Modeling of Renewable Resources in Distribution System Planning and Operation

by

Majed Alotaibi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2014

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

In recent decades, interest in placing renewable resources in conventional power systems has increased because of their ability to reduce fossil fuel consumption, which leads to the preservation of the environment. The rapid increase in employing these renewable resource-based DGs drives the system to be more dynamic, and causes many obstacles that need to be overcome. Power system planners and operators should look at the distribution system from another angle, taking into consideration the intermittent behavior of most renewable resources. Furthermore, solid models that are able to handle the uncertainty in generation levels are required.

This thesis presents a comprehensive probabilistic model for representing renewable energy resources in long term planning problems. This model utilized large historical data sets, grouping technique, and statistical analysis in order to handle the fluctuations that are caused by the variations in wind speed or solar irradiance.

In this research, renewable resources (wind and PV based DGs) ae well as dispatchable units are optimally allocated and sized using a probabilistic optimization model. This model incorporates the intermittent nature of wind speed and solar radiation into the deterministic optimal power flow equations. The variability from the load side and the uncertainty from the feeding side are considered. Genetic algorithm is used in order to minimize the annual energy losses of a distribution system.

This thesis proposes a new iterative-based optimization algorithm is proposed in order to determine the minimum number of states that can precisely describe or represent the behavior of wind speed and solar irradiance in operational planning problems. This algorithm is evaluated using a power system planning problem. The proposed algorithm takes into account the annual energy losses and the total DG penetration level and considers them as an indication of how far the proposed method's outcomes are from the actual results. Three different data groupings are applied (hourly, seasonally, and yearly) to investigate the variety of weather and electricity demands on the proposed method. The obtained results should be maintained within an acceptable limit of error which is in this thesis, 2.5%, and any violation of this limit will interrupt the algorithm sequences.

The importance of this method actually lies in its ability to reduce the complexity in reliability analysis such that the number of overall system states will be minimized when the analytical evaluation methods are utilized.

Acknowledgements

All praises are due to Allah for providing me the guidance, strength, and patience to accomplish this thesis work and my MASc degree successfully.

I would like to express my sincere gratitude and appreciation to my advisor Professor Magdy Salama for the expert guidance, endless support and encouragement, which he provided throughout my MASc journey.

I am thankful to my examining committee, Professor Ramadan A. El-Shatshat and Professor Tarek Abdelgalil for their patience in reviewing this thesis and their constructive comments, which helped me in improving the final presentation of this work.

On this occasion, I would like to extend my heartfelt gratitude to my beloved parents for their continual support, love, patience, and prayers. Many thanks to my dear brothers and sisters for their advice, support and faithful love.

I would like to acknowledge my colleagues in the Power and Energy Systems Group for their friendly discussion and suggestions.

Last but not the least, my deepest appreciation to my lovely wife, Maha, for her support, patience, understanding, encouragement during my master journey. Thanks to my beloved son, Fares, who is the source of my inspiration and strength.

This work was financially supported by King Saud University, Saudi Arabia.

Dedication

To my parents
To my brothers and sisters
To my beloved wife
To my beloved son
(with respect and love)

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 General

This chapter presents a short overview of the worldwide electricity demand growth and keen interest in utilizing renewable resources locally and globally. Section 1.3 presents the motivation behind the work done in this thesis. Moreover, this chapter will outline all objectives, and will conclude with the organization of this thesis.

1.2 Preface

Based on statistical studies, the electricity demand is anticipated to increase for the next few years. This increase reflects the widespread growth of population, infrastructures, industrial sectors, and commercial sectors. In reference [1], the electricity demand is expected to grow by 28%, from 2011 until the end of 2040, growing from $3,839$ billion KWh in 2011 to $4,930$ billion KWh in 2040. As a result of this growth, power generation utilities are responsible to meet the demand, so the expansion in generation capacities is considered as one of the duties of the planning engineers. Fig 1.1 shows the continuous growth of electricity sales in the united states which track the continuous growth of generation capacity,

where the base value of generation capacity at year 1949 is 902.78 GW and the base value of electricity sales at the same year is 19340.6 GWhrs. Besides the growth in generation capacity, the carbon dioxide emissions are increased as well, and the additional amount of greenhouse gas emissions can be inferred by the consuming of fossil fuels in the generation process. For instance, the CO2 emissions in the US have increased by 10% from 1990 till 2011 [2].

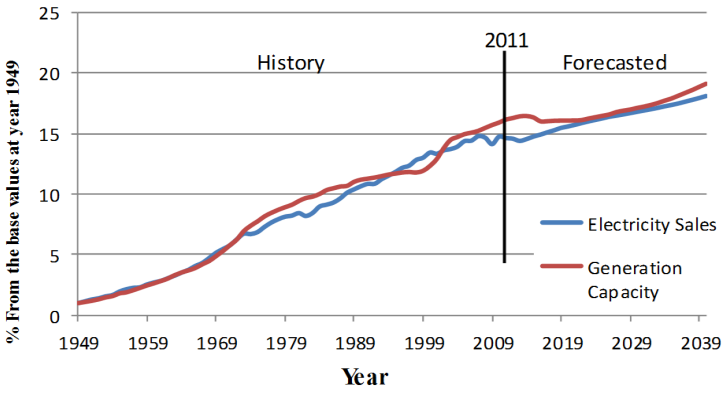

Figure 1.1: The Growth of Electricity Sales and Generation Capacity.

However, in recent decades, the interest in utilizing renewable resources in conventional power system networks has dramatically increased. The reason behind this is due to the cleanness and inexhaustibility of these resources, as well as their ability to support the existing grid. Approximately, 19% of the world's energy consumption is met by renewable resources and this was by the end of 2011 [3]. Despite the expensive installation cost of most renewable sources, the rapid growth of the installation capacity of these technologies is still ongoing. As an example, the global installed capacity of solar PV grew by 42% reaching 100 GW; in addition, the global installed capacity of wind has increased by 19% reaching 284 GW, and that was only for the last five years (2007-2012) [3]. Fig 1.2 shows the different kinds of renewable resources and their percentage increase for the last five years globally.

Canada strives to promote and develop these technologies, and it offers many incentives

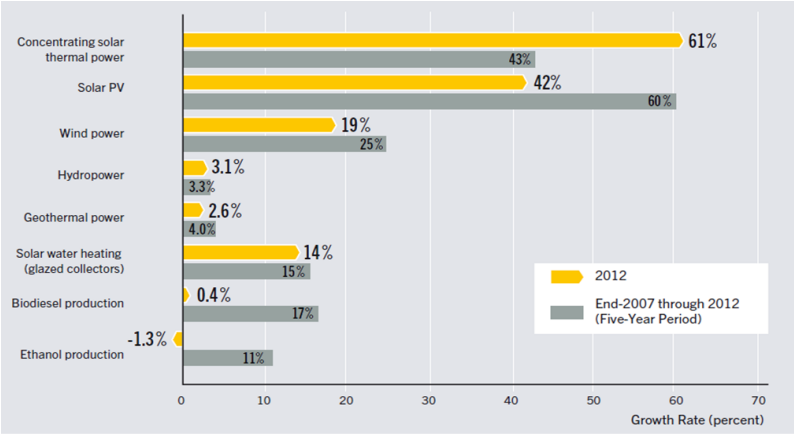

Figure 1.2: The Growth Rate of the Installed Capacity of Renewable Resources for the Past Five Years [3].

Table 1.1: Total Installed Capacity for Each Provence in Canada

Provence	ON	QC	AB	BC	NS	NB
Installed Capacity (MW)	2470.5	2398.3	1120.3	389.7	329.9	294
Provence	MB	SK	PE	NL	NT	YT
Installed Capacity (MW)	258.4	198.4	173.6	54.7	9.2	0.81

to facilitate investment for both companies and individuals in this area. Until the year of 2013, the capacity of wind farms in Canada is up to $7,698$ MW, enough to feed about 3% of the total electrical energy consumption of Canada, which is equivalent to almost 2 million homes [4]. The breakdown of the total generation capacity is illustrated in table 1.1 by each provence, where the installed capacity reflects to the ratings of the installed turbines at each provence. These farms are scattered in rural areas, and there are still many ongoing construction projects throughout the country. This interest is not limited to wind power only, but extends to the field of solar energy, including solar photovoltaic systems. The province of Ontario, one of the largest and most important provinces in Canada, leads the investment in PV power. In 2012, the installed capacity of PV systems has increased by 267 MW, resulting in a cumulative capacity of 645 MW [5].

1.3 Motivation

Because of the fluctuation in the output power of wind turbines and photovoltaic modules, the degree of uncertainty from the generation side is increasing. The variation of wind speed and solar irradiance could result in technical problems such as frequency deviation and bus voltage violation, and these concerns could be extended to reach the economic side such as the unit dispatches, electricity market, and spinning reserve [6]. The stochastic nature of wind speed and solar irradiance has forced researchers to incorporate statistical based analysis and some devices such as storage battery in their studies in order to tackle these issues. The stochastic analysis and the storage devices play key roles in the modern power system grids in the presence of renewable resources leading the grids to be "*smarter*". In fact, power system planning of renewable resources can be classified into two paths [7]. The first one is called short term planning which includes as an example unit commitment problems and storage scheduling. Wind speed forecasting and prediction techniques are usually preferred for this class. However, for long term planning problems, analyzing large sets of data to indicate the behavior of these variables (wind speed and solar irradiance) is mostly desirable. These data are often fitted to suitable distribution functions for the investigated sites.

Indeed, planning studies are essential in daily power engineers activities, and it should take into consideration the reliability and restoration costs beside the capital and running costs. Actually, incorporating reliability analysis in planning and even in operating states are magnificent in assuring secure and adequate power services to consumers. One of the challenges that face reliability analysis is the numerous system states which redirect the calculations to utilize simulation techniques rather than analytical methods. These obstacles come from the fact that overall system states are represented by an exponential equation which is a composite of the number of system components and the states of each component as shown in equation 1.1.

$$s = (m)^n \tag{1.1}$$

Where, m is the number of states for each component, n is the number of system components, and s is the system's overall states.

So, the motivation behind the work done in this thesis is to find a suitable algorithm that could minimize the number of wind and PV states that represent the intermittent nature of wind speed and solar irradiance so that the complexity in reliability and planning analysis could be handled. This target can not be achieved easily, since it needs more detailed analysis to prove the concept applied in this work.

1.4 Thesis Objectives

The objectives of this thesis are as follows:

1- Modeling the renewable resources in power distribution planning by analyzing a large amount of historical data and fitting it to an appropriate distribution density function.

2- Allocating these renewable resources in distribution networks using an optimal probabilistic power flow to minimize the annual energy losses considering the uncertainty caused by wind speed and solar irradiance as well as the variation from the load side.

3- Proposing a new algorithm to find the minimum number of wind and PV states that are able to precisely describe the behavior of wind speed and solar irradiance for the site under investigation, while reducing the complexity in reliability and planning analysis as well as minimizing the computational time for the calculations.

1.5 Thesis Organization

The organization of this thesis is as follows:

Chapter 2 provides background information for the distributed generation. In this chapter, a literature review for handling uncertainties from wind and PV production for long term planning problems is presented. The state of the art for conventional and renewable generators representation states is discussed in this chapter. This chapter is concluded by the previous work in sizing and allocation of distributed generation in power

networks since the planning problem that has been carried out in this thesis is sizing and allocation of distributed generation.

Chapter 3 presents wind and PV models applied in this thesis. One of goodness of fitness techniques is studied and applied in this chapter as well. A K-means algorithm for data clustering is addressed. In addition, calculating the output power from wind turbines and PV modules, and the annual capacity factor for wind and PV in the site under study are investigated in this chapter.

Chapter 4 proposes a new algorithm to find the minimum number of renewable resources representing states. In this part of the thesis, renewable resources are optimally allocated and sized using a genetic algorithm to test the proposed technique.

Chapter 5 exhibits the case study followed by the results achieved by the work done in this thesis and discussion .

In last chapter, **Chapter 6**, the conclusion of this thesis and the outlines of future work are presented.

Chapter 2

Background Information and Literature Review

2.1 Introduction

In a centralized power system, electrical power essentially flows in one direction from the central power plants throughout the transmission systems and ends up at the distribution control centers to reach the end users. However, in a decentralized power grid, the generating units can be located in the distribution levels and even at the customer sides to feed their own demand or their neighboring loads, or it can be act as micro-grids. In the recent emergence of distributed generators, DGs are able to contribute in powering the grid from distribution or even transmission stages. Remote and high population density regions are considered an attractive environment for utilities in order to implement these DGs. This is because of the high cost of grid expansion which is a composite of the cost of new transformers, overhead conductors, underground cables, and infrastructure expenses. Distributed generators are providing numerous benefits for both the utilities and the end point customers. Indeed, these advantages will be constrained by some technical challenges. These advantages and limitations will be addressed in this chapter.

This chapter first will discuss the definitions of distributed generation, its advantages,

technologies, and limitations. Secondly, modeling of two kinds of the most widely used distributed generation technologies in long term planning problem will be investigated (wind and solar-based PV). Lastly, various ways to represent the states of these renewable sources-based DGs in reliability and planning problem are discussed.

2.2 Distributed Generation Definitions

Various books, associations, utilities, and papers have given different definitions of the term "*Distribution Generation*". These definitions vary in their consideration, so that some take into account the size of the DGs, while others concentrate on the location of these DGs or their technical effects such as voltage level or power quality. Among of these definitions, this subsection highlights the most commonly used definitions utilized in literature.

According to the Institute of Electrical and Electronics Engineers (IEEE), the distributed generation is defined as "*generation of electricity by facilities that are sufficiently smaller than central generating plants so as to allow interconnection at nearly any point in a power system*" [8]. This definition does not specify the size into account. Thus, Electrical Power Research Institute (EPRI) in [9] considers the rating of these DGs, by saying that they are "*generating from a few kilowatts up to* 50 *MW*". CIGRI, The International Council on Large Electricity Systems, defines these DGs as "*all generation units with a maximum capacity of* 50 *MW to* 100 *MW, that are usually connected to the distribution network and that are neither centrally planned nor dispatched*" [8]. The most general definition for distributed generation is proposed in [9], and it does not restrict the size limits of these DGs while this definition does define the locations of these small generators. It is defined as: "*the installation and operation of electric power generation units connected directly to the distribution network or connected to the network on the customer site of the meter*".

Actually, the types of the DGs as well as the owners of these DGs are disregarded in the aforementioned definitions. Nevertheless, many papers suggest that each utility should have its own definition which is depend on the conditions of the network. So, table 2.1

illustrates a suggested classification for different capacities of DGs.

Table 2.1: Various DG Capacities

Class	Capacity
Micro DGs	1 W $<$ 5 KW
Small DGs	5 KW $<$ 5 MW
Medium DGs	5 MW $<$ 50 MW
Large DGs	50 MW $<$ 300 MW

2.3 Distributed Generation Benefits

Distributed generation is able to provide the system numerous benefits [10]. Indeed, these benefits could be clustered into three categories including technical, economic, and environmental advantages. The following subsections discuss these benefits.

2.3.1 Technical Benefits

When the distributed generation units are properly located and sized in the distribution systems considering adequacy and security regulations, these devices are expected to provide a positive credit to the overall network. These technical advantages involve:
1- Reduction in power losses.
2- Improving system reliability.
3- Improving voltage levels.
4- Enhancing network security.
5- Alleviation of congestion at substations and conductors.
6- Improving system overall efficiency.

7- Improving system quality.

When the DGs are supporting the grids close to the load points, there will be a reduction in the power flow in transmission and sub-transmission systems which results in the reduction of system losses as well as reducing possible congestion with regards to equipment. Moreover, DGs can be used a back-up or main generation source which leads to enhancement of the system's reliability.

2.3.2 Economic Benefits

Economic benefits gained from installing DGs play a crucial role in reducing power system expenses for either long or short planning horizons. Thus, power system utilities are attempting to provide the electricity for all consumers at low cost. These costs involve the cost of stationary parts (fixed cost) and the running cost (operation and maintenance costs). The economic benefits of DGs could be summarized in the following points.
1- Deferring the investments due to system upgrades or expansions.
2- Reduction in operating costs that have been drawn by shaving peak demands since DGs are working during peak load intervals.
3- Minimizing the consumption of fossil fuel that leads to decreasing in energy prices.
4- Reducing the maintenance costs due to the alleviation of congestion for the devices.
5- Minimizing the cost of spinning reserve requirements.

2.3.3 Environmental Benefits

The strong motivation behind employing DGs in power networks lies in their environmental benefits. According to report illustrated in [10], carbon dioxide emissions have been dramatically decreased by 30% in only a three year period in the Danish power system. Furthermore, wind turbines, PV modules, and hydro turbines are non-polluting and have a high degree of sustainability. The environmental benefits of employing DGs are as follows:
1- Reduced carbon dioxide emissions CO2.

2- Reduced health care costs.
3- Reduced land use (depends on the location DGs).

Land use varies and is dependent on location, so it is important to minimize the land use needed to install these DGs. For example, the land use will be minimized when the PVs are installed on the rooftops of buildings instead of land areas. In the case of wind, the land use will be minimized when the wind turbines are installed on unused hilly areas instead of flat areas.

There are some challenges and issues to be overcome, and these obstacles increase when the penetration level of the installed DGs increases. Bus voltage limits, power quality issues such as harmonic injections, reverse power flow, protection schemes and coordination, short circuit currents, reactive power support, and frequency deviations are all examples of these challenges.

2.4 Distributed Generation Technologies

Depending on the type of the primary fuel source of the distributed generation, reference [7] classified these technologies into four categories. The first category is called conventional technologies. Diesel generators are usually located in remote areas, and they are a great example of conventional technologies. Advanced fossil technologies are the second type of these technologies. Advanced fossil contains fuel cells which are mainly fueled by hydrogen in electrochemical power conversion. In addition to fuel cells, micro- turbines which are fed by natural gas is another form of Advanced fossil technologies that is based on cyclic gas processing. Renewable technologies play a key role among these technologies since they are natural, sustainable, and conservative for the environment. These technologies include wind turbines, hydro turbines, photovoltaic modules, tidal systems, geothermal technologies, and solar thermal systems. The degree of uncertainty in these forms of energy is relatively high. There are some technologies that are able to increase the system's overall

efficiency, such as energy storage and combined heat and power systems, and these systems could be considered as efficient technologies.

Based on the location of the different types of distributed generation, these technologies can be divided into two levels [10]. The end point customer level is the level at which DGs are placed in the last points in the distribution systems, and most of the technologies in this level are owned by the customers and have minor effects on the grid. The site specific level is the second level, in which the DGs need large spaces in which to be employed. They are mostly owned by the utilities or government, and they tend to be less centralized than the conventional plants. The following subsections highlight the most attractive renewable resources in recent years (wind turbines and PV module based-DGs).

2.4.1 Wind Power

Wind is generated by heat differences between different areas of the earth's surface, and has been used as a source of energy for many years. The availability and usage of wind energy differs from location to location throughout the world. Examples of uses of wind power from early generations include: sailing ships, water pumping, and grain milling. Recently, the kinetic energy of the wind has grabbed global attention as a natural source to generate electricity. For this use, wind farms are scattered throughout the world to convert the wind that drives turbine blades into mechanical energy. The movement in the blades results in shaft rotation which drives a generator, and this generator converts the mechanical energy into electrical energy through a electromechanical conversion process.

2.4.2 Wind Turbines

The main function of wind turbine is to convert the wind energy into mechanical power. There are many classifications for these turbines, and these types clustered from various perspectives as follows;

A- Rotational speed: (1- Fixed speed turbines, 2- Variable speed turbines).

B- Pitch angle: (1- Fixed pitch angle turbines, 2- Variable pitch angle turbines).

C- Wind direction: (1- Upwind, 2- Downwind).
D- Axis orientation: (1- Horizontal axis turbines, 2- Vertical axis turbines).
E- Blades number: (1- One, 2- Two, 3- Three, 4-Four).
F- Type of generators: (1- Squirrel-cage generators, 2- Doubly-fed induction generators, 3- Multi-pole synchronous generators).
G- Connection to the grid: (1- Directly connected, 2- Connected through converters).

It is important to mention that here, horizontal axis wind turbines HAWTs are commonly used in the field instead of vertical ones. In addition, the fixed speed turbines operate somewhat above the synchronous speed of the generator; whereas, the variable speed turbines are designed to operate above and below the synchronous speed.

2.4.3 Wind Turbines Generators

Based on the construction of the wind energy conversion system, turbines' generators can be classified into four classes [11]:

Type A - Squirrel Cage Induction Generator

The rotational speed of this type is constant, so it is classified as a fixed speed wind turbine equipped with a squirrel cage induction generator. When the generator speed is higher than the synchronous speed, then the power will be generated. As it is shown in fig.2.1, the turbine shaft is coupled with the SCIG shaft using the gear box.

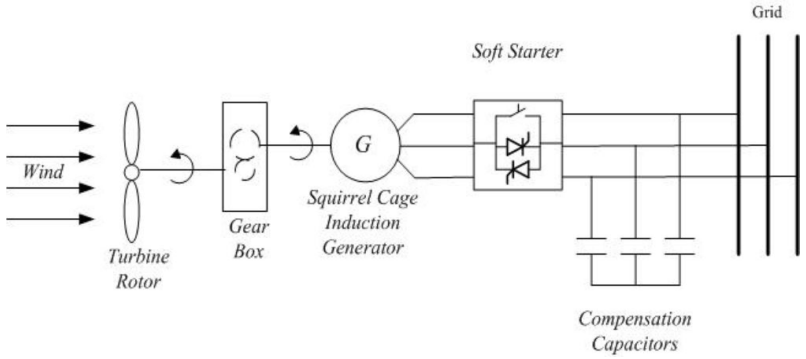

Figure 2.1: Type A - Squirrel Cage Induction Generator [11].

This type of generator is directly connected to the grid; therefore, a soft starter circuit is utilized in order to mitigate the high stating current. Capacitor banks are used to improve the power factor of these machines.

Type B - Wound Rotor Induction Generator

The rotational speed of wound rotor induction generators is variable; therefore, it is used in variable speed wind turbines. Here, the rotor is connected to variable resistance through electronic power circuit to control the rotor's speed, shown in fig.2.2. The function of the gear box is to couple the blades' shaft with WRIG shaft. The additional resistance causes efficiency losses in this type of wind generator.

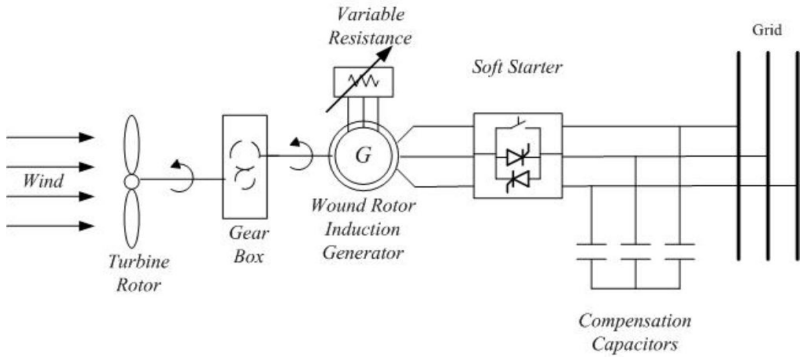

Figure 2.2: Type B - Wound Rotor Induction Generator [11].

Type C - Doubly Fed Induction Generator

This type is mostly used in open fields, and is a variable speed type. The turbine is coupled with a doubly fed induction generator as shown in fig. 2.3. At low and medium wind speed, the maximum energy output can be achieved; whereas, when the wind speed exceeds the rated speed, then the pitch control system will operate to limit the rotor speed. So, the generator will be operating at its rating if the wind speed is higher than the rated speed. Rotor speed is controlled by using what is called "back to back converters" which allow independent control for active and reactive power. For the purpose of shaft coupling, a gear box is used.

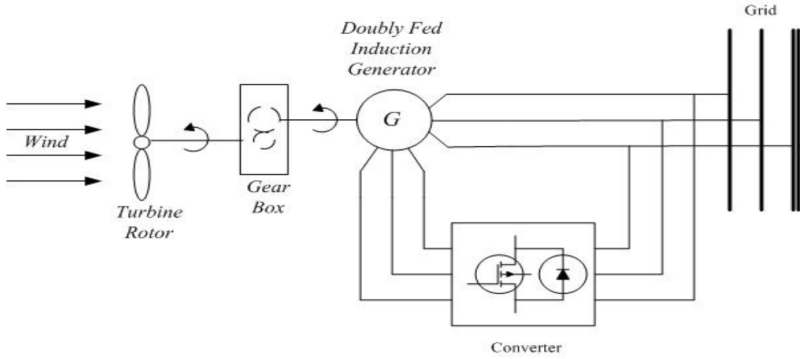

Figure 2.3: Type C - Doubly Fed Induction Generator [11].

Type D - Synchronous Generator

This type is a variable speed turbine coupled with a multi-pole synchronous generator, and utilizes a full power electronic conversion system (AC-DC and DC-AC). Similar to type C, this type achieves the maximum power output when the rotor speed is lower than the rated speed. When the rotor speed is higher than the rated speed, the control system (pitch control) will operate to maintain the output power at the rated power. In this type, active and reactive power is independently controlled as well.

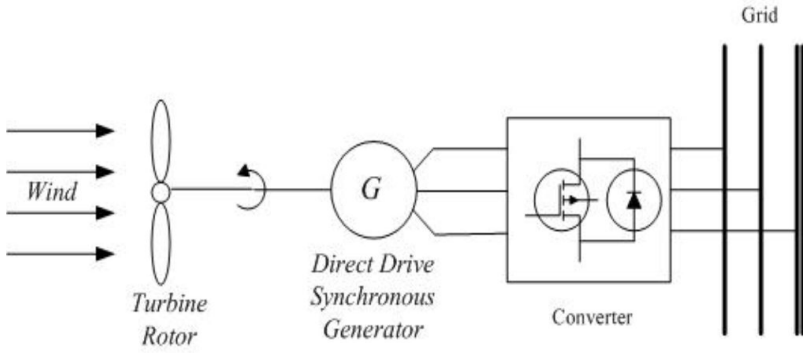

Figure 2.4: Type D - Synchronous Generator [11].

2.4.4 Photovoltaic Power

The first observation of the photoelectric effect was in 1839 when light was converted to electricity by using metal electrodes and electrolyte. This observation was later interpreted by Albert Einstein in 1904 and labeled the "*photoelectric effect theory*". Photovoltaic power conversion is a process whereby sunlight (solar irradiance) is captured by semiconductor material and converted into electrical charges (current) via solar cells. PVs were used in space programs as the first practical applications. Later, PV devices were employed in limited mundane applications such as water pumping, street illumination, and small remote area systems, since it was more cost effective for these uses. These devices were costly, because crystalline semiconductor materials is used in PV cell construction. However, from 1980 until now, the cost of photovoltaic modules has gradually declined, and since the governments encourages customers with incentive programs, the interest in implementing these modules has risen. Fig. 2.5 shows the module prices in Canada from 2000 up to 2012 [12]. Generating power from photovoltaic modules has many advantages such as, low operating and maintenance costs, zero noise due to stationary or static parts, light weights, high reliability, long lifetime operation, and short lead times for installation.

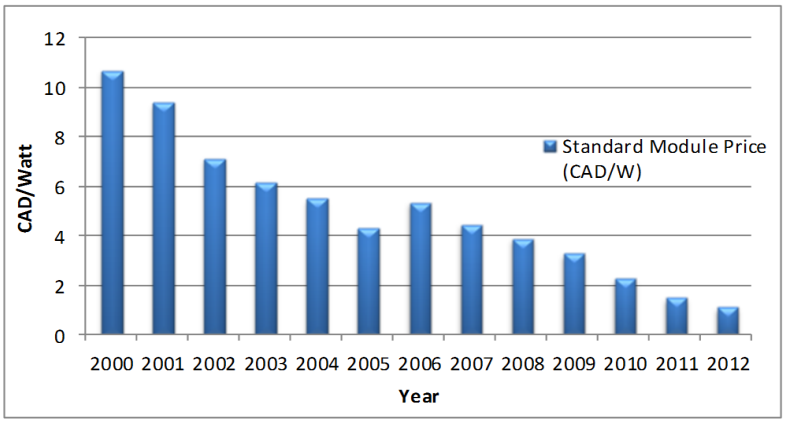

Figure 2.5: Photovoltaic Modules Prices in Canada

2.4.5 PV Cell Technologies

More than 80% of photovoltaic cells in the world are made from silicon since it is reliable and provide service for many years. The production lines are still running to produce an efficient cells with low cost of production. There are many types of silicon based photovoltaic cells, and can be divided into three categories [13]:

1- Mono-Crystalline Silicon

This type is made from very thin wafers, and these wafers are obtained from a large mono-crystal which is generated from pure melted silicon. The efficiency of this type is high when it compared to other types. The wafers are smooth and take shapes as hexagonal or pseudo-square in order to fit in the modules properly. To produce more electrical power, the crystalline silicon is doped with other elements. Mono-crystalline silicon wafer manufacturing is expensive and needs more attention and a slower production process than some other methods. The behavior of mono-crystalline silicon is uniform and predictable.

2- Poly-Crystalline Silicon

The manufacturing cost of this type is lower than mono-crystalline, and the process in production is faster and less complicated than the first one. It follows a casting process in production, and is obtained from pure melted silicon. It has low efficiency compared to mono-crystalline since the structure of the crystal is random. It is not as expensive as the mono-crystalline type.

3- Amorphous Silicon

This is the least expensive type among the different silicon classes. In this technology, very thin semiconductor materials are deposited onto μm-thick amorphous films on stainless steel rolls. The efficiency of amorphous silicon is about one-half of the mono-crystalline silicon; however, its cost of production is significantly lower than the crystalline ones. It is widely used in calculators and small watches, and it has a short lifespan. Table 2.2 presents the efficiencies and module surface areas for silicon modules.

Table 2.2: The Efficiency and Module Surface Areas for Silicon Modules [13].

PV Technology	Efficiency %	Module surface area required
Mono-Crystalline Silicon	$12 \sim 16$	$7m^2/kW$
Poly-Crystalline Silicon	$11 \sim 15$	$8m^2/kW$
Amorphous Silicon	$6 \sim 8$	$16m^2/kW$

2.4.6 PV Arrays Topologies

PV modules are composite solar cells which are connected in series to increase the voltage, or in parallel to increase the current and therefore the output power. PV modules are the basic units of photovoltaic systems. A photovoltaic panel is composed of multiple wired modules, and it is the basic unit of a photovoltaic array. These arrays then are connected to power conditioning units to convert the DC output into AC output in order to match these units with the grid system. Actually, there are different topologies of PV arrays based on its connection to the power conditioning units. These topologies are [14]:

1- Centralized Topology

In this configuration, all photovoltaic arrays are connected to one inverter, and are utilized for large PV systems with output power in the megawatts. When the inverter fails, the output power will be zero, so the reliability of this configuration is low. However, the cost of this topology is low since it is uses only one inverter.

2- Master-Slave Topology

These arrays are connected to multiple parallel inverters, and the power is shared between these inverters. So, the operation lifetime will be extended for this topology. When one inverter fails, the output power finds another path (another inverter) to pass through, so the reliability of this configuration is higher than the first one. However, this topology requires a higher cost due to the multiple inverters used.

3- String Topology

Here, one inverter is connected to each row of arrays. In this topology, the reliability is

improved due to using multiple inverters. Moreover, the efficiency is enhanced since each string or row is working independently at its maximum power point. This topology also requires high cost due to the multiple inverters used, and the power can be increased by adding many modules to each string.

4- Team Concept Topology

In this topology, each string is connected to one inverter, and multiple inverters are connected in parallel so that if one inverter fails, the other inverter will carry the power. So, the reliability and system cost in this configuration will be increased. Moreover, the efficiency is enhanced since the tracking system of each string or row is working independently to reach its maximum power point. The operation lifetime will be increased to due to the power sharing between the inverters.

5- Multi-String Topology

Each row or string will be connected to DC-DC converters, and these converters will be connected to the AC inverter. The aim of this configuration is that that maximum power will be obtained since it will be working at its maximum power tracking point, and the cost will be reduced since only one inverter is used. However, this topology affects and reduces the reliability.

6- Modular Topology

In this configuration, the inverter is coupled with the photovoltaic modules to enhance the reliability and efficiency. The cost of this topology is high compared to the other topologies; however, each module will be operated at its maximum power, so the power loss due to shading mismatch will be minimized.

2.4.7 Mode of Operation

The PV systems can be classified based on operation mode into two types. The first class is the off-grid PV system or isolated system, and this type is usually used to power remote areas, pump water, feed communication systems, or it could be used to supplement the distribution system when a fault occurs to isolate some part of it. The second class is the grid-connected system. It is directly connected to the grid through power electronic

conversion systems. Most of PV farms that are able to produce bulk power to the grid or large consumers are considered one of the grid-connected forms.

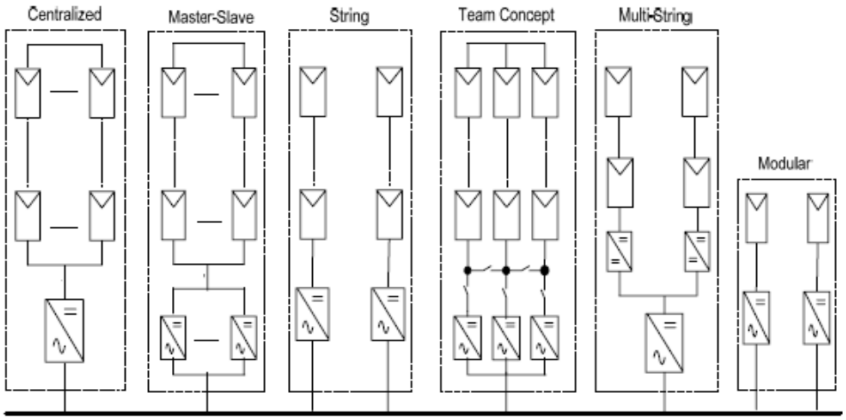

Figure 2.6: Different PV Topologies [14].

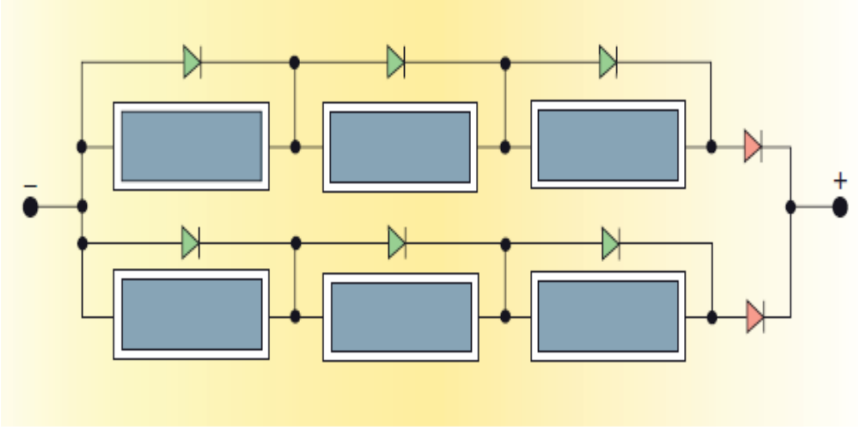

Figure 2.7: Photovoltaic Array Consisting of Six Modules [13].

2.5 Distribution Density Functions in Long Term Planning

Long term planning actually involves reliability evaluation and economic analysis. One of the power planning applications applied in decentralized power systems is allocation and sizing of distributed generation, and this problem is usually evaluated by using technical, economic, and reliability analysis. Power system planning, including renewable resource based-DGs, requires a time consuming analyzing process for large sets of historical data. These data are fitted to the suitable statistical probability density functions to represent the behavior of wind speed or solar irradiance in the investigated site. This method is efficient for long term planning problems to handle the uncertainties associated with fluctuations in wind speed or solar irradiance.

2.5.1 Popular Distribution Functions

In the following subsections, most of distribution functions used in the literature to model the intermittent nature of wind and solar will be described [15].

1- Normal Distribution

This is one of most popular continuous distributions to model the randomness of events. It utilizes two parameters which are the mean μ of sample data and the variance σ^2 to generate the probability density function. Normal distribution proves its effectiveness to model such continuous random variables, such as solar irradiance, growth rate, time, and others. Normal density function is defined by:

$$F(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(\frac{X-\mu}{2\sigma^2})}, Where - \infty \leqslant X \leq \infty \tag{2.1}$$

Where, μ is the mean of the data, and σ^2 is the variance.

2- Weibull Distribution

Weibull distribution is considered as one of the distributions able to model the complicated continues random variables. It is commonly and extensively used to model many events,

such as wind speed, reliability problems, states of failure, survival analysis, delivery estimation in industrial analysis, and extreme deviation analysis. It is driven by two parameters which are shape index k and scale index c. Its probability density function is defined by:

$$F(X) = \begin{cases} (\frac{k}{c}) * (\frac{X}{c})^{k-1} \exp[-(\frac{X}{c})^k] & X > 0 \\ 0 & , Otherwise \end{cases} \tag{2.2}$$

3- Exponential Distribution

The general expression of exponential distribution is shown below.

$$F(X) = \begin{cases} \theta e^{-\theta X} & X > 0 \\ 0 & , Otherwise \end{cases} \tag{2.3}$$

The parameter θ must be greater than 0, and it is called rate parameter. θ represents the average of the sample data. Exponential distribution is commonly used to model waiting times between recurring independent rare events, and lifetimes of devices.

4- Gamma Distribution

Gamma distribution is useful in power system reliability, especially to calculate failure time of power equipment. It is also used in queuing theory analysis as well as waiting times of providing such services. Gamma density function is described as below.

$$F(X) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} X^{\alpha-1} e^{\frac{-X}{\beta}} & X > 0 \\ 0 & , Otherwise \end{cases} \tag{2.4}$$

Where, α and β are greater than 1. Exponential distribution is a special case of gamma distribution when α equals 1. So, because of the relationship between gamma and exponential distributions, gamma can be used for several applications.

5- Beta Distribution

Beta distribution is utilized in a wide range of applications. It has been used early on to model the randomness of solar irradiance. It only occurs when the random variable is

between zero and one.

$$f(X) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} X^{\alpha-1} \\ (1-X)^{\beta-1} & 0 \leq X \leq 1, \alpha and \beta \geq 0 \\ 0 & Otherwise \end{cases} \tag{2.5}$$

α, β are the parameters of beta distribution function. μ, σ are the mean and standard deviation of the random variable, respectively.

6- Lognormal Distribution

This distribution is used for various applications. Lognormal distribution is employed for random variables that are normally distributed when the log transformation is applied to them. It is also driven by two parameters which are μ and σ. Its density function represented as below.

$$F(X) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma X} e^{\frac{(\mu - ln(X))^2}{2\sigma^2}} & X > 0 \\ 0 & , Otherwise \end{cases} \tag{2.6}$$

2.5.2 Goodness of Fitness Tests

Goodness of fitness tests are statistical based methods which indicate how well the measurements or observations fit the claimed probability distributions. These tests are a type of mathematical analysis that measures the errors and differences between the candidate distributions and the random samples. In fact, there are several ways to examine whether the proposed distribution fits or not. Here, three fitting tests applied in the literature are illustrated [16].

1- Chi-Squared Test

The Chi-Squared test is the oldest test, and it can be commonly used for any univariate distribution for which a cumulative distribution function can be defined. Data are firstly clustered into groups, and the observed frequency will be compared with the expected (claimed) frequency. The error of the Chi-Squared test can be minimized when sufficient

data are used. It is defined as below:

$$\chi^2 = \sum_{i=1}^{N} \frac{[O(i) - E(i)]^2}{E(i)} \tag{2.7}$$

Where, $O(i)$ is the observed frequency for bin i, and $E(i)$ is the expected frequency. $E(i)$ is obtained from the difference between the cumulative distribution function of bin i in the lower and upper limits. Then, depending on the degree of freedom and significance level, the result of this test is determined.

2- Kolmogorov-Smirnov Test

This method basically computes the difference between the cumulative distribution for the sample data and for the suggested distribution. It is simple and suitable for limited and unlimited data availability. The expression of $K - S$ test is described below:

$$D_n = Max \{ \mid F_n(x) - F(x) \mid \} \tag{2.8}$$

The data should be arranged from small numbers to the large numbers, e.g. T_1, T_2,.., T_i. $F(x)$ is the cumulative distribution for the expected distribution, and $F_n(x)$ is the cumulative distribution for the data which is equal $Y(i)/N$, where $Y(i)$ is the number of data that are smaller than T_i, N is the total number of data. The calculated D_n will be compared to the critical value obtained from tables, and if D_n is greater than the critical value, the hypothesis will be rejected.

3- Anderson- Darling Test

This test appears after applying some modifications on the $k - s$ test. In fact, the $K - S$ test does not give that much weight to the tail; however, the Anderson- Darling test does. The advantage of the $A - D$ test over the $K - S$ test is that the candidate distribution is entered in critical values calculations in contrast to the $K - S$ test. This will give a positive credit for the $A - D$ test since it makes more sense that the critical values will be calculated for each tested distribution. This test is defined by;

$$A^2 = -N - S \tag{2.9}$$

And

$$S = \sum_{i=1}^{N} \frac{2i-1}{N} [lnF(Y_i) + ln(1 - F(Y_{N+1+i}))] \tag{2.10}$$

2.5.3 Distribution Parameters Estimation

Most distribution systems are usually driven by two or three parameters. These parameters are estimated using different methodologies in order to obtain the maximum advantages of PDF fittings for the random data. In the literature, distribution parameters are calculated using various methods ,such as graphic method, maximum likelihood method, moment method, energy pattern factor method, empirical method, and modified maximum likelihood method. Parameter estimation methods that are commonly used in cases of estimating Weibull distribution parameters (scale c and shape k factors) are described in this section [17].

1- Graphical Method

In this method, a linear equation should be obtained from the cumulative distribution function of Weibull as shown in 2.11.

$$ln[-ln[1 - F(v)]] = k\, ln\, v\ - k\, ln\, c \tag{2.11}$$

Then, when the right hand side of the above equation is drawn versus $ln\, v$, the shape factor will represent the slope of the straight line, and the scale factor will be obtained from the intersection with $ln\, v$ axis.

2- Maximum Likelihood Method

The maximum likelihood method is an iterative-based method. It involves trying to fit weibull to time series data. Large numerical iterations are required, so the difficultly increases.

$$k = \left[\frac{\sum_{i=1}^{n} v_i^k ln(v_i)}{\sum_{i=1}^{n} v_i^k} - \frac{\sum_{i=1}^{n} ln(v_i)}{n} \right]^{-1} \tag{2.12}$$

$$c = \left(\frac{1}{n} \sum_{i=1}^{n} v_i^k \right)^{\frac{1}{k}} \tag{2.13}$$

Where, n is the number of non-zero observations, and v_i is the wind speed at period i.

3- Moment Method

The moment method is another way to find the parameters, and it is also a numerical-based analysis. It requires pre-computing for the mean speed $\bar{v}$ and standard deviation for the wind speed.

$$\bar{v} = c\Gamma(1 + \frac{1}{k}) \tag{2.14}$$

$$\sigma = c[\Gamma(1 + \frac{2}{k}) - \Gamma^2(1 + \frac{1}{k})]^{\frac{1}{k}} \tag{2.15}$$

4- Energy Pattern Factor Method

This is also called "power density method". This method is simple, easy to implement, with less execution time required. Energy pattern factor is the average of wind speed cubes over the cubes of wind speed average.E_{pf} are usually between 1.45 and 4.4 for most Weibull distributions.

$$E_{pf} = \frac{\bar{v^3}}{(\bar{v})^3} \tag{2.16}$$

$$k = 1 + \frac{3.69}{E_{pf}^2} \tag{2.17}$$

$$\bar{v} = c\Gamma(1 + \frac{1}{k}) \tag{2.18}$$

5- Empirical Method

This method is a special type of moment method. It depends on calculating the mean and standard deviation in advance. This method provides good estimations if and only if the shape factor is between 1 and 10 ($1 \leq k \leq 10$).

$$k = (\frac{\sigma}{\bar{v}})^{-1.086} \tag{2.19}$$

$$c = \frac{\bar{v}}{\Gamma(1 + \frac{1}{k})} \tag{2.20}$$

2.6 State Representation of Wind and PV Powers

Conventional power generators are mostly represented by two states (up and down). Up state is the state at which the generator is working at its rated capacity, and down state is the state at which the power generator is out of the service and its power is zero. Later, a derated state is added to the conventional representation, and describes a case when there are some partial fails in some of the auxiliary equipments and the generators are continue to generate but with reduced power.

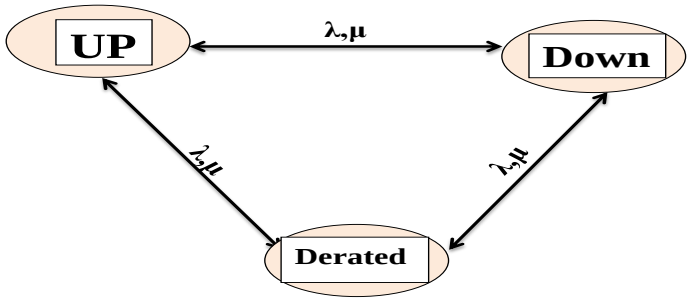

Figure 2.8: Convectional Generator States.

However, multi-state wind and PV unit models are usually incorporated in distribution system planning or operation problems. These problems are mainly reliability analysis, where the selection of derated states of wind and PV units are very sensitive and handling fluctuations in operation analysis, where the derated states play an important role for the control systems. The selection of these states is very crucial since it is a tradeoff between the accuracy of the results and the complexity in the analysis. The complexity comes from the fact that the states of any system is the number of component states powered by the number of system's components $(m)^n$. In the literature, there are three main techniques to select the derated states that represent the behavior of wind speed and solar irradiance, and these techniques are described in the following subsections [6, 18–21].

2.6.1 Equivalent States Method

In reference [18], depending on the maximum wind speed and how many intervals are required, wind speed is divided into multiple equal intervals. The power density function in the equivalent states method is divided into multiple equal bands. The size of each band depends on the number of required intervals and the standard deviation ($\frac{10\sigma}{N_b}$). The value of each state is represented by the midpoint of each interval (MB_i) which is demonstrated in the equation below.

$$MB_i = \begin{cases} \mu + (\frac{10\sigma}{N_b})(i - 0.5N_b) & \text{for } even N_b \\ \mu + (\frac{10\sigma}{N_b})(i - 0.5(N_b + 1)) & \text{for } odd N_b \end{cases} \tag{2.21}$$

$$P_i = \frac{NS_i}{N_y} \tag{2.22}$$

Where, P_i is the probability of interval i, μ is the mean value of wind speed or solar irradiance, NS_i is the number of simulated wind speed or solar irradiance at interval i, and N_y is the total of wind speed or solar irradiance data. The disadvantage of this method is that it requires large numbers of intervals to acquire accurate results.

2.6.2 Apportioning Method

This method is extensively used in reliability analysis, when the capacity outage rate for the components is needed to be generated [21]. The apportioning method tries to reduce the number of derated states. This method is based on dividing the residence times of the actual derated states between the up or down states and pre-assigned states. The probability of the pre-assigned states will be contributed by the probabilities of absorbed states. This method does not give more details in how to find the pre-assigned states, and it does not mention any information regarding the stopping criterion for the apportioning process.

2.6.3 Clustering-Based Method

Clustering is basically applied to a large number of observations for the purpose of placing these observations that have common attributes or are relatively homogenous in groups. There are two types of clustering techniques; the first one is supervised clustering, and the other type is unsupervised clustering. The difference between them is that in case of supervised clustering, the number of the centroids are considered as an input for the algorithm; however, in case of unsupervised clustering which is the most useful in the previous works, the algorithm will determine the number of the centroids. The state value is represented by the cluster centroid. In reference [19], the K-means algorithm is utilized to cluster the wind speed observation into many clusters in order to address the wind in system operation. The K-means algorithm will be described intensively in the next chapter. There is another way of clustering which has been used in [20]. This cluster-based method is depending on the output power curve of wind speed. When the wind speed is below the cut in speed, this state is considered to be the first state with a value of zero. If the wind speed is on or above the rated speed, this state is considered to be the last state, with a value of 100% of turbine's capacity. The derated states are determined by time step 1 m/s from a cut in speed up to rated speed, and the value of each state is the average of wind speed between each time increment.

2.7 Previous Work in DG Allocation and Sizing

Much work has been done in the literature for allocating distributed generation in the network. Depending on the development of the system, losses could reach up to 20% from the total power generation for some utilities [10]. Therefore, one of the significant objectives for placing DGs in a distribution system is to reduce overall system losses. The authors in [22] and [23] employed analytical techniques in order to minimize the feeder losses, and this method was based on the 2/3 rule. Besides, the study that has been done in [24] used what is called loss factor for the same purpose while in [25] the exact power losses expression has been taken into account. In the same context, Tabu search has been applied

in reference [26] to figure out the magnificence of DGs in losses reduction.
By conceding multiple DGs and load levels, the multi-objective function has been mathematically modeled and solved in [27] using mixed integer non-linear programming, and the cost of the losses was presented among the cost functions. In addition, genetic algorithm proved its accuracy and effectiveness in solving optimization problems. A study performed in [28] employed GA in different load models for optimally DG allocation. Furthermore, the authors in [29] used a particle swarm optimization model to reduce the energy losses while reactive power losses and feeder loading are considerable. It is important to mention here that most of the optimization models considered some practical constraints, and these constraints vary form one application to another. To illustrate, these restrictions include the thermal limit of the feeders, the voltage at each bus, the maximum reverse power, short circuit level, and others.
In [20], the authors optimally allocate a mixture of renewable resources in distribution system to minimize the annual energy losses. Handling the uncertainty produced from the intermittent nature of wind speed and solar irradiance is presented. A heuristic optimization model solved using decision theory is proposed as well in [30]. Moreover, hourly online reconfiguration taking into consideration the time varying nature of load is presented in [31], and interestingly, reconfiguration reveals the effectiveness of this method on loss reduction. In [32], multi-period AC power flow method has been illustrated with adjustment of the power factor of DGs to find the optimal mix of renewables. The work in [33] performs a method which is able to compute and analyze the fluctuation in annual energy losses, comparing various DG technologies and penetrations.

2.8 Summary

In this chapter, the concept of distributed generators in modern power system is demonstrated as well as the technical, economic, and environmental benefits of these devices. The different technologies of wind turbines and PV modules are widely explained in order to give an idea about the selected wind turbines and PV modules in this thesis. Since the problem in this thesis is considered as long term planning problem, the appropriate way to

model the intermittent nature of wind speed and solar irradiance is by analyzing a large historical data and fitting these data into suitable probability density functions. So, the different types of density functions that have been commonly used in the literature and their parameters estimation are defined in the chapter. The different techniques that have been used to define the number of wind and solar states are presented. These states will be incorporated in the deterministic planning problem as a multi-state model that will combine the probabilistic nature of the wind and solar variation into the deterministic DG allocation and sizing problems.

Chapter 3

Modeling of Renewable Resources and Load

3.1 Introduction

As discussed before, the employment of renewable resources in the power grid has been a keen interest for the utilities, governments, and consumers. This awareness comes from the fact that these renewable resources are able to provide enormous benefits to the grid, the environment, utilities, and customers. However, dealing properly with these renewable resources is crucial due to major fluctuations in their strength and availability. Therefore, robust models and representation equations are essential in order to indicate the randomness that is caused by variations in wind speed or solar irradiance. Modeling of these resources primarily depends on the application of these technologies. Chronological time series models are efficient for short term planning such as generation scheduling and market bidding. In contrast, probabilistic models using probability density functions prove their effectiveness in long term planning such as reliability assessments and resources allocation. This chapter is designated to model the behavior of wind speed and solar irradiance using the proper probability density functions. The selected distributions are tested using the K-S test. Moreover, in this chapter, the models that are utilized to calculate the output

power from wind turbines and photovoltaic modules are demonstrated. Unsupervised clustering technique (K-means) is described in this chapter in order to define the states of wind or PV power, and these states will be used in the proposed method in chapter 4.

3.2 Data Preparation

The study period of the system is one year, so it has been divided into 12 months, and a group of three months represents one season. In order to model the intermittent behavior of wind speed and solar irradiance, a typical day is created to represent each within a particular season. Moreover, this season is divided into 24 hours, and each hour represents the solar irradiance and wind speed within that particular hour for the whole season. Table 3.1 presents an example of how the wind speed is prepared in order to be included in the analysis. In theis table, the historical data of wind speed is assigned to their suitable time of occurrence.

This model allows making hourly correlations between the generation of intermittent resources and the load for the entire year. Five years of historical data is used for wind speed and solar irradiance, and each season is assumed to be ninety days. As result, the data of each hour for a particular season is 450.

Table 3.1: Example for Data Preparation of Wind Speed

	Winter				**Spring**				**Summer**				**Fall**			
H\D	**1**	**2**		**90**	**1**	**2**		**90**	**1**	**2**		**90**	**1**	**2**		**90**
Hour 1	1,2,2,0,1,2,m/s				4,4,4,5,3,6,5,2,m/s				1,2,0,1,2,4,m/s				6,5,8,9,10,12,m/s			
Hour 2	4,2,2,2,1,3,m/s				4,2,4,5,6,3,5,2,m/s				.				8,7,8,5,10,12,m/s			
.	.				..				4,4,4,5,3,6,5,2,m/s				..			
Hour 24	1,2,0,1,2,4,m/s				4,4,3,5,6,6,2,2,m/s				6,5,8,9,10,12,m/s				8,7,8,9,5,12,m/s			

3.3 Wind Speed Model

From the historical data analysis above, it is obvious that the data would be divided into 96 groups (24 hours X 4 seasons). For each group, the parameters of any distribution are different. Therefore, based on the "*Kolmogorov Smirnoff*" (K-S) algorithm illustrated in [34], five distribution systems that are commonly used in the literature to model the behavior of wind speed are tested to select the most fitted one to the historical data among them. These distributions are Weibull, Normal, Rayleigh, Gamma, and Lognormal.

Firstly, the mean (μ) and standard deviation (σ) for each group of data are obtained. Secondly, the parameters of each distribution for each group are calculated using predefined formulas; equations 3.2, 3.3, 3.5, and 3.6 are examples of these formulas (they are obtained from empirical methods that are discussed in chapter 2). Thirdly, the cumulative distribution function (CDF) is performed for the distribution function in the particular group and discrete events in the same group. Then, the absolute aggregated error which is equal to the summation of the differences between the actual cumulative distribution function for discrete events and the cumulative distribution function that acquired through step 3 is computed. Eventually, the PDF that has minimum aggregated error is chosen to represent wind behavior in this specific site. After performing the above mentioned procedures, the Weibull distribution function is selected since it was the one of best fit for most of the hours.

The Weibull distribution function is driven by two parameters which are shape index k and scale index c. The equations of the Weibull function and their parameters are illustrated below as [20].

$$f(v) = (\frac{k}{c}) * (\frac{v}{c})^{k-1} * \exp[-(\frac{v}{c})^{k}] \tag{3.1}$$

$$k = (\frac{\sigma}{v_m})^{-1.086} \tag{3.2}$$

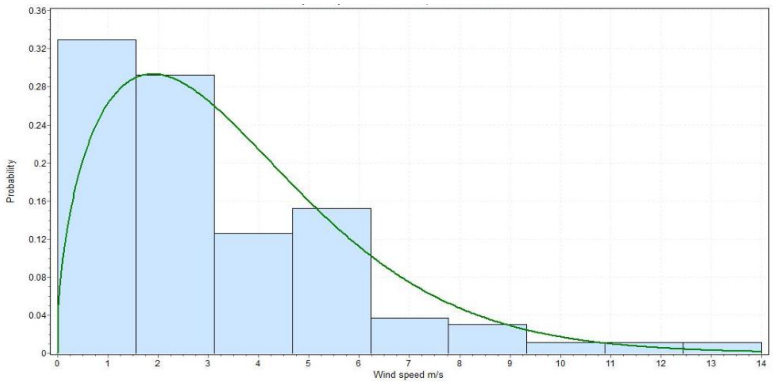

Figure 3.1: Wiebull PDF for One Sample Data

$$c = \frac{v_m}{\Gamma(1 + \frac{1}{k})} \tag{3.3}$$

Where f(v) is the Weibull distribution function of v, v is wind speed m/s, and (k , c) are the shape factor and scale factor, respectively. Average wind speed is v_m m/s ,and σ is the standard deviation. Weibull pdf mimics the behavior of wind speed. Fig.**??** shows the fitting of Weibull distribution to a group of historical data.

3.4 Solar Irradiance Model

In the same way as wind speed modeling, solar irradiance is represented in 96 groups and for each group of data, the Kolmogorov Smirnoff (K-S) algorithm is performed to find the best fitting PDF. In contrast, in some hours when the solar radiation is equal to zero, there are no PDFs obtained and the output power from PV modules is considered to be zero. A Beta distribution function is selected since it was the most fitted one for most of the hours.

The Beta distribution equation that has been used in this research is presented below, and the equations for calculating its parameters are shown as well [20].

$$f(s) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} * s^{\alpha-1} * \\ (1-s)^{\beta-1} & 0 \leq s \leq 1, \alpha and \beta \geq 0 \\ 0 & Otherwise \end{cases} \tag{3.4}$$

$$\beta = (1-\mu) * (\frac{\mu * (1+\mu)}{\sigma^2} - 1) \tag{3.5}$$

$$\alpha = \frac{\mu * \beta}{1-\mu} \tag{3.6}$$

Where f(s) is the Beta distribution function of s, s represents solar irradiance kW/m^2,and α, β are parameters of the Beta distribution function. μ, σ are the mean and standard deviation of irradiance, respectively.

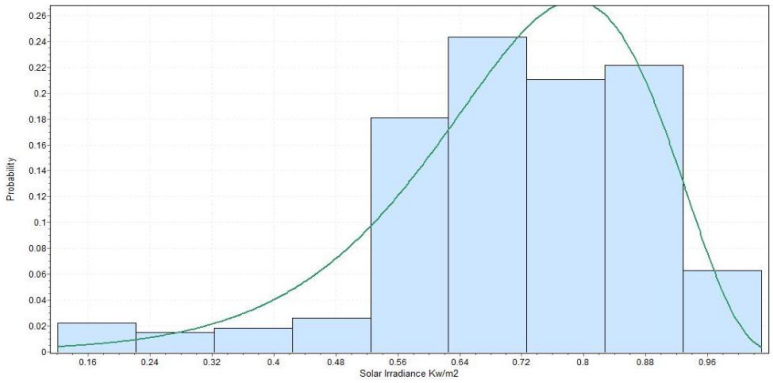

Figure 3.2: Beta PDF for One Sample Data

Fig.3.2 shows the fitting of Beta distribution to a group of historical data. It can be seen from fig.3.3 the difference between the CDF of discrete samples and the selected distribution function.

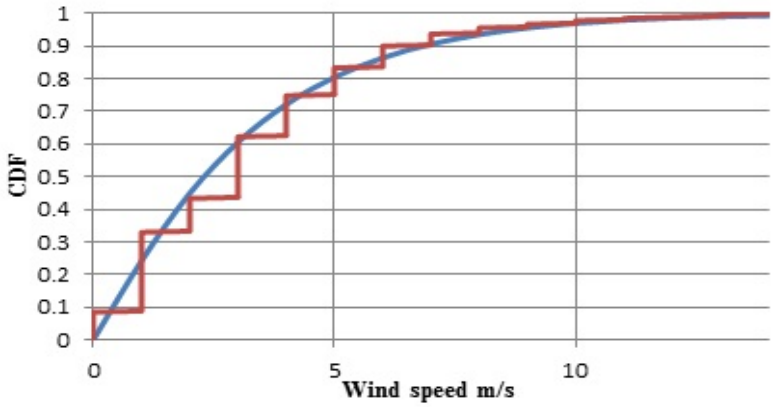

Figure 3.3: Cumulative Distribution Function

3.5 Dispatchable Units Modeling

The dispatchable and controllable units are considered to feed the system by a constant power input which represents its rated capacity at any time.

3.6 Load Modeling

To accommodate the fluctuation in the load, IEEE RTS in [35] is applied, and since the season is represented by 24 hours, 96 demand points will shape the fluctuation in the load. Every demand point is equal to the averaged value of 90 demand readings for a particular hour in a particular season.

3.7 Unsupervised Clustering Technique and Probability Calculations

When the probability distribution functions for all time segments are defined, these PDFs have to be divided into states to be incorporated in calculating the wind and PV output

power as well as power flow equations. Data clustering is mainly used here to provide the initial states that could represents the output power from renewables.

Indeed, clustering is basically applied to a large number of observations for the purpose of placing those observations that have common attributes or are relatively homogenous into groups [36]. One of the biggest advantages of clustering is that by analyzing limited numbers of observations instead of dealing with entire observations, this method is actually preferred. To quantify this partitioning in the data, the proximity of the observations is measured as a distance. Intuitively, the data that have the same pattern would be distinguished from their closeness to each other. Even though there are many clustering techniques which are based on distance or correlation measurements, K-means algorithm is mostly preferred due to its simplicity and popularity [19]. Hence, in this paper K-means is carried out to initially cluster the observations. In the following, k-means algorithm and its application are discussed.

3.7.1 K-means Algorithm

Fundamentally, this algorithm is attempting to minimize the average squared Euclidean distance calculated from one or more group members. For each cluster, there are two important elements. The first one is the center of the cluster which is called centroid or seed, and this centroid represents the core of the cluster. The second element is called the members or agents of the cluster. It is important to mention that a clustering algorithm is based on the sorting approach and the results could vary, yet these results could be enhanced using one of the global optimal techniques [19].

In the following, K-means procedures for assigning 1 by M observations into C clusters are demonstrated [37].
1- Select the initial seeds (centroids) of the clusters randomly.
2- Based on the proximity to the centroids, all the data would be assigned to the nearest seed.
3- Compute the distance between the clusters members and clusters centroid. The centroids of any cluster are calculated as 3.7 in Euclidean Space.

$$m_i = \frac{1}{|C_j|} \sum_{x_i \in c_j} X_i \tag{3.7}$$

Where, $|C(_j)|$ is the number of data points (observations) in cluster $C(_j)$. The distance from one data point X_i to a mean (centroid) m_j is calculated as follow;

$$dist(X_i, m_j) = \|X_i - m_j\| = \sqrt{(X_{i1} - m_{j1})^2 + (X_{i2} - m_{j2})^2 + ... + (X_{ir} - m_{jr})^2} \tag{3.8}$$

4- Calculate the sum of squared error (SSE),

$$SSE = \sum_{j=1}^{k} \sum_{X \in C_j} dist(X, m_j)^2 \tag{3.9}$$

C_i is the j_{th} cluster, m_j is the centroid of cluster C_j (the mean vector of all the data points in C_j), and $dist(X, m_j)$ is the distance between data point X and centroid m_j.

5- If there is no reallocating of the data or the agents into a new cluster or it is minimum as well as the sum of squared error is minimum, this means the results are converged and processing stops.

6- If the results did not converge, then recalculate the seeds and reassign the agents, and go to step 2.
Figures 3.4 and 3.5 show a piece of an actual wind and PV output power with 44 measurements and the curve obtained after clustering into nine clusters (states) for wind and eight clusters for PV output.

3.7.2 Probability Calculation

When a probability distribution function is created for wind and solar, then wind and solar output power have been initially divided into x and y states, respectively. This division

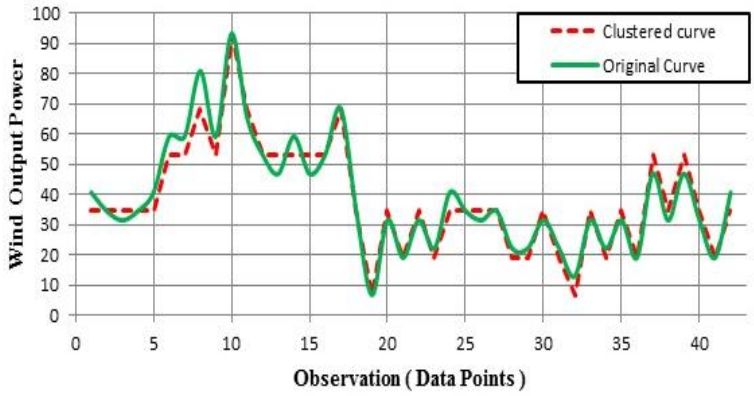

Figure 3.4: Wind Actual and Clustered Power

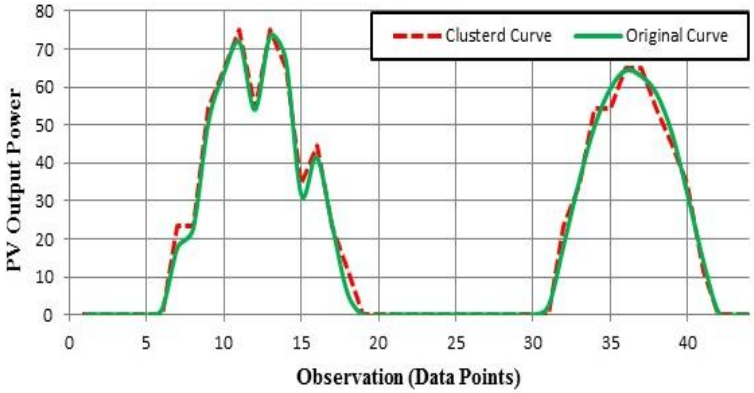

Figure 3.5: PV Actual and Clustered Power

is based on one of the unsupervised clustering techniques which is called K-means. As soon as the observations from wind and solar output are clustered, the boundaries of each cluster (state) are obtained. So, representing the states by the intervals which describe the speed limits and solar irradiance for each state are applied. The probability for each state can be obtained using the integral equations below.

$$P(v_a \leq V \leq v_b) = \int_{v_a}^{v_b} f(v).dv \tag{3.10}$$

$$P(s_a \leq S \leq s_b) = \int_{s_a}^{s_b} f(s).ds \tag{3.11}$$

It is important to mention here that in case of combining Wind and PV based DGs, the total number of states is equal to the multiplication of wind and solar states, and the probability of each state will be equal to the product of wind probability and solar probability at that corresponding state. This is because wind speed and solar irradiance are independent events.

3.8 Calculating Wind Turbine and PV Module

The output power from Wind turbines can be obtained directly from what is called wind power curves, and these curves are provided to the customers in the form of figures or equations. Likewise, PV modules have specific characteristics and these data are also provided.

3.8.1 Wind Output Power

The main function of a wind turbine is to convert the wind's kinetic energy into mechanical energy. Typically, wind turbine output power is a function of wind average speed $\bar{v}^3$, air density ρ, and the rotor swept area A. So, the aerodynamic power obtained from the wind is given by the following equation.

$$P_{wind} = \frac{1}{2}\rho A\bar{v}^3 \tag{3.12}$$

A wind power curve can be classified into 3 regions; no wind power region, de-rated region, and rated power region. No wind power region is defined as the region at which the wind

Figure 3.6: Wind Turbine Power Curve

speed is below the cut in speed of the turbine. Before the cut in speed, the turbine's blades are not able to overcome the produced friction since there is insufficient torque. When the wind speed goes above the cut in speed, turbine's output power increases dramatically to reach the rated power of the turbine. As a protective action to protect the rotor from the high forces in the turbine's structure, the wind turbine will shut down when the wind speed goes beyond the cut out speed of the turbine.

Since the wind speed is divided into states, the wind output power for each state will be acquired from the equation below [20].

$$P_w(v_{ax}) = \begin{cases} 0 & 0 \leq v_{ax} \leq v_{ci} \\ P_{rated} * \frac{v_{ax}-v_{ci}}{v_r-v_{ci}} & v_{ci} \leq v_{ax} \leq v_r \\ P_{rated} & v_r \leq v_{ax} \leq v_{co} \\ 0 & v_{ax} \geq v_{co} \end{cases} \tag{3.13}$$

Where v_{ci}, v_r, v_{co} : cut-in speed, rated speed, and cut-off speed of the wind turbine, respectively; $P_w(v_{ax})$:output power during state x, and v_{ax} : average speed of state x.

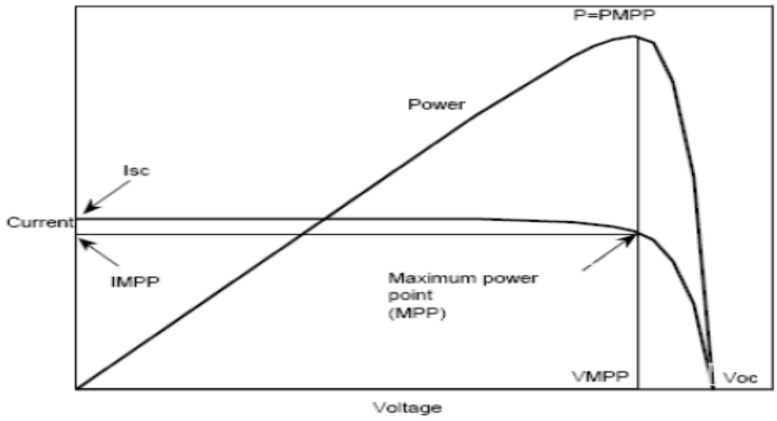

Figure 3.7: I-V Characteristics for a PV Module [7].

3.8.2 PV Modules Output Power

PV cells have the sole responsibility to convert the solar radiation into electric DC current throughout the photoelectric effect. Rating of these modules is computed under standard tests. The connection of PV modules whether in series or parallel, play an important role to scale up the voltage or the current which leads to a rise in ratings of these modules. The main parameters to determine PV power are PV module characteristics which are provided from the manufacturer, solar irradiance in the site, and the ambient temperature in the site. I-V characteristics of PV modules which are provided from the manufacturer differ from one radiation level to another and from one temperature level to another. The relationship between the solar irradiance and the short circuit current is proportional; however, the relationship between the voltage and module temperature is inverse. I-V characteristics can be graphically presented as shown in 3.7. Because of solar irradiance is divided into states, the PV output power for each state will be acquired from equations below [20].

$$Tc_y = T_A + s_{ay}(\frac{N_{OT} - 20}{0.8}) \tag{3.14}$$

$$I_y = s_{ay}[I_{sc} + K_i(T_c - 25)] \tag{3.15}$$

$$V_y = V_{oc} - K_v * Tc_y \tag{3.16}$$

$$FF = \frac{V_{MMP} * I_{MMP}}{V_{oc} * I_{sc}} \tag{3.17}$$

$$Ps_y(S_{ay}) = N_m * FF * V_y * I_y \tag{3.18}$$

Where Tc_y: cell temperature C during state y; T_A: ambient temperature C; K_v: voltage temperature coefficient V/ C; K_i: current temperature coefficient A/C; N_OT: nominal operating temperature of cell in C; FF: fill factor; N_m: number of modules; I_sc: short circuit current in A; V_{oc}: open circuit voltage in V; I_{MPP}: current at maximum power point in A; V_{MPP}: voltage at maximum power point in V; $Ps(s_{ay})$: output power during state y, and s_ay : average irradiance of state y.

3.9 Selecting Wind Turbine and PV Module

Wind turbines and PV modules that will be used in this paper are selected based on calculating the average annual capacity factor [38, 39]. The average output power of a turbine or module divided by its rated power is called the capacity factor. The wind turbine and PV modules that have a high annual capacity factor are selected among deferent types of wind turbines and PV modules.

$$CF = \frac{\sum_{t=1}^{96} \sum_{s=1}^{n} P_{out_s} * Prob_s}{P_{rated}} \tag{3.19}$$

Where t: time segments ; CF: annual capacity factor; s: state; P_{out_s}: Wind or PV output power in state s; $Prob_s$: probability of state s, and P_{rated}: rated power of wind turbine or PV module.

The average seasonal capacity factors CF for wind power and PV power are shown respectively in fig.3.8 and fig.3.9.

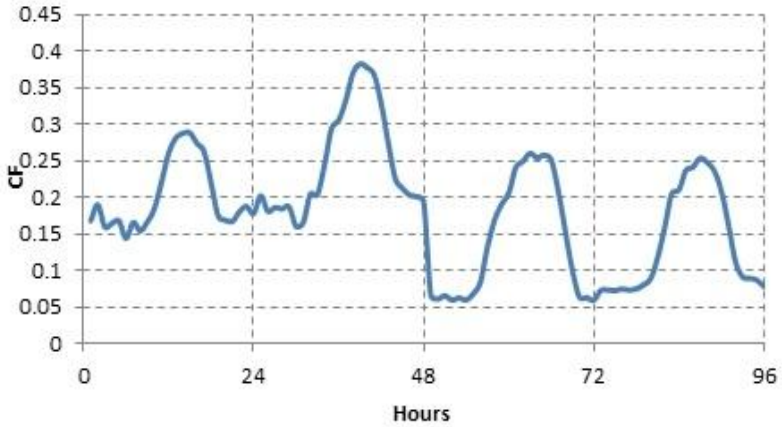

Figure 3.8: Seasonal Wind Capacity Factor

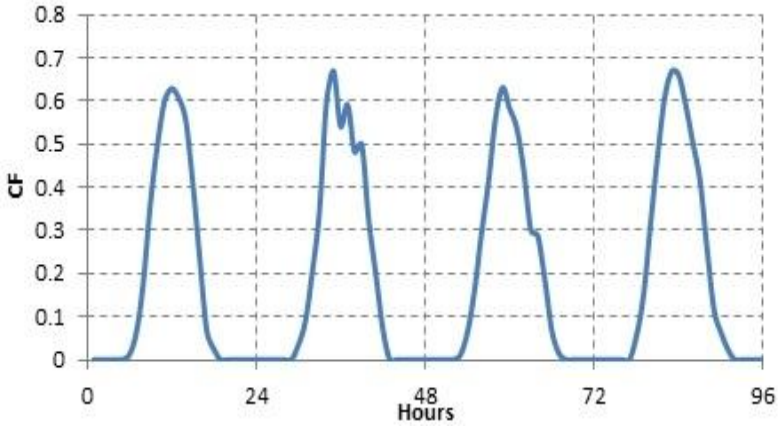

Figure 3.9: Seasonal PV Capacity Factor

3.10 Summary

This chapter presented the models of the renewable resources that have been used in this thesis. It outlined the preparation and analysis of large amounts of historical data to be suitably utilized in planning problems. The probability density functions that have used to model the behavior of wind speed and solar irradiance are discussed. The goodness of fitness technique (K-S) is illustrated and employed to properly fit the data into the proper density

function. Moreover, an unsupervised clustering technique (k-means) is implemented in this chapter to find the centroids of clustered data which are then used to find the represented states of wind and PV power. Turbine power curve and I-V characteristics of PV modules are explained to calculate the output power from these renewable sources. Lastly, capacity factor calculations are carried out to select the most suitable turbines or modules in such a location as to provide more power than others.

Chapter 4

Renewable Resources Allocation and Sizing in Distribution System Networks

4.1 Introduction

Traditionally, the electrical power is produced from the conventional generators and then transmitted to the ultimate customer throughout the transmission systems in one way direction. However, this concept is no longer accepted in the presence of distributed generation since the power may flow in two directions which leads the system to be more dynamic. Therefore, the improper selection of the location and sizing of these distributed generators may worsen the network. So, it is crucial to model the stochastic nature of wind speed and solar radiation in a good manner and then implement a comprehensive probabilistic optimization model to allocate and size these renewable resources. There are many operation constraints that need to be satisfied in the planning process. These constrains involve voltage levels at each bus, line capacity, maximum reverse power, and others. In this chapter, a mathematical problem formulation that takes into account the probabilistic nature of wind speed and solar irradiance will be presented. This model can be used for

various objective functions, such as minimizing power losses, reducing system's expenses, and maximizing system's profits.

4.2 Planning Process

The process of the allocating and sizing intermitted-based DGs is quite difficult and requires much effort in the data analysis stage before incorporating the analyzed data into the intelligent optimization tools. The procedures will start by analyzing the historical data of wind speed and solar irradiance and fitting these data to the probability density functions to model the intermittent behavior. In the same time, the demand of the system will be analyzed in such away as mentioned in chapter 3 to be in sequence with the variation in the generation level. After that, the probability density functions will be divided into many states in order to include these states with the optimal power flow equations (K-means clustering-based algorithm is used in this research to define the states). An intelligent optimization tool will be applied to solve this problem. Genetic algorithm is used because of its powerful ability to solve the complected problems.

4.3 Problem Formulation

One of the essential problem in the distribution system is renewable resources allocation and sizing due to the randomness of the generating units. To handle the uncertainty in the generation level of renewables, a probabilistic formulation is applied. This formulation accommodates the modeling of wind and PV since it performs matching or combining the load side and generation side. This formulation is incorporated with deterministic optimal power flow equations; this means that each equation is represented by composed probabilistic equations equal to the number of determined states. In other words, after computing power losses for state s, these losses are weighted based on the occurrence probability of this state in hour h from the entire year. Taking into consideration the constraints of the planning problem, DGs will be placed in the distribution system to

minimize the overall annual losses without any violation at any state. In this study, all the buses are considered to be residential as well as the DGs are working at unity power factor. Wind and solar profiles are assumed to be same at each bus for the purpose of simplicity.

4.3.1 Objective Function

The objective function is to allocate renewable resources based DGs and controllable DGs in order to minimize the overall power losses while all operating constraints are met. As mentioned previously, the planning horizon is one year, and this year is represented by four seasons. Each season is represented by 24 four hours, and these 24 hours describe the 2160 hours in its corresponding season on the basis of each season having 90 days. So, the objective function can be expressed as follows:

$$min \sum_{t=1}^{96} \sum_{s=1}^{T} P_{loss_s} * 90 * Prob_s \tag{4.1}$$

Where P_{loss_g}: Power losses at state s; $Prop_s$: probability of combined occurrence of wind, solar, and dispatchable in state s; t:time segment; and T: total number of states.

4.3.2 Constraints

Power Flow Equations

$$P_{slack} + C_{sW} * P_{rated_{W_i}} + C_{sPV} * P_{rated_{PV_i}} + Prated_{Dis_i} \tag{4.2}$$

$$-P_{D_i} = \sum_{j=1}^{N} V_{s,i} * V_{s,j} * Y_{ij} * cos(\theta_{ij} + \delta_{s,j} - \delta_{s,i})$$

$$Q_{slack} - Q_{D_i} = \sum_{j=1}^{N} V_{s,i} * V_{s,j} * Y_{ij} * sin(\theta_{ij} + \delta_{s,j} - \delta_{s,i}) \tag{4.3}$$

Where $P_{rated_{W_i}}$: Wind rated power at bus i; $P_{rated_{PV_i}}$: PV rated power at bus i; $Prated_{Dis_i}$: Dispatchable rated power at bus i; C_{sW}: The percentage of wind output

power from its rated in state s; C_{sPV}: The percentage of PV output power from its rated in state s; P_{D_i}: Power demand at bus i in time segment t; Y_{ij}: The magnitude of admittance matrix element (Y_{ij}) of the branch between node i and node j; θ_{ij}: Tha angle of admittance matrix element (Y_{ij}) of the branch between node i and node j; $\delta_{s,j}$: voltage angle of bus j in state s, and $\delta_{s,i}$: voltage angle of bus i in state s.

Branch Current Equation

$$I_{ij,s} = |Y_{ij}| * [(V_{s,i})^2 + (V_{s,j})^2 - 2 * (V_{s,i})(V_{s,j}) * cos(\delta_{s,j} - \delta_{s,i})]^{0.5} \tag{4.4}$$

Where $I_{ij,s}$ is the current flows between bus i and j in state s.

Power Loss Equation

$$P_{loss_s} = 0.5 * \sum_{i=1}^{N} \sum_{j=1}^{N} G_{ij} * [(V_{s,i})^2 + (V_{s,j})^2 - 2 * (V_{s,i})(V_{s,j}) * cos(\delta_{s,j} - \delta_{s,i})] \tag{4.5}$$

Where G_{ij} is the real part of the admittance matrix element Y_{ij} of the branch between node i and node j.

Voltage Limits at Each Bus Except Slack Bus

$$V_{min} \leq V_{s,i} \leq V_{max} \tag{4.6}$$

Slack Bus Voltage and Angle (at Bus 1)

$$V_{s,1} = 1.0 \quad \delta_{s,1} = 0.0 \tag{4.7}$$

Feeder Capacity Limit

$$0 \leq I_{s,ij} \leq I_{ij_{max}} \tag{4.8}$$

Where $I_{ij_{max}}$ is the maximum permissible current in the feeder between bus i and j.

Maximum Penetration of DGs in the System

According to technical specifications of Hydro One in [40] the maximum penetration level for DGs in a distribution system is equal to the summation of 60% of maximum MVA rating of the single transformer and the minimum station load. This constraint is based on the maximum power reserve that substation equipment can handle.

$$\sum_{i=1}^{N} P_{rated_{W_i}} + \sum_{i=1}^{N} P_{rated_{PV_i}} + \sum_{i=1}^{N} Prated_{Dis_i} \leq 0.3 \sum_{i=1}^{N} P_{D_i} + 0.6\ MVA_{_{single_{Transformer}}} \quad (4.9)$$

Maximum Penetration on Each Bus [40]

$$P_{rated_{W_i}} + P_{rated_{PV_i}} + P_{rated_{Dis_i}} \leq 10\ MW \quad (4.10)$$

Discrete Size of DGs

For wind and dispatchable DGs, one more constraint should be added which represents the discretization in these two types of DGs. Unlike wind and dispatchable-based DGs, PV output power can be arranged by adding or removing some modules.

$$P_{rated_{W_i}} = a * \sum_{a=0}^{m} P_{rated\ of\ wind\ turbine} \qquad a = 0, 1, 2, ..., m \quad (4.11)$$

$$P_{rated_{Dis_i}} = a * \sum_{a=0}^{m} P_{rated\ of\ disp\ unit} \qquad a = 0, 1, 2, ..., m \quad (4.12)$$

4.4 Application of Genetic Algorithm in Optimal Allocation Problems

A genetic algorithm is a meta-heuristic searching algorithm that has positive credit in solving complicated problems [41]. It actually mimics the natural evolution process. In

GA population, a large number of chromosomes (initial candidate solutions) are generated. Each string or each chromosome is composed of a number of genes. The most important parameters of GA are initial population, fitness function, and genetic operators (i.e. selection, crossover, and mutation). At any generation, the fitness function will be evaluated for each string, and then these strings will be ranked based on its evolution to the objective function. If it is a maximization problem, the chromosomes that have high cost functions have a great opportunity to be selected as parents for the offspring chromosomes since they have high probability for copying the best results. This process is called the selection process for reproduction. To produce a new offspring string from two individuals, a crossover step takes place, and the selected parents exchange some parts of their corresponding string to set up a new chromosome that carries the parent's characteristics. A mutation is performed by changing some binary digits from 1 to 0 and vice versa to produce a new solution from an individual. These procedures will be repeated until the fitness function converges; then the chromosomes are decoded to present the final variables of the objective function.

Before beginning GA processing, the strings should be prepared in such a way as to be compatible with GA format. Since the chromosome represents a candidate solution and the problem is locating and sizing for DGs, the chromosome will be composed of multiple genes and each genes contains a vector from two components. The first component represents the location in which DGs are allocated, and the second component represents the size of the connected DG. The number of components in a given chromosome is equal to the number of candidate locations for allocating DGs multiplied by two. This is can be shown in table 4.1.

Table 4.1: Representation of Chromosome and its Genes

Gene 1		Gene 2				Gene N	
binary sitting *loc*#1	binary sizing *loc*#1	binary sitting *loc*#2	binary sizing *loc*#2			binary siting *loc*#*N*	binary sizing *loc*#*N*

4.5 Case Study

The system that has been used in this study, as shown in fig.4.1, is the well-known 33 bus radial distribution system presented in [42]. The total peak system demand is 4.37 MVA and base voltage is 12.66 KV. The main source that feeds the system is the substation which is connected at bus 1. All loads are assumed to be residential loads. The load is assumed to follow IEEE RTS system as [35], and some modification in this model is presented in a load modeling subsection in chapter 3. Selecting candidate buses criteria is beyond the scope of this work since it requires data based on many technical, governmental, environmental, and economic issues. In this thesis, candidate locations for connecting DGs are bus 5, 8, 12, 18, 25, and 29. The maximum current carrying capability for each feeder is mentioned in [42]. The wind turbine that has the highest capacity factor among the different turbines is selected, and it has a rating of 100 kW. The specifications of the wind turbine is illustrated in table 4.2. Moreover, dispatchable DGs have ratings of 100 kW for each as well. PV module specifications are presented in table 4.3. Four different scenarios are presented in this work to assess the probabilistic model. These scenarios include wind based DGs, PV based DGs, wind and PV based DGs; and wind, PV, and dispatchable based DGs. In evaluating the objective function, one more scenario is added to determine the base case results.

4.5.1 Optimal Sizing and Allocation Outcomes

Based on the states of wind and PV output power that has been acquired from the clustering method as shown in 4.4 and the probabilistic optimal power flow model, the results of allocating and sizing renewable based DGs are illustrated in table 4.5. The clustering technique shows that wind power can be represented by 11 states while 10 states are enough to model PV behavior. This problem formulation is tested and applied for four scenarios. Regardless of the type of DG in each scenario, whether it is a single type or a combined type, the objective function is affected by integrating these DGs in the network. In the first scenario, the annual energy losses are reduced by 26.86% from its base case

Table 4.2: Wind Turbine Specifications [43]

Parameter	**Discription**
Rotor Diameter	20 m
Number of Blades	3
Working Wind Speed	325 m/s
Cut-in Wind Speed	3 m/s
Rated Wind Speed	12 m/s
Rated Output Power	100 kW
Rated Output Voltage	DC 460 V
Generator Type	3 Phase Permanent Magnet Generator

Table 4.3: PV Module Specifications [44]

Parameter	**Discription**
nominal power(+/-5%)	75.0 W
Voltage at PMaX	46.9 V
Current at PMaX	1.6 A
Open circuit Voltage	60.1 V
Short circuit current	1.82 A
Temperature Coefficient of Voc	(-0.2%/C)
Temperature Coefficient of Isc	(+0.04%/C)
Nominal cell operating temperature	43 C

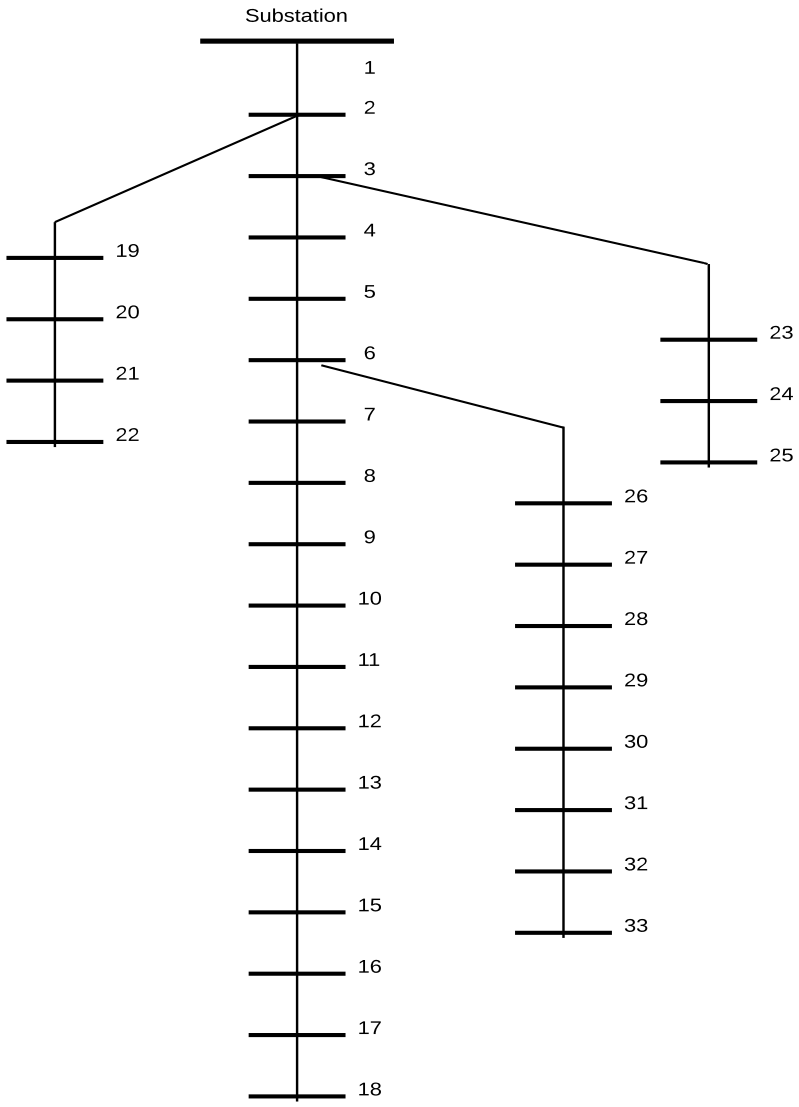

Figure 4.1: Well-Known 33 Distribution Bus System

when wind turbines with a 3.8 MW total rating (penetration level) are connected to the grid. In case of PVs only, annual energy losses are reduced by 26.86% when 2.9 MW rated PV modules are utilized. It is clear from these two scenarios that wind turbines are higher in loss reduction than PV modules, and this can be interpreted by the fact that

Table 4.4: Wind and PV States

Wind Output Clusters						
Cluster	**1**	**2**	**3**	**4**	**5**	**6**
Centriod	0	5.32	12.734	22.34	35.29	46.64
Cluster	**7**	**8**	**9**	**10**	**11**	
Centriod	58.83	69.82	80.12	91.77	100	
PV Output Clusters						
Cluster	**1**	**2**	**3**	**4**	**5**	**6**
Centriod	0	7.32	17.76	29.18	36.93	48.23
Cluster	**7**	**8**	**9**	10		
Centriod	56.28	62.44	69.94	75		

almost 40%-45% from the annual output power from PVs is zero since there is no radiation during the night. When wind turbines and PV modules are combined, the results reveal that the cost function is reduced by almost 31.75%. The share of wind turbines and PV modules from 3.79 MW DG penetration is 63.5% and 36.5%, respectively. Due to the high uncertainty in wind and PV output power, dispatchable DGs provide a significant loss reduction over the wind and PV based DGs. This is because the output power from dispatchable units is considered to be fixed throughout the planning horizon. Therefore, case four shows that approximately half of the annual energy losses can be deducted by incorporating three types of DGs. 40.87% wind form total penetration, 29.15% PV form total penetration, and 29.97% dispatchable form total penetration are mixed in order to achieve this significant loss reduction. In comparison with scenario three, dispatchable units in case four contribute an almost 20% reduction in annual energy losses over using only wind and PV based DGs.

To assure that the probabilistic optimization solution does not violate any of the operation constraints, voltages at each bus, line loading, and the reverse power are tested in most effective states. Most effective states include two cases; the first case is when the

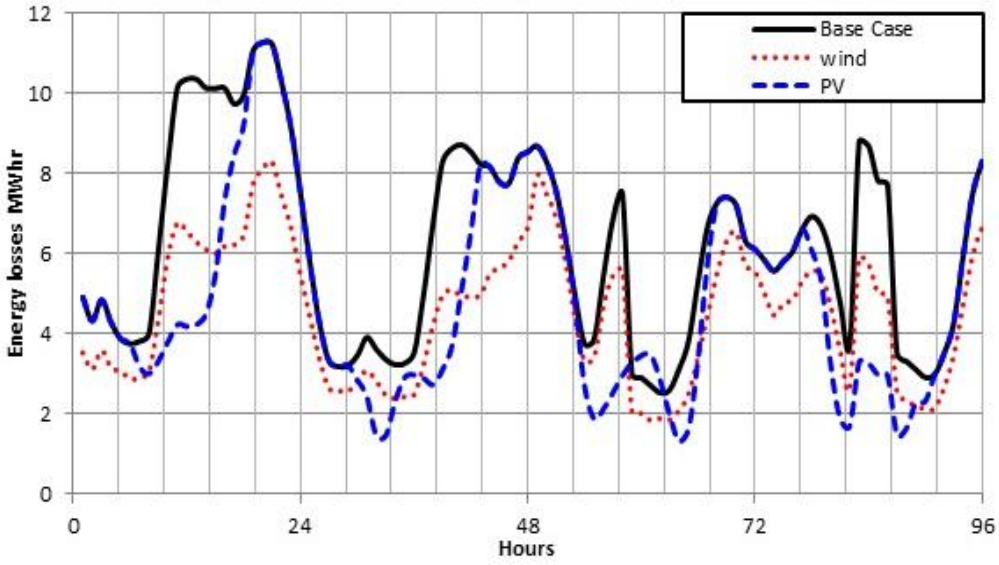

Figure 4.2: Reduction in Energy Losses in Scenario 1 and 2

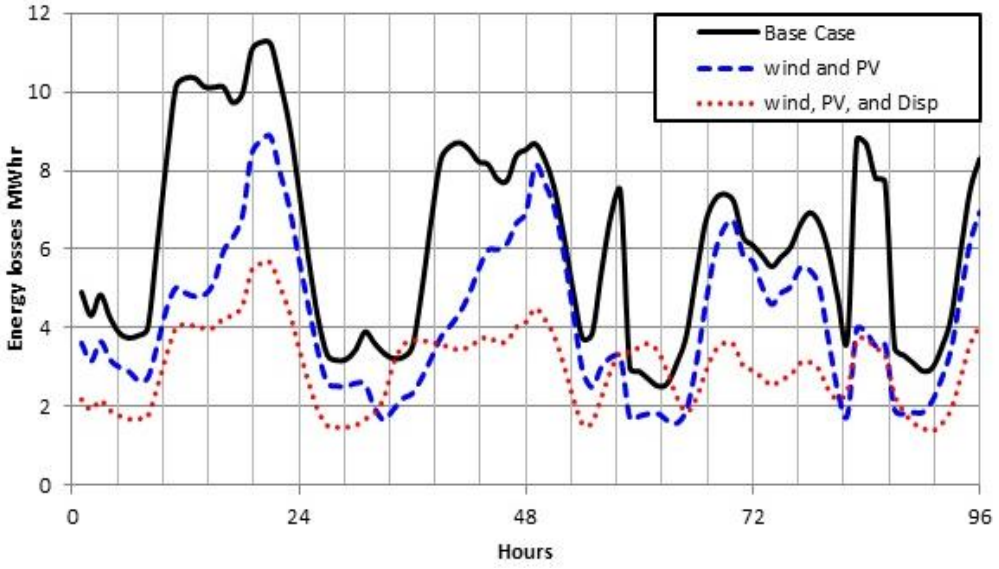

Figure 4.3: Reduction in Energy Losses in Scenario 3 and 4

Table 4.5: Placement and Sizing for Various DG Technologies

	Candidate Location	No DG	Wind	PV	Wind-PV		Wind-PV-Disp.		
DG1	Bus# 5	0	0.8	0.62	0.1	0.67	0	0	0.6
DG2	Bus# 8	0	0.1	0.1	0.2	0	0	1.011	0
DG3	Bus# 12	0	0.9	0.7	0.9	0	0.1	0	0.1
DG4	Bus# 18	0	0.1	0.09	0	0.13	0	0.05	0
DG5	Bus# 25	0	0.6	0.63	0.4	0.11	0.6	0	0
DG6	Bus# 29	0	1.3	0.76	0.8	0.48	0.8	0.01	0.4
AEL(MWhr)		594.6	434.88	466.1	405.77		290.44		
Losses Reduction %		0	26.86	21.6	31.75		51.15		
DG Penetration(MW)		0	3.8	2.9	2.4	1.39	1.5	1.07	1.1
DG total installation %		0	100	100	63.5	36.5	40.87	29.15	30

wind and PV outputs are at the rated power, and the load is at the minimum level. The second case is when there is no power from wind and PV, and the load at the high level. The results show that the voltage at each bus is improved when the wind and PV at rated power are connected to the grid with no violation in each voltage at each bus. In addition, the loading at each line in case 1 is still kept below the rated capacity of each line. In most feeders, the integration of renewable resources reduce the contingency of these feeders. Power reserve is tested as well, and it is below the maximum reverse power in case 1; whereas, it is zero in case 2. Figures 4.4, 4.5, and 4.6 show these constraints in both cases.

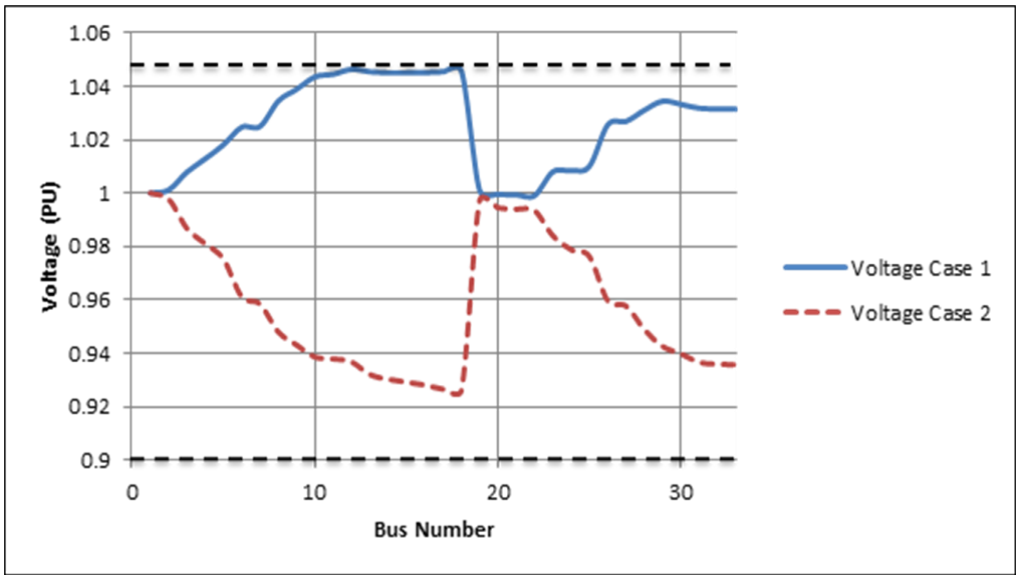

Figure 4.4: Reduction in Energy Losses in Scenario 1 and 2

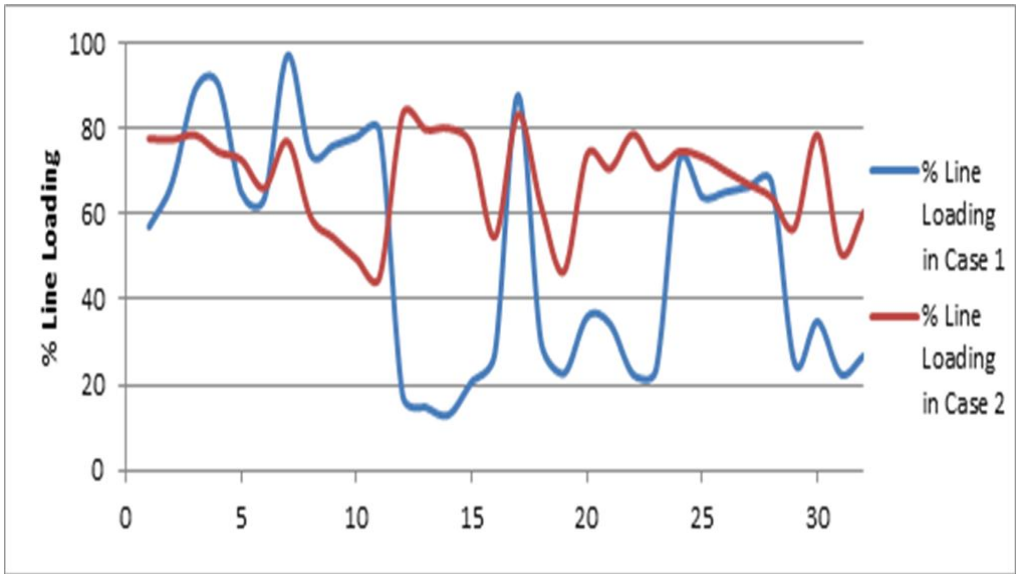

Figure 4.5: Reduction in Energy Losses in Scenario 1 and 2

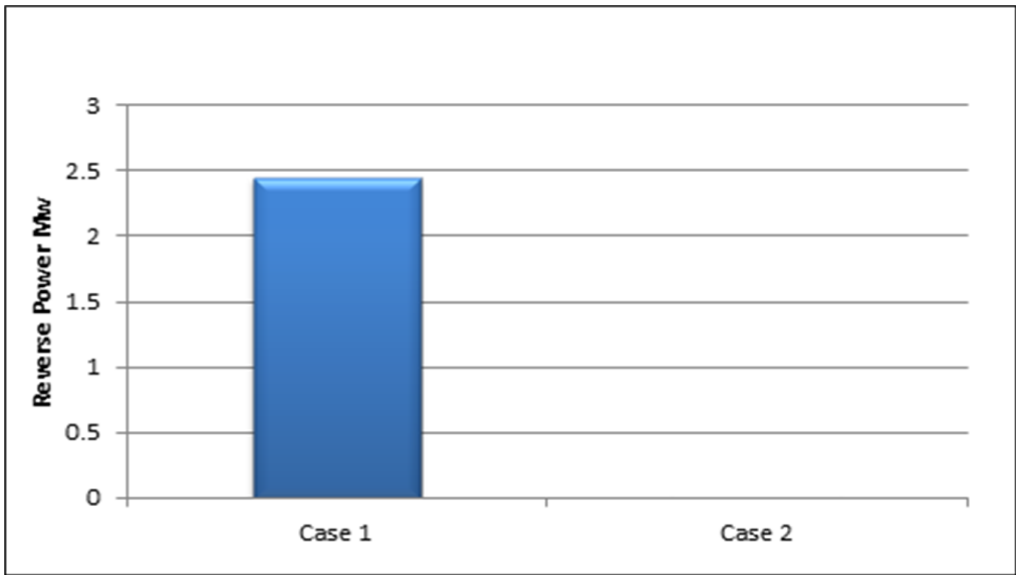

Figure 4.6: Reduction in Energy Losses in Scenario 1 and 2

4.6 Conclusion

In this chapter, a comprehensive probabilistic optimization model is presented. This model is aimed to optimally allocate and size the renewable resources in conventional distribution system to minimize the annual energy losses. The results showed that renewable resources are able to contribute magnificently in the energy losses reduction. Genetic Algorithm is used to solve the optimization problem incorporated with optimal power flow equations, and technical constraints have been considered, such as the thermal limit of the feeders, voltage limits at each bus, and the maximum penetration level of these DGs in the distribution system. All of the constrains are satisfied in each state including the the most extreme states.

Even though the probabilistic model proves its effectiveness and robustness in the planning problem, the complexity in analyzing a large number of states is high. In addition, the computational analysis takes long time to provide the answers. Therefore, finding a method that can reduce the number of representing states is very magnificent. The next chapter

proposes a new algorithm to determine the minimum number of representing states. This method can be used in reliability analysis to reduce the number of overall system states and in operation analysis to reduce the time of the analysis.

Chapter 5

Minimum Representing States for the Renewable Resources

5.1 Introduction

As explained earlier in chapter 2, in order to extract the output power from wind and PV based DGs, wind and PV output power should be divided into many states as probabilistic model. These states vary and depend on accuracy and complexity of selected states in power flow, reliability, and optimization calculations. Large wind or PV states mean that the overall system states will be so huge that the analytical reliability evaluation is almost impossible to perform, and the operation analysis that requires a fast processing (computational) time will take longer. Therefore, the motivation behind this work is to reduce the overall system states and execution time while the accuracy in results is maintained.

5.2 State Reduction Using Rounding Technique

In this thesis, the states are initially selected based on a K-means algorithm. However, these states can be truncated by rounding the states that have low probability of occurrence to the states that have high probability of occurrence. The reduction of these states gives a positive credit since it considerably saves on execution time of the calculations. A large number of states means more execution time and more complexity in reliability and planning analysis, so determining the minimum number of states that represent the output power from renewable resources while maintaining the error in the results (objective function and DG penetration) within an acceptable range is proposed in this section. The main valuable contribution of this algorithm is considerable reduction in wind and PV states as well as computing time while keeping the accuracy of the results high.
A rounding technique has been used in this study due to its flexibility and simplicity. This technique basically rounds up or rounds down the initial states to the nearest newly assigned states [45]. The rounding technique follows these general equations:

$$P(C_j) = \frac{C_k - C_i}{C_k - C_j} P(C_i) \tag{5.1}$$

$$P(C_k) = \frac{C_i - C_j}{C_k - C_j} P(C_i) \tag{5.2}$$

For state i which falls between the rounded states k and j. Rounding increments have the main contribution in rounding error. This method basically depends on the linear sharing of the previous states into the new rounded states. Fig. 5.1 shows a graphical representation for the rounded method.

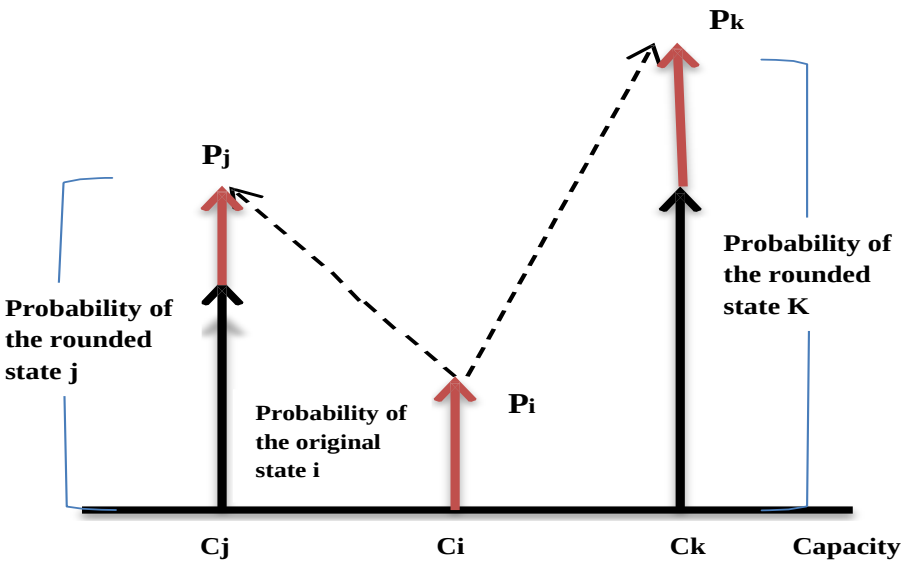

Figure 5.1: Representation for Rounding Method

5.3 The Proposed Algorithm for Finding Minimum States

Since the previous algorithms for finding the states of wind speed and solar irradiance consume more time in calculation, the idea of determining minimum states that could represent these two parameters and reduce the computational time has seen the light. The advantages of this algorithm are that this method is fast, starts from a solid base (clustering technique), and provides accurate results. The proposed algorithm is a step ahead from the clustering technique since any cluster method is dealing with the first level of statistical analysis (original data), while the proposed method incorporates the second level of statistical analysis (low probability, high probability, mean, and variance). The proposed method is used for different clustering topologies (hourly, seasonally, and yearly). Optimal DG placement is an important problem in power system long term planning, and this problem is formulated in order to test the robustness of the proposed method. DG sizing and placement are done also for the different time scales.

As mentioned before, the main two tested points are the objective function (Annual energy losses) and the penetration level of DGs, and they should adhere to an acceptable limit of error while the process of minimizing the states is being applied. The procedures of this algorithm follow these steps:

1- Read the historical data from wind speed and solar irradiance (it depends here on the test type; whether it is hourly, seasonally, or on a yearly basis).

2- Extract wind output power and solar output power from the wind power curve and IV characteristics of PVs, respectively.

3- Divide wind and PV output powers into many clusters by using K-means algorithm.

4- After clustering, and based on the centroid and boundaries of a cluster, wind speed and solar irradiance is divided into states (No. state = No. cluster).

5- Define the probability of each state for each time segment in the hourly/seasonally/yearly model using equations 3.10 and 3.11.

6- Build set A which contains all the states and their probabilities.

7- Run the probabilistic optimization model using Genetic Algorithm.

8- Store the objective function and DG penetration (Value X1, for Annual Energy Losses and Value Y1, for DG penetration).

9- Construct set B (set B = All the rounded staes except in the first iteration, set B = set A) and let X2=X3, Y2=Y3 except in the first iteration, X2=X1, Y2=Y1.

10- Perform the state reduction technique for the lowest probability state (rounding technique) using equations 5.1 and 5.2.

11- Construct set C which includes the rounded states and their probabilities.

12- Run the probabilistic optimization model using Genetic Algorithm.

13- Store the objective function and DG penetration (Value X3, for Annual Energy Losses and Value Y3, for DG penetration).

14- If the error between the values X1 and X3, or the error between the values Y1 and Y3 are greater than or equal 2.5%, set B is considered to be the minimum states that represent the behavior of wind speed and solar irradiance and value X2 and Y2 are the annual energy losses and DG penetration, respectively.

15- If the error between the values X1 and X3, or the error between the values Y1 and Y3

are smaller than or equal 2.5%, set B is equal set C , X2 is the annual energy losses, and Y2 is the DG penetration, and go to step 9.
The selection of the threshold is based on that the default margin of the error for the random data is plus or minus 2.5% for a 95% confidence level [46]. So, this threshold is selected to keep the confidence level of the annual energy losses and DG penetration high. These steps are clearly illustrated by the flowchart in fig 5.2. After performing the proposed technique, minimum states that effectively represent wind speed and solar irradiance with high accuracy and low computational time are obtained.

5.4 Case Study

The outcomes of the proposed method are extensively discussed in this section. Since the proposed algorithm follows certain steps, the findings from this method will be preceded by the results of the clustering technique and the probabilistic optimization model. Therefore, the outcomes will be discussed in subsections based on their priority in processing.

5.4.1 Clustering Technique Outcomes

As mentioned previously, the K-means algorithm is carried out in order to determine the initial states of wind speed and solar irradiance. Figures 3.4 and 3.5 show a piece of an actual wind and PV output power with 44 measurements and the curve obtained after the clustering in nine clusters (states) for wind and eight clusters for PV output.

Hourly Clusters

After performing K-means algorithm for each time step (96 points of data), the results reveal that there are 68 hours converged at 6 states (clusters); whereas, 28 hours converged at 7 states. In fact, these states vary from one hour to another, and this makes the accuracy of the results high since each hour is dealing with its own states obtained from the statistical history. Some of the selected states for selected hours are presented in table 5.1. State

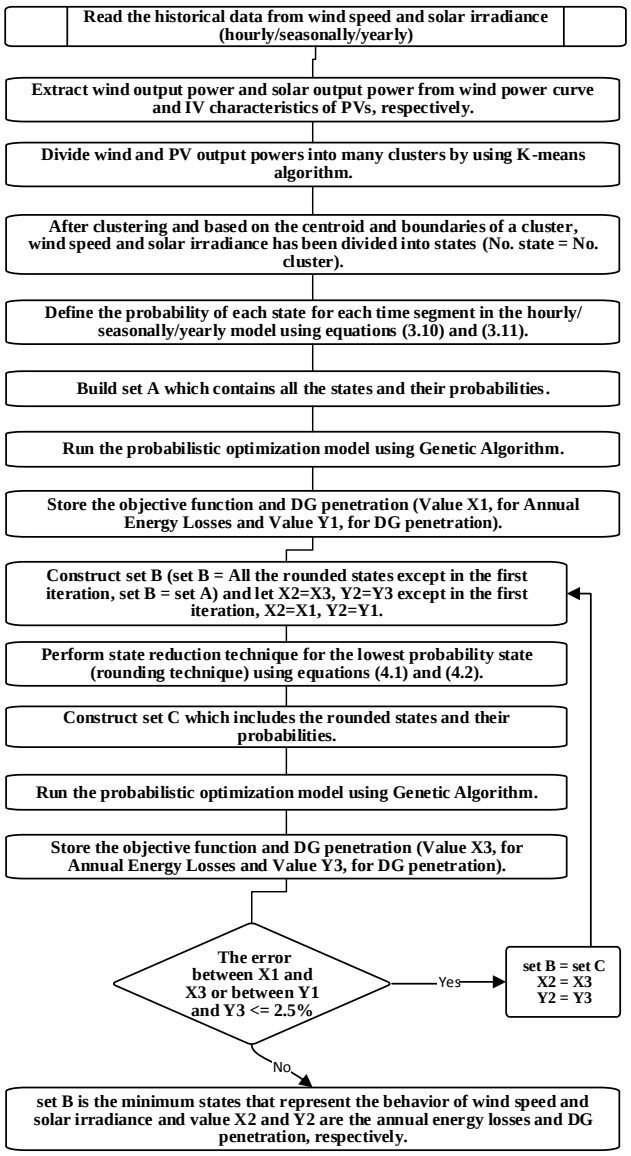

Figure 5.2: Flowchart of the Proposed Algorithm

limits are also acquired, and they are substituted in the integration limits in equations 3.10 and 3.11 to calculate the probability of each state. Table 5.2 and 5.3 present the probability of each state and their boundaries.

Table 5.1: Selected States for Selected Hours

State\Hour	**12**	**15**	**36**	**43**	**61**	**92**
1	0	0	1.800693	0	2.632318	0
2	6.849799	7.680765	8.154222	0.202073	21.1617	4.949407
3	14.8152	24.87701	18.8988	18.63378	30.27426	9.298181
4	36.44715	39.15854	39.88	22.15089	46.7599	18.33106
5	42.64068	47.85357	49.43949	32.66909	69.29094	24.31704
6	65.66151	67.03327	55.61958	49.43949	87.55241	35.4102

Seasonal Clusters

K-means results for seasonal clustering show that in case of wind, all of the various seasons (Fall, Winter, Spring, Summer) could be historically represented by 9 states to model the fluctuation in wind output power. However, in the case of PV, 8 states are adequate to represent the behavior of the output power from PV modules. The centroids of each cluster for the Fall season are presented in table 5.2 for wind and table 5.3 for PV. In addition, the probability of each cluster or state is clearly demonstrated in these two tables.

Yearly Clusters

When all historical data are clustered, the wind output power can be divided into 11 states to precisely represent the intermittent nature of wind speed. However, after clustering all historical solar irradiance data for the site under investigation, the output power from PV modules is fair enough to be represented by 10 states. The probability of each state

Table 5.2: One Sample of Wind States and Their Probabilities

State	Rated Power %	Wind Speed Level (m/s)		Probability
		Lower Limit	Upper limit	
1	0	0	3	0.3043
2	7.12	3.01	4.28	0.2261
3	20.64	4.29	5.42	0.1759
4	35.22	5.43	6.91	0.1597
5	52.58	6.92	8.54	0.0872
6	67.42	8.55	9.59	0.0243
7	79.07	9.6	10.62	0.011
8	91.41	10.63	11.9	0.00516
9	100	12	25	0.001734

Table 5.3: One Sample of PV States and Their Probabilities

State	Rated Power %	Solar irradiance level (Kw/m2)		Probability
		Lower Limit	Upper limit	
1	0	0	0.11	0.01299
2	11	0.12	0.19	0.02911
3	23.35	0.2	0.301	0.06164
4	34.19	0.302	0.42	0.12033
5	44.36	0.43	0.54	0.20699
6	54.10	0.55	0.67	0.28469
7	65.10	0.68	0.82	0.23741
8	75	0.83	1	0.05

is calculated using the integral equations. These probabilities depend basically on the boundaries of each state.

Table 5.4: Wind and PV States for Yearly Representation

Wind Output Clusters						
Cluster	**1**	**2**	**3**	**4**	**5**	**6**
Centriod	0	5.32	12.734	22.34	35.29	46.64
Cluster	**7**	**8**	**9**	**10**	**11**	
Centriod	58.83	69.82	80.12	91.77	100	
PV Output Clusters						
Cluster	**1**	**2**	**3**	**4**	**5**	**6**
Centriod	0	7.32	17.76	29.18	36.93	48.23
Cluster	**7**	**8**	**9**	10		
Centriod	56.28	62.44	69.94	75		

5.4.2 Proposed Algorithm Outcomes

The proposed technique for finding minimum states that represent wind and solar behavior in power planning and operation problems is carried out, and interesting results are obtained. Tables 5.7, 5.6, and 5.5 demonstrate these findings.

Hourly Findings

Firstly, when the wind turbines are connected only to the grid, the algorithm showed that the wind output power can be represented by 3 states for each hour while the results are still close to the original states results. The error in annual energy losses is less than 2.5% as shown in fig 5.3 while the reduction in execution time is 36.16% from the original states computational time. In addition, the level of penetration did not violate the limit of error when the states are truncated to five states. When the representing states are reduced by more than 3 states, a significant mismatch between these states will occur, and the added values to state probabilities throughout the eliminated states are not enough to compensate

for this mismatch.

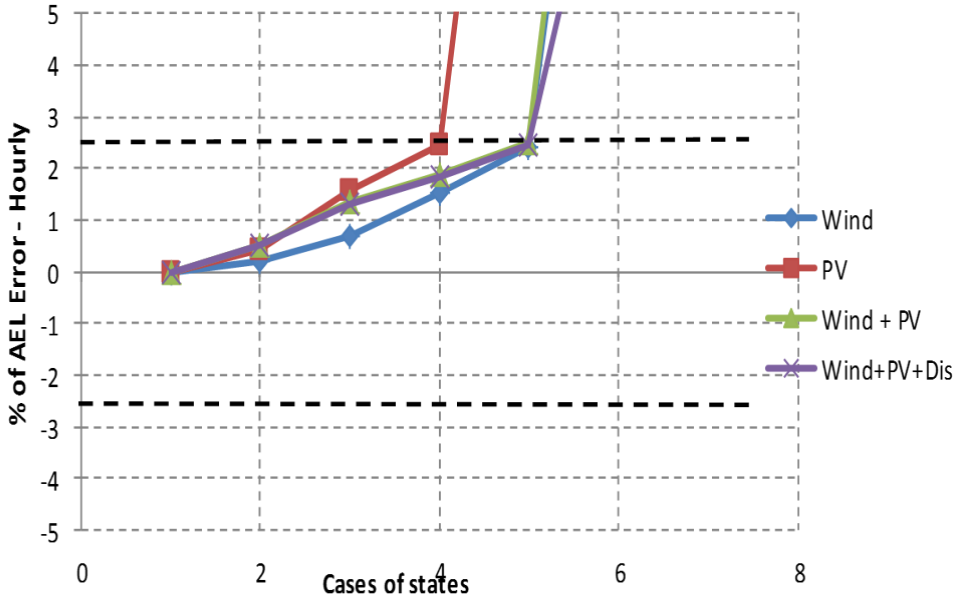

Figure 5.3: Error in Annual Energy Losses for Hourly Clustering

In addition, the results reveal that in the case of PV, the output power from PV modules can be represented by 3 states as well with a very small difference compared to 6 state outcomes. Interestingly, almost 34.27% of the execution time is deducted by representing PV output power by 3 states. As it is shown in fig 5.4, the error in penetration level did not exceed the impermissible limits and it is kept within the safe area. As expected, this algorithm will prove its strength in the case of mixed renewable sources, and this is because in the case of wind and solar, the states that represent the combination between them are large.

Compared to the results obtained from the initial clustering technique, 9 states are able to represent the hourly behavior of combined wind and PV power with precision in results. The previous states were 42 states which are the product of seven states of wind and six

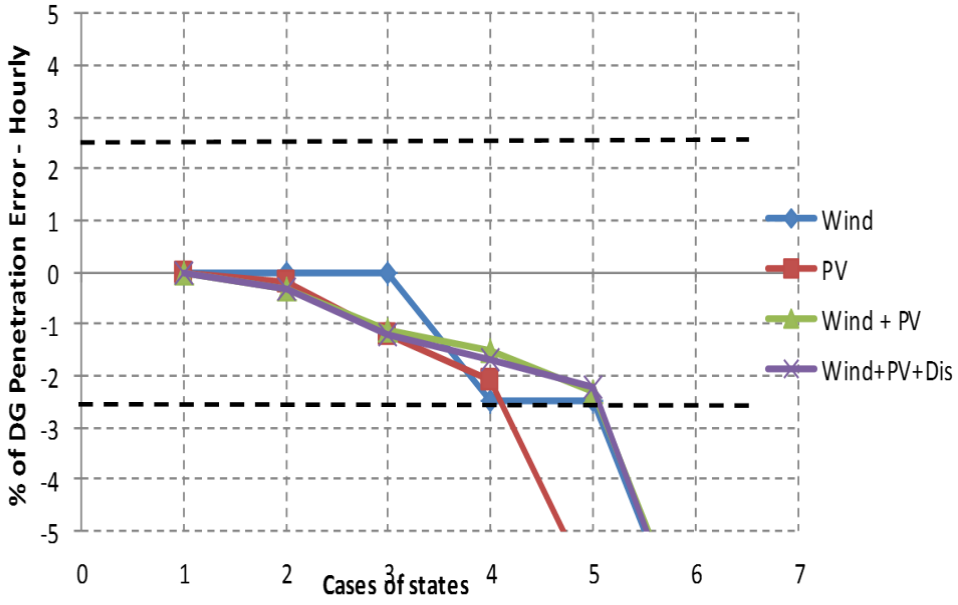

Figure 5.4: Error in DG Penetration for Hourly Clustering

states of PV. Indeed, the new 9 states are a result of multiplying 3 states of wind and 3 states of PV. The computational time is reduced considerably, by more than two thirds of the original case (42 states). Furthermore, in the last scenario, the minimum states that represent the hourly behavior of renewable resources with conventional distributed generation are 9 states. As well, the penetration level was still kept in the permissible region. The computational time for performing the optimization problem is reduced by approximately 81.11% compared to the original execution time. Interestingly, the results showed that the locations of the distributed generation have not affected by using this methodology which gives a credit to this method.

Seasonal Findings

In the case of wind turbines only being plugged into the network, the intermittent nature of wind speed can be modeled throughout 5 states seasonally. Annual energy losses did

Table 5.5: Proposed Algorithm Results - Hourly Clustering

DG	Wind					
Case - States	1 - 7	2 - 6	3 - 5	4 - 4	5 - 3	6 - 2
AEL (MWhr)	416.38	417.173	419.189	422.85	426.75	491.594
Penetration(MW)	4	4	4	3.9	3.9	3.7
AEL error(%)	0	0.19	0.67	1.53	2.43	15.3
Penetration error (%)	0	0	0	-2.5	-2.5	-7.5
Execution time reduction(%)	0	9.435	17.77	26.965	36.16	44.88
DG	PV					
Case - States	1 - 6	2 - 5	3 - 4	4 - 3	5 - 2	6 - 1
AEL (MWhr)	443.25	445.164	450.32	454.476	529.57	N/A
Penetration(MW)	3.3	3.29341	3.26087	3.23213	3.10734	N/A
AEL error(%)	0	0.43	1.57	2.47	16.3	N/A
Penetration error (%)	0	-0.2	-1.2	-2.1	-6.2	N/A
Execution time reduction(%)	0	11.6	21	34.27	48.7	N/A
DG	Wind and PV					
Case - States	1 - 42	2 - 30	3 - 20	4 - 12	5 - 9	6 - 6
AEL (MWhr)	387.84	389.789	393.108	395.311	397.703	465.315
Penetration(MW)	3.83	3.81778	3.78833	3.7734	3.74389	3.56943
AEL error(%)	0	0.5	1.34	1.89	2.48	16.65
Penetration error (%)	0	-0.32	-1.1	-1.5	-2.3	-7.3
Execution time reduction(%)	0	39.41	64.51	75.09	83.24	94.81
DG	Wind, PV, and Disp.					
Case - States	1 - 42	2 - 30	3 - 20	4 - 12	5 - 9	6 - 6
AEL (MWhr)	274.84	276.304	278.432	280.02	281.8	332.736
Penetration(MW)	3.76	3.74763	3.71542	3.69715	3.67906	3.49572
AEL error(%)	0	0.53	1.29	1.85	2.47	17.4
Penetration error (%)	0	-0.33	-1.2	-1.7	-2.2	-7.56
Execution time reduction(%)	0	38.41	63.81	72.91	81.11	91.21

not violate the acceptable error limit when the wind is represented by 5 states as shown in fig 5.5. Moreover, wind penetration levels were not affected when the states are truncated to five states. There is a considerable saving in computational time, almost 35%, when the states are reduced to 5 states.

However, the results reveal that in the case of PV, the output power from PV modules can be represented by 4 states, while the accuracy obtained is equivalent to an 8 state outcome. If the representing states go below 4 states, a significant mismatch between these states will be introduced. The computational time is obviously drops by 47.3%, which is approximately half of the original states (8 states) execution time. 2.84 MW from PV modules which is the penetration level of these modules when the states are minimized to 4 states. Fig 5.6 shows the penetration level of PV and wind based DGs is located within the acceptable zone when the states are reduced.

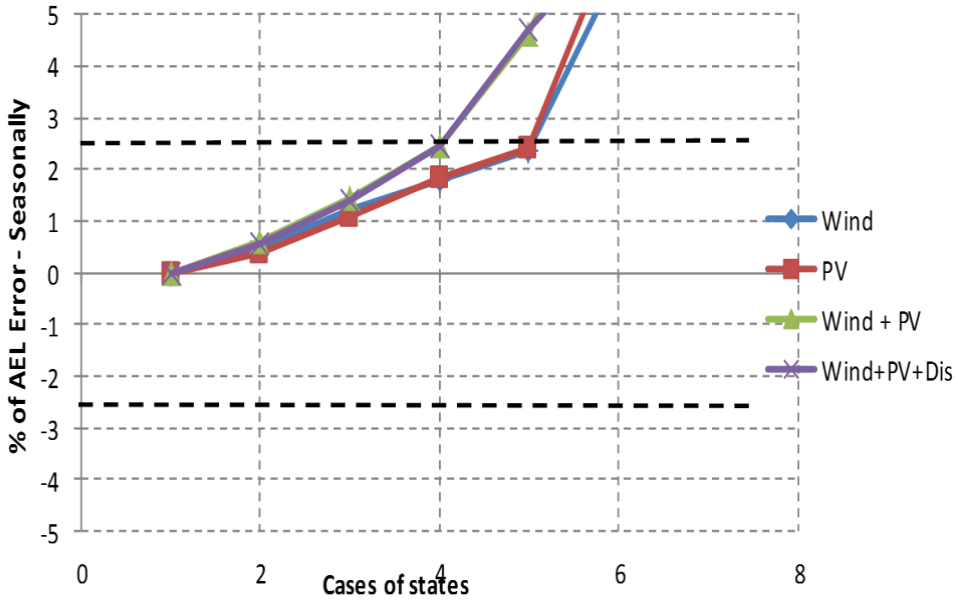

Figure 5.5: Error in Annual Energy Losses for Seasonal Clustering

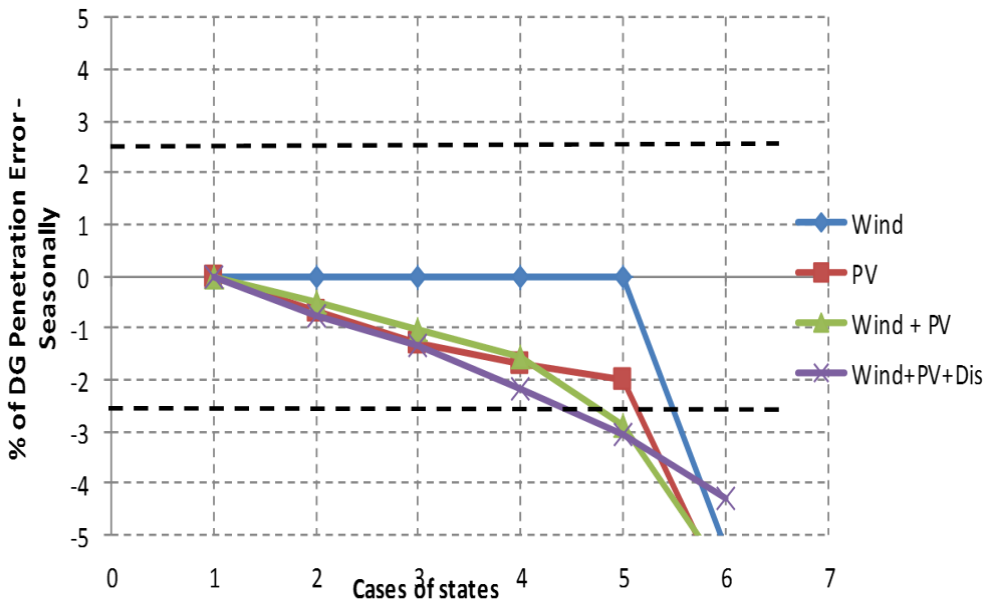

Figure 5.6: Error in DG Penetration for Seasonal Clustering

In the case of that wind and PV are both connected to the system, 72 representing states can be reduced to 20 states to combine and precisely model the fluctuation of wind speed and solar irradiance. The previous states were 72 states which are the product of nine states of wind and eight states of PV. However, the new 20 states are a result of multiplying 5 states of wind and 4 states of PV. The computational time is reduced considerably, by 73% from the original case (72 states). Furthermore, in the last scenario, the minimum states that represent the behavior of renewable resources with dispatchable units are 20 states. As well, the penetration level was still maintained in the allowable region. The computational time for performing the optimization problem is reduced by nearly 72% from the original execution time. The locations of the DGs are kept same, and there are no changes in these locations after performing the reduction method.

Table 5.6: Proposed Algorithm Results - Seasonal Clustering

DG	Wind					
Case - States	1 - 9	2 - 8	3 - 7	4 - 6	5 - 5	6 - 4
AEL (MWhr)	434.88	436.934	440.296	442.761	445.346	462.147
Penetration(MW)	3.8	3.8	3.8	3.8	3.8	3.6
AEL error(%)	0	0.47	1.23	1.78	2.35	5.9
Penetration error (%)	0	0	0	0	0	-5.26316
Execution time reduction(%)	0	8.335	16.67	25.865	35.06	43.78
DG	PV					
Case - States	1 - 8	2 - 7	3 - 6	4 - 5	5 - 4	6 - 3
AEL (MWhr)	466.14	467.918	471.325	474.829	477.651	499.614
Penetration(MW)	2.9	2.88	2.86	2.85	2.842	2.72
AEL error(%)	0	0.38	1.1	1.83	2.41	6.7
Penetration error (%)	0	-0.67	-1.3	-1.7	-2	-6.2
Execution time reduction(%)	0	10.2	19.6	32.87	47.3	59.43
DG	Wind and PV					
Case - States	1 - 72	2 - 56	3 - 36	4 - 20	5 - 16	6 - 9
AEL (MWhr)	405.77	408.301	411.657	416.004	425.335	442.497
Penetration(MW)	3.79	3.771	3.75	3.73	3.68	3.58
AEL error(%)	0	0.62	1.43	2.46	4.6	8.3
Penetration error (%)	0	-0.5	-1.05	-1.58	-2.9	-5.8
Execution time reduction(%)	0	38.2	63.3	73.88	82.03	93.6
DG	Wind, PV, and Disp.					
Case - States	1 - 72	2 - 56	3 - 36	4 - 20	5 - 16	6 - 9
AEL (MWhr)	290.44	292.105	294.534	297.704	304.764	317.421
Penetration(MW)	3.67	3.642	3.62	3.59	3.557	3.512
AEL error(%)	0	0.57	1.39	2.44	4.7	8.5
Penetration error (%)	0	-0.76	-1.36	-2.17	-3.07	-4.3
Execution time reduction(%)	0	37.4	62.8	71.9	80.1	90.2

Yearly Findings

At the beginning, when the wind turbines are only connected to the grid, the algorithm showed that the wind output power can be represented yearly by 6 states while the results are still close to the original states results (11 states). The error in annual energy losses is less than 2.5% as shown in fig 5.7 while the reduction in execution time is 42% of the original states computational time. In addition, the level of penetration did not violate the limit of error as well when the states are truncated to six states.

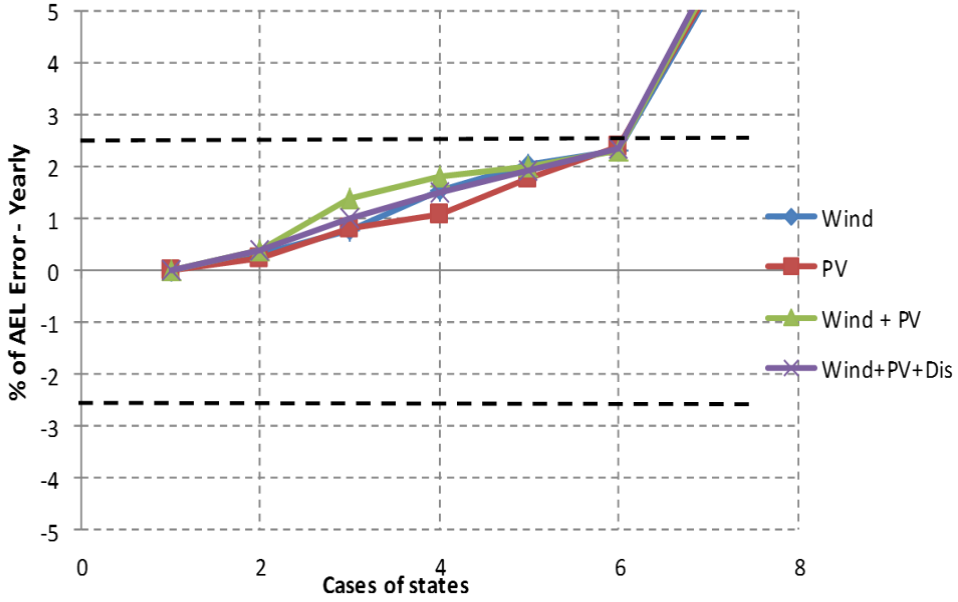

Figure 5.7: Error in Annual Energy Losses for Yearly Clustering

In addition, the results reveal that in the case of PV, the output power from PV modules can be represented by 5 states as well, with very slight difference compared to 10 state outcomes. Interestingly, almost 58.4% of the execution time is deducted by representing PV output power by 5 states. As it is shown in fig 5.8, the error in penetration level did

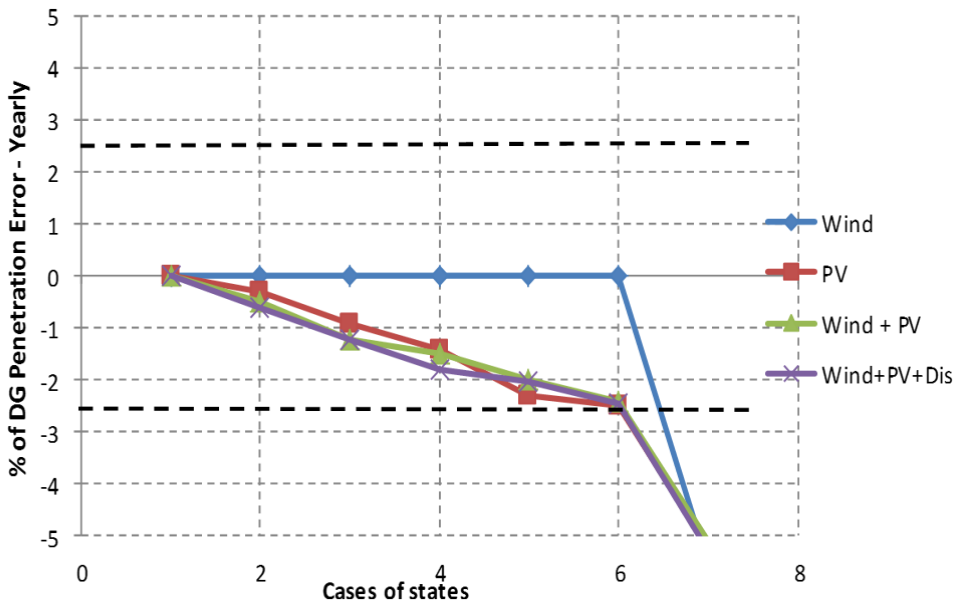

Figure 5.8: Error in DG Penetration for Yearly Clustering

not exceed the impermissible limits and it is kept within the safe area.
Compared to the results obtained from the initial clustering technique, 30 states are able to represent the yearly behavior of combined wind and PV powers with precision in results. The previous states were 110 states which are the product of eleven states of wind and ten states of PV. Indeed, the new 30 states are a result of multiplying 6 states of wind and 5 states of PV. The computational time is reduced considerably, by 92% from the original case (110 states). Furthermore, in the last scenario, the minimum states that represent the yearly behavior of renewable resources with conventional distributed generation are 30 states. As well, the penetration level was still kept in the permissible region. The computational time for performing the optimization problem is reduced by approximately 89% from the original execution time. DGs locations have not affected when the reduction method is performed.

Table 5.7: Proposed Algorithm Results - Yearly Clustering

DG	Wind						
Case-States	1-11	2-10	3-9	4-8	5-7	6-6	7-5
AEL (MWhr)	452.25	453.6	455.8	459.2	461.6	462.8	477.5
Penetration(MW)	3.6	3.6	3.6	3.5	3.6	3.6	3.4
AEL error(%)	0	0.31	0.796	1.53	2.03	2.3	5.3
Penetration error (%)	0	0	0	0	0	0	-5.5
Execution reduction(%)	0	7.285	15.62	24.815	34.01	42.73	50.31
DG	PV						
Case-States	1-10	2-9	3-8	4 -7	5-6	6-5	7-4
AEL (MWhr)	480.2	481.3	484.3	485.4	488.9	492.1	507.6
Penetration(MW)	2.83	2.82	2.8	2.7	2.7	2.76	2.69
AEL error(%)	0	0.23	0.83	1.07	1.78	2.4	5.4
Penetration error (%)	0	-0.3	-0.9	-1.4	-2.29	-2.49	-5.2
Execution reduction(%)	0	9.17	18.57	31.84	46.27	58.4	62.47
DG	Wind and PV						
Case-States	1-110	2-72	3-64	4-49	5-36	6-30	7-20
AEL (MWhr)	424.34	426.1	430.2	432.25	433	434.3	449.03
Penetration(MW)	3.63	3.6	3.58	3.57	3.55	3.54	3.45385
AEL error(%)	0	0.41	1.38	1.83	2	2.3	5.5
Penetration error (%)	0	-0.5	-1.2	-1.48	-2	-2.41	-5.1
Execution reduction(%)	0	37.16	62.26	72.84	80.99	92.56	93.32
DG	Wind, PV, and Disp.						
Case - States	1 - 110	2 - 72	3 - 64	4 - 49	5 - 36	6 - 30	7 - 20
AEL (MWhr)	307.84	309.076	311.012	312.528	313.898	315.281	326.102
Penetration(MW)	3.52	3.49901	3.47792	3.45742	3.45064	3.43549	3.34283
AEL error(%)	0	0.4	1.02	1.5	1.93	2.36	5.6
Penetration error (%)	0	-0.6	-1.21	-1.81	-2.01	-2.46	-5.3
Execution reduction(%)	0	36.37	61.77	70.87	79.07	89.17	92.51

5.4.3 The Impact of Different Time Representations on Annual Energy Losses

In order to determine whether the annual energy losses are affected by different time representations or not, various time scales are examined which are: hourly, seasonally, and yearly. The results reveal that in the case of wind based DGs, the error between the yearly and seasonal representation for annual energy losses is 3.84%, and the error is 4.44% between the seasonal and hourly representation. As a result, the error between hourly and yearly representation reaches 8.28%. Therefore, the seasonal representation is preferable to represent the annual behavior of wind speed and solar irradiance due to its ability to provide an accurate results which are very close to the most accurate representation (hourly representation), as well as its ability to make the analysis of the problems easier and faster. Figure 5.9 shows the annual energy losses for various time representations while table 5.8 presents the percentage error between these representations.

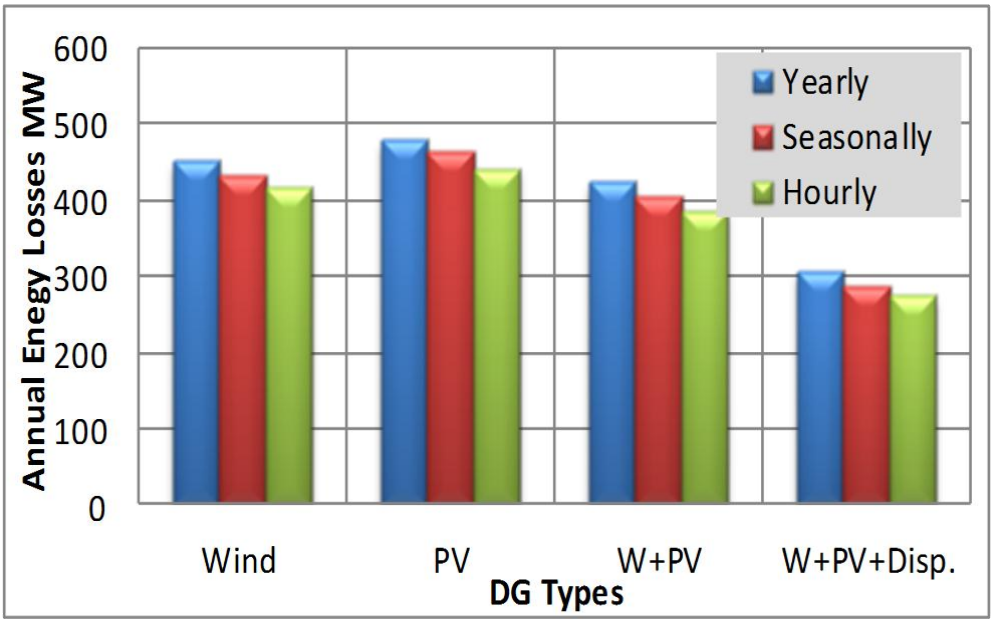

Figure 5.9: Different Time Representations for Annual Energy Losses

Table 5.8: Percentage % of Errors for Different Representations

Wind				**Wind+PV**			
Rep.	**Hour**	**Season**	**Year**	**Rep.**	**Hour**	**Season**	**Year**
Hour	-	4.44	8.28	**Hour**	-	4.62	9
Season	4.44	-	3.84	**Season**	4.62	-	4.37
Year	8.28	3.84	-	**Year**	9	4.37	-
PV				**Wind+PV+Disp.**			
Rep.	**Hour**	**Season**	**Year**	**Rep.**	**Hour**	**Season**	**Year**
Hour	-	5.16	9.1	**Hour**	-	5.67	11.32
Season	5.16	-	2.94	**Season**	5.67	-	5.65
Year	9.1	2.94	-	**Year**	11.32	5.65	-

5.4.4 Comparing Between the Proposed Method and a Supervised Clustering Method

As it is known from the previous chapter, the proposed technique is initiated by using an unsupervised clustering method where the number of the centroids is unknown. However, supervised clustering algorithms should take into account the number of required seeds as an input for the algorithm. Therefore, it is very important to compare the proposed method with the supervised clustering in order to evaluate the robustness of the proposed method. The seasonal case is enforced to be represented by 9, 8, 7, 6, and 5 states for the wind. As well, PV power is enforced also to be represented by 8, 7, 6, 5, and 4 states. After that, the results are compared to the proposed method outcomes. As shown in figures 5.10 and 5.11, the proposed method has an advantage over the supervised clustering method since the wind power and PV power can be represented by lower states compared with the supervised clustering. Supervised clustering violates the limit which is 2.5% from the base value, when the wind is represented below 7 states and the PV by 7 states, as well.

In contrast, the proposed method is able to represent the wind and PV power by 5 and 4 states, respectively. This is due to the fact that the proposed method goes beyond the clustering itself and it is dealing with the second level of statistical analysis (low probability and high probability) while the supervised cluster is treating the bulk of the data which is defined as the first layer of the statistics.

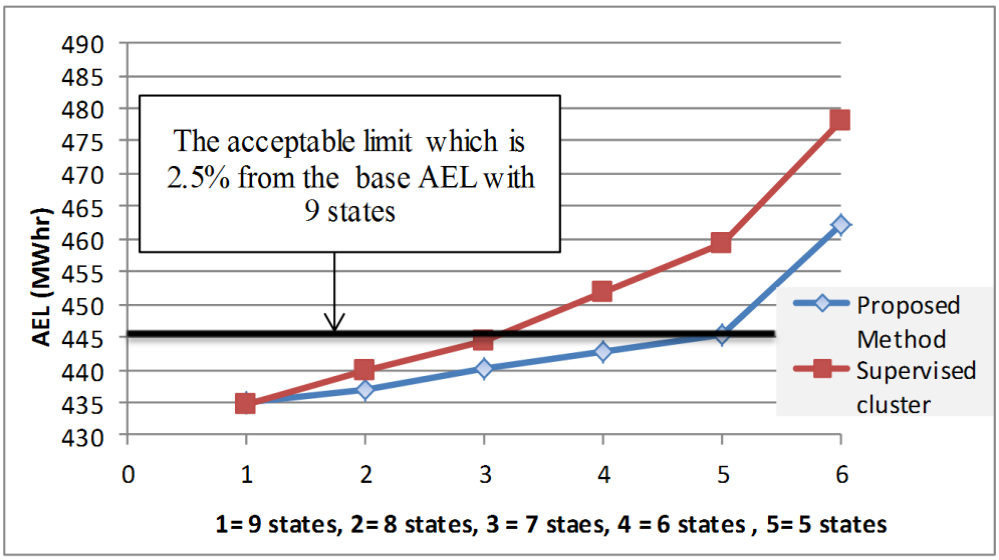

Figure 5.10: Proposed Method vs. Supervised Clustering for Wind Based DGs

5.5 Conclusion

The proposed algorithm which is designed to find the minimum number of states which represent the intermittent nature of wind speed and solar radiation is presented in this chapter. These results are divided into many subsections since the proposed method follows certain steps. It began with the initial states that are obtained from the k-means clustering method. Then, the results from the allocation and sizing for different types of DGs are presented. Annual energy losses and total DG penetration is used in the proposed

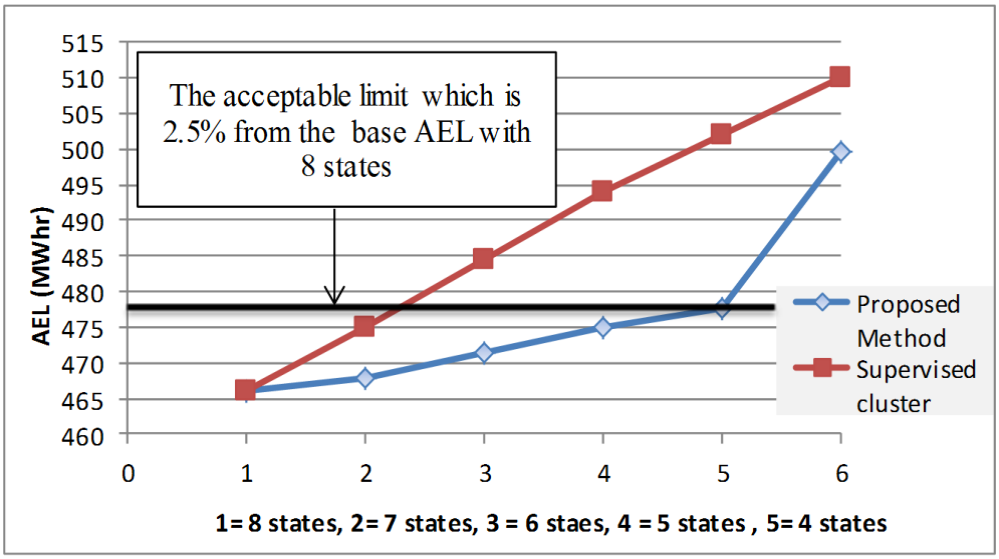

Figure 5.11: Proposed Method vs. Supervised Clustering for PV Based DGs

method to evaluate the robustness of the method. The intersecting outcomes from the proposed method are exhibited with different time representations. After that, an assessment analysis is done to identify the differences between various time representations. Lastly, the proposed algorithm is compared with a supervised clustering method to evaluate the strength of the presented technique.

Chapter 6

Conclusions and Future Work

6.1 Summary of the Thesis

The rapid growth in utilizing renewable resources as an alternative energy source to cope with increasing demand and to feed remote areas has played a vital role in recent decades. The reason behind this change in the traditional system is due to the benefits that have been gained from the transition from the centralized power system to a decentralized structure. Moreover, the global concern about the emissions from the fossil fuel burning leads governments to issue considerable regulations to mitigate these emissions through many incentives to employ renewable resources. Many planning and operation problems should be investigated carefully to tackle the randomness inherent in the intermittent nature of wind speed and solar radiations.

This thesis presents a methodology to handle and model the uncertainties produced from the intermittent behavior of wind speed and solar irradiance in long term planning problems. In addition, the state of the art in selecting the represent states for a multi-model is presented. This method requires analyzing a large set of historical data to examine the nature of the site under investigation. Many planning problems can utilize this model in many aspects.

A multi-state probabilistic optimization model is presented in this research to optimally

allocate and size the renewable resources in the distribution system to minimize the annual energy losses. A genetic algorithm has been used as an intelligent optimization solver. The results revealed the ability of the DGs in reducing the losses in the grid. The presented model made a combination between the variability in the load side as well as in the generation side. However, the large number of representing states raised an issue regarding the complexity in the analysis stages and it requires large computational time. So, it is essential to fined a method that can reduce the number of representing states while the accuracy of the results are still in an acceptable limit.
A new algorithm is presented in this thesis to tackle the issue of the large number of states. This method aimed to fined the minimum representing states of the wind speed and the solar irradiance, and it proves its usefulness and effectiveness in the results. Different time representations are tested in order to define the suitable number of states at each time segment. Considerable savings in the processing time can be clearly seen from the outcomes of this method. Many cases are tested to identify the effects of multiple time representations in the final results.

6.2 Main Contributions of Thesis

The main contributions that have been achieved and presented in this thesis are as follows:

1- Robust models for renewable resources (wind and PV) in planning and operation problems are presented. In these models, preparation and analyzing of a large set of historical data for the purpose of implementing these data into the planning problems was well investigated.

2- This thesis proposed an algorithm to determine the minimum number of representing states that represent the intermittent behavior of wind speed and solar radiation in power planning problems. This method proves its strength over the traditional methods that define the number of states. The proposed method mitigates the difficulties in reliability analysis and operation problems that require minimum effective states.

3- A comprehensive probabilistic optimization planning model is presented in this thesis. This model allocates and sizes the renewable distributed generation in the distribution

system to be suitable for many objective functions. Moreover, this thesis describes the way that genetic algorithm structure can handle the probabilistic optimization model.

4- An evaluation methodology is addressed to test the impact of various time representations in planning problems. These time representations are: hourly, seasonally, and yearly.

6.3 Future Work

The work presented in this thesis can be extended to address some suggestions such as:

1- Apply the proposed method for reliability analysis with renewable resources to reduce the number of overall system states.

2- Apply the probabilistic planning model to address the effect of renewable resources on power planning problems such as network reinforcements.

3- The proposed method and the probabilistic model can be used to allocate different smart grid devices like DGs, PEV, and storage systems to accommodate multiple objective functions.

Appendix A

33 Bus Distribution System Data

Table A.1: Well-Known 33 Bus Distributions System Data - A

Bus		**Feeder Parameters**		**Bus Loads**	
From	**To**	**R (p.u.)**	**X(p.u.)**	**P (MW)**	**Q (MVAr)**
1	**2**	0.005753	0.002932	0.06782	0.040692
2	**3**	0.03076	0.015667	0.061038	0.027128
3	**4**	0.022836	0.01163	0.081384	0.054256
4	**5**	0.023778	0.01211	0.040692	0.020346
5	**6**	0.051099	0.044112	0.040692	0.013564
6	**7**	0.01168	0.038608	0.135641	0.06782
7	**8**	0.106779	0.077061	0.135641	0.06782
8	**9**	0.064264	0.04617	0.040692	0.013564
9	**10**	0.065138	0.04617	0.040692	0.013564
10	**11**	0.012266	0.004056	0.030519	0.020346
11	**12**	0.02336	0.007724	0.040692	0.023737
12	**13**	0.091592	0.072063	0.040692	0.023737
13	**14**	0.033792	0.04448	0.081384	0.054256
14	**15**	0.036874	0.032818	0.040692	0.006782
15	**16**	0.046564	0.034004	0.040692	0.013564
16	**17**	0.080424	0.107378	0.040692	0.013564
17	**18**	0.045671	0.035813	0.061038	0.027128
2	**19**	0.010232	0.009764	0.061038	0.027128
19	**20**	0.093851	0.084567	0.061038	0.027128
20	**21**	0.02555	0.029849	0.061038	0.027128
21	**22**	0.04423	0.058481	0.061038	0.027128
3	**23**	0.028152	0.019236	0.061038	0.03391
23	**24**	0.056028	0.044243	0.284845	0.135641
24	**25**	0.055904	0.043743	0.284845	0.135641

Table A.2: Well-Known 33 Bus Distributions System Data - B

Bus		**Feeder Parameters**		**Bus Loads**	
From	**To**	**R (p.u.)**	**X(p.u.)**	**P (MW)**	**Q (MVAr)**
6	**26**	0.012666	0.006451	0.040692	0.016955
26	**27**	0.017732	0.009028	0.040692	0.016955
27	**28**	0.066074	0.058256	0.040692	0.013564
28	**29**	0.050176	0.043712	0.081384	0.047474
29	**30**	0.031664	0.016128	0.135641	0.406922
30	**31**	0.060795	0.060084	0.101731	0.047474
31	**32**	0.019373	0.02258	0.142423	0.06782
32	**33**	0.021276	0.033081	0.040692	0.027128

Appendix B

IEEE-RTS System Data

Table B.1: IEEE-RTS System - Daily load in Percent of Weekly Peak

Day	Peak Load
Monday	93
Tuesday	100
Wednesday	98
Thursday	96
Friday	94
Saturday	77
Sunday	75

Table B.2: IEEE-RTS System - Hourly Peak Load in Percent of Daily Peak

Season	**winter weeks**		**summer weeks**		**spring/fall weeks**	
Weeks	**1 -8 & 44 - 52**		**18 -30**		**9-17 & 31 - 43**	
Hour	**Wkdy**	**Wknd**	**Wkdy**	**Wknd**	**Wkdy**	**Wknd**
12-1 am	67	78	64	74	63	75
2-1	63	72	60	70	62	73
2-3	60	68	58	66	60	69
3-4	59	66	56	65	58	66
4-5	59	64	56	64	59	65
5-6	60	65	58	62	65	65
6-7	74	66	64	62	72	68
7-8	86	70	76	66	85	74
8-9	95	80	87	81	95	83
9-10	96	88	95	86	99	89
10-11	96	90	99	91	100	92
11-12	95	91	100	93	99	94
12-13	95	90	99	93	93	91
13-14	95	88	100	92	92	90
14-15	93	87	100	91	90	90
15-16	94	87	97	91	88	86
16-17	99	91	96	92	90	85
17-18	100	100	96	94	92	88
18-19	100	99	93	95	96	92
19-20	96	97	92	95	98	100
20-21	91	94	92	100	96	97
21-22	83	92	93	93	90	95
22-23	73	87	87	88	80	90
23-24	63	81	72	80	70	85

Table B.3: IEEE-RTS System - Weekly Peak Load in Percent of Annual Peak

Week	Peak Load	Week	Peak Load
1	86.2	27	75.5
2	90	28	81.6
3	87.8	29	80.1
4	83.4	30	88
5	88	31	72.2
6	84.1	32	77.6
7	83.2	33	80
8	80.6	34	72.9
9	74	35	72.6
10	73.7	36	70.5
11	71.5	37	78
12	72.7	38	69.5
13	70.4	39	72.4
14	75	40	72.4
15	72.1	41	74.3
16	80	42	74.4
17	75.4	43	80
18	83.7	44	88.1
19	87	45	88.5
20	88	46	90.9
21	85.6	47	94
22	81.1	48	89
23	90	49	94.2
24	88.7	50	97
25	89.6	51	100
26	86.1	52	95.2

References

[1] U. E. I. Administration, "Annual energy outlook 2013," May 2013, http://www.eia.gov/forecasts/aeo/MT_electric.cfm.

[2] U. E. P. Agency, "Overview of greengouse gases," September 2013, http://www.epa.gov/climatechange/ghgemissions/gases/co2.html.

[3] J. Sawin *et al.*, "Renewables global status report: 2013 update," *Paris, France: REN21 Secretariat*, 2013.

[4] "Canadian wind energy association," available online at http://www.canwea.ca/.

[5] "Natural resouces canada," available online at http://www.nrcan.gc.ca/home.

[6] M. Elnashar, "Enabling high wind penetration in electrical grids," Ph.D. dissertation, University of Waterloo, 2011.

[7] Y. M. Atwa, "Distribution system planning and reliability assessment under high dg penetration," Ph.D. dissertation, University of Waterloo, 2010.

[8] K. Purchala, R. Belmans, L. Exarchakos, and A. Hawkes, "Distributed generation and the grid integration issue," *KULeuven, Imperial College London*, 2006.

[9] T. Ackermann, G. Andersson, and L. Söder, "Distributed generation: a definition," *Electric power systems research*, vol. 57, no. 3, pp. 195–204, 2001.

[10] R. Viral and D. Khatod, "Optimal planning of distributed generation systems in distribution system: A review," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 7, pp. 5146–5165, 2012.

[11] T. H. M. El-Fouly, "Wind farms production: Control and prediction," Ph.D. dissertation, University of Waterloo, 2007.

[12] J. H. Y. P. D. H. L. D.-B. P. Luukkonen, P. Bateman, "National survey report of pv power applications in canada 2012," June 2013. [Online]. Available: http://www.cansia.ca/sites/default/files/201306_cansia_2012_pvps_country_report_long.pdf

[13] P. A. Lynn, *Electricity from sunlight: an introduction to photovoltaics*. Wiley. com, 2011.

[14] W. Omran, "Performance analysis of grid-connected photovoltaic systems," Ph.D. dissertation, University of Waterloo, 2010.

[15] Y. D. Arthur, K. B. Gyamfi, and S. Appiah, "Probability distributional analysis of hourly solar irradiation in kumasi-ghana," *International Journal of Business and Social Research*, vol. 3, no. 3, pp. 63–75, 2013.

[16] V. YILMAZ and H. E. CELIK, "A statistical approach to estimate the wind speed distribution: The case of gelibolu region," 2008.

[17] P. A. Costa Rocha, R. C. de Sousa, C. F. de Andrade, and M. E. V. da Silva, "Comparison of seven numerical methods for determining weibull parameters for wind energy generation in the northeast region of brazil," *Applied Energy*, vol. 89, no. 1, pp. 395–400, 2012.

[18] R. Karki, P. Hu, and R. Billinton, "A simplified wind power generation model for reliability evaluation," *Energy conversion, IEEE Transactions on*, vol. 21, no. 2, pp. 533–540, 2006.

[19] A. P. Leite, C. L. Borges, and D. M. Falcao, "Probabilistic wind farms generation model for reliability studies applied to brazilian sites," *Power Systems, IEEE Transactions on*, vol. 21, no. 4, pp. 1493–1501, 2006.

[20] Y. Atwa, E. El-Saadany, M. Salama, and R. Seethapathy, "Optimal renewable resources mix for distribution system energy loss minimization," *Power Systems, IEEE Transactions on*, vol. 25, no. 1, pp. 360–370, 2010.

[21] R. Billinton and Y. Li, "Incorporating multi-state unit models in composite system adequacy assessment," *European transactions on electrical power*, vol. 17, no. 4, pp. 375–386, 2007.

[22] H. L. Willis, "Analytical methods and rules of thumb for modeling dg-distribution interaction," in *Power Engineering Society Summer Meeting, 2000. IEEE*, vol. 3. IEEE, 2000, pp. 1643–1644.

[23] C. Wang and M. H. Nehrir, "Analytical approaches for optimal placement of distributed generation sources in power systems," *Power Systems, IEEE Transactions on*, vol. 19, no. 4, pp. 2068–2076, 2004.

[24] T. Gözel and M. H. Hocaoglu, "An analytical method for the sizing and siting of distributed generators in radial systems," *Electric Power Systems Research*, vol. 79, no. 6, pp. 912–918, 2009.

[25] P. Mahat, W. Ongsakul, and N. Mithulananthan, "Optimal placement of wind turbine dg in primary distribution systems for real loss reduction," *Proceedings of Energy for Sustainable Development: Prospects and Issues for Asia, Phuket*, 2006.

[26] K. Nara, Y. Hayashi, K. Ikeda, and T. Ashizawa, "Application of tabu search to optimal placement of distributed generators," in *Power Engineering Society Winter Meeting, 2001. IEEE*, vol. 2. IEEE, 2001, pp. 918–923.

[27] W. El-Khattam, K. Bhattacharya, Y. Hegazy, and M. Salama, "Optimal investment planning for distributed generation in a competitive electricity market," *Power Systems, IEEE Transactions on*, vol. 19, no. 3, pp. 1674–1684, 2004.

[28] D. Singh and K. Verma, "Multiobjective optimization for dg planning with load models," *Power Systems, IEEE Transactions on*, vol. 24, no. 1, pp. 427–436, 2009.

[29] A. J. Ardakani, A. K. Kavyani, S. Pourmousavi, S. Hosseinian, and M. Abedi, "Siting and sizing of distributed generation for loss reduction," *International Carnivourous Plant Society*, 2007.

[30] G. Celli and F. Pilo, "Mv network planning under uncertainties on distributed generation penetration," in *Power Engineering Society Summer Meeting, 2001*, vol. 1. IEEE, 2001, pp. 485–490.

[31] E. Lopez, H. Opazo, L. Garcia, and P. Bastard, "Online reconfiguration considering variability demand: Applications to real networks," *Power Systems, IEEE Transactions on*, vol. 19, no. 1, pp. 549–553, 2004.

[32] L. F. Ochoa and G. P. Harrison, "Minimizing energy losses: Optimal accommodation and smart operation of renewable distributed generation," *Power Systems, IEEE Transactions on*, vol. 26, no. 1, pp. 198–205, 2011.

[33] V. M. Quezada, J. R. Abbad, and T. G. S. Roman, "Assessment of energy distribution losses for increasing penetration of distributed generation," *Power Systems, IEEE Transactions on*, vol. 21, no. 2, pp. 533–540, 2006.

[34] D. C. Montgomery and G. C. Runger, *Applied statistics and probability for engineers*. Wiley. com, 2010.

[35] J. Pinheiro, C. Dornellas, M. T. Schilling, A. Melo, and J. Mello, "Probing the new ieee reliability test system (rts-96): Hl-ii assessment," *Power Systems, IEEE Transactions on*, vol. 13, no. 1, pp. 171–176, 1998.

[36] M. Ramezani, C. Singh, and M.-R. Haghifam, "Role of clustering in the probabilistic evaluation of ttc in power systems including wind power generation," *Power Systems, IEEE Transactions on*, vol. 24, no. 2, pp. 849–858, 2009.

[37] I. H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005.

[38] S. H. Jangamshetti and V. G. Rau, "Site matching of wind turbine generators: a case study," *Energy Conversion, IEEE Transactions on*, vol. 14, no. 4, pp. 1537–1543, 1999.

[39] Z. M. Salameh, B. S. Borowy, and A. R. Amin, "Photovoltaic module-site matching based on the capacity factors," *Energy Conversion, IEEE Transactions on*, vol. 10, no. 2, pp. 326–332, 1995.

[40] Tech. Rep., distributed Generation Technical Interconnection Requirements Interconnections at Voltages 50kV and Below, available as http://www.hydroone.com/Generators/Pages/\TechnicalRequirements.aspx.

[41] A. Abou El-Ela, S. Allam, and M. Shatla, "Maximal optimal benefits of distributed generation using genetic algorithms," *Electric Power Systems Research*, vol. 80, no. 7, pp. 869–877, 2010.

[42] D. Singh and R. Misra, "Effect of load models in distributed generation planning," *Power Systems, IEEE Transactions on*, vol. 22, no. 4, pp. 2204–2212, 2007.

[43] Wind Energy Resources, Wind Turbine Generators. Available [Online]: http://www.wind-energy-resources.com/wer_100kw_wind_turbine.html.

[44] First Solar, PV Modules. Available [Online]: http://www.belectric.com/fileadmin/MASTER/pdf/datasheet/First_Solar_Datasheet_FS_Series_3.pdf.

[45] R. Billinton and R. N. Allan, *Reliability evaluation of engineering systems: concepts and techniques*. Plenum Press New York, NY, 1983.

[46] An introduction to Statistics. Available [Online]: http://www.millnet-consulting.com/review/statistics/#margin_of_error.

[47] I.-S. Bae, J.-O. Kim, J.-C. Kim, and C. Singh, "Optimal operating strategy for distributed generation considering hourly reliability worth," *Power Systems, IEEE Transactions on*, vol. 19, no. 1, pp. 287–292, 2004.

[48] Independent Electricity System Operator. Available [Online]: http://www.ieso.ca/.

[49] M. F. Shaaban, Y. M. Atwa, and E. El-Saadany, "Dg allocation for benefit maximization in distribution networks," *Power Systems, IEEE Transactions on*, vol. 28, no. 2, pp. 639–649, 2013.