed in Table 6-6 with those of Table 6-4 leads to the following statements:

1- Dumping part of the generated power to limit the power fluctuations will lead to loss of profit. This loss depends on the fluctuation limit imposed by the network operator.

2- For high fluctuation limits, dumping the power leads to less loss in profit as compared to the case when a BS system with the parameters given in Table 6-1 is installed. On the other hand, for lower fluctuation limits, the use of a BS system leads to lower loss in profit.

3- For the case when it is required to generate constant power from the system (fluctuation limit = 0), dumping the power leads to 51% loss in the profit. This large loss is due to the fact that during periods of high fluctuation, the PV/DL system has to generate a constant power equal the lowest power during this period, while any additional generation is lost in the dump load.

Table 6-6 Reduction of power fluctuations using a dump load

Fluctuation Limit (kW/10min.)	Profit with dump load, $R_{PV/DL}$ ($)	Change in profit, $\Delta R\%$	Dump load power rating (kW)
±400	20,188,999	-14.7	6754
±250	18,946,440	-19.9	7354
±100	15,957,593	-32.6	7954
0	11,587,790	-51.0	8532

6.5 Smoothing the Power Fluctuations by Operating Below the Maximum Power Point

Another method that can be used to reduce the fluctuations in the output power of the PV system is to curtail the power by operating below the MPP rather than dumping the power using a dump load. This method does not require any additional installations, as only the control strategy of the PCU needs to be modified. In this case, the curtailed

power, P_C, can be treated in the same way as the dumped power, P_D, and, thus, Equations (6-16) to (6-21) can be still used but by setting $\delta = 0$ and replacing the symbols P_D and $R_{PV/DL}$ with P_C and $R_{PV/PC}$, respectively. The results in Table 6-7, obtained from solving the optimization problem for the same fluctuation limits, show that the loss in profit is less than in the case of using a dump load. This is mainly due to the savings in the capital cost of the dump load. The loss in profit is also less than in the case when a BS system is used (Table 6-4) except for the case when the generated power from the PV system is required to be constant.

Table 6-7 Reduction of fluctuations by operating below the MPP

Fluctuation Limit (kW/10min.)	**Profit with power curtailment ($), $R_{PV/PC}$**	**Change in profit, $\Delta R\%$**
±400	21,539,799	-9.0
±250	20,417,240	-13.7
±100	17,548,393	-25.9
0	13,294,190	-43.8

Based on the above results, a logical extension is to investigate the combination of a BS system and the curtailment of power by operating the PCU below the MPP. In this case, the original optimization problem, Equations (6-1) to (6-13), can be applied with minor modifications. The first modification is to account for the curtailed power in the power balance constraint, Equation (6-4), which is replaced by:

$$P_G(i,j,k) = P_{PV}(i,j,k) + P_B(i,j,k) - P_C(i,j,k) \qquad \forall i,j,k \qquad (6\text{-}22)$$

The second modification is in the non-negativity constraint, Equation (13), which becomes:

$$P_{B\max}, E_{B\max}, E_B, P_G, P_C \geq 0 \tag{6-23}$$

The results displayed in Table 6-8 and Figure 6-5 show that installing a BS system and curtailing the power by operating below MPP lead to the least loss in profit for all power fluctuations limits. Moreover, the power and energy ratings of the battery are highly reduced as compared to the case when power is not curtailed. Thus, combining the two methods leads to the most economical solution for the system owner. In this case, the fluctuations are reduced by charging and discharging the battery. When the power limit of the battery is reached, the excess power is curtailed by operating the PCU below the MPP. Figure 6-6 shows the power profiles for different system components in a cloudy day when the power fluctuations are limited to ±250kW/10min. It can be seen that the curtailed power is zero except when the power being charged to the battery reaches the power capacity limit of the battery (1099 kW) between 10:00 and 10:10 am and between 11:00 and 11:10 am.

Table 6-8 Effect of changing the power fluctuation limit

Fluctuation Limit (kW/10min.)	**Profit with battery and power curtailment, $R_{PV/BS\text{-}PC}$ ($)**	**Change in profit, $\Delta R\%$**	**Battery power rating (kW)**	**Battery energy rating (kWh)**
±400	21,819,402	-7.8	675	817
±250	21,149,109	-10.6	1099	1514
±100	19,794,563	-16.4	1910	3279
0	17,859,939	-24.5	2946	5881

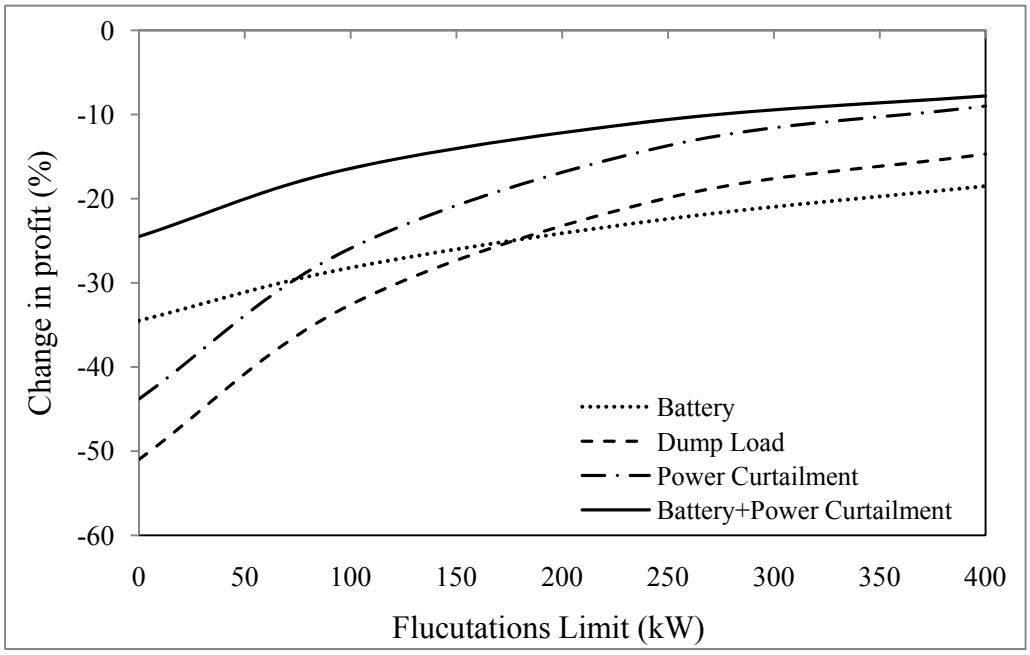

Figure 6-5 Percentage change in profit for different methods

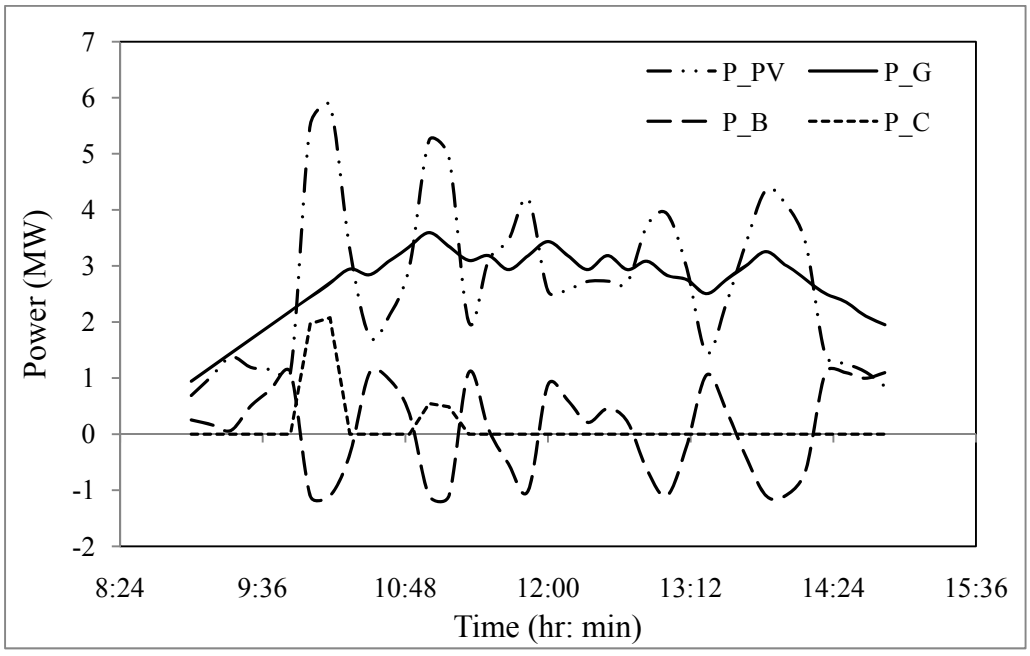

Figure 6-6 Power profiles for different system components

6.6 Incentives for the PV System Owner

The results presented in the previous sections show that all the investigated methods can effectively reduce the power fluctuations of the large PV systems. However, the profit of selling the PV energy to the grid obtained from any of the methods is less than that obtained in the case when no limits on the power fluctuations are imposed. Thus, the utilities have to cooperate with the system owner to compensate for the profit losses. A possible way of cooperation is by providing additional incentives to the system owner. These incentives can be in the form of increasing the price of PV energy during the periods when the fluctuations are to be limited. The increase in price depends mainly on the amount of profit lost, and thus, on the method applied to reduce the power fluctuations. The lowest increase in the selling price is expected to correspond to the case when a BS system is operated while curtailing the output power of the PV arrays by operating below the MPP. Table 6-9 shows the selling prices required to overcome the loss in profit due to adopting any of the methods that can limit the power fluctuations. The selling price is based on a nominal selling price of $0.42/kWh and a fluctuation limit of ± 250kW/10min from 9:00 am to 3:00 pm.

Table 6-9 Price of selling the PV energy for the different methods

Method	***Selling Price ($/kWh)***	***Profit, R ($)***	***Change in profit, ΔR%***
No fluctuation limits	0.42	23,666,808	0
Battery	0.54	23,796,871	≈ 0
Dump load	0.55	23,709,281	≈ 0
Power curtailment	0.51	23,714,592	≈ 0
Battery and Power curtailment	0.48	23,702,874	≈ 0

The results of Table 6-9 show that using batteries to limit the fluctuation to ±250kW/10min. may lead the utility to pay an additional 0.12 $ (28.6%) for each kWh during the period of operation of the battery. In case of using a dump load, the utility would have to pay an additional $0.13/kWh (30%), which is slightly higher than that in the case of using a BS system. If the power is curtailed by operating below the MPP the utility should pay an additional $0.09/kWh (21%), which is a better option. On the other hand, if the system owner installs a battery and curtails the power by operating below the MPP, the utility will pay only an additional $0.06/kWh (14%), which is the best option for the utility.

6.7 Summary and Conclusions

This chapter examined the economical aspect of using different methods to smooth the output power of PV systems. The results for all studied cases show that there is a loss of profit from selling PV energy to the grid irrespective of the method adopted. When a battery storage system is used, the expensive NaS batteries prove to be superior to the cheap LA batteries from a long-term economical point of view due to the higher efficiency and longer lifetime of NaS batteries.

Usually, the battery operates during the period of high generation of the PV system, around noon. Extending the period of operation on both sides of the noon hour does not affect the power rating of the battery. This indicates that the most severe power fluctuations occur around noon. However, upon changing the fluctuation limit, it is observed that the loss in profit increases as the power fluctuation limit decreases. This is mainly because setting a low fluctuation limit increases both power and energy ratings of the battery.

Dumping the excess power to satisfy the power fluctuations constraint, and power curtailment by operating below MPP, prove to result in lower profit loss than using a battery when larger power fluctuations are allowed, with power curtailment being the

superior option. On the other hand, combining a battery storage system and power curtailment is found to be the most economical solution. In all cases, offering incentives in the form of higher PV energy sale price during the period where the fluctuations are limited seems to be the only way to overcome the loss in profit and encourage the PV system owner to provide smoother power.

Chapter 7

Summary, Contributions, and Future Work

7.1 Summary

The main objective of this thesis is to help increase the penetration level of grid-connected PV systems. To achieve this objective, the thesis proposed a new method that can accurately analyze the impacts of large centralized grid-connected PV systems on the operation of the electric network, especially due to the fluctuation of power generated from these systems. Moreover, the thesis investigated the economical aspects of different strategies that can mitigate the adverse impacts of these power fluctuations.

In Chapter 2, an overview of the main components of grid-connected PV systems was presented while highlighting the recent research activities carried out in this field. Weather stations usually measure the global irradiance on a horizontal surface, and thus, models are required to estimate the irradiance on the tilted surface of PV panels. The accuracy of any of these models depends on the location where the PV system is being installed. Thus, it is important to choose a suitable model for the case under consideration. Also, there are different models that can be used to calculate the DC output power from the PV cells. These models can be divided into two main categories; detailed and simplified models. Detailed models attempt to represent the physics of the PV cell and are usually suitable for studies that require detailed cell information. On the other hand, simplified models can provide a direct estimate of the maximum power generated

from the PV cell at certain operating conditions. Thus, these models are suitable for studies that aim to identify the impacts of PV systems on the electric network. However, it is important to examine the accuracy of using any of the simplified models by comparing their outputs with those obtained from one of the detailed models.

The main purpose of Chapter 3 was to discuss, in details, the positive and negative impacts of PV systems on the performance of the electric network. Some of the benefits of PV systems can include: power loss reduction, improvement of the voltage profile, and reduction in the maintenance and operating costs of some equipment installed in the electric network. In order for these benefits to be effective a number of conditions should be satisfied, such as strategic placement and proper sizing of the PV system, and suitability of the output power profile of the system. In some cases PV systems can have negatives impacts on the operation of the network. These impacts can be divided into: 1) impacts on the generation side, 2) impacts on transmission and sub-transmission networks, and 3) impacts on distribution networks. However, current trends show that large centralized PV systems are being installed in distribution networks. Thus, these networks should be the focus of studies that aim to assess the impacts of PV systems on the performance of the network.

Chapter 3 also highlighted the inadequacy of the existing methods used to assess the impacts of power fluctuations generated from PV systems. For example, methods based on the deterministic approach consider specific loading and generation scenarios, and thus, the obtained results are valid only for the specified operating conditions. Moreover, these methods do not consider the temporal information in the data which means that the impacts of power fluctuations cannot be assessed. Methods based on the probabilistic approach can provide accurate information about the performance of the network if the random variables are modeled properly. However, these methods do not consider the temporal information in the data, and thus, the impacts of power fluctuations cannot be investigated. Finally, methods based on chronological simulations consider the fluctuations in the output power of the PV system. However, these methods require

performing extensive simulations, especially if long historical time series data with high time resolution is used, which is necessary for obtaining accurate results about the impacts of power fluctuations. Moreover, using a huge amount of data in the analysis makes it difficult to obtain comprehensive results that can be used efficiently to assess the performance of the network.

Based on the discussions presented in Chapter 3, the layout of a new method, the clustering-based method, was introduced in Chapter 4. One of the advantages of the new method is its ability to use the long historical time series data in an efficient and intelligent way that helps reduce the amount of analysis while preserving the temporal information of the data. The main idea of the proposed method is to divide the long PV power time series into segments that can be grouped together according to their profiles, such that the segments with close profiles are grouped in one cluster. Accordingly, a representative segment for each cluster is used to replace all other segments of the cluster in the analysis.

After presenting the layout of the proposed method, Chapter 4 also introduced the details of different stages of the method. In the first stage, the models that are used to convert the temperature and irradiance obtained from weather stations into AC power generated from the PV system were presented and their accuracies were examined. In the second stage, the techniques that can be used to divide the PV power time series into segments were explained. These methods depend mainly on the loading profile of the electric network, which is considered an advantage for the proposed method. In the third stage, the feature extraction stage, the use of two different sets of features was proposed. The first set of features is the sampled data of each segment while the second set is the conversion coefficients obtained from applying the Principal Component Analysis. In the fourth stage, three different clustering algorithms that can be used to group together power segments with similar profiles were introduced.

The use of different feature extraction techniques, different clustering algorithms, as well as different cluster representatives, generates different results from the power flow

analysis. Thus, one of the important steps in the proposed method is to compare different techniques used in these stages to be able to choose the technique that is most suitable for the data set in hand. To achieve this task, the cluster validity indices that can be used for evaluating different clustering algorithms were presented in Chapter 5. The results obtained from applying these indices showed their tendency to prefer the clustering algorithm that generates more clusters with single segments. This preference does not comply with the main purpose of the clustering process in the proposed method. Thus, a new index that is more suitable for utility analysis was proposed and examined. One of the advantages of the proposed index is its ability to provide information about using a certain group of cluster representatives to represent the power segments contained in the clusters. Moreover, the values of the index can be used as a guideline to choose a suitable number of clusters.

In the second part of Chapter 5, the clustering-based method was used to investigate the impacts of power fluctuations generated from a large PV system on the performance of a distribution feeder. The method was also used to estimate the annual energy loss in the feeder, and hence, was applied to identify the suitable size and site of the PV system. The results obtained in this section proved the ability of the proposed method to accurately assess the performance of the network in the presence of large PV systems.

After introducing and applying the method that can be used to analyze the impacts of power fluctuations generated from PV systems, Chapter 6 examined different strategies that can reduce these fluctuations. The strategies considered in this chapter were: 1) the use of battery storage systems, 2) the use of dump loads, 3) curtailing the PV power by operating the power conditioning unit below the maximum power point, and 4) the use of battery storage systems while curtailing the PV power. The main focus in the analysis was to examine the impact of implementing each of these strategies on the profit gained by the PV system owner. To achieve this task, an LP optimization problem was formulated and solved to estimate the maximum profit gained by the system owner. The results obtained from applying all four strategies showed that there was a loss in profit

from selling the PV energy as compared to the case when the power fluctuations of the PV system were not smoothed out. The least loss in profit corresponded to the case when a battery storage system was used while curtailing some of the output power from the PV system. To overcome the loss in profit, it was suggested to offer incentives, in the form of higher PV energy sale price, to encourage the PV system owner to provide smoother power.

7.2 Main Contributions of the Research

The main contributions of the research presented in this thesis can be summarized in the following:

1- This thesis presented a comprehensive overview of the current status of grid-connected PV systems while highlighting the research activities in different areas related to this field. Moreover, the thesis highlighted the importance of studying the impacts of large centralized PV systems on distribution networks. This is because current trends show that sizes and locations of installation of these systems are expected to highly impact these networks.

2- The thesis also showed the need to develop a new method to analyze the impacts of installing large PV systems on the operation of the electric network, especially due to power fluctuations generated from these systems. This is mainly due to the fact that existing methods used to study the impacts of PV systems either do not include the temporal information of the data in the analysis, and thus, fail to provide any details about power fluctuations, or are difficult to apply in cases where long historical time series data with high resolution is used.

3- The clustering-based method proposed in this research is a novel method that overcomes the drawbacks of existing methods used to evaluate the performance of PV systems. The method utilizes the available data in an efficient and intelligent manner, and thus, generates comprehensive results that can help the

system operator better understand the behavior of the PV system. Some advantages of the proposed clustering-based method are:

a. The method groups the data in a way that facilitates the discovery of important profiles of power segments generated from the PV system with potential impact on the performance of the electric network.

b. The method can evaluate the performance of the electric network due to the installation of the PV system with high accuracy and with a highly reduced number of simulations.

c. Clustering of PV power segments facilitates performing statistical analysis to estimate the probability of occurrence of segments with potential impact on the performance of the electric network. This in turn can help in predicting periods in which these segments are most likely to occur.

d. In cases where there are more than one loading profile for the electric network, using the cluster representatives in power flow simulations facilitates studying the impacts of possible combinations of different generation profiles of the PV system with different load profiles of the electric network.

4- This research recommended the use of a new index for comparing the clustering algorithms and feature extraction techniques used in the clustering-based method. The proposed index has the following advantages:

a. The index can be used to predict the deviation of the power flow results using a reduced data set based on a specific number of clusters, from those obtained using the full data. This in turn helps in choosing a suitable number of clusters.

b. The proposed index facilitates the choice of cluster representatives that can best represent the data. This is because the index has the ability to

compare the results obtained using different cluster representatives, which is not the case for other validity indices used to compare clustering algorithms.

c. Unlike the internal validity indices, the proposed index does not use the compactness and separation of the formed clusters to evaluate the goodness of the clustering. Thus, the index does not have the tendency to prefer clustering algorithms that produce higher number of clusters with single segments.

5- The thesis investigated the use of four strategies that can mitigate the impacts of power fluctuations generated from PV systems. One of the main advantages of the analysis presented in this part of the thesis is the focus on the economical aspects of using each of the proposed strategies. Accordingly, the results of this study can help both utilities and PV system owners to choose the most economical solution for smoothing out the output power generated from PV systems.

7.3 Scope of Future Research

This thesis establishes a new direction for research related to PV systems and other intermittent renewable energy sources. Based on the research presented in this thesis, some of the studies that can be carried out in the future are summarized in the following:

1- Investigate the performance of the clustering-based method with different loading profiles of the network and with different network topologies.

2- Examine the use of other sets of features and different grouping techniques in order to improve the accuracy of the results obtained from using the cluster representatives in the power flow analysis.

3- Apply the clustering-based method in studies related to asset management such as those carried out to predict the lifetime of voltage regulators. Other studies that

can also utilize the proposed method are those related to studying the impacts of very large PV systems on the scheduling of generating units.

4- Use the clustering-based method to analyze the impacts of wind farms on the performance of the electric network and comparing the results with those obtained from large PV systems.

5- Incorporate the hourly variable energy prices in the optimization model used for estimating the profit gained from selling the PV energy after implementing the methods used to smooth out the fluctuations in the output power of the PV system.

6- Investigate the effect of the severity of PV power fluctuations on the lifetime of the battery storage system.

Appendix A

Steps for Applying the Principal Component Analysis (PCA)

The steps of applying PCA in the Feature Extraction stage are as follows:

1- Considering the PV power segments in each category, organize all segments in an $m_x \times n_x$ matrix $\boldsymbol{X}$ (pattern matrix), where m_x is the number of data points of the PV power for each segment (features) and n_x is the number of segments (patterns).

2- Calculate the mean value of the patterns for each feature; then, subtract the mean from the patterns of each feature. The result is an $m_x \times n_x$ matrix $\boldsymbol{Z}$ with features having zero mean.

3- Calculate the $m_x \times m_x$ covariance matrix $\boldsymbol{Cov}$ using Equation (A-1); then, calculate the m_x eigenvectors and the m_x eigenvalues. Let the i^{th} eigenvalue be λ_i.

$$\boldsymbol{Cov}_{m_x \times m_x} = \frac{1}{n_x}\left(\boldsymbol{Z}_{\boldsymbol{m_x \times n_x}} \times \boldsymbol{Z}^{\boldsymbol{T}}_{\boldsymbol{m_x \times n_x}}\right) \qquad \text{(A- 1)}$$

4- Sort the columns of the eigenvector matrix in descending order of the eigenvalues, i.e., the first column should contain the eigenvector corresponding to the highest eigenvalue and the second column should contain the eigenvector corresponding to the second highest eigenvalue, and so on.

5- Plot the sorted eigenvalues and choose an appropriate number of the highest L eigenvalues. This plot is called the SCREE plot [114]. Another method to choose L is to select the number of eigenvalues that retains a certain ratio, ξ, of the variance of the original feature set in the new feature space. This can be obtained by:

$$\frac{\sum_{i}^{L} \lambda_i}{\sum_{i}^{m_x} \lambda_i} \geq \xi \qquad \text{(A- 2)}$$

6- The corresponding L eigenvectors are the principal components. Hence, form an $m_x \times L$ matrix, $\boldsymbol{V}$, containing the first L PCs.

7- Find the $L \times n_x$ projection matrix $\boldsymbol{A}$ that projects the original data onto the new basis. The matrix contains the transformation coefficients and can be obtained as:

$$\boldsymbol{A}_{\boldsymbol{L}\times\boldsymbol{n}_x} = \boldsymbol{V}^{\boldsymbol{T}}_{\boldsymbol{L}\times\boldsymbol{m}_x} \times \boldsymbol{Z}_{\boldsymbol{m}_x\times\boldsymbol{n}_x} \qquad \text{(A- 3)}$$

8- To reconstruct the data in the reduced dimensional space, Equation (A-4) is used.

$$\boldsymbol{X}'_{\boldsymbol{m}_x\times\boldsymbol{n}_x} = \boldsymbol{F}_{\boldsymbol{m}_x\times\boldsymbol{n}_x} + \boldsymbol{V}_{\boldsymbol{m}_x\times\boldsymbol{L}} \times \boldsymbol{A}_{\boldsymbol{L}\times\boldsymbol{n}_x} \qquad \text{(A- 4)}$$

Equation (A-4) can be written for each segment as:

$$x'_j(t) = f(t) + \sum_{i=1}^{L} a_{ij}\, v_i(t) \qquad \forall j \qquad \text{(A- 5)}$$

In (A-4) and (A-5),

$\boldsymbol{X}'$ is the $m_x \times n_x$ matrix of the reconstructed data;

$\boldsymbol{F}$ is the $m_x \times n_x$ matrix containing the mean of each feature;

$x'_j(t)$ is the j^{th} reconstructed segment;

m_x is the original number of features in each segment;

n_x is the number of segments;

t is the time at which the PV power is calculated during the day;

$f(t)$ is the mean of the samples at each time t;

L is the number of significant principle components;

$v_i(t)$ is the i^{th} principal component; and

$a_{i,j}$ is the i^{th} transformation coefficient for the j^{th} segment.

After obtaining the matrix $\boldsymbol{A}$ that contains the transformation coefficients of all samples, these coefficients can be used in the clustering stage to group segments that have close profiles, as explained before.

Appendix B

Comparison of the Clustering Algorithms for Different Data Sets

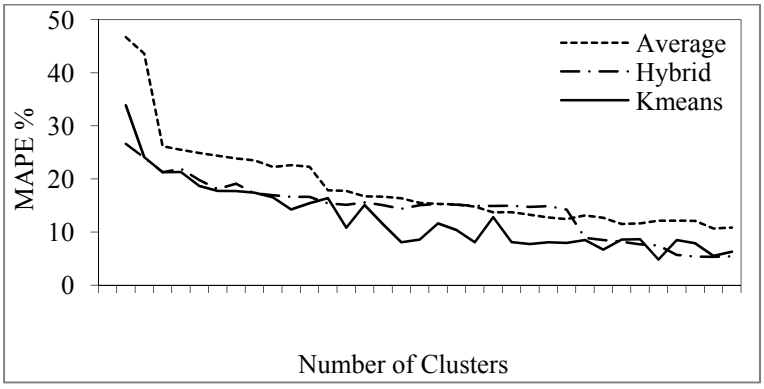

Figure B-1 *MAPE* index for the summer season of the 1-year data set

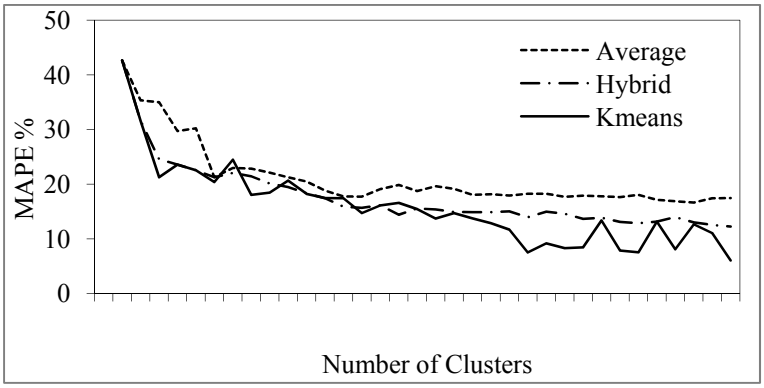

Figure B-2 *MAPE* index for the winter season of the 1-year data set

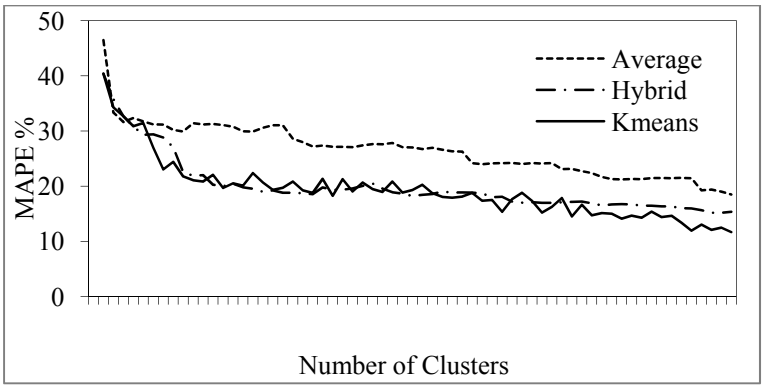

Figure B-3 *MAPE* index for the spring/fall season of the 1-year data set

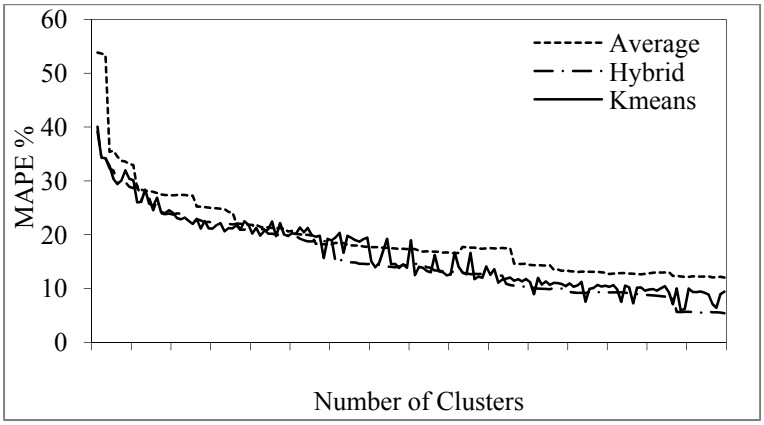

Figure B-4 *MAPE* index for the summer season of the 5-year data set

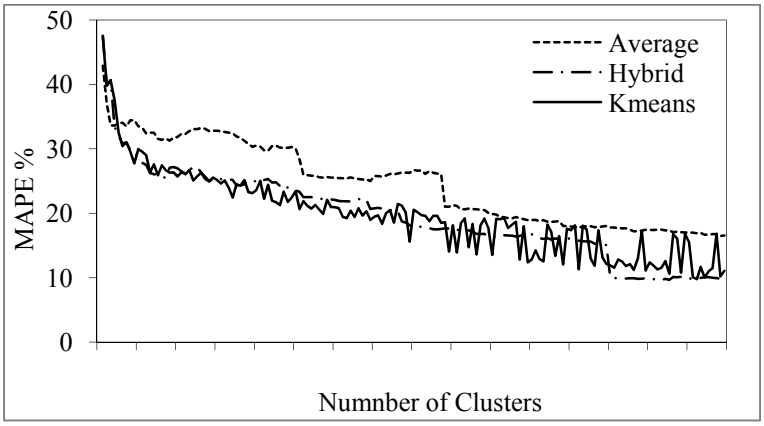

Figure B-5 *MAPE* index for the winter season of the 5-year data set

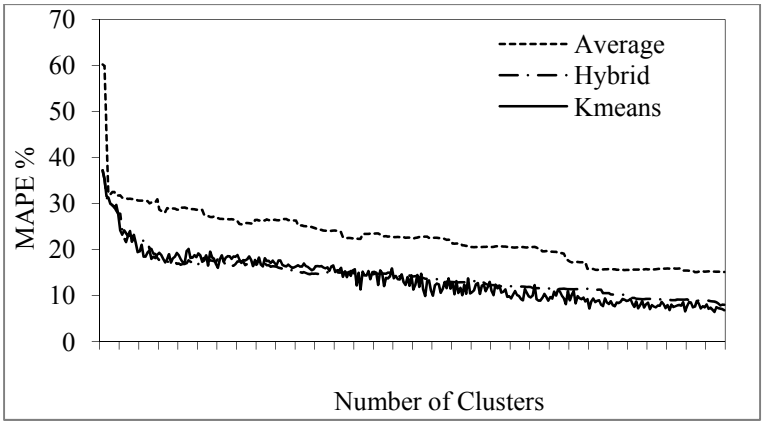

Figure B-6 *MAPE* index for the spring/fall season of the 5-year data set

Appendix C
Distribution Feeder Data

Table C-1 Substation Data

Name	Node	Phase	R (pu)	X (pu)	angle	V_{phase} (kVA)	S_{base} (kVA)	R (Ω)	X (Ω)	L (mH)
SS	1	ABC	0.00055	0.01715	0	16	10,000	0.04224	1.31712	3.494

Table C-2 Transformers Data

Name	From	To	Vph-prim (kV)	Vph-sec (kV)	Tap setting	Phase	S_{ph} (kVA)	R (pu)	X (pu)	connection
T1	7	8	16	4.8	1	ABC	1200	0.01	0.0595	DY -30
T2	15	16	16	16	1.0125	ABC	5000	0.001	0.0725	YY
T3	20	21	16	4.8	1	B	250	0.001	0.044	YY
T4	40	41	16	4.8	1	ABC	1200	0.01	0.0565	DY -30
T5	42	38	16	1	1	ABC	4333	0.0075	0.075	YY

Table C-3 Line data

Name	From	To	Length (km)	Phase	R (Ω /km)	X (Ω /km)	Charge (µs/km)
L1	1	2	5.7	ABC	0.1691	0.4182	3.954
L2	2	3	1.01	ABC	0.1691	0.4182	3.954
L3	2	4	0.4	ABC	0.1691	0.4182	3.954
L4	4	5	0.38	ABC	0.1691	0.4182	3.954
L5	5	6	0.13	ABC	0.1691	0.4182	3.954
L6	5	7	0.17	ABC	0.1691	0.4182	3.954
L7	7	9	0.26	ABC	0.1691	0.4182	3.954
L8	9	10	0.14	ABC	0.1691	0.4182	3.954
L9	9	11	0.38	ABC	0.1691	0.4182	3.954
L10	11	12	0.56	ABC	0.1691	0.4182	3.954
L11	12	13	0.3	ABC	0.1691	0.4182	3.954
L12	12	14	3.33	ABC	0.1691	0.4182	3.954
L13	14	15	1.03	ABC	0.1691	0.4182	3.954
L14	16	17	1.08	ABC	0.1691	0.4182	3.954
L15	17	18	1.64	ABC	0.1691	0.4182	3.954
L16	18	19	0.47	ABC	0.1691	0.4182	3.954
L17	19	20	0.47	ABC	0.3481	0.4685	3.7571
L18	21	22	0.96	ABC	1.3919	0.4788	3.5971
L19	19	23	0.19	ABC	0.3481	0.4685	3.7571
L20	23	24	1.94	ABC	0.3481	0.4685	3.7571
L22	24	25	2.45	ABC	0.3481	0.4685	3.7571
L23	24	26	1.63	ABC	0.3481	0.4685	3.7571
L24+L26	26	27	1.2	ABC	0.5523	0.4852	3.6035
L27	26	28	2.12	ABC	0.3481	0.4685	3.7571
L28	28	29	0.73	ABC	0.5523	0.4852	3.6035
L29+L30	29	30	0.75	ABC	0.5523	0.4852	3.6035

Table C-4 Line data (Cont'd)

Name	From	To	Length (km)	Phase	R (Ω /km)	X (Ω /km)	Charge (µs/km)
L31	28	31	2.54	ABC	0.3481	0.4685	3.7571
L32	23	32	0.36	ABC	0.2765	0.4586	3.828
L33	32	33	0.26	ABC	0.2765	0.4586	3.828
L34	33	34	3.58	ABC	0.5523	0.4852	3.6035
L35	33	35	0.77	ABC	0.2765	0.4586	3.828
L36	35	36	2.08	ABC	0.3481	0.4685	3.7571
L37	35	37	4.51	ABC	0.2765	0.4586	3.828
L38	37	38	3.24	ABC	0.1691	0.4182	3.954
L39	38	39	0.3	ABC	0.1691	0.4182	3.954
L40	39	40	0.5	ABC	0.1691	0.4182	3.954

Table C-5 Lengths of line segments connecting different nodes in the feeder

Connected Nodes	1,4	4,15	15,23	23,38	23,31	38,40
Length of line in km	5.7	6.11	3.38	8.23	9.44	0.5

Table C-6 Load Data

Name	Node	P(A) (kW)	Q(A) (kVAR)	P(B) (kW)	Q(B) (kVar)	P(C) (kW)	Q(C) (kVar)
M1	4	2083.99	684.97	2083.99	684.97	2083.99	684.97
M2	4	50.66	16.65	50.66	16.65	50.66	16.65
M3	4	3.16	1.04	3.16	1.04	3.16	1.04
M4	6	62.64	35.50	62.64	35.50	62.64	35.50
M5	6	238.38	135.10	238.38	135.10	238.38	135.10
M6	8	1062.41	349.20	1062.41	349.20	1062.41	349.20
M7	10	192.00	169.33	192.00	169.33	192.00	169.33
M8	13	6.33	0.00	6.33	0.00	6.33	0.00
M9	14	6.65	2.19	6.65	2.19	6.65	2.19
M10	14	47.50	15.61	47.50	15.61	47.50	15.61
M11	14	53.84	17.70	53.84	17.70	53.84	17.70
M12	14	7.91	2.60	7.91	2.60	7.91	2.60
M13	22	15.84	5.21	15.84	5.21	15.84	5.21
M14	25	96.59	31.75	96.59	31.75	96.59	31.75
M15	27	50.66	16.65	50.66	16.65	50.66	16.65
M16	30	64.91	21.34	64.91	21.34	64.91	21.34
M17	31	47.50	15.61	47.50	15.61	47.50	15.61
M18	31	41.16	13.53	41.16	13.53	41.16	13.53
M19	31	34.84	11.45	34.84	11.45	34.84	11.45
M20	31	49.09	16.13	49.09	16.13	49.09	16.13
M21	34	68.09	22.38	68.09	22.38	68.09	22.38
M22	37	15.84	5.21	15.84	5.21	15.84	5.21
M23	37	19.00	6.24	19.00	6.24	19.00	6.24
M24	41	722.00	237.31	722.00	237.31	722.00	237.31
M25	36	26.91	8.85	26.91	8.85	26.91	8.85
M26	23	3.16	1.04	3.16	1.04	3.16	1.04

Appendix D

MAPE Calculated for the 22% data set

Table D-1 *MAPE* calculated for the active and reactive powers of the 22% data set

Season	Spring/Fall	Summer	Winter
Number of clusters: 22% of the full data set			
MAPE in P_{PV_freq}	***11.90***	***13.40***	***15.95***
MAPE in $P_{1\text{-}2\_freq}$	12.25	20.62	13.08
MAPE in $P_{19\text{-}23\_freq}$	18.92	18.18	15.76
MAPE in $P_{37\text{-}38\_freq}$	18.90	21.91	13.73
MAPE in $Q_{1\text{-}2\_freq}$	5.71	5.65	8.97
MAPE in $Q_{19\text{-}23\_freq}$	9.05	9.34	9.43
MAPE in $Q_{37\text{-}38\_freq}$	11.90	7.87	7.92
MAPE in $P_{1\text{-}2\_av}$	0.25	0.02	0.04
MAPE in $P_{19\text{-}23\_av}$	-8.81	0.27	1.50
MAPE in $P_{37\text{-}38\_av}$	-1.48	-0.60	-0.39
MAPE in $Q_{1\text{-}2\_av}$	0.02	0.01	0
MAPE in $Q_{19\text{-}23\_av}$	0.03	0	0.02
MAPE in $Q_{37\text{-}38\_av}$	0.05	0	0.02
MAPE in $P_{1\text{-}2\_max}$	0	0	0
MAPE in $P_{19\text{-}23\_max}$	0	0	0
MAPE in $P_{37\text{-}38\_max}$	0	0	0
MAPE in $Q_{1\text{-}2\_max}$	0	0	0
MAPE in $Q_{19\text{-}23\_max}$	1.77	1.88	0
MAPE in $Q_{37\text{-}38\_max}$	1.97	1.99	0

Table D-2 *MAPE* calculated for the voltages of the 22% data set

Season	Spring/Fall	Summer	Winter
Number of clusters: 22% of the full data set			
MAPE in P_{PV_freq}	***11.90***	***13.40***	***15.95***
MAPE in V_{4_freq}	0	0	0
MAPE in V_{23_freq}	0	0.23	0
MAPE in V_{38_freq}	0	0.55	1.01
MAPE in V_{41_freq}	1.50	9.86	0.28
MAPE in V_{4_av}	0	0	0
MAPE in V_{23_av}	0.01	0	0
MAPE in V_{38_av}	0.01	0	0
MAPE in V_{41_av}	0.01	0	0
MAPE in V_{4_max}	0	0	0
MAPE in V_{23_max}	0.02	0	0.01
MAPE in V_{38_max}	0.10	0.04	0
MAPE in V_{41_max}	0.11	0	0

Appendix E

Comparison Between Different Storage Technologies

Table E-1 Comparison between storage technologies

[131][132][133][135][145][146][135][147][148][149] [154][155]

Technology	Lifetime	Order of Cycle Lifetime	Cycle Efficiency	Time specifications	Advantage	Disadvantage	Suitability for PV systems
Pumped Hydro	50 years	-	70% - 85%	Access: few minutes Operation: few hours	- Well established technology - Effective in shifting the peak time of generation	- Used only for large applications - Requires large land area	Low
Flywheel	20 years	10^6 cycles	80% - 85%	Access: milliseconds Operation: few minutes to an hour	- Low maintenance - Long life - High efficiency	- Low energy density for conventional flywheels - High cost - Safety issues due to high rotational speed	Medium
SMES	20 years	10^3 cycles	95%	Access: milliseconds Operation: few seconds	- Very fast response - Very high efficiency	- Operates at low temperatures - Low energy density - Hazards due to magnetic fields	Low
Ultracaps	10 years	10^6 cycles	90%	Access: milliseconds Operation: few minutes	- long cycle life - High power density	- Low energy density - Expensive	Low-Medium

Table E-1 **Comparison between storage technologies (cont'd)**

Technology	**Lifetime**	**Order of Cycle Lifetime**	**Cycle Efficiency**	**Time specifications**	**Advantage**	**Disadvantage**	**Suitability for PV systems**
Lead-Acid Batteries	5 years	10^3 cycles	75% - 85%	Access: milliseconds Operation: minutes to few hours	- Cost Effective - Mature Technology	- Low lifetime - Frequent maintenance required	Medium
Lithium-Ion Batteries	6-10 years	10^3 cycles	95%	Access: milliseconds Operation: minutes to several hours	- High Efficiency - High energy density - Long cycle life	- Expensive - Safety issues	Medium
Sodium-Sulphur Batteries	15 years	10^3 cycles	90%	Access: milliseconds Operation: minutes to several hours	- High Efficiency - High energy density - Long cycle life when deeply discharged - Low maintenance	- Relatively expensive - Requires heaters - Environmentally hazardous materials	High
Flow Batteries	10 years	10^3 cycles	70% - 80%	Access: milliseconds Operation: minutes to several hours	- Low maintenance - Independent power and energy specifications	- Low energy density as compared to NaS batteries	High

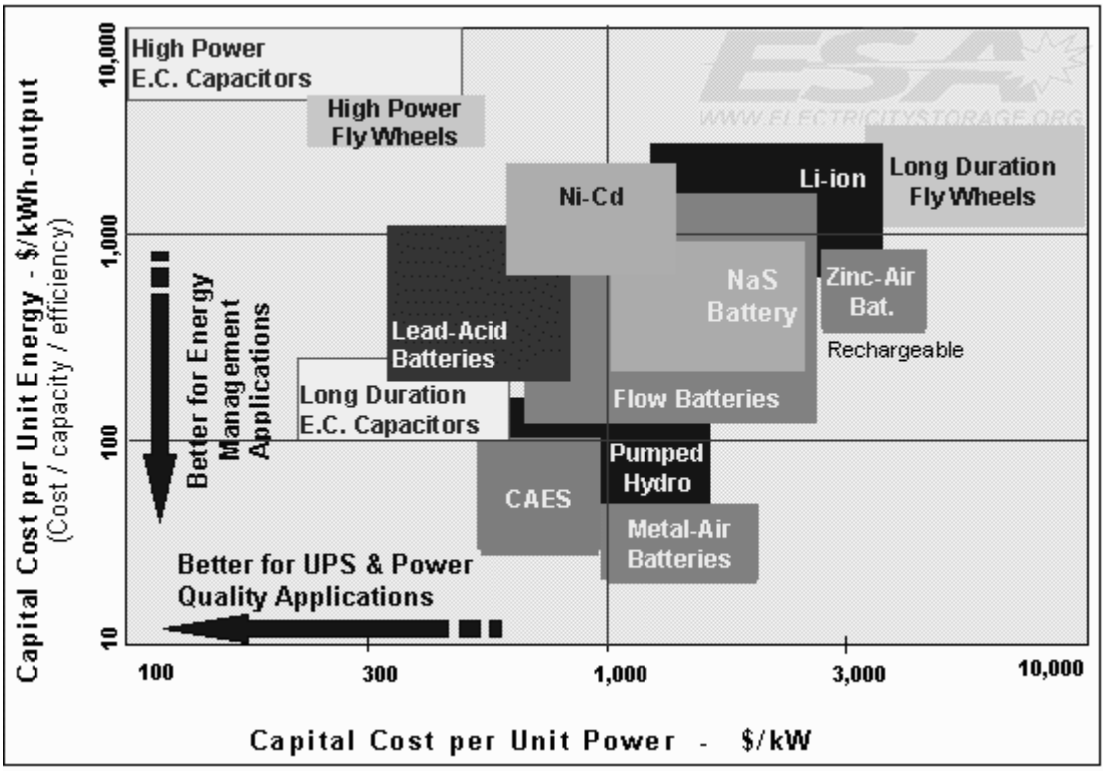

Figure E-1 Capital costs of different storage technologies [146]

References

[1] "International Energy Outlook, 2008", Energy Information Administration, Available online: http://www.eia.doe.gov/oiaf/ieo/index.html. Retrieved February 2009.

[2] "Meeting Future Electricity Demand Will Increase Emissions of Some Harmful Substances", Report to Congressional Requesters prepared by the United States General Accounting Office, October 2002, available online: http://www.gao.gov/new.items/d0349.pdf. Retrieved August 2009

[3] "Unleashing the Potential of On-Grid Photovoltaics in Canada. An Action Plan to make PV an Integral Component of Canada's Energy Future," Report prepared for Industry Canada by The Delphi Group, 2003. Available online: http://www.ic.gc.ca/epic/site/rei-ier.nsf/en/h_nz00017e.html. Retrieved February 2009

[4] "Trends in Photovoltaic Applications. Survey report of selected IEA countries between 1992 and 2007," Available online at: http://www.iea-pvps.org. Retrieved February 2009.

[5] E. Vartiainen, "A new approach to estimating the diffuse irradiance on inclined surfaces," Renewable Energy, Vol. 20, 1999.

[6] Volker Quaschning "The Sun as an Energy Resouce, Technology Fundamentals" Renewable Energy World magazine, May 2003, pp. 90-93.

[7] S. Roy, "Optimal planning for utility generation by photovoltaic sources spread across multiple sites," IEEE Transactions on Energy Conversion,vol.21, no.1, pp. 181-186, March 2006.

[8] N. Kawasakia, T. Oozekia, K. Otania, K. Kurokawaa, "An evaluation method of the fluctuation characteristics of photovoltaic systems by using frequency analysis," Solar Energy Materials & Solar Cells Vol. 90, 2006.

[9] A. Woyte, R. Belmans, J. Nijs, "Fluctuations in instantaneous clearness index: Analysis and statistics," Solar Energy, Vol 81, 2007.

[10] J. Ehnberg, M. Bollen, "Simulation of global solar radiation based on cloud observations," Solar Energy, Vol 78, 2005.

[11] J. Polo, L. Zarzalejo, L. Ramirez, B. Espinar "Iterative filtering of ground data for qualifying statistical models for solar irradiance estimation from satellite data," Solar Energy, Vol. 80, 2006.

[12] F. Vignola, P. Harlan, R. Perez, M. Kmiecik, "Analysis of satellite derived beam and global solar radiation data," Solar Energy, Vol 81, 2007.

[13] A. Moreno-Munoz, J. de la Rosa, R. Posadillo, V. Pallares, "Short term forecasting of solar radiation," IEEE International Symposium on Industrial Electronics, 2008. ISIE 2008.

[14] E. Crispim, P. Ferreira, A. Ruano, "Solar radiation prediction using RBF Neural Networks and cloudiness indices," International Joint Conference on Neural Networks, 2006. IJCNN '06.

[15] A. Mellit, M. Menghanem, M. Bendekhis, "Artificial neural network model for prediction solar radiation data: application for sizing stand-alone photovoltaic power system," IEEE Power Engineering Society General Meeting, 2005.

[16] D. M. Chapin, C. S. Fuller, and G. L. Pearson, "A New Silicon p-n Junction Photocell for Converting Solar Radiation into Electrical Power," Journal of Applied Physics, Vol. 25, Issue 5, May 1954.

[17] A. Goetzberger, and V. U. Hoffmann, "Photovoltaic Solar Energy Generation," Springer Series in Applied Science, 2005.

[18] R. A. Messenger, and J. Ventre, "Photovoltaic Systems Engineering," 2nd editiob, CRC press, 2004.

[19] http://www.solarbuzz.com/moduleprices.htm. Retrieved February 2009.

[20] J. Szlufcik, S. Sivoththaman, J.F. Nlis, R.P. Mertens, R. Van Overstraeten, "Low-cost industrial technologies of crystalline silicon solar cells ," Proceedings of the IEEE , vol.85, no.5, pp.711-730, May 1997

[21] A. Breeze, "Next generation thin-film solar cells," IEEE International Reliability Physics Symposium, 2008. IRPS 2008.

[22] Hans. Rauschenbach, "Solar Cell Array Design Handbook," Van Nostrand Reinhold, Ltd., 1980.

[23] J. W. Bishop, "Computer simulation of the effects of electrical mismatches in the photovoltaic cell interconnection circuits," Solar Cells 1988, Vol. 25, No 1, Pages: 73 – 89.

[24] J. A. Duffie, W. A. Beckman, "Solar Engineering of Thermal Processes," second edition John Wiley & Sons Inc., New York. 1991.

[25] J.A.Gow, C.D.Manning, "Development of a photovoltaic array model for use in power-electronics simulation studies," IEE Proceedings- Electric Power Applications, Vol. 146, Issue 2, March 1999.

[26] E. van Dyk, E. Meyer, "Analysis of the effect of parasitic resistances on the performance of photovoltaic modules," Renewable Energy, Vol 29, 2004.

[27] W. Xiao, W. Dunford, A. Capel, "A Novel Modeling Method for Photovoltaic Cells," IEEE 35th Annual Power Electronics Specialists Conference, 2004. PESC 04.

[28] W. De Soto, S.A. Klein, W.A. Beckman, "Improvement and validation of a model of PV array performance," Solar Energy Journal, Vol 80, No 1, 2006.

[29] Liu Shengyi, R.A. Dougal, "Dynamic multiphysics model for solar array," IEEE Transactions on Energy Conversion, Vol. 17, Issue 2, June 2002.

[30] M. AbdulHadi, A. Al-Ibrahim, G. Virk, "Neuro-Fuzzy-Based Solar Cell Model," IEEE Transactions on Energy Conversion, Vol. 19, no. 3, September 2004.

[31] W. Xiao, M. Lind, W. Dunford, A. Capel "Real-Time Identification of Optimal Operating Points in Photovoltaic Power Systems," IEEE Transactions On Industrial Electronics, Vol. 53, No. 4, August 2006.

[32] J. V. Paatero, and P. D. Lund, "Effects of large-scale photovoltaic power integration on electricity distribution networks," Renewable Energy Journal, Volume 32, Issue 2, February 2007, Pages 216-234

[33] M. Veerachary, "PSIM Circuit-Oriented Simulator Model for the Nonlinear Photovoltaic Sources," IEEE Transactions on Aerospace and Electronic Systems, Volume 42, Issue 2, April 2006.

[34] M. E. Ropp, M. Begovic, A. Rohatgi, "Determination of the curvature derating factor for the Georgia Tech Aquatic Center photovoltaic array," Twenty-Sixth IEEE Photovoltaic Specialists Conference, 1997.

[35] S.B. Kjaer, J.K. Pedersen, F. Blaabjerg, "A review of single-phase grid-connected inverters for photovoltaic modules" IEEE Transactions on Industry Applications, Vol. 41, Issue 5, 2005.

[36] Y. Xue; L. Chang; S. B. Kjaer, J. Bordonau, T. Shimizu, "Topologies of single-phase inverters for small distributed power generators: an overview," IEEE Transactions on Power Electronics, , Vol.19, 2004.

[37] G.R. Walker, P.C. Sernia, "Cascaded DC-DC converter connection of photovoltaic modules," IEEE Transactions on Power Electronics, Vol.19, No.4, 2004.

[38] C. Cecati, A. Dell'Aquila, M. Liserre, "A Novel Three-Phase Single-Stage Distributed Power Inverter," IEEE Transactions On Power Electronics, Vol. 19, No. 5, September 2004.

[39] B. Ho, H. Chung, "An Integrated Inverter With Maximum Power Tracking for Grid-Connected PV Systems," IEEE Transactions On Power Electronics, Vol. 20, No. 4, July 2005.

[40] N. Femia, G. Petrone, G. Spagnuolo, M. Vitelli "Optimization of Perturb and Observe Maximum Power Point Tracking Method," IEEE Transactions On Power Electronics, Vol. 20, No. 4, July 2005.

[41] J. Park, J. Ahn, B. Cho, G. Yu, "Dual-Module-Based Maximum Power Point Tracking Control of Photovoltaic Systems," IEEE Transactions On Industrial Electronics, Vol. 53, No. 4, 2006.

[42] N. Mutoh, M. Ohno, T. Inoue "A Method for MPPT Control While Searching for Parameters Corresponding to Weather Conditions for PV Generation Systems," IEEE Transactions On Industrial Electronics, Vol. 53, No. 4, August 2006.

[43] T. Esram, P.L. Chapman, "Comparison of Photovoltaic Array Maximum Power Point Tracking Techniques," IEEE Transaction on Energy Conversion, vol.22, no.2, pp.439-449, June 2007.

[44] T. Shimizu, M. Hirakata, T. Kamezawa, H. Watanabe, "Generation control circuit for photovoltaic modules," IEEE Transactions on Power Electronics, vol.16, no.3, pp.293-300, May 2001.

[45] K. Kobayashi, I. Takano, Y. Sawada, "A study on a two stage maximum power point tracking control of a photovoltaic system under partially shaded insolation conditions," IEEE Power Engineering Society General Meeting, 2003, vol.4, no., pp.-2617 Vol. 4, 13-17 July 2003.

[46] S. Busquets-Monge, J. Rocabert, P. Rodriguez, S. Alepuz, J. Bordonau, "Multilevel Diode-Clamped Converter for Photovoltaic Generators With Independent Voltage Control of Each Solar Array," IEEE Transactions on Industrial Electronics, , vol.55, no.7, pp.2713-2723, July 2008.

[47] F. Blaabjerg, R. Teodorescu, M. Liserre, A.V. Timbus, "Overview of Control and Grid Synchronization for Distributed Power Generation Systems," IEEE Transactions on Industrial Electronics, vol.53, no.5, pp.1398-1409, Oct. 2006.

[48] J. Kwon, K. Nam, B. Kwon, "Photovoltaic Power Conditioning System With Line Connection," IEEE Transactions On Industrial Electronics, Vol. 53, No. 4, August 2006.

[49] B. Blazic, I. Papic, "Advanced control of a converter used for connection of photovoltaic modules," IEEE Power Engineering Society General Meeting, 2006.

[50] "IEEE Recommended Practice for Utility Interface of Photovoltaic (PV) Systems," IEEE Std 929-2000.

[51] "IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems," IEEE Std 1547-2003 , vol., no., pp. 0_1-16, 2003.

[52] H. Zeineldine, "Distributed generation micro-grid operation: control, protection and electricity market operation," Ph.D. dissertation, Dept. Electrical Engineering, University of Waterloo, Canada, 2006.

[53] M. Ropp, M. Begovic, A. Rohatgi, G. Kern, R. Bonn, S. Gonzalez "Determining the Relative Effectiveness of Islanding Detection Methods Using Phase Criteria and Nondetection Zones," IEEE Transactions On Energy Conversion, Vol. 15, No. 3, September 2000.

[54] L. Lopes, H. Sun "Performance Assessment of Active Frequency Drifting Islanding Detection Methods," IEEE Transactions On Energy Conversion, Vol. 21, No. 1, March 2006.

[55] A. Woyte, R. Belmans, J. Nijs, "Testing the Islanding Protection Function of Photovoltaic Inverters," IEEE Transactions On Energy Conversion, Vol. 18, No. 1, March 2003.

[56] L. Lopes, Z. Yongzheng "Islanding Detection Assessment of Multi-Inverter Systems With Active Frequency Drifting Methods," IEEE Transactions on Power Delivery, Vol. 23, Issue 1, Jan. 2008.

[57] S. Alepuz, S. Busquets-Monge, J. Bordonau, J. Gago, D. Gonzalez, J. Balcells, "Interfacing Renewable Energy Sources to the Utility Grid Using a Three-Level Inverter," IEEE Transactions on Industrial Electronics, vol.53, no.5, pp.1504-1511, Oct. 2006.

[58] T.F. Wu, C.H. Chang, Y.K. Chen, "A multi-function photovoltaic power supply system with grid-connection and power factor correction features," IEEE 31st Annual Power Electronics Specialists Conference, 2000. PESC 2000, vol.3, no., pp.1185-1190 vol.3, 2000

[59] M.I. Marei, E.F. El-Saadany, M.M.A. Salama, "A novel control algorithm for the DG interface to mitigate power quality problems," IEEE Transactions on Power Delivery, vol.19, no.3, pp. 1384-1392, July 2004.

[60] Wu Libo; Zhao Zhengming; Liu Jianzheng, "A Single-Stage Three-Phase Grid-Connected Photovoltaic System With Modified MPPT Method and Reactive Power Compensation," IEEE Transactions On Energy Conversion, , Vol.22, No.4, 2007.

[61] C.V Nayar,.; M. Ashari,; W.W.L. Keerthipala, “A grid-interactive photovoltaic uninterruptible power supply system using battery storage and a back up diesel generator,” IEEE Transaction on Energy Conversion, vol.15, no.3, pp.348-353, Sep 2000.

[62] J. Carrasco, L. Franquelo, J. Bialasiewicz, E. Galván, R. Guisado, M. Prats, J. León, N. Moreno-Alfonso, "Power-Electronic Systems for the Grid Integration of Renewable Energy Sources: A Survey," IEEE Transactions On Industrial Electronics, Vol. 53, No. 4, August 2006.

[63] "Innovative Electrical concepts," Report IEA – PVPS T7-7: 2001. Available online: http://www.iea-pvps.org/ Retrieved: August 2009

[64] A. Pregelj, M. Begovic, A. Rohatgi, "Impact of inverter configuration on PV system reliability and energy production," Twenty-Ninth IEEE Photovoltaic Specialists Conference, 2002.

[65] M. Calais, J. Myrzik, T. Spooner, V. Agelidis, "Inverters for single-phase grid connected photovoltaic systems-an overview," IEEE 33rd Annual Power Electronics Specialists Conference, 2002.

[66] J. Myrzik, M. Calais, "String and module integrated inverters for single-phase grid connected photovoltaic systems - a review," IEEE Bologna Power Tech Conference Proceedings, 2003, vol.2.

[67] Walid A. Omran, Mehrdad Kazerani, and M. M. A. Salama, "Impacts of Large Grid-Connected PV Systems," 2009 CIGRE Canada conference on Power Systems, (October 4-6, 2009), Toronto, Canada.

[68] "Revised FIT Program Rules, Standard Definitions and Price Schedule - Draft July 10, 2009," Available online: http://www.powerauthority.on.ca/fit/. Retrieved August 2009.

[69] B. H. Chowdhury, and A. W. Sawab, "Evaluating the Value of Distributed Photovo1aic Generations in Radial Distribution Systems," IEEE Transactions on Energy Conversion, Vol. 11, No. 3, September 1996.

[70] T. Hoff ,and D.S. Shugar, "The Value of Grid-Support Photovoltaics in Reducing Distribution System Losses," IEEE Transactions on Energy Conversion, Vol. 10, No. 3, September 1995.

[71] T. Hoff, H. J. Wenger, and B. K. Farmer, "The Value of Grid-Support Photovoltaics in Providing Distribution System Voltage Support," in Proc. American Solar Energy Society Annual Conf., San Jose, CA, 1994.

[72] M. M. El-Gasseir, K. P. Alteneder, J. Bigger, "Enhancing Transformer Dynamic Rating through Grid Application of Photovoltaic Arrays," Proceedings of the 23rd IEEE PV Specialists Conference, May 1993.

[73] W. T. Jewell, R. Ramakumar, "The effects of moving clouds on electric utilities with dispersed PV generation," IEEE Transactions on Energy Conversion, Vol EC-2, Issue 4, Dec. 1987.

[74] J. Boland, M. Dik, "The level of complexity needed for weather data in models of solar system performance," Solar Energy, Vol. 71, issue 31, 2001.

[75] J. F. Jockell, S Rahman, "Application of High Resolution Insolation Data for Photovoltaic System Design Analysis," IEEE Proceedings Southeastcon '89. 'Energy and Information Technologies in the Southeast'. 9-12 April 1989, Page(s):1430 – 1435, vol.3.

[76] R. Gansler, S. Klein, W. Beckman, "Investigation of minute radiation data," Solar Energy Vol. 55, 1995.

[77] G. Vijayakumar, M. Kummert, S. Klein, W. Beckman, "Analysis of short-term solar radiation data". Solar Energy 79 (2005) pp. 495–504.

[78] C. Craggs, E. M. Conway and N. M. Pearsall, "Statistical investigation of the optimal averaging time for solar irradiance on horizontal and vertical surfaces in the UK". Solar Energy 68 (2000) pp. 179–187.

[79] S. M. Chalmers, M.M. Hitt, J.T. Underhill, P.M. Anderson, P.L. Vogt, R. Ingersoll, "The effect of PV power generation on utility operation," IEEE Transactions on Power Apparatus and Systems Volume PAS-104, Issue 3, March 1985.

[80] W. T. Jewell, T.D. Unruh, "Limits on cloud induced fluctuation in PV generation," IEEE Transactions on Energy Conversion, Volume 5, Issue 1, March 1990.

[81] S. Rahman, M Bouzguenda, "A model to determine the degree of penetration and energy cost of large-scale utility interactive photovoltaic systems". IEEE Transactions on Energy Conversion, Vol 9, Issue 2, Jun. 1994.

[82] B. H. Chowdhury, and S. Rahman "Is central station photovoltaic power dispatchable," IEEE Transactions on Energy Conversion, Vol. 3(4), 747-754, Dec. 1988.

[83] W. T. Jewell, R. Ramakumar, S.R. Hill, "A study of dispersed PV generation on the PSO system". IEEE Transactions on Energy Conversion, Vol 3, Issue 3, Sep. 1988.

[84] Yun Tiam Tan, and Daniel S Kirschen, “Impact on the Power System of a Large Penetration of Photovoltaic Generation,” IEEE Power Engineering Society General Meeting, 2007. 24-28 June 2007 Page(s):1 – 8.

[85] G. A.Vokas, A. V.Machias, “Harmonic voltages and currents on two Greek islands with photovoltaic stations: study and field measurements,” IEEE Transaction on Energy Conversion, vol.10, no.2, pp.302-306, Jun 1995.

[86] E. Vasanasong, E. D.Spooner, “The prediction of net harmonic currents produced by large numbers of residential PV inverters: Sydney Olympic Village case study,” Proceedings. Ninth International Conference on Harmonics and Quality of Power, 2000. vol.1, no., pp.116-121 vol.1, 2000.

[87] A. R. Oliva, J. C.Balda, “A PV dispersed generator: a power quality analysis within the IEEE 519,” IEEE Transactions on Power Delivery, vol.18, no.2, pp. 525-530, April 2003.

[88] J.H.R. Enslin, P.J.M. Heskes, “Harmonic interaction between a large number of distributed power inverters and the distribution network,” IEEE Transactions on Power Electronics, vol.19, no.6, pp. 1586-1593, Nov. 2004.

[89] A. Kotsopoulos, P.J.M. Heskes, M.J. Jansen, “Zero-crossing distortion in grid-connected PV inverters,” IEEE Transactions on Industrial Electronics, vol.52, no.2, pp. 558-565, April 2005.

[90] E. C. Kern Jr., E. M. Gulachenski, G. A. Kern, “Cloud effects on distributed photovoltaic generation: slow transients at the Gardner, Massachusetts photovoltaic experiment,” Experiment," IEEE Transactions on Energy Conversion, Vol. 4, pp. 184-190, 1989.

[91] A. Gross, J. Bogensperger, and D. Thyr, “Impacts of large scale photovoltaic systems on the low voltage network,” Solar Energy, Vol. 59, Issue 4-6, 1997.

[92] S. Conti, S. Raiti, G. Tina, and U. Vagliasindi, “Study of the Impact of PV Generation on Voltage Profile in LV Distribution Networks,” Proc. IEEE Power Tech, Porto, Portugal, Vol. 4, Sept. 10-13, 2001.

[93] A. Canova, L. Giaccone, F. Spertino, M. Tartaglia, "Electrical Impact of Photovoltaic Plant in Distributed Network," IEEE Transactions on Industry Applications, vol.45, no.1, Jan.-Feb. 2009.

[94] B. H. Chowdhury, "Effect of central station photovoltaic plant on power system security," Proc. of 21st IEEE Photovoltaic Specialist Conference, Kissimmee, FL, May, 1990.

[95] Yahia Baghzouz, "Voltage Regulation and Overcurrent Protection Issues in Distribution Feeders with Distributed Generation – A Case Study," Proceedings of the 38th Hawaii International Conference on System Sciences – 2005.

[96] V. H. M. Quezada, J. R. Abbad, and T. G. San Román, "Assessment of Energy Distribution Losses for Increasing Penetration of Distributed Generation," IEEE Transactions on Power Systems, VOL. 21, NO. 2, MAY 2006.

[97] I. Abouzahr, and R. Ramakumar, "An Approach to Assess the Performance of Utility-Interactive Photovoltaic Systems," IEEE Transactions on Energy Conversion, Vol. 8, No. 2, June 1993.

[98] Aleksandar Pregelj, Miroslav Begovic, and Ajeet Rohatgi, "Quantitative Techniques for analysis of large data sets in renewable DG," IEEE Transactions on Power Systems, Vol. 19, No. 3, August 2004.

[99] F. A. Viawan, F. Vuinovich, and A. Sannino, "Probabilistic Approach to the Design of Photovoltaic Distributed Generation in Low Voltage Feeder," 9th International Conference on Probabilistic Methods Applied to Power Systems, Stockholm, Sweden - June 11-15, 2006.

[100] Badrul H. Chowdhury, "Optimizing the integration of PV systems with electrical utilities," IEEE Transactions on Energy Conversion, vol. 7, No. 1, March 1992.

[101] A. Woyte, V. Van Thong, R Belmans, J. Nijs, "Voltage Fluctuations on Distribution Level Introduced by Photovoltaic Systems". IEEE Transactions on Energy Conversion, Vol 21, Issue 1, Mar. 2006.

[102] Walid A. Omran, Mehrdad Kazerani, and M. M. A. Salama, "A Clustering-Baed Method for Quantifying the Effects of Large On-Grid PV Systems", Accepted for publication in the IEEE Transactions on Power Delivery.

[103] M. Iqbal, "An introduction to solar radiation," Academic Press Canada, 1983.

[104] T. Muneer, "Solar radiation and daylight models," Elsevier, Oxford, 2004.

[105] G. Notton, C. Cristofari, M. Muselli, P. Poggi, "Performance evaluation of various hourly slope irradiation models using Mediterranean experimental data of Ajaccio," Energy Conversion and Management, Vol. 47, 2006.

[106] T.M. Klucher, "Evaluation of models to predict insolation on tilted surfaces," Solar Energy, Vol. 23, 1979.

[107] Solar Radiation Research Laboratory (BMS) available online at: http://www.nrel.gov/midc/srrl_bms. Retrieved August 2009.

[108] E.I. Ortiz-Rivera, F.Z. Peng, "Analytical Model for a Photovoltaic Module using the Electrical Characteristics provided by the Manufacturer Data Sheet," IEEE 36th Power Electronics Specialists Conference, 2005. PESC '05, vol., no., pp. 2087-2091, 11-14 Sept. 2005.

[109] V.H.M. Quezada, J.R. Abbad, T.G.S. Roman, "Assessment of energy distribution losses for increasing penetration of distributed generation," IEEE Transactions on Power Systems, vol.21, no.2, May 2006.

[110] Hans. Rauschenbach, "Solar Cell Array Design Handbook," Van Nostrand Reinhold, Ltd., 1980.

[111] PC1D Ver 5.6, 2001, University of New South Wales, Australia, 2001.

[112] http://www.solarserver.de/solarmagazin/anlage_0606_e.html. Retrieved, September 2009.

[113] A. Preglej, "Impact of Distributed Generation on Power Network Operation," Ph.D. dissertation, Dept. Electrical and Computer Eng., Georgia Institute of Technology, 2003.

[114] I. T. Jolliffe, "Principal Component Analysis". Springer-Verlag New york, Inc., Second Edition, 2002.

[115] C. Chatfield & A. J. Collins, "Introduction to Multivariate Analysis". Chapman and Hall, New York, 1980.

[116] Lindsay I. Smith, " A tutorial on Principal Component Analysis", February 26, 2002. http://csnet.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf

[117] T. Babnik, R. Aggarwal, P. Moore, "Data mining on a transformer partial discharge data using the self-organizing map," IEEE Transactions on Dielectrics and Electrical Insulation, Volume 14, Issue 2, April 2007 Page(s):444 – 452.

[118] K.K. Anaparthi, B. Chaudhuri, N.F. Thornhill, B.C. Pal, "Coherency Identification in Power Systems Through Principal Component Analysis," IEEE Transactions on Power Systems, Volume 20, Issue 3, Aug. 2005 Page(s):1658 – 1660.

[119] G. Chicco, R. Napoli, F. Piglione, "Comparisons among clustering techniques for electricity customer classification," IEEE Transactions on Power Systems, Volume 21, Issue 2, May 2006 Page(s):933 – 940.

[120] A.K. Jain; M.N. Murty; P.J. Flynn, "Data clustering: A review". ACM Computing Surveys, Vol. 31, No. 3, September 1999.

[121] P. Tan, M. Steinbach, V. Kumar, "Introduction to Data Mining," Pearson Education Inc., 2006.

[122] A.K. Jain, R. C. Dubes, “Algorithms for clustering data,” Printece Hall Advanced References series, New Jersy, 1988.

[123] M. Halkidi, Y. Batistakis, M. Vazirgiannis, "Cluster validity methods: Part I," SIGMOD Record 31 (2), 2002.

[124] P.J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," Journal of Computational and Applied Mathematics (20), 1987.

[125] D. Davies, D. Bouldin, "A cluster separation measure," IEEE Transactions on Pattern Analysis and Machine Intelligence. 1 (4), 1979.

[126] A. Bensaid, L. O. Hall, J. Bezdek, L. P. Clarke, M. L. Silbiger, J. A. Arrington, and R. F. Murtagh, “Validity-guided (Re)clustering for image segmentation,” IEEE Transactions on Fuzzy Systems, vol. 4, May 1996.

[127] Walid A. Omran, M. Kazerani, and M. M. A. Salama, “Evaluation of Clustering Algorithms used for Studying the Performance of Electric Networks in the Presence of Large PV Systems”, Submitted to IEEE Transactions on Power Systems.

[128] H. Sadaat, "Power System Analysis", Second Edition, McGraw-Hill Higher Education, 2002.

[129] Walid A. Omran, Mehrdad Kazerani, and M. M. A. Salama, “A Study of the Impacts of Power Fluctuations Generated from Large PV Systems”, 2009 IEEE PES/IAS Conference on Sustainable Alternative Energy (Sept 28 – 30, 2009), Valencia, Spain.

[130] Walid A. Omran, M. Kazerani, and M. M. A. Salama, “Investigation of methods for Reduction of Power Fluctuation of Large Grid-Connected PV Systems," submitted to IEEE Transactions on Energy Conversion.

[131] EPRI Technical Report, "EPRI-DOE Handbook of Energy Storage for Transmission and Distribution Applications," 2003.

[132] E. Spahic, G. Balzer, B. Hellmich, W.Munch, , "Wind Energy Storages - Possibilities," IEEE Power Tech, 2007.

[133] M. Glavin, W. Hurley, "Ultracapacitor/ battery hybrid for solar energy storage," 42nd International Universities Power Engineering Conference, 2007.

[134] J. Carrasco, L. Franquelo, J. Bialasiewicz, E. Galván, R. Guisado, M. Prats, J. León, N. Moreno-Alfonso, "Power-Electronic Systems for the Grid Integration of Renewable

Energy Sources: A Survey," IEEE Transactions On Industrial Electronics, Vol. 53, No. 4, August 2006.

[135] S. Choi, K. Tseng, D. Vilathgamuwa, T. Nguyen, "Energy storage systems in distributed generation schemes," 2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008.

[136] H.A.M. Maghraby, M.H. Shwehdi, G.K. Al-Bassam, "Probabilistic assessment of photovoltaic (PV) generation systems," IEEE Transactions on Power Systems, , vol.17, no.1, pp.205-208, Feb 2002.

[137] Bagen & R. Billinton, "Incorporating well-being considerations in generating systems using energy storage," IEEE Transactions on Energy Conversion, vol.20, no.1, March 2005.

[138] M. Kolhe, "Techno-Economic Optimum Sizing of a Stand-Alone Solar Photovoltaic System," IEEE Transactions on Energy Conversion, vol.24, no.2, pp.511-519, June 2009.

[139] C.V. Nayar, M. Ashari, W.W.L. Keerthipala, "A grid-interactive photovoltaic uninterruptible power supply system using battery storage and a back up diesel generator," IEEE Transactions on Energy Conversion, vol.15, no.3, Sep 2000.

[140] T. E. Hoff, R. Perez, R. M. Margolis, "Maximizing the value of customer-sited PV systems using storage and controls," Solar Energy 81 (2007).

[141] Wei-Fu Su; Shyh-Jier Huang; Chin-E Lin, "Economic analysis for demand-side hybrid photovoltaic and battery energy storage system," IEEE Transactions on Industry Applications, vol.37, no.1, Jan/Feb 2001.

[142] P. Poonpun, W.T. Jewell, "Analysis of the Cost per Kilowatt Hour to Store Electricity," IEEE Transactions on Energy Conversion, vol.23, no.2, June 2008.

[143] Bo Lu, M. Shahidehpour, "Short-term scheduling of battery in a grid-connected PV/battery system," IEEE Transactions on Power Systems, vol.20, no.2, May 2005.

[144] D. Ton, C. Hanley, G. Peek, and J. Boyes, "Solar Energy Grid Integration Systems – Energy Storage (SEGIS-ES)," SANDIA Report, SAND2008-4247, Unlimited Release, July 2008.

[145] M. R. Patel, "Wind and Solar Power Systems," Boca Raton, FL: CRC, 1999.

[146] Electricity Storage Association: http://electricitystorage.org

[147] J. McDowall, "Status and Outlook of the Energy Storage Market," IEEE Power Engineering Society General Meeting, 2007.

[148] G. Corey, "Batteries for stationary standby and for stationary cycling applications part 6: alternative electricity storage technologies," IEEE Power Engineering Society General Meeting, 2003.

[149] "California ISO integration of renewable resources report (Draft)," August, 2007. Available online: http://www.caiso.com

[150] G. J. Thuesen, W. J. Fabrycky, "Engineering Economy, Eighth Edition" Prentice-Hall International Series in Industrial and Systems Engineering, 1993.

[151] GAMS Distribution 23.0, "A user's guide," GAMS Development Corporation, 2009.

[152] Ali Nourai, "Installation of the First Distributed Energy Storage System (DESS) at American Electric Power (AEP). A Study for the DOE Energy Storage Systems Program" SANDIA REPORT, SAND2007-3580, Unlimited Release, 2007.

[153] D.P Jenkins, J. Fletcher, D. Kane, "Lifetime prediction and sizing of lead-acid batteries for micro-generation storage applications," IET Renewable Power Generation, vol.2, no.3, 2008.

[154] Bent Sørensen, "Renewable Energy: Its physics, engineering, use, environmental impacts, economy and planning aspects," Third Edition. Elsevier Science, 2004.

[155] R. Schainker, "Executive overview: energy storage options for a sustainable energy future," IEEE Power Engineering Society General Meeting, 2004.