11-1-2014

Solar Power Deployment: Forecasting and Planning

Mohana Shandal Alanazi

University of Denver

Follow this and additional works at: https://digitalcommons.du.edu/etd

Part of the Power and Energy Commons

Recommended Citation

https://digitalcommons.du.edu/etd/8

This Thesis is brought to you for free and open access by the Graduate Studies at Digital Commons @ DU. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ DU. For more information, please contact jennifer.cox@du.edu, dig-commons@du.edu.
SOLAR POWER DEPLOYMENT: FORECASTING AND PLANNING

A Thesis

Presented to

The Faculty of the Daniel Felix Ritchie School of Engineering and Computer Science

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

By

Mohana Alanazi

November 2014

Advisor: Dr. Amin Khodaei
Abstract

The rapid growth of Photovoltaic (PV) technology has been very visible over the past decade. Recently, the penetration of PV plants to the existing grid has significantly increased. Such increase in the integration of solar energy has brought attention to the solar irradiance forecasting. This thesis presents a thorough research of PV technology, how solar power can be forecasted, and PV planning under uncertainty.

Over the last decade, the PV was one of the fastest growing renewable energy technologies. However, the PV system output varies based on weather conditions. Due to the variability and the uncertainty of solar power, the integration of the electricity generated by PV system is considered one of the challenges that have confronted the PV system. This thesis proposes a new forecasting method to reduce the uncertainty of the PV output so the power operator will be able to accommodate its variability. The new forecasting method proposes different processes to be undertaken before the data is fed to the forecasting model. The method converts the data sets included in the forecasting from non-stationary data to a stationary data by applying different processes including: removing the offset, removing night time solar values, and normalization. The new forecasting method aims to reduce the forecasting error and analyzes the error effect on the long term planning through calculating the payback period considering different errors.
Acknowledgements

I would like to acknowledge and express my deep gratitude to all persons who supported me to complete this report. Special thanks and appreciation is to my academic advisor Dr. Amin Khodaei for his guidance, enthusiastic encouragement and useful critiques of this research work. Also, I would like to thank the thesis defense committee members: Dr. Ali Besharat, Dr. Mohammad Matin, and Dr. Jun Zhang for their valuable time spent to review the thesis report and provide all the useful comments. Thanks and appreciation to all staff members in the Department of Electrical and Computer Engineering for their support during my master program. Moreover, I would like to thank my parents and my family for their endless support during the research. Thanks and appreciation is to the National Renewable Energy Laboratory (NREL) and the National Climatic Data Center (NCDC) for their support in getting all required data for this research.
Table of Contents

Chapter One: Introduction ........................................................................................................ 1
  1.1 Solar photovoltaic (PV) .................................................................................................. 2
    1.1.1 PV Technologies ................................................................................................. 3
      1.1.1.1 First-generation PV ....................................................................................... 3
      1.1.1.2 Second-generation PV ................................................................................... 5
      1.1.1.3 Third-generation PV ....................................................................................... 8
      1.1.2 PV Benefits ....................................................................................................... 11
  1.2 Solar Power Versus Solar Irradiance ............................................................................. 12
    1.2.1 Direct normal irradiance (DNI) ............................................................................ 13
    1.2.2 Diffuse horizontal irradiance (DHI) .................................................................... 14
    1.2.3 Global horizontal irradiance (GHI) ..................................................................... 14
  1.3 Current State of Art ....................................................................................................... 15
  1.4 Solar Power Challenges ............................................................................................... 16
    1.4.1 PV Modules Capital Cost ..................................................................................... 17
    1.4.2 Solar Energy Variability ..................................................................................... 17
    1.4.3 Solar Power Integration ....................................................................................... 19

Chapter Two: Forecasting ..................................................................................................... 20
  2.1 Why We Need Forecasting? ........................................................................................ 21
  2.2 Literature Review of Forecasting Methods .................................................................. 22
    2.2.1 Physical Method ................................................................................................ 23
      2.2.1.1 Numerical Weather Prediction (NWP) .............................................................. 23
      2.2.1.2 Satellite and Cloud Imagery Model ................................................................. 27
    2.2.2 Statistical Method .............................................................................................. 28
      2.2.2.1 Time Series Models ...................................................................................... 28
      2.2.2.2 Persistence Model ......................................................................................... 35
    2.2.3 Hybrid Models ................................................................................................ 36
  2.3 Understanding the Forecasting Error .......................................................................... 36
  2.4 Neural Network Toolbox ............................................................................................ 37
    2.4.1 Nonlinear Autoregressive with External Inputs (NARX) ..................................... 38
    2.4.2 Nonlinear Autoregressive (NAR) ...................................................................... 38
    2.4.3 Nonlinear Input-Output ....................................................................................... 39
  2.6 Different Stages of Data Processing ......................................................................... 41
    2.6.1 Data Preprocessing ............................................................................................ 42
      2.6.1.1 Removing Offset ............................................................................................ 42
      2.6.1.2 Normalization ............................................................................................... 43
      2.6.1.3 Removing Nighttime Hours ........................................................................... 43
2.6.2 Data Processing ................................................................. 44
2.6.3 Data Post-Processing .......................................................... 45
  2.6.3.1 Denormalization ......................................................... 45
  2.6.3.2 Adding Nighttime Hours .............................................. 45
Chapter Three: PV System Economic Analysis .................................. 47
  3.1 Economic Aspects ............................................................... 48
    3.1.1 The Total Cost of PV System ........................................... 48
    3.1.2 Net Cash Flow .............................................................. 49
    3.1.3 Simple Payback Period .................................................. 50
    3.1.4 Net Present Value (NPV) ................................................ 50
  3.2 Value of Lost Load (VOLL) .................................................. 51
  3.3 Electricity Price ................................................................. 53
    3.3.1 Fixed Electricity Price .................................................... 53
    3.3.2 Variable Electricity Price ................................................ 54
  3.4 Sensitivity Analysis ............................................................ 54
  3.5 PV vs. CO2 Emission ........................................................... 55
Chapter Four: Case Study .............................................................. 56
  4.1 Data Preparation and Preprocessing ........................................ 57
    4.1.1 Global Horizontal Irradiance (GHI) .................................... 57
    4.1.2 Meteorological Data ....................................................... 58
    4.1.3 Clear Sky GHI ............................................................... 59
    4.1.4 Electricity Price ............................................................. 59
    4.1.5 Normalization ............................................................... 60
    4.1.6 Removing Offset ............................................................ 60
    4.1.7 Removing Nighttime Hours .............................................. 61
  4.2 Forecasting Process ............................................................ 63
    4.2.1 GHI Forecasting Using only Historical Data ......................... 63
    4.2.2 GHI Forecasting Using Meteorological Data as input ............... 67
    4.2.3 Forecasting Fixed Electricity Price .................................... 68
    4.2.4 Forecasting Electricity Real Time Price ............................... 69
  4.3 Economic Analysis .............................................................. 70
    4.3.1 Total PV System Cost ..................................................... 71
    4.3.2 Total PV System Energy Value ......................................... 72
      4.3.2.2 Energy Value with Actual Data ..................................... 72
      4.3.2.3 Energy Value with Large Error .................................... 73
      4.3.2.4 Energy Value with Small Error .................................... 74
    4.3.3 Reliability Benefit ........................................................ 75
      4.3.3.1 Reliability and CAIDI .............................................. 75
4.4 Results .................................................................................................................. 77
  4.4.1 Payback Period Without Error ........................................................................ 77
    4.4.1.1 Payback Period with Fixed Price ............................................................. 77
    4.4.1.2 Payback Period with Variable Price ......................................................... 78
  4.4.2 Payback Period With Large Error ................................................................. 78
    4.4.2.1 Payback Period With Fixed Price ............................................................. 78
    4.4.2.2 Payback Period With Variable Price ......................................................... 78
  4.4.3 Payback Period With Small Error ................................................................. 79
    4.4.3.1 Payback Period With Fixed Price ............................................................. 79
    4.4.3.2 Payback Period With Variable Price ......................................................... 79
  4.5 Discussions ........................................................................................................ 80

Chapter Five: Summary .......................................................................................... 82
  5.1 Summary ............................................................................................................ 82
  5.2 Conclusion ......................................................................................................... 82
  5.3 Recommendations ............................................................................................ 83
  5.4 Future Work ...................................................................................................... 84

References ............................................................................................................... 85

Appendix ................................................................................................................... 89
List of Figures

Figure 1 Electrons flow from the n to p side, through the load ........................................ 2
Figure 2 Solar PV System ........................................................................................................ 2
Figure 3 Production process for typical commercial crystalline silicon cells. .................... 4
Figure 4 Cross section of an amorphous silicon pin cell in (10^-9) .................................... 6
Figure 5 CdTe PV cell. ............................................................................................................ 7
Figure 6 Fresnel lenses are used to focus sunlight onto a small solar cell. ...................... 9
Figure 7 Trend shows how efficiency over time for different PV technologies.......... 11
Figure 8 illustrates how PV helps in saving environment from CO2 ........................... 12
Figure 9 Solar-radiation components .................................................................................. 14
Figure 10 Global Map that shows average GHI. ................................................................. 15
Figure 11 New U.S generation capacity in 2012 vs. 2013 ............................................ 16
Figure 12 Difference between variability and uncertainty. ............................................. 20
Figure 13 NAM GHI (W/m2) Forecast for April 10th, 2010 ........................................ 27
Figure 14 (a) ACF and PACF for AR(1),(b) MA(1). (c) ARMA(1,1) .................................. 33
Figure 15 Neural network Structure. .................................................................................. 35
Figure 16 Neural Network Start GUI. ................................................................................ 38
Figure 17 (a) Net structure for NARX.(b) NAR. (c) Nonlinear input-output .......... 39
Figure 18 Neural Network loading input and target window ........................................ 40
Figure 19 Neural Network training window ..................................................................... 41
Figure 20 Flowchart For GHI forecasting process .......................................................... 46
Figure 21 PV System average Installed Price by Sector ................................................. 49
Figure 22 Estimated Average VOLL for ERCOT. ........................................................... 52
Figure 23 Hourly GHI data from 1996 to 2010 for Denver International Airport ........ 57
Figure 24 Hourly GHI data for four days. ......................................................................... 58
Figure 25 Clear Sky GHI vs. Cloudy day GHI ................................................................. 59
Figure 26 GHI deviation between clear sky GHI and historical day GHI .................. 61
Figure 27 Daytime hourly GHI in loudness and clear day ............................................. 62
Figure 28 The startup window for time series NN. ................................................................. 64
Figure 29 The input is fed as target .................................................................................. 64
Figure 30 The data divided to three different process ....................................................... 65
Figure 31 The structure selection of NN........................................................................... 65
Figure 32 Training Process ............................................................................................... 66
Figure 33 Actual and forecasted GHI using historical data only ....................................... 66
Figure 34 NN structure for forecasting GHI .................................................................... 67
Figure 35 Actual and forecasted GHI using historical and weather data ......................... 68
Figure 36 Fixed price actual and forecasted data ................................................................ 69
Figure 37 Variable price actual and forecasted .................................................................. 69
List of Tables

Table 1 Comparison of the three thin-film technologies ........................................ 8
Table 2 NWP models configurations ........................................................................ 27
Table 3 Sunrise and Sunset time and Daytime hours ................................................. 62
Table 4 MAPE for GHI and Price forecasting ............................................................. 70
Table 5 PV system design specification ..................................................................... 70
Table 6 Payback periods for different CAIDI incidents ............................................. 76
Table 7 Payback periods for different scenarios ....................................................... 79
## Nomenclature

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>ARMA</td>
<td>Autoregressive Moving Average Model</td>
</tr>
<tr>
<td>CAIDI</td>
<td>Customer Average Interruption Duration Index</td>
</tr>
<tr>
<td>CPV</td>
<td>Concentrated Photovoltaic</td>
</tr>
<tr>
<td>CSP</td>
<td>Concentrated Solar Power</td>
</tr>
<tr>
<td>DHI</td>
<td>Diffuse Horizontal Irradiance</td>
</tr>
<tr>
<td>DNI</td>
<td>Direct Normal Irradiance</td>
</tr>
<tr>
<td>DOE</td>
<td>The U.S. Department of Energy</td>
</tr>
<tr>
<td>ECMWF</td>
<td>Europe Center Medium Range Weather Forecast</td>
</tr>
<tr>
<td>EPIA</td>
<td>European Photovoltaic Industry Association</td>
</tr>
<tr>
<td>GFS</td>
<td>Global Forecast System</td>
</tr>
<tr>
<td>GHI</td>
<td>Global Horizontal Irradiance</td>
</tr>
<tr>
<td>MAPE</td>
<td>Mean Absolute Percentage Error</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Square Error</td>
</tr>
<tr>
<td>NAM</td>
<td>North American Mesoscale Model</td>
</tr>
<tr>
<td>NREL</td>
<td>National Renewable Energy Laboratory</td>
</tr>
<tr>
<td>NWP</td>
<td>Numerical Weather Prediction</td>
</tr>
<tr>
<td>PV</td>
<td>Photovoltaic</td>
</tr>
<tr>
<td>RRTM</td>
<td>Rapid Radiative Transfer Model</td>
</tr>
<tr>
<td>WMO</td>
<td>World Metrological Organization</td>
</tr>
</tbody>
</table>
Chapter One: Introduction

Solar Power is becoming an important sector in energy generation. This is due to many reasons such as global warming and the increase of pollution caused by CO2 emissions from fossil fuel based generation methods such as coal and gas. Solar energy is a clean energy source, as it is emissions free.

Driven by the above-mentioned reasons, solar power has grown so fast over the past decade, and its growth is anticipated to continue in the future. The utilization of the PV system has moved from small-scale residential to large-scale industrial deployments.

Over the last decade, PV was one of the fastest growing renewable energy technologies. At the end of 2011, a total capacity of 65 GW solar PV was installed globally, while in 2000, around 1.5 GW PV capacity was installed. Over half of that capacity was installed in Germany and Italy, followed by Japan, Spain, the United States and China [1].

This thesis is organized in five chapters. Chapter 1 provides an overview of the PV including: explanations of how this technology works, description of state-of-the-art PV technology, and evaluation of existing challenges. Chapter 2 presents in detail one of the major challenges in solar power, which is forecasting. This chapter discusses different methods used for forecasting PV irradiance. Chapter 3 discusses the economic analyses of PV deployment. Chapters 4 provides comprehensive case studies for long-term solar power forecasting and planning. Chapter 5 includes a summary and recommendations.
1.1 Solar photovoltaic (PV)

Solar PV systems directly convert solar energy into electricity. That is achieved by converting the energy in photon (light) to an electric voltage using photovoltaic. The electron of such material (photovoltaic) is freed from its atom by a photon that has a short wavelength and high energy in sunlight. When nearby electric field is added, the electrons move to a metallic contact. As a result an electric current is created [2].

![Figure 1](image1.png)

Figure 1 Electrons flow from the n to p side, through the load [2]

![Figure 2](image2.png)

Figure 2 Solar PV System [3]
The demand on renewable resources such as solar power has increased recently. As a result, the solar PV is considered as one of the fastest growing renewable technologies. Solar power is expected to play a major role in the global electricity generation mix. The PV system is a highly reliable and is utilized by different sectors, where it is within the reach of residential, commercial, and industrial customers [4].

1.1 PV Technologies

PV technologies have been classified into three different generations. They are classified based on the basic materials that have been used and the size of the PV. More technologies will be developed in the near future due to the high growing demand on solar energy. A photovoltaic system is composed of PV cells that are put together to form a PV module. More PV modules are grouped together to form a PV array. The PV module consists of different auxiliary components such as an inverter, a control system and a battery storage system. In today’s market, different semiconductor materials are used by different PV technologies. Based on the used materials and the level of commercial market, PV technologies are classified into three generations [4].

1.1.1 First-generation PV

The first generation is made of crystalline silicon (C-Si). Around 80-90% of the today’s market is represented by this technology. C-Si modules come in two main different categories. One is made of single crystalline (sc-Si) and the other comes in multi-crystalline (mc-Si). The efficiency of this technology ranges between 16% and 19% [4] [5].
Our earth’s crust is rich with Silicon, which is a semiconductor material with energy gap of 1.1eV and suitable for PV applications. Crystalline silicon has been used widely in the PV industry and therefore it dominates the current PV market. The current production of solar cells will be deployed from small KW-scale to several hundreds of MW in the near future. The process of producing crystalline silicon solar cell is shown in Figure 3. The manufacturing process comprises in four steps [4]:

- Silicon Production,
- Wafer Production,
- Cell Production, and
- Model Assembly.

Figure 3 Production process for typical commercial crystalline silicon cells [6]
Depending on how Si wafer is manufactured, the crystalline silicon cells are categorized to three types [4]:

- Single Crystal (sc-Si), Monocrystalline (Mono c-Si), which is dominant in the silicon technology.
- Polycrystalline (Poly c-Si), sometimes referred to as multi-crystalline (mc-Si), and
- EFG ribbon silicon and silicon sheet-defined as film growth (EFG ribbon-sheet c-Si).

1.1.1.2 Second-generation PV

The second generation of PV is based on thin-film technology. It depends mainly on thin layers of photosensitive materials. This technology has low cost production and its efficiency ranges from 12% to 20%. Around 15% of the global PV module market is represented by thin film PV [7]. The thin film PV technology includes three main types [7].

- Amorphous silicon (a-Si and a-Si/μc-Si);
- Cadmium Telluride (Cd-Te); and
- Copper-Indium-Selenide (CIS) and Copper-Indium-Gallium-Diselenide (CIGS).

Thin film solar cells require less semiconductor material, which is up to 99% less materials than crystalline solar cells. Thin film solar cells have very lower cost compared to the sc-Si module [4].
1.1.1.2.1 *Amorphous silicon solar cells*

Amorphous silicon cells are thinly layered and arranged spontaneously. They represent almost all of today’s thin film technology. They are the most developed and widely known thin film solar cells. Figure 4 shows a cross section of a simple a-Si cell.

![Cross section of an amorphous silicon pin cell](image)

**Figure 4 Cross section of an amorphous silicon pin cell in (10^{-9}) [2]**

The amorphous silicon PV module has an efficiency that ranges from 10% to 12%. The amorphous silicon solar cell is less expensive than a crystalline-silicon based solar cell. It uses only 1% of the silicon that would be used to produce a crystalline-silicon based solar cell. Amorphous silicon solar cells can be used in various ways due to their flexibility and can be rolled up and stored when not in use. The main disadvantage of the amorphous silicon cell is the degradation of its performance over time caused by the sun. This problem causes a reduction in the power output, which reaches up to 35%. On the other hand, thinner layers can provide better stability and less power output reduction by increasing the electric field strength across the material [4]. Amorphous silicon solar cells show a wide window for improvement. For instance, they have lower
efficiency. Amorphous silicon solar panels also need to cover a larger surface area compared to crystalline-based solar panels for the same PV system capacity [8].

1.1.1.2.2 Cadmium Telluride

Cadmium Telluride PV modules are considered to have the lowest cost among thin-film modules. CdTe PV cells have efficiency of around 15%. They are considered to be the most eco-friendly solar panels. The production of such panels consumes less energy. The estimated manufacturing cost for Cadmium Telluride thin-film PV cells is under $0.75/W. The main two materials used to manufacture Cadmium Telluride PV cells are cadmium and tellurium. The problem of Cadmium is its toxicity, which could limit the use of this raw material. Therefore, the use of Cadmium during manufacturing should be monitored and controlled to protect the health of workers. The other raw material, Telluride, is produced in lower quantities. Figure 5 shows how the layers of CdTe cells are arranged [4].

![Figure 5 CdTe PV cell](image-url)
1.1.1.2.3 Copper Indium Selenide (CIS) and (CIGS)

At around 20% efficiency, Copper Indium Selenide (CIS) and Copper Indium Gallium-Diselenide (CIGS) PV modules have achieved the highest efficiency among the thin-film PV modules. The CIGS PV panel does not use crystalline silicon like the first PV generation. The Copper Indium Selenide (CIS) and the Copper-Indium-Gallium-Diselenide (CIGS) have become the most accepted thin-film solar technologies due to their low cost and high efficiency [4]. Table 1 summarizes the differences between the three types of thin-film solar technologies.

<table>
<thead>
<tr>
<th>Technology</th>
<th>Efficiency</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| a-Si        | 10% -12%   | • Mature technology  
|            |            | • Flexible to fit different applications | • Lower efficiency  
|            |            |                      | • Degrades over time and output will be reduced |
| CdTe       | 15%        | • Low manufacturing cost | • Medium efficiency  
|            |            |                      | • Toxicity material (Cd) |
| CIGS       | 20%        | • High efficiency Among thin film PV | • Uniformity challenge on large-area |

Table 1 Comparison of the three thin-film technologies.

1.1.1.3 Third-generation PV

Third generation PV systems refer to different technologies that are under demonstration, to be commercialized, and novel concepts under development. Examples of third generation technologies include concentrated PV (CPV) and organic solar PV as well as some novel concepts under development.
1.1.1.3.1 Concentrated Photovoltaic (CPV)

The CPV uses optical devices, such as mirrors or lenses, which concentrate direct solar radiation onto a high efficiency small cell that is made of semiconductor materials (see Figure 6). The CPV requires less photovoltaic material to focus the direct sunlight. It uses high efficiency but expensive cells that are economically feasible due to the less space they require. The optical system of CPV uses standard materials, which make CPV less dependent on the silicon supply and less expensive cells [10]. The CPV are classified based on the degree of sunlight concentration. First is a low degree concentration that ranges from 2-10 suns (sun= 1000W/m^2). Second classification is a medium concentration that reaches up to 100. Third is a high degree concentration, which concentrates direct sunlight up to 1000 suns. To track the direct sunlight over the day, a tracking system is used to increase the efficiency. The mirrors or lenses permanently face the sun through a single or double axis tracking system. Due to high concentrated sunlight, cooling systems (passive or active) are needed for the CPV. In the passive cooling system, the cell is positioned on a high thermal conductivity ceramic substrate. The active cooling system utilizes the liquid metal as cooling fluid [4].

Figure 6 Fresnel lenses are used to focus sunlight onto a small solar cell [10]
CPV modules with silicon-based cells have an efficiency that ranges from 20% to 25%, while commercial CPV cells based on multi-junction solar cells have reached up to 35% efficiency. To maximize the performance of CPV modules, they need to be installed in regions where clear sky and high direct irradiation (DNI) exist since CPV relies on direct sunlight [4].

1.1.1.3.2 Organic Solar PV

Organic solar PV cells are made from organic or polymer materials. Although they consume low energy production, they also have low efficiency. Because they consume low energy production, they are inexpensive. Organic solar PV (OPV) modules have an efficiency range from 6-10%. In addition to low efficiency, OPV cells have stability issues over time. Manufacturing costs of OPV cells are in decline and expected to be as low as $0.5/W by 2020, which will allow them compete with other PV technologies [11].

1.1.1.3.3 Other Novel and Emerging PV Concepts

Some of the novel and emerging PV concepts are introduced to achieve ultra-high efficiency through using new advanced materials that maximize sunlight conversion to electricity. These concepts are under development and rely on using quantum or super lattice technologies [4]. Different approaches are ongoing to increase the energy absorption of current active materials. One approach is the quantum, which endeavors to enable more solar PV output current and voltage trade-off. Another approach is maximizing the energy capture of current solar cells by relying on up/down converters.
The current research and development aim to increase cell efficiency above 25% by 2015 [12].

The following chart presents the differences in the efficiency between the three different PV generations.

![Figure 7 Efficiency over time for different PV technologies [13]](image)

### 1.1.2 PV Benefits

A number of significant benefits can be offered by PV technologies that include [14]:

- Solar energy is an abundant resource that is available everywhere, and it’s free to everyone.
- Solar PV systems are eco-friendly. They produce no harmful or toxic gases that
pollute air, such CO2. Solar PV contributes actively to reduce global warming problems. See Figure 8.

- No fuel cost is included in PV plants unlike other conventional power plants such as coal or gas. They have low operation and maintenance (O&M) costs.
- Solar PV has helped rural areas (especially in developing countries) by providing electricity.
- Although solar energy is variable, during peak electricity demand driven by cooling in summer, PV has a high coincidence with such peak demand.
- PV modules can be recycled. Therefore, the materials that are included in PV production such as silicon and glass can be reused, which helps the environment and reduce energy needed to produce those materials.

Figure 8 illustrates how PV helps in saving environment from CO2 [14]

1.2. Solar Power Versus Solar Irradiance

The first step to understand the PV system is to know how much sunlight is available. Solar irradiance is the measurement of irradiance that is produced by the sun to the earth in the form of electromagnetic radiation. Solar irradiance is expressed as a
power density ($W/m^2$). Solar power is the available solar irradiance received by the solar conversion system multiplied by the system’s total effective collector area ($W/m^2 \times m^2 = W$) [15].

Solar irradiance is variable and hence forecasting is very important to predict the solar energy amount to ensure the stability of the power grid. Also, many factors contribute to the process of estimating solar energy conversion such as system conversion performance and factors related to the environment.

There are three important fundamental components of solar irradiance that are taken into consideration when calculating the received solar irradiance. The World Meteorological Organization (WMO) provides detailed guidelines on how these components are being measured in addition to the used instruments. Figure 9 illustrates these three fundamental components [15].

1.2.1 Direct normal irradiance (DNI)

Direct normal irradiance is the amount of solar radiation received at the horizontal earth surface in a direct path from the sun without any atmospheric losses. This quantity of radiation is a very important component to concentrating solar thermal installations such as Concentrated Solar Power (CSP) and Concentrated Photovoltaic (CPV) systems.
1.2.2 Diffuse horizontal irradiance (DHI)

Diffuse horizontal irradiance is the amount of solar radiation received at the horizontal earth surface in an indirect path from the sun as it has been scattered by air molecules, aerosol particles, cloud particles, or other particles.

1.2.3 Global horizontal irradiance (GHI)

Global Horizontal Irradiance is the total amount of radiation received by the surface horizontal to the ground. It consists of both Direct Normal Irradiance (DNI) and Diffuse Horizontal Irradiance (DHI). Since this kind of radiation includes other radiation, it is going to be considered in the research that will be introduced in chapter four.

The three fundamental components of solar irradiance can be related to each other using the following equation:

\[ GHI = DHI + DNI \times \cos(Z) \quad (1) \]

Figure 9 Solar-radiation components [15]

Where \( Z \) is the solar zenith angle. Solar zenith angle is the angle between the direction of the sun and the zenith that is the overhead direction [16]. Figure 10 shows the
average GHI around the world. Middle East countries, Africa, Australia, and part of North and South America have high GHI compared with other parts of the world.

Figure 10 Global Map that shows average GHI [17]

1.3 Current State of Art

In 2013, the cumulative installation capacity of PV systems reaches up to 138.9 GW with new installation of 38.4 GW [18]. Around 56% of the world PV market was in Asia. The highest PV installation was in China which installed 11.8 GW of PV in 2013. China was followed by Japan which installed 6.9 GW while the U.S. installed 4.8 GW. In 2013, Germany is placed on the top European PV market with installation of 3.3 GW [18].

In 2013, the U.S. achieved a remarkable trend in solar power. PV installation increased by 41% in 2013 compared to 2012. This resulted in solar power taking the second place in generating electricity, coming only after natural gas. More than half of 2013 installations in the U.S. occurred in California with an installation capacity around
2745.8 MW. Arizona was in the second place with a total installation capacity around 700.7 MW. In tenth place, Colorado had an installation capacity around 55.9 MW. There is now a total of 12.1 GW of PV and 918 MW of CSP operating in the United States. The installation cost of PV fell by 15% compared to 2012. New growth in PV installation is expected to be around 29% in 2014, reaching nearly 6 GW. Most of the new installations will be encountered under residential sector [19]. Figure 11 compares the electricity generation in the U.S. in 2012 vs. 2013.

![Figure 11 New U.S. generation capacity in 2012 vs. 2013][19]

### 1.4 Solar Power Challenges

There are a lot of obstacles that confront solar power and prevent it from becoming the primary source of energy. High capital cost, inconsistent availability and integration with the existing grids have traditionally prevented solar power from being highly utilized.
1.4.1 PV Modules Capital Cost

Solar panels are made from semiconductor materials that are relatively expensive to be manufactured. Conventional power sources have gone through different developments over decades to reduce generation cost. However, solar power has a long way to go. A number of different technologies are being developed (e.g. CPV, OPV, and novel concepts) to reduce the manufacturing cost, as it accounts for about a third of the capital cost, and increase the efficiency. Over the last two years, a 60% reduction in PV module price has occurred. By 2020 more than 40% reduction in PV module cost will take place [4]. Many governments have offered incentives to help in solar power cost reduction, such as Feed-in Tariffs (FIT) and tax credits. FIT, in brief, is an energy supply policy that guarantees a long contract period of 15-20 years for renewable energy producers. The main goal of FIT is to form a robust market for renewable energy sources, which will also promote the rapid development of renewable energy resources and consequently will result in lower technology costs [20]. Applying the FIT policy will make the PV more attractive to investors and customers and result in more clean energy production. It also mitigates greenhouse gas emissions. European Photovoltaic Industry Association (EPIA) has estimated the growth of PV power to be 4.9-9.1% of the global electricity generation by 2030 and it will reach up to 17-21% of electricity share by 2050 [12].

1.4.2 Solar Power Variability

Solar power is an abundant source. However, clouds and weather conditions could affect the availability of the sunrays. Solar energy produces a variable generation as
Electricity production varies in different time scales based on the availability of the energy source i.e. solar radiation. Some features of solar availability are predictable, for example, sunrise and sunset, while other features are not predictable such as intermittent cloud cover. Not only renewable energy sources are variable, different aspects related to power system such as electricity or load demand vary from time to time and the system operator has to handle such changes. All these types of variability must be handled and managed properly by the power system operator [21].

Variability of solar energy makes the integration of solar power to the existing grid and responding to load demand a very challenging task especially at high penetrations. On the other hand, solar variability has less impact at low penetrations because load variability is higher than solar variability [22]. Forecasting in different time horizons is needed to predict the variability and availability of solar power. Chapter two will address different types of forecasting methods that could be applied to forecast solar power. Solar power has reached a high level of generation and penetration that made many entities consider solar power forecasting as an essential component for operating power system. Solar power forecasting will help utility companies in different ways such as maintaining fewer operating reserves, helping power system operators to monitor solar power conditions and be ready for any rapid fluctuations in power output. Also, energy storage systems are considered as one of the potential solutions that could help in dealing with solar power variations. Energy storage systems can mitigate the solar power variability by shifting the power in a reliable way.
1.4.3 Solar Power Integration

With the rapid growth of renewable energy markets, integrating solar with the existing power systems at high penetration levels has raised many issues. The current electricity grid was built a long time ago. It needs to be modernized to accommodate renewable energy sources. The U.S. department of energy (DOE) is now supporting and funding different aspects of renewable energy sources through the SunShot program. SunShot supports different strategies that would help high penetrations of solar electricity on the national power grid. System integration is one of the challenges that the U.S. DOE is tracking to provide all the support and funding. Such support is to increase solar penetration in the nation’s electrical grid and create very reliable, safe, cost effective, and widely deployed solar [23].

Solar variability is also a challenge that has an impact on solar integration. As discussed before, forecasting and energy storage systems are potential solutions that will help improve the reliability of solar integration. Other challenges can be addressed in different aspects, such as grid performance and reliability, communications, and power electronics. As the solar power integration has exponentially increased in the last few years with the existing grid, power electronic has played an important role to make this integration more reliable and safe. Also, power electronic has gone through different developments to accommodate the large scale of solar penetrations and to smooth power flow from large-scale PV plants. Moreover, new manufacturing materials could reduce the integration problem [24].
Chapter Two: Forecasting

Large-scale renewable energy plants have experienced a significant boost in the last few years. The oil price changes and the threat of climate changes, i.e. the global warming, have expedited the utilization of renewable energy resources. Due to the variability and the uncertainty of these renewable resources (solar and wind), the integration of the electricity generated by these resources is a challenge [25]. It is good to differentiate between the variability and uncertainty, when discussing the planning and the operation of a power plant. The variability and the uncertainty are inherent characteristics of renewable resources. In this case, the variability means the changes in the power output of the solar system because of the intermittency in the solar power, while the uncertainty is the inability to predict in advance the timing and the magnitude of these changes. Forecasting aims to reduce the uncertainty of the solar output so the power operator will be able to accommodate its variability. Figure 12 below describes the difference between the variability and the uncertainty [22] [17].

Figure 12 Difference between variability and uncertainty[22]
The prediction of the solar radiation is a fundamental input to different solar applications. Many factors such as solar radiation and weather conditions can influence the output of the PV system. So to enable a good, reliable, and safe integration of the PV system to the grid, it is essential to be ready for any sudden changes in the PV output due to solar power’s changeable and random characteristics. Therefore, the accurate forecasting is necessary. The solar forecasting can be evaluated for different time horizon. The short term forecasting starts from one hour to several hours ahead. The midterm forecasting is from several hours to a week ahead. The long term forecasting is from one week to several years ahead. This chapter discusses the benefits of the forecasting and some methods that are being used to forecast solar radiations in different time scales.

2.1 Why We Need Forecasting?

Forecasting is not only essential for variable generations such as solar and wind. It is also useful in the load forecasting. In addition, some energy economics quantities such as the electricity price could be forecasted. Forecasting will help the power system in different areas such as control, unit commitment, security assessment, planning, and grid integration [26]. Variable generation forecast will have good implications by [27]:

- Minimizing penalties and charges due to imbalance of generated power.
- Providing a good knowledge of real time and day ahead energy market trading.
- Helping to carry out reliable operation and maintenance planning.
The intermittency of renewable energy resources creates deviations in the scheduled power output, which exposes power providers to imbalance charges and penalties, and increases operation cost. However, power forecast helps to reduce these charges and penalties, and lowers the operation cost by estimating the time and the magnitude of output deviations, and allows power operators to find different solutions to accommodate such deviations. Also, the variability and the uncertainty can make the power provider more conservative in a bidding market, while power forecast provide a good knowledge of the real time and day ahead market. Such knowledge will reduce the opportunity costs of being too cautious during the bidding time. As mentioned before, the solar irradiance forecasting is significant for integrating PV plants into the electrical grid. Low error solar irradiance forecasting helps the power system operators to optimize their output production and prepare a proper strategy to overcome any deviations in the production, which will reduce any additional costs [27] [28].

2.2 Literature Review of Forecasting Methods

Forecasting methods can be categorized into three different methods: Physical, Statistical, and Hybrid methods. The physical method is based on the numerical weather prediction (NWP) or atmosphere by using physical data such as temperature, pressure, and cloud cover. The statistical model utilizes the historical data to train the models and is considered as a mathematical model. The hybrid method is a combination of physical and statistical approaches [28].
2.2.1 Physical Method

Physical models tend to be good for long term forecasting. They utilize physical weather data such as temperature, clouds, and pressure. In this section, two physical models will be addressed in detail: the numerical weather prediction and the satellite sky imagery.

2.2.1.1 Numerical Weather Prediction (NWP)

The numerical weather prediction is based on the physics of the atmosphere. The NWP uses current observations of the weather and processes this data to predict the future states of the weather using super computers. Through a process called assimilation, current weather states are processed to produce outputs of temperatures, irradiance, wind, and hundreds of meteorological elements. The NWP is good for one day to multi-days ahead horizon. Thus, it is a useful tool for different utility applications, such as the scheduling of solar power plants. The transient variations in clouds, which are considered the major obstacles for solar irradiance at the ground, are predicted in the NWP. After the assimilation of current observations, the NWP forecasts the future conditions and then the error is corrected based on the previous performance by a statistical post-processing. The process of the numerical weather predication can be briefly descried as follow. First, different sources such as satellites and ground observations collect the initial states of atmosphere. The key source of the NWP error is “data-assimilation”, which is a complex process. This is due to the fact that sources measure different quantities of current states over different volumes of a space and that creates an error in the measurement. Second, the main important equations of atmosphere, such as dynamics equations, Newton’s
second law for fluids flow, thermodynamics equations, and radiative transfer equations are integrated and solved. Finally, the statistical post-processing step where the output of the NWP is manipulated using a trial and error after simulation, in order to compare the outputs with observations and find the statistical relation, and hence correct the error. NWP models can be classified into global models and regional models. In the global model, the simulation of the behavior of the atmosphere is carried out in a global or worldwide scale. The regional (Mesoscale) model simulates the behavior of the atmosphere for an area such as a continent or a country [15]. There are different NWP models, and only three different NWP models will be addressed in this chapter: two global models and one regional model.

2.2.1.1 The European Centre for Medium-Range Weather Forecasts (ECMWF)

The European Centre for Medium-Range Weather Forecasts (ECMWF) is an independent intergovernmental European organization supported by 34 states. ECMWF provides up to 15 days forecast in addition to the monthly and seasonal forecast. It utilizes supercomputers that reside in its headquarters in Reading, Shinfield Park, UK. Its computers are considered one of the largest meteorology computers, which contain the largest archive of the NWP data. The ECMWF is commercially available for a fee and the link can be found in [29].

The ECMWF can forecast thin-layer clouds due to its high vertical resolution. Such type of clouds can have a large impact on solar irradiance. The horizontal grid spacing for the ECMWF is approximately 16 km. The ECMWF has a spatial resolution of 0.25°E-W by 0.25°N-S. The output is made available to the public every 3 hours. The
output is written to a permanent storage. The ECMWF provides the forecast output from a day ahead, which is helpful for scheduling electric power up to 15 days. An aerosol that consists of small particles like dust is prognosed by the ECMWF. The aerosol affects solar radiations by scattering or observing them, especially the direct normal irradiance (DNI), which is important to the CPV and CSP. The ECMWF uses the rapid radiative transfer model (RRTM) and McRad for radiative-transfer parameterizations. Radiative-transfer parameterizations calculate the absorption by gases and scattering by particles. Radiations are classified based on their wavelengths into a shortwave and longwave. The shortwave has a wavelength less than 4 µm, and it originates from the sun, and is called “solar”. It is sometimes referred to the global horizontal irradiance (GHI). The longwave has a wavelength greater than 4 µm and originates from earth atmosphere and is called “thermal” [15] [30].

2.2.1.2 The Global Forecast System (GFS)

The Global Forecast System (GFS) is a global model forecast published by the National Oceanic and Atmospheric Administration (NOAA) through National Centers for Environmental Prediction (NCEP) in the United States. The output of the GFS provided by NCEP is available to the public for free. The horizontal grid spacing for GFS is approximately 50 km and has a spatial resolution of 0.5° E-W by 0.5° N-S. The output is made available to the public every 3 hours. The GFS provides the forecasted output from a day ahead that is helpful for scheduling electric power up to 8 days. The GFS provides atmospheric and land-soil variables such as temperature, and precipitation, to soil moisture. The GFS is a combination of four different models that are “an atmosphere
model, an ocean model, a land/soil model, and a sea ice model” [31]. The GFS gridded data are available at National Operational Model Archive and Distribution System. The GFS diagnoses cloud fraction. Cloud fraction is the partial volume of a grid that is occupied by cloud. Cloud fraction is useful in estimating the probability of PV panel shading. With regard to aerosol, the GFS considers a static climatology for aerosol calculation that can be a source of error and will be corrected in the post-processing stage. Moreover, the GFS uses the same model used by the ECMWF for radiative-transfer parameterizations (RRTM) [15].

2.2.1.3 North American Mesoscale Forecast System (NAM)

The North American Mesoscale Forecast System (NAM) is a regional forecast model run by the National Centers for Environmental Prediction. The NAM is also available to the public for free. The horizontal grid spacing for GFS is approximately 12 km with a spatial resolution of 0.113° E-W by 0.111° N-S. The NAM model makes the output available to the public hourly. The NAM provides forecast output that is helpful for scheduling electric power up to 36 hours. It is run four times daily at 00, 06, 12, and 18 UTC. The NAM assumes that cloud fraction at a given time is either overcast or clear which is inaccurate approximation. For radiative-transfer parameterizations the NAM also uses RRTM. The NAM provides dozens of weather parameters such as temperature, precipitation, lightning, and turbulent kinetic energy. Over 125 outputs from NAM are reported where the shortwave radiation is called (GHI) at the surface. Figure 13 shows the NAM GHI Forecast for April 10th, 2010 [15].
Figure 13 NAM GHI (W/m²) Forecast for April 10th, 2010 [30]

Table 2 below summarizes the differences between the three NWP models.

<table>
<thead>
<tr>
<th></th>
<th>ECMWF</th>
<th>GFS</th>
<th>NAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial Resolution</td>
<td>0.25°</td>
<td>0.5°</td>
<td>0.11°</td>
</tr>
<tr>
<td></td>
<td>16 km</td>
<td>50 km</td>
<td>12 km</td>
</tr>
<tr>
<td>Output time</td>
<td>3 h</td>
<td>3 h</td>
<td>1 h</td>
</tr>
<tr>
<td>Interval</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forecast Duration</td>
<td>15 d</td>
<td>8 d</td>
<td>36 h</td>
</tr>
<tr>
<td>Cloud Fraction</td>
<td>Prognosed</td>
<td>Diagnosed</td>
<td>Overcast/Clear</td>
</tr>
<tr>
<td>Aerosol</td>
<td>Prognostic</td>
<td>Climatological</td>
<td>-</td>
</tr>
<tr>
<td>Radiative-Transfer</td>
<td>RRTM/McRad</td>
<td>RRTM</td>
<td>RRTM</td>
</tr>
<tr>
<td>GHI variable Name</td>
<td>Surface solar radiation downwards</td>
<td>dswrfsc</td>
<td>dswrfsc</td>
</tr>
</tbody>
</table>

Table 2 NWP models configurations [15]

2.2.1.2 Satellite and Cloud Imagery Model

The satellite and cloud imagery based model is a physical forecasting model that analyzes clouds. The satellite imagery can provide a great knowledge about the
cloudiness with high spatial resolution. The high spatial resolution satellite has the potential to derive the required information on cloud motion. Understanding the cloud motion helps to forecast cloud positions and hence solar irradiance. Moreover, cloud covers and cloud optical depth are the most significant parameters that have an influence on solar irradiance at the surface. The satellite and cloud imageries are processed in order to characterize clouds and detect their variability and then forecast the global horizontal irradiance up to 6 hours ahead. This model works by determining the cloud structures during earlier recorded time steps. Understanding the structure of the clouds and their positions will help in predicting solar irradiance [32].

Under low sun elevations, low irradiance conditions, and high spatial variability the errors of satellite and cloud images can increase significantly. In [33] the author has achieved 17% RMSE for half hour cloud index forecast and 30% RMSE for 2 hours forecast [33].

2.2.2 Statistical Method

The statistical method is a mathematical model that uses historical data to predict future values. It’s referred to a statistical because it utilizes mathematical equations to identify the patterns and trends. Statistical models can be persistence models or time series models that include auto-regressive (AR), moving average (MA), or both (ARMA).

2.2.2.1 Time Series Models

Time series models are based on the historical data. Time series is defined as a sequence of observations measured over time, such as the hourly, daily or weekly. It is
stochastic process as the observations could be random. Time series techniques share a common characteristic in that they focus only at the patterns of data. To forecast a time series, these patterns should be identifiable and predictable. The analysis of time series helps to select the suitable model for forecasting. Generally, time series are described as follow [34]:

$$X(t) = T(t) + S(t) + R(t) \quad t = ...-1,0,1,2,...$$  \hspace{1cm} (2)

Where $T(t)$ is the trend term, $S(t)$ is the seasonal term, and $R(t)$ is the random or noise term.

Trend is a continuing pattern of increase or decrease over a period of time. It could be a straight line or a curve. Seasonality is the repeated pattern in time series. It is the fluctuation that occurs over a period of time, and the last term is the noise that has a random fluctuation. It represents the random error that affects the time series caused by an external source. Noise represents the pattern that has not occurred consistently in the past [35]. These patterns can be identified using the autocorrelation functions ACFs and partial correlations PACFs. Autocorrelation is defined as the cross-correlation of a time series and lagged version of itself over a time. It is similar as plotting the cross-correlation between two time series; however, here the same time series is used twice. The partial autocorrelation (PACF) is the partial correlation between a time series and its lagged version over time [36]. Under this section, four models will be addressed: AR, MA, ARMA, and NN.

2.2.2.1 Autoregressive Model (AR)

Autoregressive is a statistical method that uses stochastic process to predict future values based on past values. It is one of the most popular time series models.
Autoregressive model can be represented in different orders. For instance, AR(1) is a first order autoregressive model that current value depends on the immediately preceding value. AR(2) is a process where the future value is based on the previous two values.

Thus an autoregressive model of order AR(p) can be written as follows [37]:

\[ X_t = \phi_1 * X_{t-1} + \phi_2 * X_{t-2} + \cdots + \phi_p * X_{t-p} + \epsilon_t = \sum_{i=1}^{p} \phi_i * X_{t-i} + \epsilon_t \]  

(3)

Here \( \phi_i \) is autoregression coefficient, \( X_{t-i} \) is previous value, and \( \epsilon_t \) is white noise with zero mean.

The autocorrelation functions can sometimes tell the type and the order of the time series. In Autoregression processes, the autocorrelation function (ACF), which can be generated by Matlab using the command “autocorr”, decays to zero as \( i \to \infty \) at a geometric rate. While the partial correlation function (PACF), which can be generated by Matlab using the command “parcorr”, cuts off at lag \( p \) where the values after \( p \) become zero. The order of the autoregressive model is \( p \) where the cut off occurs [37].

There are a number of techniques to find the autoregression coefficients. The most known method is the least squares, which is based on Yule-Walker equations [36]. Matlab has a built-in toolbox that would perform these calculations to find the suitable order and coefficients of a time series given the past value.

2.2.1.2 Moving Average Model (MA)

The moving average model is a time series model that uses current and previous values of white noise to predict current value. Moving average model with a finite order is considered stationary. A moving average model of order MA(q) can be written as follow [37].
\[ X_t = \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_q \varepsilon_{t-q} + \varepsilon_t = \sum_{j=1}^{q} \beta_j \varepsilon_{t-j} + \varepsilon_t \quad (4) \]

Here \( \beta_j \) is moving average coefficient, \( \varepsilon_{t-j} \) is previous value, and \( \varepsilon_t \) is current value of white noise that has zero mean and constant variance and generates random uncorrelated variables. The white noise is created form the forecast errors or residuals when observations become available [37].

The ACF and PACF functions in moving average model behave in the opposite way to the autoregressive model. Here, the autocorrelation function cuts off at lag \( q \) where the values after \( q \) become zero, and the partial correlation function decays gradually to zero as \( j \to \infty \). The order of the moving average model is \( q \) where the tail off occurs [37].

2.2.2.1.3 Autoregressive Moving Average Model (ARMA)

The autoregressive and moving average models can be combined to form the ARMA model. The ARMA is a very useful model to understand and predict future value of a time series that is usually a correlated time series. The ARMA contains of two parts, autoregressive part and moving average part. It usually referred as the ARMA\((p,q)\) where \( p \) and \( q \) are the order of AR and MA, respectively. The Autoregressive Moving Average model can be represented mathematically as follow [37]:

\[ X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t X_t + \beta_1 \varepsilon_{t-1} + \cdots + \beta_q \varepsilon_{t-q} + \varepsilon_t = \]

\[ \sum_{i=1}^{p} \phi_i X_{t-i} + \varepsilon_t + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j} + \varepsilon_t \quad (5) \]

Here \( \phi_i \) is the autoregression coefficient, \( \beta_j \) is the moving average coefficient, and \( \varepsilon_t \) is the white noise. Typically, this model requires a large amount of historical data.
The ARMA model was created in 1951 by Peter Whittle and thoroughly developed and explained by Box-Jenkins in 1971 and is referred sometimes to the Box-Jenkins model [38]. The ARMA model is considered a very flexible stochastic process, as it can represent different types of time series by using different orders. The ARMA model requires the time series to be stationary. More information about stationary time series can be found in [39].

In the ARMA model, both ACF and PACF tail off gradually. It is more difficult to determine the order of the ARMA model $p$ and $q$ by a visual inspection of the correlations plots. However, Matlab has a toolbox called System Identification Toolbox that can help estimating both: the order of the ARMA and their coefficients [37]. Also, the estimation of the autoregressive model’s coefficients can be found using Yule-Walker relations. The Newton Raphson algorithm can be applied to find the moving average coefficients as illustrated in [40].

Figure 14 shows how the ACF and PACF can be applied to determine the model and its order.
2.2.2.1.4 The Artificial Neural Networks Method (ANN)

The artificial neural network is a conventional method that utilizes the idea of the biological neural system in the human brain. The human brain consists of a number of interconnected processing elements called nodes or neurons. These nodes receive an input signal and information from other nodes or externally, and process them through an activation process. Then, the node generates an output signal that is sent to other nodes. The neural networks model is considered as a nonlinear statistical model to understand the complex relation between inputs and outputs and to find the patterns of data. The ANN consists of three layers: input, hidden, and the output layers [42]. The ANN model was implemented in Matlab under the Neural Network Toolbox, which will be discussed in detail later in this chapter.
In the ANN, no assumptions are needed in the underlying process, which relates input and output variables. Simply, the ANN maps between the input and output variables through elements called neurons. Neurons are organized in three layers. First layer contains neurons that receive in inputs. Second layers are called hidden layers and they are in between the input and output layers and contain the hidden neurons. The third layer contains the neurons that produce the outputs. The training process of the ANN goes as follows. The neuron in the hidden layer receives the weighted sum of the inputs and transforms it to output nodes by applying the activation function. Once the ANN structure is established, the training process undergoes until it finds the weights that minimize the error such as mean square error (MSE). There are a number of training methods such as Back Propagation (BP), Levenberg- Marquardt (LM), and quasi-Newton [42]. The objective of the mathematical model of NN is to find the optimal solution of the following equations [43]:

\[ \min E(w, v, \theta, \gamma) = \frac{1}{N_1} \sum_{k=1}^{N_1} \sum_{t=1}^{N} \left[ y_k(t) - \bar{y}_k(t) \right]^2 < \varepsilon_1 \]  \hspace{1cm} (6) 

\[ \bar{y}_k(t) = \sum_{j=1}^{p} v_{jk} \ast f \left[ \sum_{i=1}^{m} x w_{ij} + \theta_j \right] + \gamma_t \]  \hspace{1cm} (7) 

\[ f(x) = \frac{1}{1 + e^{-x}} \]  \hspace{1cm} (8) 

Here \( x \) is the training or input data, \( y_k \) is the actual output and \( \bar{y}_k \) is the forecasted output, \( w_{ij} \) and \( v_{jk} \) are the weights between neurons. The objective of the first equation is to minimize the error between the actual and forecasted output given the input, weights, and number of hidden nodes and layers. The second equation is to predict the output, and
the last equation to produce the output from the hidden layer to nonlinear values using a hyperbolic tangent sigmoid function [43].

There are a lot of different neural networks. For instance, the feed forward neural network (FFNN), the radial basis function neural network (RBFNN), and the recurrent neural network (RNN). In the learning in the FFNN, information is transmitted in one direction between each hidden layer. There is a radial basis function in the RBFNN for each hidden layer. In the learning process in RNN, information is transmitted from the hidden layer to the input layer using a feedback structure [44]. Figure 15 shows a typical structure of NN. In this report ANN model will be used to forecast GHI as going to be introduced in chapter 4.

Figure 15 Neural network Structure [43]

2.2.2.2 Persistence Model

The persistence model is the simplest way for forecasting. It basically predicts the future value, assuming it is same as the previous value.

\[ X_{t+1} = X_t \quad (9) \]

It is also known as the naive predictor. It can be used to give a clue to compare with other methods. The persistence model is good when the weather patterns change very little. The persistence model shows a high error result for forecasting more than one hour. A study was conducted in [45] to forecast the solar power output in a laboratory
level microgrid using two methods, the ARMA and the persistence method. The study shows that for more than one hour forecast, the persistence forecast error was reduced by 17.62% when using ARMA model. So, persistence model is good in a very short term forecast.

2.2.3 Hybrid Models

Hybrid models merge two forecasting techniques to improve the forecast accuracy. They are also known as combined models. The basic idea of hybrid models is to overcome any deficiency of using an individual model, such as regression models, and to take the advantage of each individual model, and combine them to reduce forecast errors. For instance, the NWP model can be combined with the ANN by feeding the outputs from the NWP as input to the ANN models. In [46] a hybrid model was developed by using the satellite imaging as inputs to ANNs. Also, some studies have combined an ANN method with a wavelet forecasting method. Hybrid models can combine linear models, nonlinear models, or both linear and nonlinear models. Many studies have showed that integrated forecast methods outperform individual forecast [32].

2.3 Understanding the Forecasting Error

In order to evaluate the performance of a forecast model, the error has to be calculated. Understanding the forecast error tells how much to trust the forecast, and reevaluate the forecasting methods in case of a high error forecast. Different methods can be used to calculate the accuracy of forecasting models such as the mean absolute error
(MAE), the mean absolute percentage error (MAPE), the mean square error (MSE), and the root mean square error (RMSE). The mean absolute error (MAE) states the difference between the actual and the forecasted value. However, in MAE the size of error is not always obvious. To deal with such problem, the MAPE can represent the error in percentage terms [47]. The MSE calculates the average of the squares difference between the true value and the estimated value. The mathematical representation of these errors are as follows [45]:

$$MAE = \frac{1}{N} \sum_{t=1}^{N} |A(t) - F(t)|$$

(10)

Here $A(t)$ is the actual value, $F(t)$ is the forecasted value, and $N$ is the number of values

$$MAPE = \frac{1}{N} * \sum_{t=1}^{N} \left| \frac{A(t) - F(t)}{A(t)} \right| * 100$$

(11)

$$MSE = \frac{1}{N} * \sum_{t=1}^{N} |A(t) - F(t)|^2$$

(12)

$$RMSE = \sqrt{\frac{1}{N} * \sum_{t=1}^{N} |A(t) - F(t)|^2}$$

(13)

2.4 Neural Network Toolbox

The NN toolbox in Matlab helps modeling very complicated nonlinear systems. It supports learning with feed forward (FF), radial basis (RB), and dynamic networks. The neural network toolbox can be used in different applications, such as data fitting, pattern recognition, time series prediction, and dynamic system modeling and control. The toolbox can be called by typing in the command window “nnstart” and the following window will show.
In this report, we will discuss the dynamic time series NN model that will be used in forecasting nonlinear time series. The time series tool can be called directly by typing “ntstool” in the window command. Time series tool can be used to train and solve three different nonlinear series problems.

2.4.1 Nonlinear Autoregressive with External Inputs (NARX)

Under this tool, a nonlinear time series problem with an input is solved. The forecasted values depend on $d$ past values of the time series $y(t)$ and the input $x(t)$ as well [48].

$$y(t) = f(y(t-1), ..., y(t-d), x(t-1), ..., x(t-d)) \quad (14)$$

2.4.2 Nonlinear Autoregressive (NAR)

In this model, only one time series is solved. In order to predict the future values, only the $d$ past values are considered. This form of prediction is referred to nonlinear autoregressive, and it was mentioned earlier and can be represented as follows [48]:

$$y(t) = f(y(t-1), ..., y(t-d)) \quad (15)$$
2.4.3 Nonlinear Input-Output

In this form of NN tool, two series are involved; however, the values of $y(t)$ are predicted using the $d$ past values of $x(t)$.

$$y(t) = f(x(t-1), ..., x(t-d)) \quad (16)$$

The first model NARX will provide better result than the third model, but in case the past values of $y(t)$ are not known, we can use the nonlinear input-output model [48].

![Figure 17 (a) Net structure for NARX. (b) NAR. (c) Nonlinear input-output [48]](image)

Let assume that NARX form was selected. In order to feed the input and output (target) data, they must be loaded from the Matlab workspace as shown below.
After the input and the target are loaded, the data is divided to three time steps: training, validation, and testing. The default distribution in Matlab is 70%, 15%, 15%, respectively. Then, the structure of the neural network is developed where the number of hidden neurons and the number of delays are entered. The number of delays refers to how many $d$ past values should be considered to predict the future values. The selection of these parameters is a challenge. Many techniques were proposed to find the optimal NN structure. However, these techniques are difficult to implement. Moreover, these methods will not guarantee the optimal forecasting solution for each type of problem. In most cases, trial and error is used to find the better value for hidden neurons and delay [42]. After that, the training algorithm is selected. Three training algorithms are available by the GUI: Levenberg Marquardt, Bayesian Regularization, and Conjugate Gradient. Each algorithm consumes different training time. Bayesian Regularization requires more time but may provide a good result. Then, select train and the training process starts as shown in the Figure 19.
The training process will continue until the validation check stops. All neural network training parameters such as epoch and gradient can be adjusted using Matlab code editing window. Once the training stops, the error is calculated and shown in next window. In case the result is not good, the training process can be repeated. Different hidden neuron numbers and training algorithm can be selected to improve the error. Finally, the Matlab GUI will provide different options such as generating the forecasted values, and simple or advanced script for the trained data [48].

2.6 Different Stages Data Processing

This section describes how the data is processed and forecasted. The GHI is the main component considered during the PV system output forecasting. The data of the
GHI was collected for fifteen years. The data are provided by NREL in [49]. The selected data is from 1996 to 2010 for Denver International Airport. In addition to the GHI historical data, the clear sky GHI was also collected. The clear sky data represents the maximum GHI that could be received during a clear sky day. The variability for solar irradiance does not exist during the clear sky day. The clear sky GHI was also collected from NREL, and it is available in [49]. Also, four different weather data were collected for the same site. The data includes cloud cover, temperature, wind speed, and dew temperature. The collected data is for fifteen years for the same period of the GHI data from 1996 to 2010. The weather data is available in [50]. The data is provided by NCDC. All these data are processed at different stages.

2.6.1 Data Preprocessing

The data preprocessing is a process that occurs before the data fed into the NN toolbox in Matlab. The preprocessing includes: removing the offset, normalizing, and removing nighttime GHI data.

2.6.1.1 Removing Offset

Under this process, the historical data of GHI is subtracted from the clear sky GHI in order to get the data, which represents the GHI scattered by cloudiness or other factors. The resulted data from deviation is the data going to be forecasted. The resulted GHI data from deviation is a function of time and location as the clear sky GHI and reflects all meteorological data that affects solar irradiance, such as cloudiness and aerosol. During
the forecasting process, other meteorological data is fed to the model to predict the future GHI. The following equation is used during data preprocessing and after data processing.

\[ GHI_{\text{deviation}} = GHI_{\text{CSK}} - GHI_{\text{Hist}}. \]  

(17)

2.6.1.2 Normalization

One of the techniques used in data preprocessing is the data normalization. The data normalization ensures the quality of the data before it is fed to the model. Some data set has extreme values that could lead to distorting the forecasted result. Normalization is performed in the inputs to distribute the data consistently and scale them to an acceptable range before feeding them into the model. The data will be normalized to a minimum and maximum [0,1] or [-1,1] [51]. The GHI data set is normalized to [0,1]. This accomplished by dividing the hourly historical data over the hourly maximum GHI, which is the hourly clear sky GHI. Meteorological data is also normalized to [-1,1], except cloud cover data, which already ranges from 0 to 1.

2.6.1.3 Removing Nighttime Hours

The solar radiation varies during the daytime, and it is zero during the nighttime. Therefore, part of data preprocessing is to remove the nighttime hours when the GHI is zero. Thus only daytime GHI hourly data are processed. This is accomplished by developing a Matlab code that keeps the daytime hours knowing the sunrise and sunset times. The sunrise and sunset times shift over the year. So, the total daytime hours are shorter during the winter season and longer during the summer season. The daytime hours, sunrise, and sunset times are fixed for each specific day for the same location over the years. This preprocess is also applied to other input data such as temperature and
cloud cover. After forecasting is completed another Matlab code will put back all the removed nighttime hours with GHI values equal to zero and end up with twenty-four hours in a day. The complete list of sunrise and sunset times can be found in the appendix [52].

2.6.2 Data Processing

In the forecasting process, the neural network toolbox is used to predict the future GHI and the electrical price. The GHI data is forecasted in to different ways and the error is compared. In the first method GHI is forecasted using historical GHI data. In the second model, the meteorological data is fed as input to forecast GHI data. The electrical price data for both fixed and variable options is forecasted using only historical data fed as input to the NN model. The neural network toolbox can be called in Matlab using the command “ntstool”. The NAR model is used to forecast GHI with historical data while NARX is used to forecast GHI using weather and historical GHI data. The historical and other weather data are fed to the model as input whereas the actual data is fed as a target. The NN structure is established by choosing the number of hidden neurons and delays. Once the NN structure is completed, the training process is started, and the error is calculated. If the error is not acceptable, the data is retrained after changing the NN structure. The output of the model represents the forecasted GHI data.
2.6.3 Data Post-Processing

The forecasted data represents only the daytime hours. Also, the forecasted data is in normalization form. So, two different processes have been implemented after GHI forecasted: denormalization, and adding nighttime hours.

2.6.3.1 Denormalization

The processed data is then multiplied by clear sky GHI data to produce the deviated GHI forecasted data. Then the forecasted GHI is calculated by subtracting the forecasted deviation GHI from the clear sky GHI using the following equation.

\[
GHI_{\text{forecasted}} = GHI_{\text{CSK}} - GHI_{\text{forecast\_deviation}}
\]  (18)

2.6.3.2 Adding Nighttime Hours

Here the denormalized data represents only the daytime hours. In order to have a complete data for fifteen years, a Matlab script is developed to add the nighttime hours where the GHI is zero. The sunrise and sunset times in addition to the daytime duration are used in the script to add the removed hours before. The flow chart in figure 20 describes all three stages process mentioned before.
Figure 20 Proposed Flowchart For GHI forecasting process
Chapter Three: PV System Economic Analysis

The economic analysis for a project is a fundamental requirement to investigate the viability of the project. The objective of the economic analysis is to help decision makers in evaluating the project and shape the decisions. The calculations of the economic analysis include different costs such as capital cost, maintenance and operation costs, and electricity cost.

The term feasibility means in term of engineering that a project is technically feasible or possible. The economic feasibility of a project refers to the project that can be viable with respect to the economic bases in addition to the technical feasibility [53].

In this chapter, the benefits of the long-term planning will be discussed in the first section. In the second section, some of economic aspects such as capital cost, payback period, and net present value will be introduced. These economic aspects are very important to evaluate the economic feasibility of a project. In section three, the concept of value of lost load (VOLL) will be discussed and how customers are being benefit from the VOLL. Section four is an overview of two different electricity prices: fixed and real-time prices. The last section covers the sensitivity analysis.
3.1 Economic Aspects

There are fundamentals concepts when evaluating the project feasibility. These concepts include: capital cost, payback period, discount rate, and net present value. The investment in renewable energy systems requires an initial payment, and the return of such initial cost will be from the future incomes. The first step to assess these investments is to understand the changes in the value of the money over the time in addition to other important concepts involved in energy economics.

3.1.1 The Total Cost of PV System

The total cost of a PV system includes the cost of the PV module, installations, labor, and balance of system (BOS). The balance of system refers to all other components in the PV system except the PV module. It includes inverters, switches, wirings, controller, and other auxiliary components. The cost of the PV system is measured in dollar per watt ($/W). The cost of the PV system varies based on the module type, size of the system, and the type of the PV owner such as residential, commercial, and utility scale. The PV system can be categorized based on the end user into four groups. Residential PV systems do not exceed 20kW and are usually roof mounted. Large-scale commercial PV systems are typically less than 1MW, and are placed on commercial buildings such as hospitals and universities. Third is the utility scale PV system which is usually more than 1MW and ground mounted. The last type is off grid applications that are different in sizes from small-scale for remote homes to large systems that are connected off the grid [4].
The figure below shows the capital cost of the PV system and how it declined over the last few years. For example, the PV system price for nonresidential sector has declined from $5 /W at the end of 2011 to almost $3.5 /W in the last quarter of 2013.

3.1.2 Net Cash Flow

The net cash flow represents the difference between the revenue and the expenses. The revenue covers the value of the generated energy from a PV system. While the expenses represents the outgoing money because of the operation and maintenance (O&M) costs [54]. The O&M costs are usually considered very small in the PV system. In [55], NREL has estimated the O&M costs for a 1MW PV system around $19,000/yr.

The total revenue can be calculated as follows:

\[
Revenue = \sum_y \left( \sum_h k_y \cdot E_h \cdot p_h \right)
\]

Here \( y \) represents the year, \( h \) is the hour, \( p_h \) is the electricity price, and \( E \) is the generated energy in kilowatt-hour (kWh).
3.1.3 Simple Payback Period

The simple payback period is one of the most common ways to evaluate the economic value of a project. It is the period of time required to cover the initial cost, i.e., the capital cost, of a PV system. The payback period is simply the ratio of the capital cost over the annual saving. The shorter the payback period of a project is the more attractive to invest in the project. The payback period can be represented as follows [54]:

$$PB = \frac{\text{Capital Cost ($)}}{\text{Annual Saving ($/y)}}$$ \hspace{1cm} (20)

3.1.4 Net Present Value (NPV)

The net present value (NPV) converts the future value of money to its present value [2]. For example, one dollar now is not as valuable as one dollar after 10 years. The NPV is a useful tool in the economic analysis. In the NPV calculation, the discount rate is taken in consideration. The discount rate works as a measure of the rate that could be earned as money has been out in an investment. It is also called the interest rate. There are different methods to forecast the discount rate. However, these methods are not considered in the scope of this thesis and the discount rate is assumed to be 10%. In order to convert the future money $F$ into a present value $P$, the following equation is used [2]:

$$P = \frac{F}{(1+d)^y-1}$$ \hspace{1cm} (21)

Here $y$ is the number years and $d$ is the discount rate. The term $1/(1+d)^{y-1}$ referred as present worth and can be incorporated to all future cash flows as $\kappa_y$. 

During the economic analysis of a PV system, the value of the generated energy over a period of time is converted to a present value, in order to investigate the viability
of the system. In most cases, the investment cost or the capital cost is a loan provided by a financial institution. The load is usually paid annually over a specific period. In order to calculate the net present value annuity, the following equation is used [56]:

$$NPV_{\text{Annity}} = C \times \left[ \frac{1-(1+d)^{-n}}{d} \right]$$

(22)

Here C is the cash flow per period and n is the number of payments.

3.2 Value of Lost Load (VOLL)

The value of lost load (VOLL) is an important measure in the electricity market. It can be defined as the price of unserved energy to customers due to any disruption [57]. The VOLL varies depending on the type of customers, the duration of the outage, and the time of the outage. The VOLL is a useful tool in different studies such as planning and operation. In the planning, the cost benefit of an investment in the generation, transmission, and distribution can be evaluated by using the VOLL to estimate the cost of customer outages. In operation side, different independent system operators (ISOs), such as Midwest Independent Transmission System Operator (“MISO”), have used the VOLL to inform their recourse adequacy and pricing algorithms. A study was conducted for Electric Reliability Council of Texas (ERCOT) to estimate the VOLL based on the type of customers. The VOLL is measured in ($/MWh). The VOLL for a developed industrial sector is from $9,000/MWh to $45,000/MWh. For residential customers the VOLL is the lowest and can be from $0/MWh to $17,967/MWh. For small commercial and industrial customers, the VOLL ranges from $3,000/MWh to $53,907/MWh. It is clear that small commercial and industrial customers have the highest VOLL. This is due to
that fact that small commercial and industrial customers require more labor and capital and more likely are not prepared for operational risks. Figure 21 shows the average VOLL by customer type [58].

Figure 21 Estimated Average VOLL by customer type [58]

Installing the PV system can provide reliability and a safe operation to meet the outages that could occur from time to time and hence avoid any payments because of the VOLL. In chapter four, the reliability benefits that the PV system could provide will be introduced by showing how much a PV system will serve to overcome any disruption. The benefit is the VOLL multiplied by the energy served by a PV system at the time of the disruption. Reliability benefit can be represented as follows [57]:

\[
\text{Reliability benefits} = \sum_y \sum_{dy} \sum_h \kappa_y \cdot VOLL_{y,dy,h} \cdot E_{y,dy,h} \quad (23)
\]

Here \( \kappa_y \) is the present worth coefficient, \( dy \) is the index for the day, \( y \) is the year, \( h \) for the hour, and \( E \) is the served energy by a PV system.
3.3 Electricity Price

The electricity price is the amount of the money per KWh that customer pays to an electricity provider to supply the electricity to the end user. The electricity providers take into account different factors to calculate the total generated cost of the electricity. The average electricity price in the United States varies from 7 ¢/kWh to 34 ¢/kWh, with the maximum in Hawaii [59].

Utilities often offer multiple electricity rate options. Understanding the rate structure and choosing the suitable electricity rate for businesses with photovoltaic installations will maximize the benefits of the PV generation. Here, two different rate options will be discussed: fixed rate and real time (variable) price [59].

3.3.1 Fixed Electricity Price

The fixed electricity rate is a fixed price that utility will charge the customer based on an agreed fixed rate. In this rate option, the utility and customer sign a long-term contract with a fixed electricity rate. The customer pays a fixed bill at every end of the month during the term of the agreement. The advantage of this rate option is that customer knows how much to pay at the end of the month and can plan for it. Also, if the utility price rate rises higher than the fixed rate due to any factor, the customer is still obligated to purchase the electricity at price lower than the market price. However, if the utility price falls below the market price, the customer could pay for electricity at rate higher than the market value. Moreover, the fixed rate contract usually penalizes for a fee in case of any earlier termination of contract [58].
3.3.2 Variable Electricity Price

The variable electricity rate varies hourly based on the electricity market price. The rate fluctuates with market conditions. The customer pays less when the energy price goes down, and higher when the rate goes up. The variable electricity price is also referred as the real time price. In the real time prices, the customers receive information about the electricity cost at all times. Customers must pay the same offered price at any given time. Customers have to adjust their electricity usage accordingly to avoid high rates and shift their usage to the time when rate is low. Comed offers a real time pricing and day ahead pricing at its website where customers can check the hourly electricity rate and shift the electricity usage accordingly [60].

3.4 Sensitivity Analysis

The sensitivity analysis is to test the impact of changing key variables on particular dependent variables. For example, changing the fuel cost to determine the effect on the cash flow. Thorough the sensitivity analysis, different scenarios can be created to check how changes in one variable could impact the target variable. The sensitivity analysis is considered one of the ways to shape the outcome of the decisions. It helps the analyst to determine which parameters are the key factors that affect the target variable. In the engineering economy, the sensitivity analysis is defined as the study of the economic effect of different uncertain variables on the economics of an investment. For instant, sensitivity analysis shows how the economic payback period is sensitive to
uncertain inputs. In chapter four, the sensitivity analysis will be conducted to evaluate how the payback period changes over the uncertain value such as forecasting error. It will show how the NPV of the cash flow will be affected based on the electricity price option [61].

3.5 PV vs. CO2 Emission

The United States Environmental Protection Agency (EPA) has provided the way to calculate how much CO2 emissions can be avoided by the reduced electricity kWh. The U.S EPA uses the emissions and generation resource integrated database to develop the green power equivalency calculator. The calculator can be used to show how renewable energy resources could mitigate CO2 emissions from other conventional power plants. The emission factor for CO2 is as follows [62].

\[6.9827 \times 10^{-4} \text{ metric tons of CO2/kWh} \quad (24)\]

For example, a 200 kW PV system generates based on PVWatts calculator provided by NREL [63] around 300,768 kWh annually. To calculate the amount of CO2 avoided by installing such system:

\[
\text{CO2 avoided} = 6.9827 \times 10^{-4} \text{ metric tons of CO2/kWh} \times 300,768 \text{ kWh/year} \\
= 200 \text{ ton Co2/year}
\]
Chapter Four: Case Study

In this chapter, a thorough analysis will be conducted to what was mentioned in earlier chapters through a case study. The study starts with collecting fifteen years of hourly global horizontal irradiance (GHI), electricity variable price, and meteorological data. All data will be preprocessed for the forecasting process. Different neural network methods are used to forecast the GHI and electricity price. The error will be evaluated and different methods will be used to reduce the forecasting error. After that, the economic analysis will be conducted in the PV system with use of the forecasted GHI data and prices in order to determine the PV system viability.

The last part of the study is to conduct the sensitivity analysis in the PV system by using different forecasted GHI and electricity prices with different errors. The analysis will reveal how error could impact the payback period of the system. Finally, outcome results will be discussed, and recommendations to improve the result of the case study will be presented.
4.1 Data Preparation and Preprocessing

This section describes where the data was collected from and how the data is preprocessed for forecasting. The data that will be covered under this section includes GHI, meteorological data, and price data. Also, different preprocesses are involved such as normalization and deviation.

4.1.1 Global Horizontal Irradiance (GHI)

As mentioned before, the GHI represents the sum of other irradiance components: the DNI and DHI. Therefore, the GHI is the main component considered during the PV system output forecasting. The data of the GHI was collected for fifteen years. The data is provided by NREL in [49]. The data selected is from 1996 to 2010 for Denver International Airport. Figures 23 and 24 below show the hourly GHI data for fifteen years and four days, respectively. The GHI is represented in \( W/m^2 \).

![Hourly GHI data from 1996 to 2010 for Denver International Airport](image)

Figure 23 Hourly GHI data from 1996 to 2010 for Denver International Airport
Four different types of weather data were collected for the same site. The data includes cloud cover, temperature, wind speed, and dew temperature. The data collected is for fifteen years for the same period of GHI data from 1996 to 2010. The data provided by the NCDC is available in [50]. The cloud cover ranges from 0 to 1 where 0 is clear sky and 1 is overcast. Temperature is measured in Fahrenheit (F). Dewpoint temperature refers to the temperature at which water vapor is no longer held by air and some water vapor must be condensed into liquid water. It is also measured in Fahrenheit (F). Wind speed is measured in (m/s).
4.1.3 Clear Sky GHI

Clear sky data represents the maximum GHI that could be received during a clear sky day. The variability for solar irradiance does not exist during a clear sky day. The plot of GHI during a clear sky day represents a smooth GHI reading. Clear sky GHI was collected from NREL and it is available in [49]. It is worth mentioning that clear sky GHI is the same irradiance at the same time and location. Figure 24 below shows the plot for one day GHI during a clear sky day and a cloudy day.

![Figure 25 Clear Sky GHI vs. Cloudy day GHI](image)

4.1.4 Electricity Price

The electricity price or tariff is the amount of money that costumers have to pay for their usage of power. As mentioned before two electricity rate options are going to be addressed in this study. The first plan is a fixed price option where the price is locked during the contact duration. The fixed price is assumed to be 8¢/kWh for the first year and the contract states that every year there will be a 10% increment from the previous
value. The contract is a long-term contract for 15 years. The fixed price data will be developed based on the contract criteria and then this data is used to forecast these prices again using the NN toolbox. The second price option is the variable price. With the variable rate, the electricity price changes in hourly basis. The real time prices collected from Comed are available in [60].

4.1.5 Normalization

One of the techniques used in data preprocessing is data normalization. Data normalization ensures the quality of data input before it is fed to the model. Some data sets have extreme values that could lead to distorting the forecasted result. Normalization is performed to the inputs to distribute the data consistently and scale it to an acceptable range before feeding it into the model. The data will be normalized to a minimum and maximum [0,1] or [-1,1] [51]. The GHI data set is normalized to [0,1]. This is done by dividing the hourly historical data over the hourly maximum GHI, which is the hourly clear sky GHI. Meteorological data is normalized to [-1,1] except cloud cover data that ranges from 0 to 1.

4.1.6 Removing Offset

Under data deviation, the historical data of GHI is subtracted from clear sky GHI in order to get the data that represents GHI scattered due to cloudiness or other factors. The resulted data from deviation is the data that will be used for forecasting. The resulted GHI data from deviation is a function of time and location as the clear sky GHI and reflects all meteorological data that affect solar irradiance such as cloudiness and aerosol.
During the forecasting process other meteorological data is fed to the model to predict the future GHI that here represents obstacles for GHI to be received at the ground surface. The forecasted deviated GHI is subtracted from the clear sky data again to get the historical GHI that is main component for PV system output power. The following equation is used during data preprocessing and after data processing.

\[ GHI_{\text{deviation}} = GHI_{CSK} - GHI_{Hist}. \]

Figure 26 GHI deviation between clear sky GHI and historical day GHI

4.1.7 Removing Nighttime Hours

The solar radiation varies during the daytime and is zero during the nighttime. Therefore, part of data preprocessing is to remove the nighttime hours when the GHI is zero, where only daytime GHI hourly data are processed. This is accomplished by developing a Matlab code that keeps the daytime hours knowing the sunrise and sunset times. The sunrise and sunset times is shifted over the year. So, the total daytime hours are shorter during the winter season and longer during the summer season. The daytime
hours, sunrise and sunset times are fixed for each day for the same location over the years. This preprocess is also applied to other input data such as temperature and cloud cover. After forecasting is completed another Matlab code will put back all the removed nighttime hours with GHI values equal to zero and end up with twenty-four hours in a day. For example, table 3 shows selected days with their sunrise and sunset times and daytime duration. These data can be found in [52].

<table>
<thead>
<tr>
<th>Day</th>
<th>Sunrise</th>
<th>Sunset</th>
<th>Daytime hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>22-Jan</td>
<td>7:15</td>
<td>17:06</td>
<td>9 h, 51 min</td>
</tr>
<tr>
<td>18-Jul</td>
<td>5:46</td>
<td>20:24</td>
<td>14 h, 38 min</td>
</tr>
<tr>
<td>3-Nov</td>
<td>6:30</td>
<td>16:54</td>
<td>10 h, 24 min</td>
</tr>
</tbody>
</table>

Table 3 Sunrise and Sunset time and Daytime hours

![Figure 27 Daytime hourly GHI in cloudiness and clear day](image)

Figure 27 Daytime hourly GHI in cloudiness and clear day
4.2 Forecasting Process

In the forecasting, NN toolbox is used to predict future GHI and electricity price. The GHI data is forecasted in different ways and the error is compared. In the first method, GHI is forecasted using historical GHI data. In the second model, the meteorological data is fed as input to forecast GHI data. The electrical price data for both fixed and variable options is forecasted using only historical data fed as input to neural network model. The neural network toolbox can be called in Matlab using the command “ntstool”.

4.2.1 GHI Forecasting Using only Historical Data

As mentioned before, the GHI data is preprocessed before fed to the model by developing Matlab code that does the following. First, the data was gone through deviation process. Then, the nighttime hours with zero GHI are removed. Next, the data is normalized to [0,1]. After that, the GHI data is forecasted using Nonlinear Autoregressive (NAR) model in the NN toolbox. The GHI data represents the hourly GHI data for fifteen years and arranged as row matrix. In appendix, sample of GHI data is provided. The following steps illustrate how the forecasting is done through NN toolbox. First the time series NN toolbox is called by typing the command “ntstool”. Then the NAR model is selected as shown below.
Preprocessed and normalized GHI data with total number of 70590 hours is fed as target in model as shown below.

Then, the data is divided to three parts: training, validation, and testing. The default data distribution is 70%, 15%, and 15%, respectively. [49]
After that, the structure for the neural network is developed. The number of hidden nodes and of delays is set to 10 and 2, respectively.

It is clear that the chosen equation to produce output from hidden layer is sigmoid function. Finally, the training algorithm is chosen and the training process starts to train and validate the data and finally produce the forecasted value. The MSE is calculated during the training process. This error represents the error for preprocessed data. The final error is shown after that forecasted data is post processed.
The processed data is then multiplied by clear sky GHI data to produce the deviated GHI forecasted data. Then the forecasted GHI is calculated by subtracting the forecasted deviation GHI from the clear sky GHI using the following equation.

\[ GHI_{\text{forecasted}} = GHI_{\text{CSK}} - GHI_{\text{forecast deviation}} \]

The calculated MAPE is around 19.8738%. The plot for both actual and forecasted GHI is shown below. Sample of the resulted data is found in the appendix.

Figure 32 Training Process

Figure 33 Actual and forecasted GHI using historical data only
4.2.2 GHI Forecasting Using Meteorological Data as input

Four different weather data were collected for the same site. The data includes cloud cover, temperature, wind speed, and dew point. The data is provided by NCDC. The NN architecture for is shown below.

All weather data is preprocessed before fed into the model. The inputs are normalized and then night hours are removed. The meteorological data fed to the model as inputs and preprocessed GHI fed as target. The numbers of neurons nodes and delays are 8 and 1, respectively. The produced output GHI data is then post processed. The calculated MAPE is 10.3460%. The error has been significantly reduced here by almost 10%.
4.2.3 Forecasting Fixed Electricity Price

The fixed price is in the first year is 3.24 €/kWh that is the average price of the first year of real time price. The long-term contract states that the price increases at an average of 10% each year. The Matlab script is developed to produce the electric price for over fifteen years. The produced fixed price is fed as target in the NN toolbox using NAR model. The forecasted data shows an error of 1.3132 %. This small error due to the fact the pattern of the fixed price is recognizable and the model can easily predict the next price.
4.2.4 Forecasting Electricity Real Time Price

The real time prices RTP are obtained from Comed [60] for one year. Then the real time prices are exposed 10% increment annually. The developed RTPs are then fed to NN toolbox as target using NAR model. The output data has MAPE of 21.684%. This error is due to variation of the RTP.
Table below summarizes the MAPE for the previous forecasted data.

<table>
<thead>
<tr>
<th>Data Forecasted</th>
<th>MAPE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GHI without weather data</td>
<td>19.874</td>
</tr>
<tr>
<td>GHI with weather data as input</td>
<td>10.346</td>
</tr>
<tr>
<td>Fixed Price</td>
<td>1.3132</td>
</tr>
<tr>
<td>Real Time Price</td>
<td>21.684</td>
</tr>
</tbody>
</table>

Table 4 MAPE for GHI and Price forecasting

4.3 Economic Analysis

Under this section different economic analyses will be carried out to determine the viability of a PV system. It is assumed that the PV system will be installed in a university campus and the PV arrays are rooftop mounted in campus parking. The PV system characteristics are as follow.

<table>
<thead>
<tr>
<th>PV system Capacity</th>
<th>1 MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module Type</td>
<td>Standard (crystalline Silicon)</td>
</tr>
<tr>
<td>Model efficiency</td>
<td>16 %</td>
</tr>
<tr>
<td>Required Total area</td>
<td>6250 m²</td>
</tr>
<tr>
<td>Warranty lifetime</td>
<td>30 years</td>
</tr>
</tbody>
</table>

Table 5 PV system design specification

The module nameplate size is for standard test conditions (STC): Solar irradiance of 1,000 W/m2, cell temperature of 25°C (77°F), and air mass of 1.5. In order to calculate the total area required for the PV system, the following equation is used.
\[ A(m^2) = \frac{PV \, Size \, (W)}{1000 \, W/m^2 \, \times \, Model \, efficiency} = \frac{1,000,000}{1000 \, \times \, 0.16} = 6250 \, m^2 \]

4.3.1 Total PV System Cost

Total PV system cost includes capital cost and fixed operation and maintenance costs. The capital cost consists of the PV module, installations, labor, and balance of system (BOS). The PV system is installed at commercial sector and Figure 20 shows the cost range for small commercial and industrial sectors. The PV cost is assumed to be $3.5/W and the fixed O&M cost is $19,000/year.

Total PV system cost = 3.5 $/W * 1,000,000 W = $3,500,000

PV fixed O&M cost = $19,000 /year

The total PV cost is assumed to be a loan from a bank with interest rate of 10% and payment should be made annually over 20 years. The annual cash flow for loan payments is going to be $175,000. In order to evaluate the feasibility of the PV system all future cash flow should be evaluated in present value. So, the net present value for loan payments is as follow.

\[
NPV_{CP} = C \times \left[ \frac{1-(1+d)^{-n}}{d} \right] = 175,000 \left[ \frac{1-(1+0.1)^{-20}}{0.1} \right] = $1,489,874
\]

OR
\[
NPV_{CP} = \sum_{y=1}^{20} k_y \times C_y = $1,489,874
\]

Also, the net present O&M fixed cost is represented as follow where the annual O&M is $19,000/year.
\[ \text{NPV}_{0\&M} = \sum_{y=1}^{15} \kappa_y O\text{M}_y = 75,049 \]

\[ \text{NPV}_{\text{Total\_cost}} = \text{NPV}_{C\text{P}} + \text{NPV}_{0\&M} = 1,564,923 \]

4.3.2 Total PV System Energy Value

The generated energy for the PV system over fifteen years is calculated in three different scenarios. The energy and its saving is calculated with the actual data where the error is not considered, when the error is large, and when the error is small.

4.3.2.2 Energy Value with Actual Data

The total energy generated by the PV system over fifteen years is calculated as follow.

\[ E_{\text{gen}} = \sum_{y=1}^{15} \sum_{dy} \sum_{h} E_{y,dy,h} (\text{kW/m}^2) \times \eta(0.16) \times A(6250 \text{m}^2) = 25,788,857 \text{kWh} \]

The value of this generated energy is calculated in two different scenarios. The first one is when the fixed cost option is applied and the second option is when the real time price is applied.

Energy Value_{\text{fixed\_price}}

\[ = \sum_{y=1}^{15} \sum_{dy} \sum_{h} \kappa_y \times E_{y,dy,h} (\text{kW/m}^2) \times \eta(0.16) \times A(\text{m}^2) \]

\[ \times \text{Fixed}_{\text{Price}} (\$/\text{kWh}) = 835,550 \]
Energy Value\textsubscript{RTP}

\[
= \sum_{y=1}^{15} \sum_{dy} \sum_{h} \kappa_y * E_{y,dy,h} (\text{kW/m}^2) * \eta (0.16) * A (\text{m}^2) * \text{RTP} (\$/\text{kWh})
\]

\[
= $929,714
\]

\textbf{4.3.2.3 Energy Value with Large Error}

In this case, the error is taken in consideration when calculating the energy generated from the PV system. The MAPE is relatively high when forecasting with only historical data comparing when forecasting with weather data as input. The total generated energy is calculated.

\[
E_{\text{gen}} = \sum_{y=1}^{15} \sum_{dy} \sum_{h} E_{\text{with large error},y,dy,h} (\text{kW/m}^2) * \eta (0.16) * A (6250 \text{m}^2)
\]

\[
= 19,020,183 \text{kWh}
\]

The revenue from producing this value of generated energy is also calculated in two different tariff rates.

Energy Value\textsubscript{\text{fixed,price}}

\[
= \sum_{y=1}^{15} \sum_{dy} \sum_{h} \kappa_y * E_{\text{with large error},y,dy,h} (\text{kW/m}^2) * \eta (0.16) * A (\text{m}^2)
\]

\[
* \text{Fixed Price} (\$/\text{kWh}) = $608,633
\]
Energy Value_{RTP}

\[
= \sum_{y=1}^{15} \sum_{dy} \sum_{h} \kappa_{y} \ast E_{\text{with large error,y,dy,h}} (\text{kW/m}^2) \ast \eta(0.16) \ast A(m^2)
\]

\ast \text{RTP} (\$/kWh) = $683,026

4.3.2.4 Energy Value with Small Error

In this case, the error is also taken into consideration when calculating the energy generated from the PV system. The MAPE here is low since the weather data as is used as input to the model.

The total generated energy is calculated.

\[
E_{\text{gen}} = \sum_{y=1}^{15} \sum_{dy} \sum_{h} E_{\text{with small error,y,dy,h}} (\text{kW/m}^2) \ast \eta(0.16) \ast A(6250m^2)
\]

\[
= 25,472,401\text{kWh}
\]

The revenue from producing this value of generated energy is also calculated in two different tariff rates.

Energy Value_{\text{fixed,price}}

\[
= \sum_{y=1}^{15} \sum_{dy} \sum_{h} \kappa_{y} \ast E_{\text{with small error,y,dy,h}} (\text{kW/m}^2) \ast \eta(0.16) \ast A(m^2)
\]

\ast \text{Fixed}_{\text{Price}} (\$/kWh) = $815,055

Energy Value_{\text{RTP}}

\[
= \sum_{y=1}^{15} \sum_{dy} \sum_{h} \kappa_{y} \ast E_{\text{with small error,y,dy,h}} (\text{kW/m}^2) \ast \eta(0.16) \ast A(m^2)
\]

\ast \text{RTP} (\$/kWh) = $915,242

74
4.3.3 Reliability Benefit

The value of the lost load (VOLL) is the price of unserved energy to customers due to any disruption. Figure 21 shows the range for the VOLL for different sectors. The VOLL in this study is assumed to be $30/kWh. The VOLL is applied here randomly at two different hours every year. The cost saving from VOLL over fifteen years is estimated as follow.

Reliability saving ($)
\[
\text{Reliability}_{\text{saving}}(\$) = \sum_{y=1}^{15} \sum_{dy} \sum_{h} \kappa_y \times LS_{y,dy,h}(\text{kW/m}^2) \times \eta(0.16) \times A(m^2) \times \text{VOLL}
\]
\[
= \$148,731
\]

Here LS is load shedding. The reliability benefit is also calculated in two different cases where the error is high and low.

Reliability saving, large error ($)
\[
\text{Reliability}_{\text{saving, large error}}(\$) = \sum_{y=1}^{15} \sum_{dy} \sum_{h} \kappa_y \times LS_{\text{large error},dy,h}(\text{kW/m}^2) \times \eta(0.16) \times A(m^2) \times \text{VOLL}
\]
\[
= \$114,300
\]

Reliability saving, SE ($)
\[
\text{Reliability}_{\text{saving, SE}}(\$) = \sum_{y=1}^{15} \sum_{dy} \sum_{h} \kappa_y \times LS_{\text{SE},dy,h}(\text{kW/m}^2) \times \eta \times A \times \text{VOLL} = \$146,473
\]

4.3.3.1 Reliability and CAIDI

The Customer Average Interruption Duration Index (CAIDI) is an index that measures the average amount of the time when a customer has no power. It is a reliability
measure used by electric power utilities. It can be calculated by dividing the total interruption duration of all customers by the total number of the customers [64].

In order to evaluate the affect of VOLL on the PV system investment based on different CAIDIs, the total payback period has been reevaluated using different scenarios of CAIDI. In section 4.3.3 the reliability cost provided by VOLL was estimated assuming two hours of power interruption incident. The average power that the PV system can supply to avoid payments for the VOLL at the time of the incidents is 300 kWh. The resulted payback periods were 24 and 22 for the fixed and real time price, respectively. The simulation has shown that whenever the CAIDI has increased, the total payback periods are reduced. The table below shows different CAIDI and the resulted payback period.

<table>
<thead>
<tr>
<th>CAIDI (hours/year)</th>
<th>Fixed Cost Option</th>
<th>RTP Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 (base case)</td>
<td>24</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>22</td>
<td>21</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>17</td>
</tr>
</tbody>
</table>

Table 6 Payback periods for different CAIDI incidents

The evaluation was done assuming that the incidents have occurred when the PV output is around 300kWh. The simulation was repeated assuming the two hours case incident has occurred when the PV system is at its maximum output capacity, which is 1 MWh. The result shows a significant reduction in the payback period of about 20% from
the base case. The payback periods were 19 and 18 years for the fixed and real time price, respectively.

4.4 Results

After the economic analysis has been performed, the payback period should be estimated to reveal the viability of the PV system. The payback is the period of time required to cover the PV system total cost. Payback period simply is the ratio of capital cost over annual saving.

4.4.1 Payback Period Without Error

In this calculation, the payback period is calculated considering the actual data. The calculated total revenue and reliability cost is considered to estimate the payback period. The shorter payback period project is more attractive to invest in such project. The payback period can be calculated as in (20).

The total present value of PV system cost is estimated before as $1,564,923. The benefit from the PV system comes from selling the produced energy and improved reliability. The total revenue is calculated when the PV system goes with fixed rate option and variable rate option.

4.4.1.1 Payback Period with Fixed Price

In order to calculate the annual saving from selling PV energy, the total energy revenue is divided by the number of the years that is fifteen.

\[
\text{Total Revenue} = \left(\text{Energy Value}_{\text{fixed price}} + \text{VOLL}_{\text{saving}} \ (\$)\right) / 15 = \$ 65,485 / yr
\]

\[
\text{PB}_{\text{no error, fixed price}} = \frac{\text{Capital Cost} \ (\$)}{\text{Annual Saving} \ (\$ / yr)} = \frac{\$ 1,564,923}{\$ 65,485 / yr} = 23.90 \approx 24 \text{ years}
\]
4.4.1.2 Payback Period with Variable Price

Here the real time price option is applied. The calculated payback period for real time price with actual data considered is as follow.

\[
\text{Total Revenue} = \text{Energy Value}_{\text{RTP}} + \text{VOLL}_{\text{saving}} (\$/15 = \$71,896/\text{yr}
\]

\[
\text{PB}_{\text{no, erro, RTP}} = \frac{\text{Capital Cost (\$)}}{\text{Annual Saving (\$ yr{\text{-1}})}} = \frac{\$1,564,923}{\$71,896/\text{yr}} = 21.77 \approx 22 \text{ years}
\]

4.4.2 Payback Period With Large Error

The forecasted GHI has relatively high error this case. The payback period is also calculated under two different electric tariffs.

4.4.2.1 Payback Period With Fixed Price

The payback period is calculated considering the error and fixed price as follow.

\[
\text{Total Revenue} = \text{Energy Value}_{\text{fixed, price}} + \text{VOLL}_{\text{saving}} (\$/15 = \$48,196/\text{yr}
\]

\[
\text{PB}_{\text{Large, erro, fixed, price}} = \frac{\text{Capital Cost (\$)}}{\text{Annual Saving (\$ yr{\text{-1}})}} = \frac{\$1,564,923}{\$48,196/\text{yr}} = 32.47 \approx 33 \text{ years}
\]

4.4.2.2 Payback Period With Variable Price

The payback period in this scenario is calculated considering the error for both GHI and real time prices.

\[
\text{Total Revenue} = \text{Energy Value}_{\text{RTP}} + \text{VOLL}_{\text{saving}} (\$/15 = \$53,155/\text{yr}
\]

\[
\text{PB}_{\text{Large, erro, RTP}} = \frac{\text{Capital Cost (\$)}}{\text{Annual Saving (\$ yr{\text{-1}})}} = \frac{\$1,564,923}{\$53,155/\text{yr}} = 29.44 \approx 30 \text{ years}
\]
4.4.3 Payback Period With Small Error

Under this section, the error has been reduced and the payback period is calculated also based on two different electricity rates.

4.4.3.1 Payback Period With Fixed Price

The payback period calculations are performed considering small error and fixed price option.

\[
\text{Total Revenue} = \text{Energy Value}_{\text{fixed price}} + \text{VOLL}_{\text{saving}} (\$/15 = \$64102 /\text{yr}
\]

\[
\text{PB}_{\text{Small error fixed price}} = \frac{\text{Capital Cost (\$)}}{\text{Annual Saving (\$/yr)}} = \frac{\$1,564,923}{\$64,102 /\text{yr}} = 24.4 \approx 25 \text{ years}
\]

4.4.3.2 Payback Period With Variable Price

The payback period in this scenario is calculated considering the error for both GHI and real time prices.

\[
\text{Total Revenue} = \text{Energy Value}_{\text{RTP}} + \text{VOLL}_{\text{saving}} (\$/15 = \$70,781 /\text{yr}
\]

\[
\text{PB}_{\text{Small error RTP}} = \frac{\text{Capital Cost (\$)}}{\text{Annual Saving (\$/yr)}} = \frac{\$1,564,923}{\$70,781 /\text{yr}} = 22.3 \approx 23 \text{ years}
\]

The PB periods for different scenarios can be summarized in the following table.

<table>
<thead>
<tr>
<th>Data forecasting error</th>
<th>Fixed Price</th>
<th>Real Time Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>No error</td>
<td>24</td>
<td>22</td>
</tr>
<tr>
<td>Large error</td>
<td>33</td>
<td>30</td>
</tr>
<tr>
<td>Small error</td>
<td>25</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 7 Payback periods for different scenarios
4.5 Discussions

The economic analysis carried out for the PV system under different scenarios resulted in different payback periods. The first scenario is considered under actual data where error is zero. The payback periods for the different electric price options are relatively close. When the PV system under real time price option, the payback period is the shortest and indicates that the investment of the PV system is very attractive. Since the system lifetime is considered 30 years, the total revenue for the upcoming next 8 years are considered as profits from the system. The PV system shows a little higher payback period when under the fixed price option. This is due to the fact that under real time price PV power output can be sold at high peak period during the daytime. However, when the rate of electricity is low at evening times, the PV is not producing any power since the battery storage system is not incorporated in this study. In fixed rate option, the PV system may not benefit from peak periods since the electricity price is fixed.

The second scenario is performed while the error is taken in consideration. Under this scenario, the GHI data was simulated using only historical data that results in higher MAPE. The payback periods are higher due to the fact the forecasted GHI data are smaller than the actual data by around 20%. The payback period under fixed price option is 33 years and under real time price is 30 years. Under the fixed price the payback period is more than the lifetime of the PV system. Hence, with high payback period the investment in the PV system is not attractive. With the real time price, the payback period
is 30 years that is the same lifetime of the PV system. The resulted payback period reveals that the investment of the PV system is not economical.

In the third scenario, the MAPE is reduced significantly by 50% and the resulted payback periods are 25 and 23 years for the fixed and real time price options, respectively. The resulted payback periods when the error is reduced to 10% are almost close to the actual scenario. This concludes that accurate forecasting will help estimating the viability of PV system investment, while forecasting with large error will underestimate the viability of the PV system investment. Also, the selection of the electric price option reveals that real time price option is more profitable than fixed price option.

In different scenarios, the payback period was evaluated with different number of incidents per year and the results show as the CAIDI increases, the PV system supply increases the reliability benefit and hence the payback period will be reduced. The payback period was reduced by almost 25% when the incidents have increased from two hours to six hours per year. Also, when the incident has occurred at a time where the PV system has maximum power production, a significant reduction in the payback period has been noted. Finally, we conclude that investment in PV system in an area, where the reliability of power supply is low, can be economically feasible.
Chapter Five: Summary

5.1 Summary

In this report, the long term planning has been performed for a PV system under forecasting uncertainty. A detailed review for the PV technologies was introduced in chapter one. In chapter two, a literature review of the forecasting methods was performed. Then, a case study for PV system with 1MW size was conducted. Different economic analyses have been performed under different scenarios. The viability of the investment of the PV system was revealed by calculating the payback period of the system under different scenarios. All economic analyses for different scenarios are summarized in a table under the Appendix.

5.2 Conclusion

The investment of PV system faces different barriers. One of the major barriers is the variability of the solar radiation. Such variability could result in making incorrect decisions during the economic analysis of a PV system. In this research, the PV technology has been discussed thoroughly and different technologies have been introduced. Then literature review for different forecasting methods is introduced. Forecasting aims to reduce the uncertainty of solar output so system operator will be able to accommodate its variability. The case study conducted in chapter four has revealed
that accurate forecasting will result in making better planning decisions. The viability of the PV system introduced in the case study was infeasible when the forecast error was high. However, when the forecast error was reduced by almost 10%, the viability of the PV investment was feasible. The results indicate that payback period with small error is almost close to what analysis shows with actual data. Also, results have shown that although real time electricity price forecasting has high error, the payback period doesn’t change significantly. That concludes that PV system planning is highly dependent on the uncertainty of the forecast. Finally, by installing the 1MW PV system the total avoided CO2 emission is 1,201 tons CO2/year.

5.3 Recommendations

The following recommendations can be applied to improve the planning of the PV system under uncertainty.

- Combine different forecasting methods with the performed method in order to reduce the error.
- Improve the forecasting error of the real time price by introducing different inputs such as load demand and fuel prices.
- Use different module type of PV with higher efficiency and evaluate the impact on the system viability.
5.4 Future Work

The conducted study can be extended to perform the following future work:

- The storage system can be added to the PV system in order to overcome zero output at nighttime.
- Use hybrid model to improve the accuracy of the forecasting.
- Perform short term forecasting and evaluate the error.
- Unit commitment and economic dispatch can be studied under PV short term forecast.
- Include other renewable energy sources such as wind power.
References


## Appendix

### Summary of PV System Economic Analysis

<table>
<thead>
<tr>
<th></th>
<th>Actual Data</th>
<th>Large Error</th>
<th>Small Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fixed Price</td>
<td>RTP</td>
<td>Fixed Price</td>
</tr>
<tr>
<td>Investment Cost ($)</td>
<td>1,564,923</td>
<td>1,564,923</td>
<td>1,564,923</td>
</tr>
<tr>
<td>Revenue ($)</td>
<td>835,550</td>
<td>929,714</td>
<td>608,633</td>
</tr>
<tr>
<td>Reliability ($)</td>
<td>148,731</td>
<td>148,731</td>
<td>114,300</td>
</tr>
<tr>
<td>Energy generated (kWh/year)</td>
<td>1,719,257</td>
<td>1,719,257</td>
<td>1,268,012</td>
</tr>
<tr>
<td>CO2 avoided (tons/year)</td>
<td>1,201</td>
<td>1,201</td>
<td>885</td>
</tr>
<tr>
<td>Payback period (year)</td>
<td>24</td>
<td>22</td>
<td>33</td>
</tr>
</tbody>
</table>
GHI Sample Clear Sky Data (W/m²)

<table>
<thead>
<tr>
<th>Hour</th>
<th>GHI</th>
<th>Hour</th>
<th>GHI</th>
<th>Hour</th>
<th>GHI</th>
<th>Hour</th>
<th>GHI</th>
<th>Hour</th>
<th>GHI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>12</td>
<td>488</td>
<td>24</td>
<td>0</td>
<td>36</td>
<td>415</td>
<td>48</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>13</td>
<td>428</td>
<td>25</td>
<td>0</td>
<td>37</td>
<td>305</td>
<td>49</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>14</td>
<td>315</td>
<td>26</td>
<td>0</td>
<td>38</td>
<td>158</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>15</td>
<td>163</td>
<td>27</td>
<td>0</td>
<td>39</td>
<td>24</td>
<td>51</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>16</td>
<td>25</td>
<td>28</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>52</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>29</td>
<td>0</td>
<td>41</td>
<td>0</td>
<td>53</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>30</td>
<td>18</td>
<td>42</td>
<td>0</td>
<td>54</td>
<td>18</td>
</tr>
<tr>
<td>8</td>
<td>19</td>
<td>19</td>
<td>0</td>
<td>31</td>
<td>146</td>
<td>43</td>
<td>0</td>
<td>55</td>
<td>148</td>
</tr>
<tr>
<td>9</td>
<td>153</td>
<td>20</td>
<td>0</td>
<td>32</td>
<td>296</td>
<td>44</td>
<td>0</td>
<td>56</td>
<td>300</td>
</tr>
<tr>
<td>10</td>
<td>308</td>
<td>21</td>
<td>0</td>
<td>33</td>
<td>411</td>
<td>45</td>
<td>0</td>
<td>57</td>
<td>416</td>
</tr>
<tr>
<td>11</td>
<td>425</td>
<td>22</td>
<td>0</td>
<td>34</td>
<td>472</td>
<td>46</td>
<td>0</td>
<td>58</td>
<td>478</td>
</tr>
<tr>
<td>12</td>
<td>487</td>
<td>23</td>
<td>0</td>
<td>35</td>
<td>474</td>
<td>47</td>
<td>0</td>
<td>59</td>
<td>480</td>
</tr>
</tbody>
</table>
Clear Sky Data For 15 years

Sample of real time prices from Comed.

<table>
<thead>
<tr>
<th>Time (Hour Ending)</th>
<th>Day-Ahead Hourly Price</th>
<th>Real-Time Hourly Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00 AM</td>
<td>2.24</td>
<td>2.34</td>
</tr>
<tr>
<td>1:00 AM</td>
<td>2.24</td>
<td>1.54</td>
</tr>
<tr>
<td>2:00 AM</td>
<td>1.94</td>
<td>0.24</td>
</tr>
<tr>
<td>3:00 AM</td>
<td>1.64</td>
<td>0.84</td>
</tr>
<tr>
<td>4:00 AM</td>
<td>1.24</td>
<td>0.14</td>
</tr>
<tr>
<td>5:00 AM</td>
<td>1.14</td>
<td>1.94</td>
</tr>
<tr>
<td>6:00 AM</td>
<td>1.04</td>
<td>0.34</td>
</tr>
<tr>
<td>7:00 AM</td>
<td>1.54</td>
<td>0.84</td>
</tr>
<tr>
<td>8:00 AM</td>
<td>2.24</td>
<td>0.84</td>
</tr>
<tr>
<td>9:00 AM</td>
<td>2.54</td>
<td>2.74</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>2.94</td>
<td>1.64</td>
</tr>
<tr>
<td>11:00 AM</td>
<td>3.14</td>
<td>1.94</td>
</tr>
<tr>
<td>12:00 PM</td>
<td>3.34</td>
<td>2.04</td>
</tr>
<tr>
<td>1:00 PM</td>
<td>3.44</td>
<td>3.64</td>
</tr>
</tbody>
</table>