Modeled Estimates of Solar Direct Normal Irradiance and Diffuse Horizontal Irradiance in Different Terrestrial Locations

by

Emad Abyad

uOttawa

Thesis submitted in partial fulfillment of the requirements for the

Master of Applied Science in Electrical and Computer Engineering

School of Electrical and Computer Engineering

University of Ottawa

Abstract

The transformation of solar energy into electricity is starting to impact to overall worldwide energy production mix. Photovoltaic-generated electricity can play a significant role in minimizing the use of non-renewable energy sources. Sunlight consists of three main components: global horizontal irradiance (GHI), direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI). Typically, these components are measured using specialized instruments in order to study solar radiation at any location. However, these measurements are not always available, especially in the case of the DNI and DHI components of sunlight. Consequently, many models have been developed to estimate these components from available GHI data. These models have their own merits. For this thesis, solar radiation data collected at four locations have been analyzed. The data come from Al-Hanakiyah (Saudi Arabia), Boulder (U.S.), Ma'an (Jordan), and Ottawa (Canada). The BRL, Reindl*, DISC, and Perez models have been used to estimate DNI and DHI data from the experimentally measured GHI data. The findings show that the Reindl* and Perez model outcomes offered similar accuracy of computing DNI and DHI values when comparing with detailed experimental data for Al-Hanakiyah and Ma'an. For Boulder, the Perez and BRL models have similar estimation abilities of DHI values and the DISC and Perez models are better estimators of DNI. The Reindl* model performs better when modeling DHI and DNI for Ottawa data. The BRL and DISC models show similar metrics error analyses, except in the case of the Ma'an location where the BRL model shows high error metrics values in terms of MAE, RMSE, and standard deviation (σ). The Boulder and Ottawa locations datasets were not complete and affected the outcomes with regards to the model performance metrics. Moreover, the metrics show very high, unreasonable values in terms of RMSE and σ. It is advised that a global model be developed by collecting data from many locations as a way to help minimize the error between the actual and modeled values since the current models have their own limitations. Availability of multi-year data, parameters such as albedo and aerosols, and one minute to hourly time steps data could help minimize the error between measured and modeled data. In addition to having accurate data, analysis of spectral data is important to evaluate their impact on solar technologies.

Acknowledgment

First of all, I would like to thank Allah for giving me the opportunity to pursue graduate studies and for helping me with the completion of my master's degree.

Also, I would like to express my sincere gratitude to my family, especially my mother for supporting me and encouraging me through obstacles I have faced during my studies.

I am deeply grateful to Prof. Karin Hinzer (my supervisor) and Prof. Ahmad Atieh (my co-supervisor) for their patience and effective supervision in helping me complete my master's degree.

Special thanks goes to the SUNLAB group members: Dr. Christopher Valdivia for his useful feedback during the period of my study, especially when publishing my work, Dr. Joan Haysom for providing some of the data and for helpful discussions in analyzing and processing these data, and to Lianne De La Salle and Patrick White McVey for providing me with a corrected version of the Ottawa data.

I am grateful to my close friends, Abdullah Al-Janah and Abdullah Al-Mohaimeed, for supporting me and providing me with useful tips for completing my master's degree.

I thank the following organizations for providing me with the data for different locations required to complete my master's degree: King Abdullah City for Atomic and Renewable Energy (K.A.CARE) (Al-Hanakiyah, Saudi Arabia), National Oceanic and Atmospheric Administration (NOAA) (Boulder, U.S.), Concentrating Solar Power (CSP) Services, DLR's Institute of Solar Research (Ma'an, Jordan), and the University of Ottawa (Ottawa test site) (Ottawa, Canada).

I would also like to thank Mr. David Longenecker for answering my queries regarding the treatment of Boulder data as well as providing me with meteorological data for Boulder.

Also, I am grateful to Dr. Richard Perez for providing me with the correction matrix values required for the implementation of the Perez model.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AM	Air mass
AST	Apparent solar time
AU	Astronomical unit
BRL	Boland, Ridley & Lauret
CPV	Concentrated photovoltaic
CSP	Concentrating Solar Power
DA	Day angle
DN	Day number
DNI	Direct normal irradiance
DHI	Diffuse horizontal irradiance
DISC	Direct Insolation Simulation Code
DLR	Deutsches Zentrum für Luft and Raumfahrt
DST	Daylight saving time
EIA	Energy Information Administration
EOT	Equation of time
ETR	Extraterrestrial radiation
FOV	Field of view
GHI	Global horizontal irradiance
IR	Infrared radiation
K.A.CARE	King Abdullah City for Atomic and Renewable Energy
MAE	Mean absolute error
MBE	Mean Bias error
MDMS	Meteorological data management system
MJSC	Multi-junction solar cell
NOAA	National Oceanic and Atmospheric Administration
RH	Relative humidity
PV	Photovoltaic
RRMM	Renewable resource monitoring and mapping
RSR	Rotating shadow radiometer
SERI	Solar Research Energy Institute
SPA	Solar position algorithm
TOA	Top of the atmosphere
TST	True solar time
UTC	Coordinated universal time
UV	Ultraviolet
WMO	World Meteorological Organization

List of Symbols

r	Mean sun-earth distance
I_{sc}	Solar constant
F	Spectral irradiance
λ	Wavelength
h	Planck's constant
c	Speed of light
k	Boltzmann constant
E_o	Eccentricity correction factor
δ	Solar declination angle
Z	Solar zenith angle
L_s	Standard longitude
L_c	Local longitude
α	Solar elevation angle
AZM	Azimuth angle
Td	Dew Temperature
T	Temperature
k_t	Clearness index
k_d	Diffuse fraction
k_n	Beam transmittance
H_o	Daily extraterrestrial radiation
k_{nc}	Beam transmittance under clear-sky conditions
P	Pressure station
P_o	Standard pressure
Δk_n	k_{nc} - k_n (clear condition-calculated values)
$k_{t'}$	Clearness index without the effect of the solar zenith angle
$\Delta k_{t'}$	Stability index
W	Atmospheric precipitable water
$X(k_{t'},Z,W,\Delta k_{t'})$	Four-dimensional correction matrix
$k_{b'}$	Beam transimtance without the effect of the solar zenith angle
K_t	Daily clearness index
$\boldsymbol{\psi}$	Persistence
t	Time (current hour/minute)
DNI_{coin}	DNI coincidence
Δ	Skylight brightness
ε	Cloud condition
E	Error between measured and modeled values
σ	Standard deviation
$\boldsymbol{V_{predict}}$	Predicted values
$\overline{\boldsymbol{V_{predict}}}$	Average predicted values

$\boldsymbol{V_{measr}}$	Measured values
$\overline{\boldsymbol{V_{measr}}}$	Average measured values
ρ	Correlation coefficient

Chapter 1 Introduction

Energy sources such as oil, natural gas and fossil fuel have been used extensively and have provided great benefits to humanity. The U.S. Energy Information Administration (EIA) stated that about 80% of the total energy consumption across the U.S. depends on these sources for more than 100 years [1]. One example of their use is in the generation of electricity. The huge demand on these sources as the result of the increase of the world's population presents many challenges considering that, these sources have many disadvantages. These sources will be depleted within a few human generations. Additionally, they cause negative effects such as global increase in the atmospheric temperature of the earth [2]. This last phenomenon is known as *global warming*. In order to find alternative energy sources, many efforts have focused on the utilization of renewable sources of energy such as solar, wind, biomass, and geothermal. Solar energy, where one coverts the photons from the sun hitting the Earth's surface into electric current is one of the most promising sources of energy because it is permanent, free, and an environmentally friendly source. Also, solar energy is captured using devices that are made using different elements such as silicon (Si), which represents the second plentiful element in the Earth's crust [3].

Nowadays, two common solar technologies are used worldwide: flat photovoltaic (PV) and concentrating photovoltaic (CPV) technologies. The PV technology is installed oriented at a specific angle to receive the incoming sunlight, known as global horizontal irradiance (GHI). The light is then converted to electricity. This type uses a single semiconductor junction cell usually silicon-based in its operation. CPV, in contrast, uses different types of lenses: a primary optical element to focus the incident sunlight and a secondary optical element to homogenize the sunlight on the solar cells. It consists of several junctions, i.e. a multi-junction solar cell (MJSC) in order to cover a broad range of the solar spectrum. Thus, CPV technologies are more efficient compared to PV. In 2016, the CPV technology was the highest in terms of efficiency compared to other technologies with a world-record achieved with 46 % of the incoming sunlight converted to electricity, as depicted in Figure 1-2. Part of the sunlight striking a surface is scattered due to clouds, aerosols, objects, and so on [4]. This part is known as diffuse horizontal irradiance (DHI). The CPV cell couples direct normal irradiance (DNI), which is the sunlight hitting a surface without being scattered. It represents

the largest portion sunlight compared to GHI and DHI when the sky has low amount of cloud coverage as shown in Figure 1-1. The DNI is high in the morning and evening periods while the GHI is low because of high solar zenith angle. However, around noontime, the GHI is higher than the DNI due to low solar zenith angle.

Solar energy has been used to generate electricity in order to replace traditional sources. However, assessment of the available solar radiation is required to calculate the total amount of electricity that will be generated for a location as a function of time. Also will it be more advantageous to deploy either flat-plat technology or concentrated technology? To determine this, a detailed understanding of the three solar components as a function of time of day and time of year is necessary. These components are measured by special radiometers through deployment of solar ground stations. The availability of the three components is not always possible for each deployed station. Even if the three components are available, inaccurate measurements and loss of data have a high likelihood of occurring. In case of missing data, they are substituted by different techniques. One of the methods is the implementation of models to compute DNI and DHI when GHI data are available. The concept of the majority of developed models is to find a relationship between the clearness index k_t (i.e. the degree of clarity of the sky), which ranges from 0 to 1, and a diffuse fraction k_d (i.e. degree of cloudiness (0 to 1).

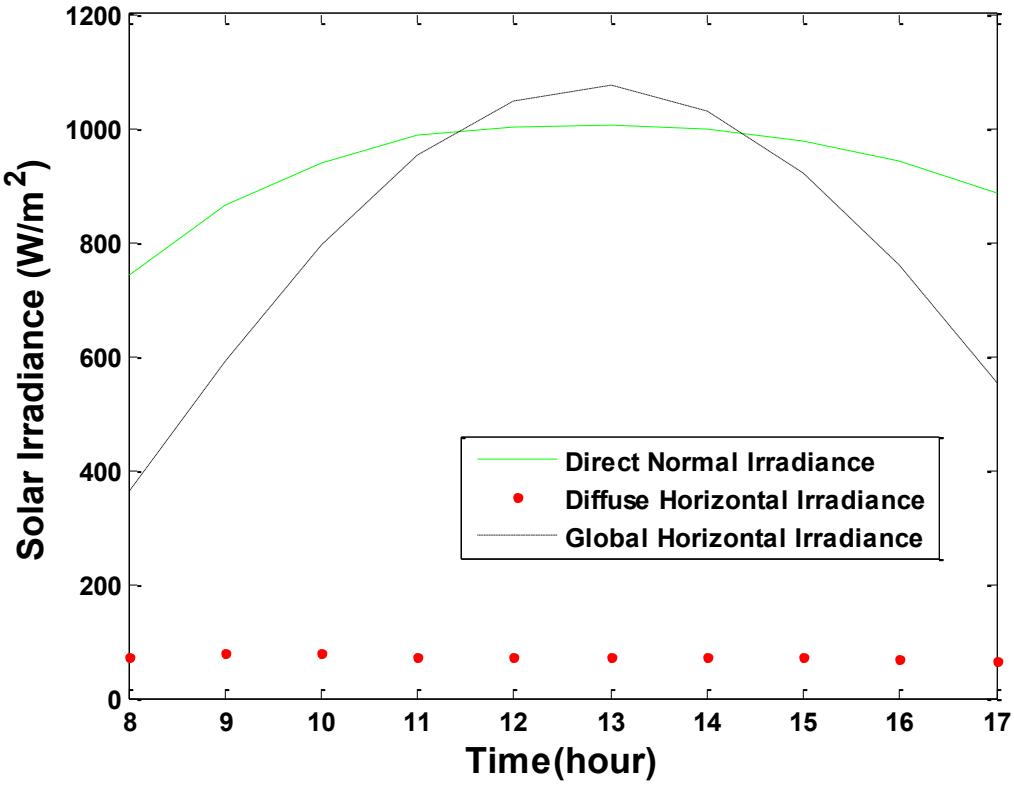


Figure 1-1 A representation of the three solar components on July/ 21/ 2014. The data are taken from Al-Hanakiyah station in Saudi Arabia [5].

Liu and Jordan [6] collected data from 98 locations in the United States and Canada to find the instantaneous intensity of DHI on days with a low amount of clouds (i.e. clear days). Also, the study aimed to find the daily diffuse radiation for different degrees of cloudiness for a variety of days. Another study was carried out by Orgill and Hollands [7] to estimate the hourly diffuse fraction on a horizontal surface. The four years of data used in this study came from Toronto, Canada. A third study was undertaken by Erbs *et al.* [8] to find the correlation between the hourly diffuse fraction and hourly clearness index. The hourly data used to establish the correlation was collected from a pyranometer, an instrument used to record GHI data, and a pyrheliometer, an instrument used to take DNI measurements, and came from four locations in the United States. The daily and monthly correlations were also developed.

In another study, Iqbal [9] used data collected from three sites in Canada and two locations in France to create correlations between the hourly global radiation and the diffuse

radiation. Unlike in the previous studies, he included another parameter- the solar elevation angle α - in addition to the clearness index to develop these correlations. Skartveit and Olseth [10] included both the hourly clearness index and the hourly solar elevation in their model to compute the diffuse fraction based on data from Bergen, Norway. In 1988, they [11] added a new parameter to their model to improve its accuracy. The new parameter was the hour-to-hour variability index, which seeks to determine the level of variation in the clearness index by looking at the previous hour and the next hour. Maxwell [12] developed the DISC model which estimates the solar direct irradiance directly without the need to compute the diffuse fraction. The data were obtained from several U.S. locations in order to develop and validate the model. The model was then corrected by Perez *et al*. [13]. The correction was applied by multiplying the output of the DISC model by a four dimensional matrix by binning the clearness index, solar zenith angle (Z), precipitable water (W) and stability index ($\Delta k_{t'}$) into different bins. Reindl *et al*. [14] included two predictors besides the clearness index and elevation angle in their model: ambient temperature (T) and relative humidity (RH). The research team developed three models based on the availability of the four predictors. Helbig [15] combined the two models developed by Reindl *et al*. where the ambient temperature and relative humidity data are not available. The combined model was the named Reindl* model. Boland *et al*. [16] developed the BRL model to establish the correlation between global radiation and diffuse radiation. They included the apparent solar time (AST), daily clearness index (K_t), persistence ($\boldsymbol{\psi}$) as new predictors in their model.

1.1 Objective of thesis

This work selects different models, briefly described in section 1.0, to test their ability to estimate either DHI, calculated from the diffuse fraction, or DNI, calculated from the beam transmittance (k_n). Data have been collected from four different locations around the world: Al-Hanakiyah (Lat. 24.85577º N, Long. 40.536º E, El. 873 m), Western region, Saudi Arabia; Ma'an (Lat. 30.17º N, Long. 35.82º E, El. 1012 m), Jordan; Boulder (Lat. 40.05º N, Long. 105 º W, El. 1584 m), Colorado, U.S.; and Ottawa (Lat. 45.42º N, Long. 75.68º W, El. 70 m), Ontario, Canada. The models selected are the Reindl*, DISC, Perez and BRL models. The goal is to assess the performance of these models for estimating the two solar components (i.e. DHI and DNI). This study makes a contribution where ground stations in

the Middle East are not plentiful compared to other regions. Also, the solar potential available in Saudi Arabia and in Jordan gives motivation to conduct this study since a majority of the models were based on the development of data from European, Canadian, and American locations. The outcome is very crucial to see if there are any models that can be applied globally and to determine advantages and disadvantages of each model.

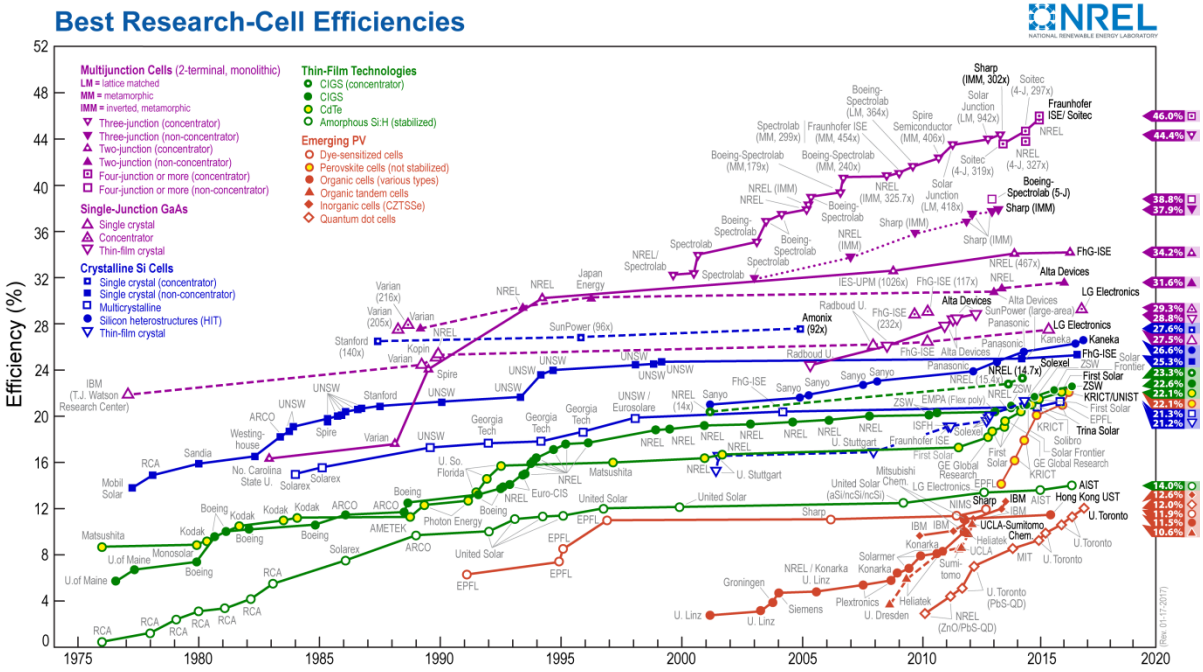


Figure 1-2 Efficiencies of different solar technologies (1975-2016) [17].

1.2 Thesis overview

This thesis is divided into five chapters. The first chapter introduces the role of renewable energy as an effective alternative to non-renewable energy sources. A comparison is made between PV and CPV solar systems in terms of efficiency. Also, the importance of using models to estimate solar radiation is highlighted. The chapter also summarizes previous studies done to model DHI and DNI components.

Chapter 2 gives an introduction to the sun and its relation to the earth. Moreover, solar geometry concepts regarding the sun-earth relations are introduced here. This chapter also covers the description of the instruments required to measure the three components of the sunlight and describes how the measurements are recorded according to the desired

component. A detailed description of the models with their relative equations used in this study is provided. These models are the Reindl*, BRL, DISC and Perez model.

Chapter 3 explains the processing of the solar radiation components for the four selected locations, which are Al-Hanakiyah, Saudi Arabia, Boulder, U.S., Ma'an, Jordan and Ottawa, Canada. A methodology that includes filling any missing data when necessary is also discussed. A description of each station is provided here along with the type of measurements available at each location. Additionally, the metrics to evaluate the models are listed in this chapter.

Chapter 4 compares the ground measured data with modeled data produced from each model. The scatter plot of modeled diffuse fraction, diffuse irradiance, and direct irradiance are plotted against measured or calculated data depending on the location. This helps interpret the weaknesses of each model. Metrics are also calculated for each model to describe how the modeled data deviate from actual ground data.

Chapter 5 summarizes and highlights the chapters and findings presented in the thesis. Additional recommendations are offered at the end of this chapter in an attempt to overcome issues related to the models developed globally.

Chapter 2 Solar Energy

This chapter summarizes basic concepts required for the computation of some solar geometry aspects. Additionally, a detailed description of the four models used in this study is introduced here. Finally, some common installed solar radiometers used to measure the GHI, DNI, and DHI are listed.

2.1 The sun

The amount of solar radiation received by the earth is crucial for making life possible on earth. For instance, the sun plays a huge role in the photosynthesis process which benefits plants, animals, and humans. The sun's radius is 6.96 $\times 10^5$ km which is about 110 times the earth's radius. The mean sun earth distance (r) is about 1.5×10^8 km, which is equivalent to one astronomical unit (AU). The minimum sun earth distance is approximately 0.983 AU (1.47×10^8 km) and the maximum is nearly (1.017 AU) (1.52×10^8 km). The minimum distance between the sun and the earth is reached on approximately January 3 (perihelion), and the maximum distance occurs on approximately July 4 (aphelion). The mean sun-earth distance is achieved on April 4 and October 5. The average power of solar radiation at top of the atmosphere (TOA), also known as the solar constant (I_{sc}), is about 1366 W/m^2; so, the total power striking the earth is about 1.73×10^{17} W and the total energy of solar radiation is about 5.46×10^{24} J [18]. The sun is approximately considered to be a blackbody, whose effective temperature is 5778 K (5505ºC or 9941F), since it absorbs radiation incident on its surface and radiates energy according to its temperature [19]. The spectral irradiance emitted by the blackbody is characterized by Plank's law:

$$F(\lambda) = \frac{2\pi \cdot h \cdot c^2}{\lambda^5\left(\exp\left(\frac{hc}{\lambda kT}\right)-1\right)}, \tag{2.1}$$

where F is the spectral irradiance (W /m^2 /μm), λ is the wavelength of the light emitted (m), T is the temperature of the blackbody (K), and h is Plank's constant (6.626196×10^{-34} J s) , c is the speed of light (2.9979250×10^{-8} m/s), and k is the Boltzmann constant (1.3806485210^{-23} J/K). The spectral irradiance is a function of both the wavelength and the blackbody temperature. Figure 2-1 illustrates the impact of temperature on radiation. The radiation and its peak shifts towards shorter wavelengths as the temperature increases. At 300K (i.e. room

temperature), no power is radiated in the visible region or near infrared range of the wavelength. At 3000K, the radiation is shifted and increased toward shorter wavelengths passing through the visible region. If the temperature increases to 6000K, the radiation range varies from the red to the violet spectrum and the light becomes white. This is the reason the sunlight appears as yellowish-white in the sky.

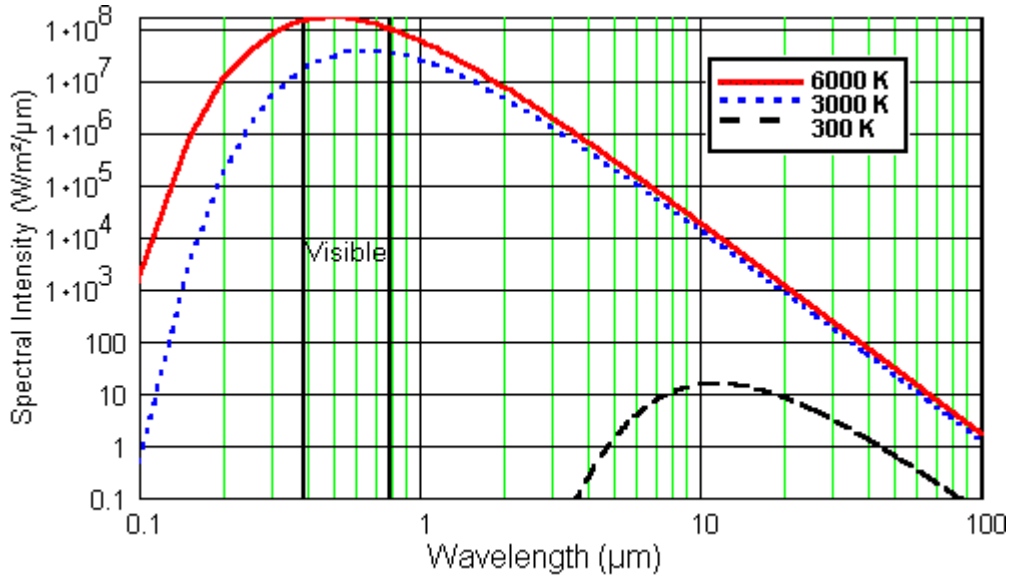


Figure 2-1 Effect of different temperatures on solar radiation [19].

The thermal radiation lies within a range of 0.2-1000 (µm). Table 2.1 shows the divisions of the spectrum of wavelength (visible and ultraviolet) into different ranges. The color section indicates the colors that can be seen by human eyes. Note that the spectral dissection of colors is based on average judgment.

Table 2.1 Spectral division of solar radiation [20].

Portion	Color	Range (μm)
Visible	Violet	0.39 – 0.455
	Blue	0.455 – 0.492
	Green	0.492 – 0.577
	Yellow	0.577 – 0.597
	Orange	0.597 – 0.622
	Red	0.622 – 0.77

Portion	Subsection	Range (μm)
Ultraviolet (UV)	Near UV	0.3 – 0.4
	Far UV	0.2 – 0.3
	Maximum UV	0. 001 – 0.2
Infrared (IR)	Near IR	0.77 - 25
	Far IR	25 - 1000

2.2 The sun-earth relationship

This section will highlight different concepts and required geometric and time parameters to determine the sun's position in relation to the earth.

2.2.1 Day number

Day number (*DN*) is required in various solar applications. Each day in one year has a particular value, usually (1-365). For example, on January 1, *DN* equals 1 and for February 15, *DN* equals 46. However, if there is a leap year cycle, the day number changes slightly, since the month of February has 29 days. For instance, on March 1, 1996, *DN* equals 61, and on March 1, 1997, *DN* equals 60. The day number is used to calculate other parameters such as equation of time (*EOT*) and solar declination angle (δ) which will be described later.

2.2.2 The eccentricity correction factor

The sun completes one rotation cycle around the earth in one year. It is known that the earth orbit is not circular (ellipse). So, the sun moves much faster. A factor, called the eccentricity correction factor, is required to compensate for the change in solar days. The eccentricity correction factor of the earth's orbit (E_0) is expressed as:

$$E_0 = (\frac{r_0}{r})^2 = 1.000110 + 0.034221 \cdot \cos(DA) + 0.001280 \cdot \sin(DA) + 0.0000719 \cdot \cos(2DA) + 0.000077 \cdot \sin(2DA), \quad (2.2)$$

where r_0 is the mean sun earth distance, *r* is the actual sun-earth distance and *DA* is the day angle in radians and is obtained as:

$$DA = 2\pi(DN - 1)/365. \tag{2.3}$$

In many engineering applications, the expression of the eccentricity correction factor is simplified to:

$$E_0 = 1 + 0.033 \cdot \cos(\frac{2\pi \cdot DN}{365}). \tag{2.4}$$

2.2.3 The solar declination angle

The plane of rotation of the earth around the sun is named the elliptic plane. The revolution of the earth around an axis is called the polar axis, which is tilted at an angle of 23.5º with reference to the elliptic plane. The inclination of the equator to the elliptic plane is 23.5º, which accounts for seasonal and diurnal changes. The seasonal change is inversely proportional to the distance between the earth and the sun. The angle of the inclination of the equator with respect to the center of the earth and the sun is called the solar declination angle (δ). The beginning of the summer season is known as summer solstice (when the solar declination angle is 23.5º) and the start of the winter season is known as winter solstice (when the solar declination angle is -23.5º). This angle is zero at the equinox, where the length of day is equal to night. As shown in Figure 2-2, the earth is inclined towards the sun in June and inclined away from the sun in December. The declination angle (δ) is expressed as:

$$\begin{aligned}\delta = (0.006918 - 0.399912 \cdot \cos(DA) + 0.070257 \cdot \sin(DA) - 0.006758 \cdot \\ \cos(2 \cdot DA) + 0.000907 \cdot \sin(2 \cdot DA) - 0.002697 \cdot \cos(3 \cdot DA) + 0.00148 \cdot \\ \sin(3 \cdot DA)) \cdot (\frac{180}{\pi}).\end{aligned} \tag{2.5}$$

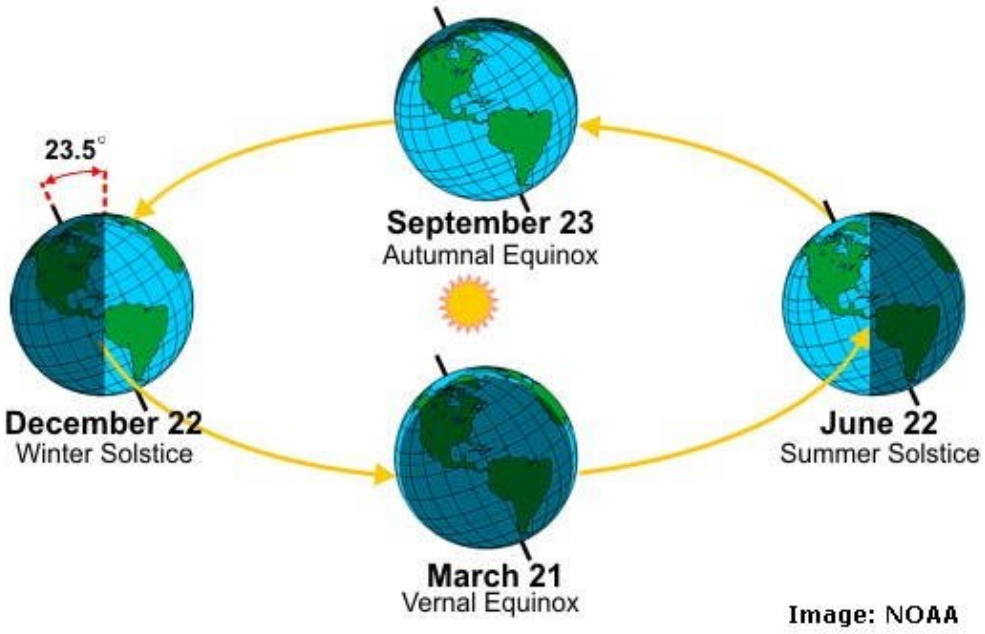


Figure 2-2 The tilt of the earth's axis with respect to the elliptical plane is 23.5º. The earth is tilted towards or away from the sun depending on the season of the year [21].

From an observer's perspective, another way to understand the concept of the solar declination angle is by imagining the earth centered in the celestial sphere with the sun rotating around the earth. Figure 2-3 shows that the North Pole and the South Pole are the points passing through the celestial sphere. The point where the equator of the earth intersects with the sun's rotation illustrates the concept of the declination angle.

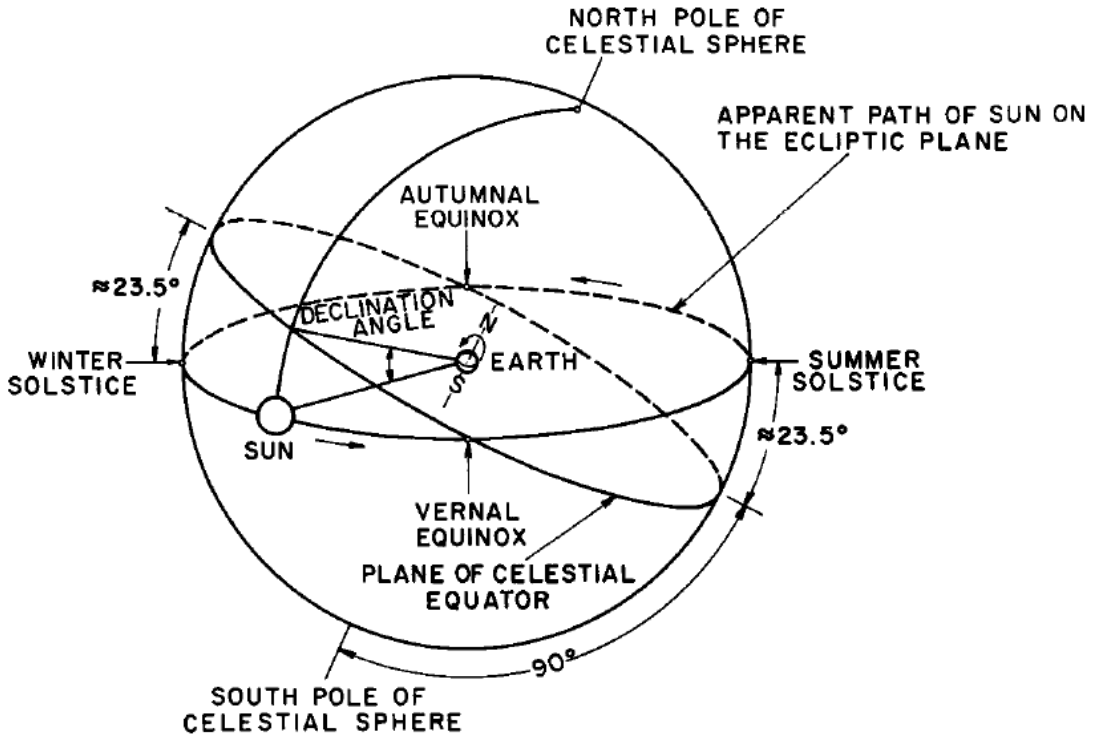


Figure 2-3 Illustration of the declination angle concept: the intersection between the earth's equator with the rotation of the sun is shown using a celestial sphere [20].

2.2.4 Equation of time

The solar day is defined as the time the sun takes to cross the local meridian until it crosses the same meridian again. The variation in the length of the solar day is caused by two factors: the area swept out by the earth on the elliptic plane is not the same when it rotates around the sun, and secondly, the plane axis of the earth is tilted with reference to the elliptic plane. The difference between solar time and clock time is characterized by the equation of time (*EOT*) (radians), which is expressed as:

$$EOT = (0.000075 + 0.001868 \cdot \cos(DA) - 0.032077 \cdot \sin(DA) - 0.014615 \cdot \cos(2 \cdot DA) \text{-} 0.04089 \cdot \sin(2 \cdot DA)) \cdot (229.18). \quad (2.6)$$

Note that the Equation (2.6) is multiplied by 229.18 to convert it to minutes. The maximum discrepancy between solar time and clock time can reach up to 16 minutes. Figure 2-4 shows the behavior of the equation of time during one year.

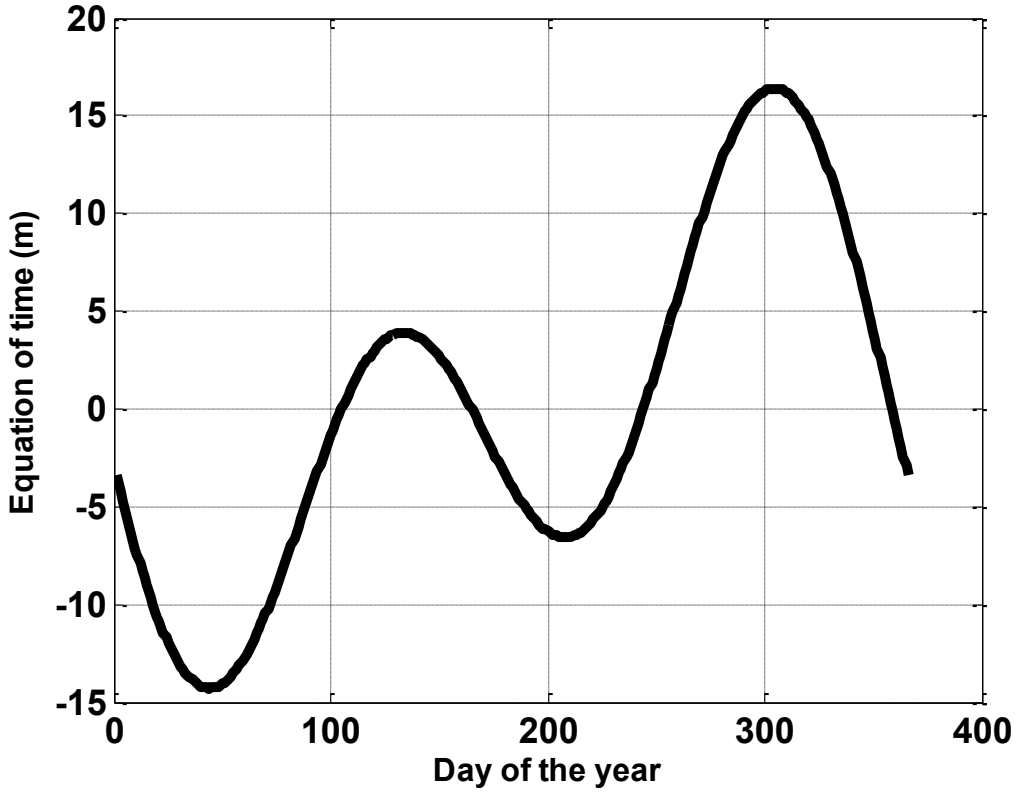


Figure 2-4 Equation of time expressed in minutes during the year.

Figure 2-5 illustrates the concept of the solar analemma, where the sun's movement deviates from its mean in the sky on the same day on a fixed location on earth. A figure eight pattern is observed.

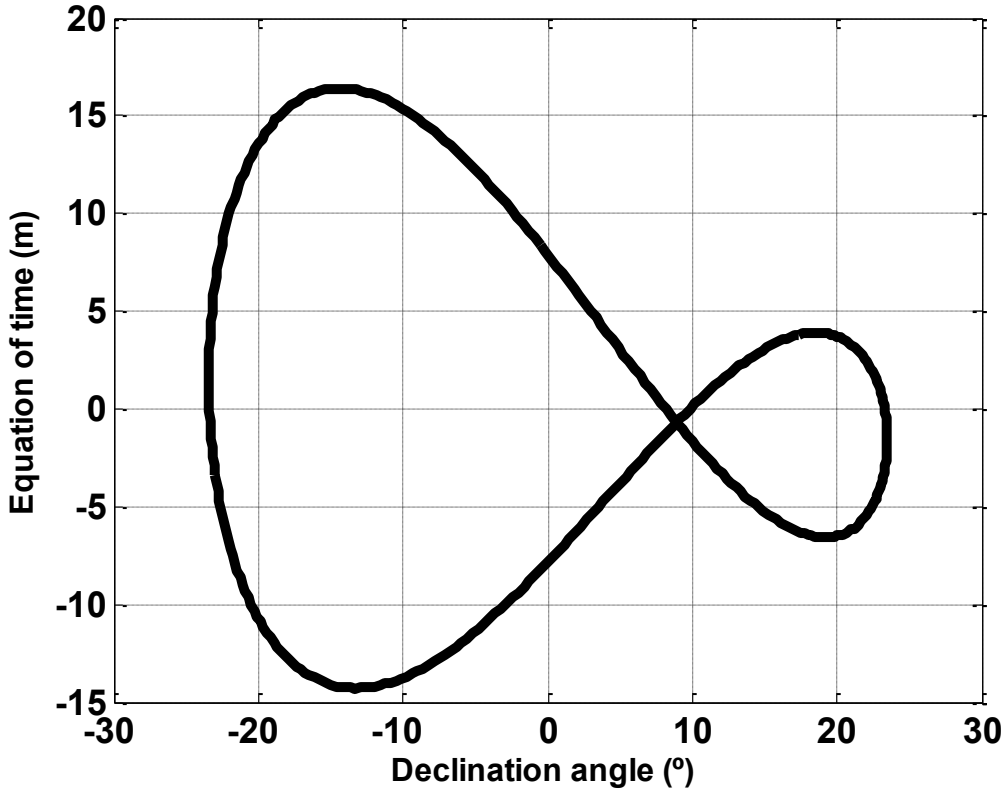


Figure 2-5 The variation of the motion of the sun from its mean motion is known as the solar analemma, where the "8" shape is seen from a particular location on earth.

2.2.5 Apparent solar time

Apparent solar time (*AST*) is also called true solar time (*TST*) where solar radiation data are recorded at this time. Also, *AST* is applied in the computation of solar geometry. The standard meridians are multiple of 15º from east to west of Greenwich meridian (0º), where a deviation of one degree represents four minutes. So, moving from one standard longitude to another means a one hour difference between different time zones. The apparent solar time is calculated as:

$$AST = local\ standard\ time \pm 4(L_s - L_c) + EOT, \tag{2.7}$$

where L_s and L_c are the standard longitude and local longitude, respectively. $4(L_s - L_c)$ is the longitude correction. Table 2.2 shows different time zones with its associated standard longitude west of Greenwich meridian. The plus and negative sign in the equation depends on the location of a city that is whether it is east or west of Greenwich meridian. The positive

sign is for cities located east of Greenwich meridian and the negative sign is for cities located west of Greenwich meridian. Daylight saving time (DST) can be added to Equation (2.7) if needed.

Table 2.2 Some time zones associated with their acronyms, their difference from Greenwich Time, and their corresponding longitude.

Time Zone	Acronym	Zone time (h)	Standard longitude
Atlantic standard Time	AST	4	60º W
Eastern standard time	EST	5	75º W
Central standard time	CST	6	90º W
Mountain standard time	MST	7	105º W
Pacific standard time	PST	8	120º W
Alaska standard time	AKST	9	135º W
Hawaii standard time	HST	10	150º W

2.2.6 Solar zenith, azimuth and elevation angle

Solar radiation, emanated by the sun, strikes the earth at different angles. So, it is crucial to know the trigonometric relationship between the position of the sun and the earth coordinates. Figure 2-6 depicts the relationship between the two concepts. Imagine if an observer on earth has a location named observer's zenith. It is junction between the observer's position and the celestial sphere, where the intersection is perpendicular to the earth (local zenith). The location opposed to the local zenith is called nadir. The zenith angle (θ_z), which will later be denoted as (Z), is defined as the angle between the local zenith and the line connecting the sun with the observer's location. Its range is between 0º to 90º. The solar latitude (α), also known as solar elevation or solar altitude, is the sun height above the observer's celestial horizon. Moreover, the solar elevation is the complement of the solar zenith angle, and ranges from 0º to 90º. The solar azimuth ($\boldsymbol{\psi}$), later denoted as (AZM), is the angle between the observer's zenith and the plane of a circle crossing the zenith and the sun. In simple terms, the solar zenith or elevation angle indicates the vertical position of the sun

while the azimuth angle identifies the horizontal motion of the sun.

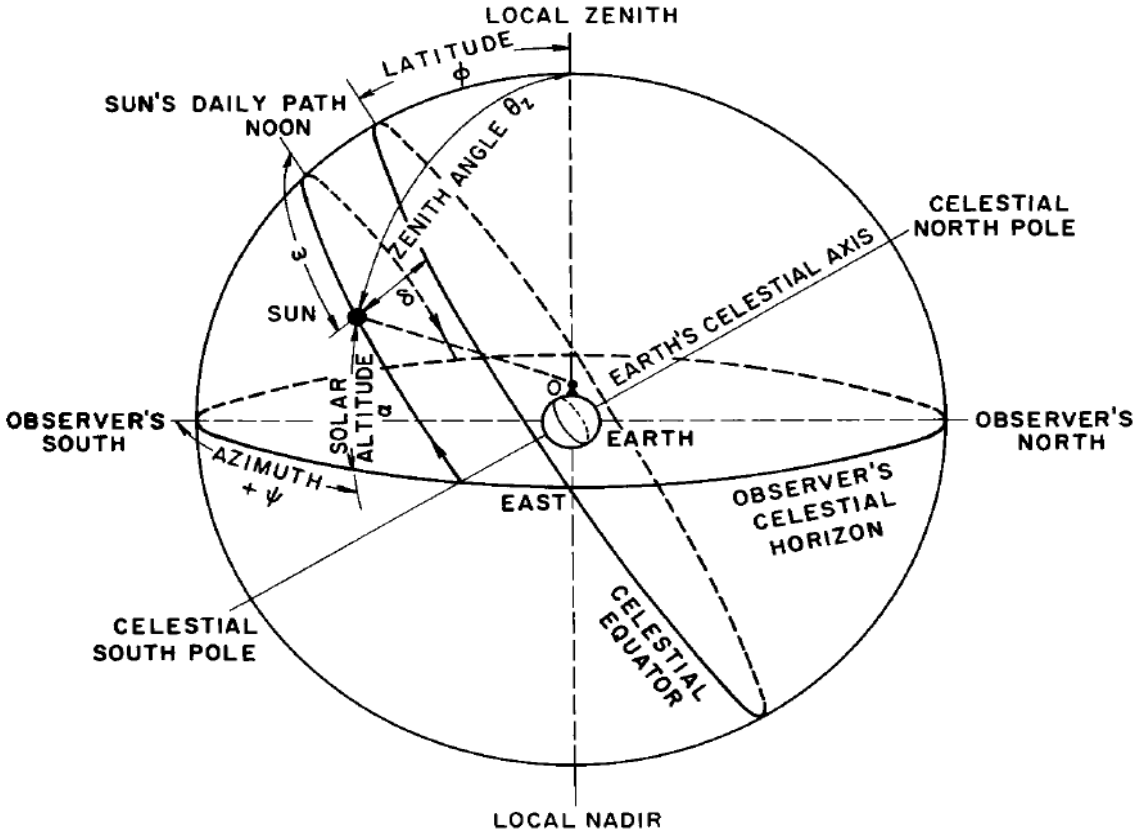


Figure 2-6 An illustration of the different solar angles: the zenith angle, elevation angle, and azimuth angle [20].

The solar azimuth angle is positive in the east direction and negative in the west direction. It ranges from 0º to ± 180º. The hour angle (ω) is the angle between the observer's meridian and solar meridian. It is an illustration of the apparent solar time (*AST*) in degrees. It alters 15º per hour. It is zero at solar noon. It is given as:

$$\omega = 15(AST - 12). \tag{2.8}$$

The sun rise hour angle (ω_s) is calculated as:

$$\omega_s = \cos^{-1}(\tan L_c \cdot \tan \delta). \tag{2.9}$$

The day length (N_d), expressed in hours (h), is given as:

$$N_d = \frac{2}{15}\cos^{-1}(\tan L_c \cdot \tan \delta). \tag{2.10}$$

The solar zenith angle (Z) is calculated as:

$$Z = \cos^{-1}(\sin\delta \cdot \sin L_c + \cos\delta \cdot \cos L_c \cdot \cos\omega). \tag{2.11}$$

The solar elevation angle (α) is computed as:

$$\alpha = 90 - Z. \tag{2.12}$$

The sun azimuth angle (AZM) is computed as:

$$AZM = \cos^{-1}\left(\frac{\sin\alpha \cdot \sin L_c - \sin\delta}{\cos\alpha \cdot \cos L_c}\right). \tag{2.13}$$

2.3 Overview of direct normal, global, and diffuse horizontal irradiance

Sunlight is comprised of three components: direct normal irradiance (DNI), diffuse horizontal irradiance (DHI), and global horizontal irradiance (GHI). Solar applications depend mainly on the type of incident solar irradiance. Concentrating photovoltaic (CPV) and concentrating solar power (CSP) systems are operated by the utilization of DNI while photovoltaic systems depend on GHI. So, it is important to describe and give a comprehensive picture of each component to understand how different solar power systems work. DNI, also named beam irradiance, is defined as the solar component striking the earth's surface without facing any obstacles such as trees, clouds, aerosols and so on. It is known that the solar radiation at the top of the atmosphere is entirely DNI and it is known as extraterrestrial radiation (ETR). Radiation emanated from the region around the sun disk is called circumsolar radiation [22]. In comparison to the DNI incident on the earth's surface, it is considered large. Part of the solar radiation goes through the atmosphere of the earth while another part is reflected back into space or reflected off the ground (i.e. ground reflected irradiance or ground albedo). Losses in the solar spectrum are due to Rayleigh scattering, ozone absorption, and aerosols extinction in the UV region while water vapor and mixed gases play a huge role in the IR region [23]. This causes a shift in the solar spectrum. The concept of DHI is opposite to DNI. DHI is the solar radiation that is diffused away from its path due to passing through various obstacles. It is scattered by clouds, molecules, and aerosols, or absorbed by gases and aerosols. Cloud plays a huge role in the amount received

by the earth. If the cloud layer is thick, no transmittance of beam radiation exists. The combination of incident DNI on the horizontal surface and hemispherical DHI is known as GHI. Figure 2-7 shows the path of each component till it reaches a surface. The equation governing the three components is given in terms of the zenith angle or elevation angle as:

$$GHI = DNI \cdot \cos(Z) + DHI, \tag{2.14}$$

$$GHI = DNI \cdot \sin(\alpha) + DHI. \tag{2.15}$$

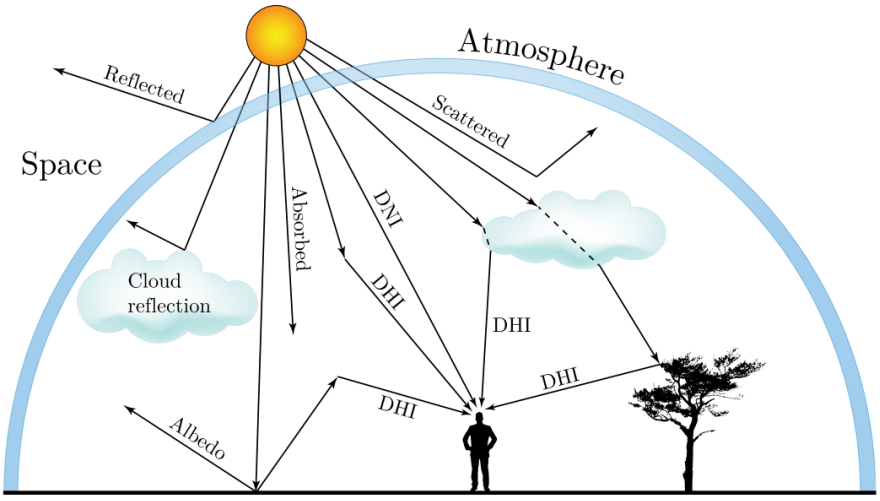


Figure 2-7 Different components of the sunlight (GHI, DNI and DHI), where each one reaches the ground through some or no obstacles. GHI is the combination of DNI on horizontal surface and DHI [23].

One fact about GHI is that its values can be larger than extraterrestrial radiation, but only for a short interval. The reason for this is that as clouds approach the sun disc, more sunlight is reflected back to instruments measuring GHI. This event is known as the *lensing effect*. Table 2.3 shows some types of surfaces associated with their reflection ratio of the sunlight.

Table 2.3 Different objects and their ratio of reflection of the sunlight [24],[25].

Surface	**Ratio of albedo**
Dark	0.1
grasslands	0.2
Fallen snow	0.8

River sand	0.43
Rock	0.12 – 0.15
Sea ice	0.36 – 0.5

2.3.1 Extraterrestrial radiation

The extraterrestrial radiation (*ETR*) is the solar radiation at the top of earth's atmosphere. The general formula of this parameter is given by:

$$ETR = I_{sc} \cdot E_0. \tag{2.16}$$

The daily extraterrestrial radiation (H_0), expressed in $MJm^{-2}\,day^{-1}$, is computed as:

$$H_0 = \frac{24}{\pi} \cdot I_{sc} \cdot E_0 [(\frac{\pi}{180}) \cdot \ \omega_s \cdot (\sin\delta \cdot \sin L_c) + (\cos\delta \cdot \cos L_c \cdot \cos\omega_s)]. \tag{2.17}$$

It is important to note that extraterrestrial radiation is not constant over a year. This is interpreted as a result of two factors. The first factor is the sunspot activities which may cause a slight variation of intrinsic solar radiation of $\pm$1.5%. The second factor is the alteration of the sun-earth distance during a year. The change in solar radiation lies within $\pm$3% [26]. Figure 2-8 shows the variation of extraterrestrial radiation throughout one full year.

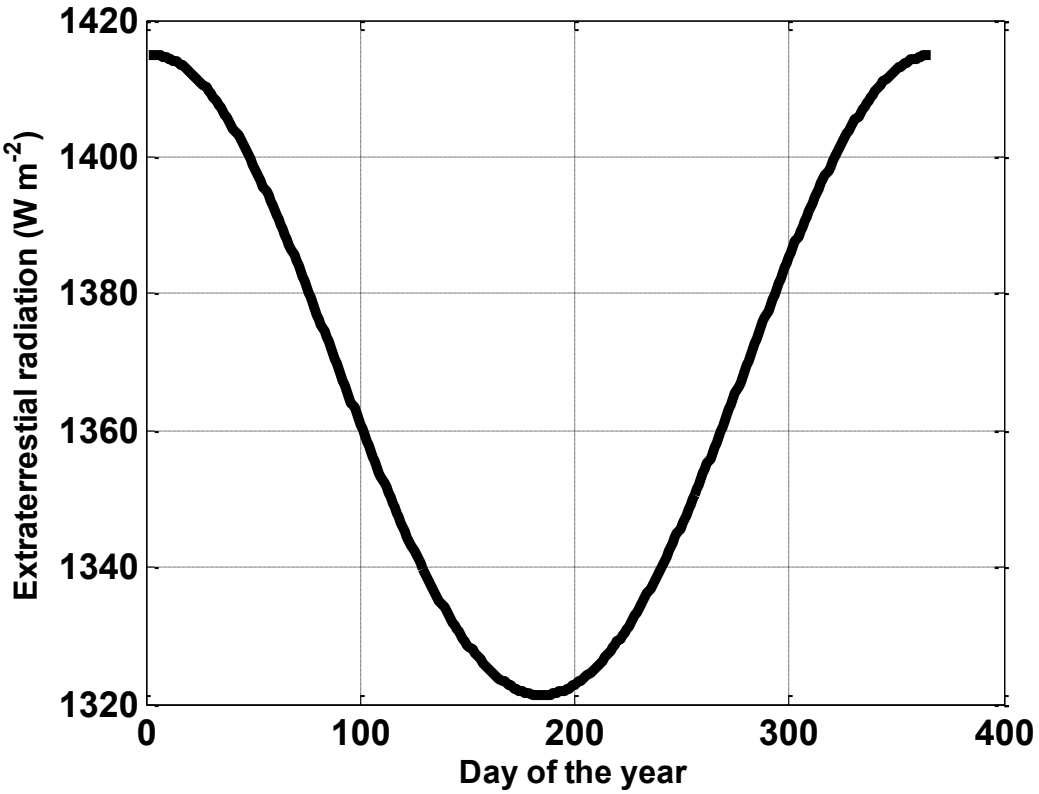


Figure 2-8 Variation of the intensity of the sunlight received by the earth during the year.

2.3.2 Air mass

Air mass (*AM*) describes the path as the sunlight passes through the atmosphere until it reaches a surface with respect to a normal axis (Z=0º), that is, when the sun is overhead. The air mass increases with a decrease in the solar altitude angle (α) or an increase in the solar zenith angle (Z) and vice versa. It is governed by the following equations:

$$AM = \csc(\alpha) = \sec(Z). \tag{2.18}$$

In space, the air mass is zero (AM 0 spectrum) since no obstacles alter the path of incident solar radiation outside the atmosphere of the earth. Equation (2.18) assumes that the surface is flat and horizontal. However, due to the curvature of the earth, the air mass value is not accurate for higher values of the zenith angle (i.e. Z > 75 º) if the above equation is used. Equation (2.19) takes into consideration the curvature of the earth.

$$AM = \frac{1}{\cos(Z)+0.50572(96.07995-Z)^{-1.6364}}. \quad (2.19)$$

Figure 2-9 depicts the principle of air mass. The standard spectrum (Am 1.5G and AM 1.5D) (Z=48.2º) are standardized by the American Society for Testing and Materials (ASTM). The reduction in solar radiation could be due to many factors such as refraction of the atmosphere, absorption by molecules in the atmosphere, and the phenomenon of Rayleigh scattering and Mie scattering.

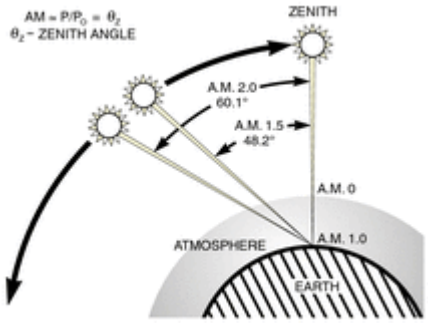


Figure 2-9 The path the sunlight takes to hit a surface. Sunlight is absorbed or scattered more for higher air masses [27].

2.3.3 Clearness index, beam transmittance, and diffuse fraction:

Measurements of GHI are available in many locations, compared to DNI and DHI [28] . One useful parameter that depends on GHI is called the clearness index (k_t). The clearness index is an indication of the clearness of the atmosphere. Its values range from 0 to 1. Low values of the clearness index mean that the sky is full of clouds (overcast sky), medium values indicate that the sky is partially cloudy, and high values mean the sky is clear. The clearness index is a scalar quantity which provides useful data since the shift in the solar spectrum can affect the performance of solar technologies, especially CPV systems. It is calculated as:

$$k_t = \frac{GHI}{ETR \cdot \cos(Z)}. \quad (2.20)$$

Also, knowing DHI helps determine weather pattern by knowing the relationship between the clearness index and the diffuse fraction (k_d). The diffuse fraction characterizes the relationship between DHI and GHI and can be found from the equation below:

$$k_d = \frac{DHI}{GHI}. \tag{2.21}$$

The range of the diffuse fraction is the same as the ratio of the clearness index. The diffuse fraction is inversely proportional to the clearness index. Therefore, if the sky is clear, there are high values of the clearness index and low values of the diffuse fraction. On the other hand, when there is a large cloud cover, the clearness index values are low and the diffuse fraction values are high. The beam transmittance (k_n) indicates the amount of direct beam radiation reaching a surface without being scattered. It is the ratio of DNI values to their corresponding extraterrestrial values:

$$k_n = \frac{DNI}{ETR}. \tag{2.22}$$

All three indices are dimensionless. The range of k_n is the same as the range of k_t and k_d. However, the beam transmittance is directly proportional to the clearness index. A simple relation between the beam transmittance and horizontal indices is given as:

$$k_t = k_n + k_d. \tag{2.23}$$

2.4 Radiometers

Many instruments have been designed and developed to measure GHI, DNI, and DHI. This section will give an overview of instruments used to measure each component.

2.4.1 Pyrheliometer

A pyrheliometer is a device used to measure DNI. Pyrheliometers should be pointed at the solar disk so that their field of view (FOV) is 5º due to the contribution of circumsolar components to DNI measurements. These instruments are mounted on automatic solar trackers to track the sun path from sunrise to the time of sunset.

Thermoelectric or photoelectric detectors are utilized so that the solar radiation (W/m^2) is transformed into electric signals expressed in microvolt (μV) dc [22]. The detectors have an optical black coating in order to permit a wide and uniform spectral response of wavelengths lying within a range of 300 nm to 3000 nm. These devices have slow response times which may vary from one to five seconds. Prevalent Si-based photoelectric detectors have a

response to wavelengths available in the visible and near-infrared region of the spectrum. In other words, they respond to wavelengths lying between 300 nm to 1100 nm. A slight change of alignments of solar trackers can measure the region around the sun disc (the circumsolar component). This component can be detected by pyrheliometers because apertures with a narrow view are installed inside these devices. To ensure highly accurate measurements of DNI, the World Radiometer Reference (WRR) has utilized electrical self-calibrating absolute cavity pyrheliometers. In terms of cost, pyrheliometers are less expensive than good quality pyranometers, which are instruments used to measure GHI.

2.4.2 Pyranometer

Like pyrheliometers, pyranometers have thermopile or photoelectric detectors; however, pyranometers have a hemispherical FOV of 180º, known as "fish-eye" [22]. These devices can be mounted horizontally to take measurements of GHI or can be inclined in order to measure tilted irradiance, vertical irradiance, or reflected irradiance. To ensure safety of the pyranometers' sensors, they are placed usually under protective quartz dome or diffuser. The latter two designs also save the detectors from weather so that they provide the sensors with optical properties that match incident solar radiation. In addition to the previous designs, the pyranometers are equipped with ventilation units to circulate air to lessen the impact of factors such as dust, snow, insects, frost, ice, or other materials which may cause contamination for the pyranometers' optics. Sometimes, shadowbands are attached to the pyranometers in order to measure the diffused light in the sky. Other devices with a complete structure (pyranometer plus shadowbands) are called rotating shadow irradiometers (RSIs), Alternate names for this last instrument are rotating shadowband radiometers (RSRs) or rotating shadowband pyranometers (RSPs).

2.4.3 Rotating shadowband radiometer

The basic operating function of rotating shadowband radiometers (RSRs) is that the shadowbands are controlled and moved by a motor. The motor drives the shadowbands across the FOV of the pyranometers' detectors. So, when the detectors are shaded, DHI is measured. GHI is measured if the detectors are not blocked by the shadowbands. Since only two out of three components (GHI and DHI) are measured, the algorithm to calculate DNI is given as:

$$DNI = \frac{(GHI - DHI)}{\cos(Z)}. \quad (2.24)$$

Two common types of RSRs are available: RSRs with continuous rotations and RSRs with discrete rotations. The shadowband blocks the pyranometer's sensor within one second. This movement of the shadowband occurs once per minute. As the shadowband shades the pyranometer's sensor, the sensor takes measurements of irradiance with a high and constant sampling rate (about 1 KHz). The latter method of measurement is named *burst or sweep*. During the rotation of the shadowband, GHI measurements are taken. A slight drop in the left side of the slope of the burst can be seen due to blocking of the circumsolar component when the shadowband is about to cover the detector. Asymmetry between the right side and left side in the slope of the burst is observed because of the angular position of the sun and the FOV of the sensor. When the center of the shadow covers the center of the sensor, DHI values are measured; however, the values of the DHI are not the final values because the shadowband is blocking a portion of the sky. The measurements when the shadow of the shadowband lies next to the RSR are related to shoulder values. They are obtained by applying curve analysis algorithm. In order to find the actual value of the DHI, the difference between the GHI values and the average of the shoulder values is added to the minimum burst's slope to get the DHI values, as depicted in Figure 2-11. The zenith angle values are derived from parameters such as an instrument's latitude, longitude, and time.

The measurement of the three components is known as RSRs with continuous rotations. Figure 2-10 shows the RSR manufactured by Irradiance Inc. The discrete rotation of RSRs differs from continuous rotations of RSRs. Only four points of the whole burst are measured by the discrete RSRs. The first step is that the GHI is measured when the shadowband is in a rest position (i.e. the shadowband lies below the sensor). Before shading the instrument's sensor, another measurement is taken. The third measurement is taken when the shadowband shades the sensor. The last measurement is obtained when the shadow of the shadowband passes the sensor. The same method is used to obtain the actual values of DHI.

Figure 2-10 Rotating shadowband manufactured by Irradiance Inc. This measures GHI, DHI , and calculates DNI [29].

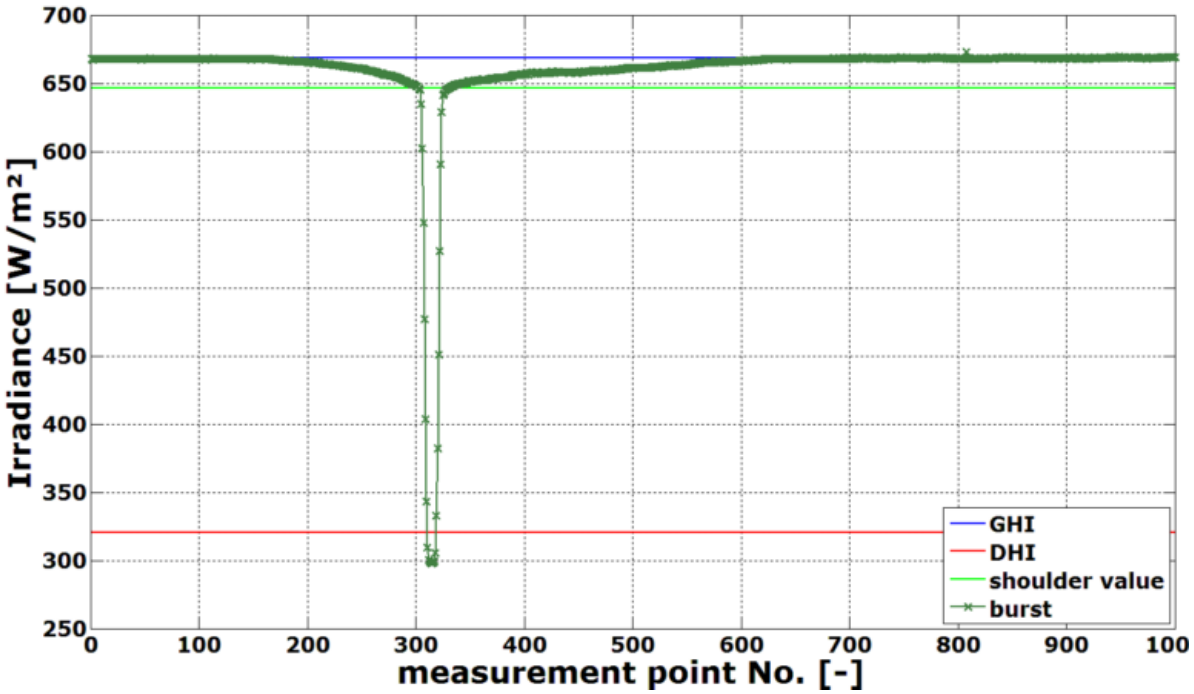


Figure 2-11 Illustration of the method used to take measurements of DHI and correct them. This method is used in RSRs due to the shading of the silicon sensor [30].

2.4.4 Specifications for different instruments measuring solar radiation components

This section includes a list of common instruments installed to measure solar radiation components.

Table 2.4 Specifications of LICOR RSR's (Irradiance Inc.)

LICOR RSR	
Head unit	LI-200SZ pyranometer
Accuracy	±2%
Enclosure material	Fiberglass
Enclosure width	30.48 cm (12 inches)
Enclosure height	35.56 cm (14 inches)
Enclosure depth	15.24 cm (6 inches)
Operating temperature	5º - 50ºC
Price	$8750

Table 2.5 Specifications of CHP1 pyrheliometer's (Kipp&Zonen)

Pyrheliometer CHP1	
ISO classification	First class
Response time	5 s
Zero offset B	±1 W/m^2
Sensitivity	7-14 $\mu V/W/m^2$
Temperature dependence of sensitivity	±0.5% (-20º - 50ºC)
Field of view	5º±0.2º
Operating temperature	-40º - 80º C
Non-linearity	±0.2%
Spectral range	200 – 4000 nm
Maximum solar irradiance	4000 W/m^2
Weight (w/o cable)	0.9 kg
Length	31.6 cm (12.4 in)
Body diameter	3.8 cm (1.5 in)

Base diameter	7.6 cm (3.0 in)

Table 2.6 Specifications of Eppley pyrheliometer's (Eppley Inc.)

Eppley Pyrheliometer	
ISO classification	Secondary standard
Response time	5 s
Zero offset B	±1 W/m^2
Sensitivity	8μV/W/m^2
Temperature dependence of sensitivity	±0.5%
Field of view	5º
Non-stability	±0.5%
Non-linearity	±0.2%
Spectral range	250– 3000 nm
Expected daily uncertainty	±1%
Expected hourly uncertainty	±1%
Expected instant uncertainty	< 5 W/m^2

Table 2.7 Specifications of CMP21 pyranometer's (Kipp&Zonen)

CMP21 pyranometer	
ISO classification	Secondary standard
Response time	5 s
Zero offset (Thermal radiation)	±7 W/m^2 (200 W/m^2)
Sensitivity	7-14 μV/W/m^2
Temperature dependence of sensitivity	±1% (-20º - 50ºC)
Field of view	5º±0.2º
Operating temperature	-40º - 80º C
Non-linearity	±0.2% (0 – 1000 W/m^2)
Non-stability (per year)	±0.5%
Directional error (1000W/m^2) (80º)	±10W/m^2
Expected daily uncertainty	±2%

Spectral range	285– 2800 nm
Maximum solar irradiance	4000 W/m^2
Width (with shield)	15 cm (5.9 in)
Height	9.25 cm (3.64 in)
Dome diameter	5 cm (2 in)
Warranty	2 years

Table 2.8 Specifications of standard precision pyranometer (SPP) (Eppley Inc.)

SPP	
ISO classification	Secondary standard
Response time	10 s
Sensitivity	8μV/W/m^2
Temperature dependence of sensitivity	±1% (-30º - 50ºC)
Operating temperature	-40º - 80º C
Non-linearity	±0.5% (0 – 2800 W/m^2)
Non-stability (per year)	±0.5%
Directional error (1000W/m^2) (80º)	±10W/m^2
Expected daily uncertainty	±1%
Expected hourly uncertainty	±2%
Expected instant uncertainty	< 10 W/m^2
Spectral range	285– 2800 nm

Table 2.9 Specifications of SolarSIM-D2 (Spectrafy)

SolarSIM-D2	
Classification	Secondary standard
Field of view	5º
Spectral measurement uncertainty	±5% (per wavelength)
DNI measurement uncertainty	±2%
Meteorological measurement	±3%

uncertainty	
Sensitivity	8μV/W/m^2
Temperature dependence of sensitivity	±1% (-30º - 50ºC)
Operating temperature	-30º - 65º C
Non-linearity	±0.5% (0 – 1366W/m^2)
Non-stability (per year)	±0.5%
Maximum irradiance	1366 W/m^2
Weight (w/o cable)	1.2 kg
Price	$ 9900

Table 2.10 Specifications of SolarSIM-G1 (Spectrafy)

SolarSIM-G1	
Classification	Secondary standard
Spectral measurement uncertainty	±5% (per wavelength)
GHI measurement uncertainty	±2%
Operating temperature	-30º - 65º C
Non-linearity	±0.5% (0 – 1366W/m^2)
Non-stability (per year)	±0.5%
Spectral range	280– 4000 nm
Maximum irradiance	1366 W/m^2
Weight (w/o cable)	1 kg
Price	$ 9900

2.4.5 Ground vs satellite measurements

Ground measurements are highly accurate in comparison to satellite measurements since they are recorded by high level instruments at a specific pinpoint of interest (i.e. the location of the ground station). These data can be available for short durations (i.e. 1 min,, 2 min., and 10 min.). However, they require frequent maintenance and cleaning to ensure highly reliable data. The cost of installation varies according to the class of the radiometers being used. As a

result, the cost of maintaining ground stations is high. There are additional costs such as cleaning instrument sensors and labour to do the work.

2.5 Studies of the correlation between the clearness index and diffuse fraction

One of the objectives of this study is to find the correlation between the hourly clearness index and the hourly diffuse fraction. Liu and Jordan [6] used datasets from 98 locations in the U.S. and Canada (19º to 55º North latitude) to predict the instantaneous intensity of the diffuse radiation under a cloud free atmosphere on a horizontal surface. They did this by knowing the relationship between the diffuse fraction and beam transmittance, or clearness index. They also estimated the long-term average hourly and daily total intensity of the diffuse radiation with various degrees of cloudiness. Orgill and Holland [7] extended the work done by Liu and Jordan [6] to develop a correlation equation for hourly diffuse radiation on horizontal from September 1967 to August 1971. The model was developed with the use of Toronto data (Lat. 42.8º N), since both the global and diffuse radiation are measured in Canada. Linear relationships were developed to predict the hourly diffuse fraction knowing the hourly clearness index for three different intervals (i.e. different degrees of cloudiness).

In 1982, Erbs *et al*. [8] followed the same approach as Orgill and Holland [7]. However, the model developed to establish a relationship between the hourly clearness index and hourly diffuse fraction was based on data from five locations in the U.S. (31.08º to 42.42º North latitude) with a short duration ranging from one year to four years.

In 1990, Reindl *et al*. [14] added three more factors in addition to the clearness index: solar elevation angle, temperature, and relative humidity. They used data from five locations in Europe and the U.S. in order to develop a correlation between the diffuse fraction and clearness index that included the three new parameters in their model. They developed three models based on the availability of the parameters required to model the diffuse fraction. Almost two decades later, Helbig [15] combined the two models developed by Reindl *et al*. [14] when weather data are not available (i.e. temperature and relative humidity). Moreover, only models where only the clearness index and solar altitude are known were combined to achieve better performance.

2.5.1 Reindl-Helbig model

Hourly data play an important factor in simulating solar energy for a short period of time for locations around the world [14]. Reindl *et al*. [14] developed three decomposition models in 1990 based on hourly data from five different locations, with at least one year of data from each site. The data were obtained from two locations in the U.S., Albany, New York (Lat. 42.7º N, Long. 73.8º W) and Cape Canaveral, Florida (Lat. 28.4º N, Long. 80.6º W), and three locations in Europe: Copenhagen, Denmark (Lat. 55.7º N, Long. 12.6º E), Hamburg, Germany (Lat. 53.5º N, Long. 10.0º E), and Valentia, Ireland (Lat. 51.95º N, Long. 10.22º W). The five sites collected measurements of GHI, DNI or DHI, ambient temperature, and wet bulb temperature (or dew point). The model was developed as a form of a piecewise first order linear model. The Reindl model aims to estimate the diffuse fraction (k_d) to calculate DHI using Equation (2.21). Then, if needed, the DNI can also be computed using Equation (2.24). The development of the model took into consideration which factors can be included in order to provide a better estimation of the diffuse fraction. As an initial step, Reindl *et al*. considered including 28 predictor variables in the model. One of the primary variables is (k_t). However, they ended up with only four variables which are hourly clearness index, solar elevation angle, ambient temperature, and relative humidity. They found that the variable k_t is crucial in the low interval (overcast sky) and middle interval (partly cloudy sky). For a high interval (clear sky), α was found to be dominant. The authors had further reduced the predictor variables to two, where k_t and α are used to fit the data, or one predictor variable, where k_t only is used. Consequently, three models were developed for use depending on the availability of ambient temperature and wet bulb temperature. Helbig [15] combined the two models proposed by Reindl *et al*. when k_t and α are only available. The equations for estimating hourly diffuse fraction k_d under different sky conditions given by Reindl *et al*. are listed below. The first set of equations is used when the ambient temperature and relative humidity are not available.

For $0 \leq k_t \leq 0.3$,

$$k_d = 1.020 - 0.254\, k_t + 0.0123 \sin(\alpha),\ k_d \leq 1.0. \tag{2.25}$$

For $0.3 < k_t < 0.78$,

$$k_d = 1.400 - 1.749\, k_t + 0.177\, sin(\alpha)\ ,\ 0.1\ \leq\ k_d \leq 0.97. \tag{2.26}$$

For $k_t \geq 0.78$,

$$k_d = 0.486\, k_t - 0.182\, sin(\alpha),\ k_d \geq 0.1. \tag{2.27}$$

The second set of equations computes k_d as a function of k_t only.

For $0 \leq k_t \leq 0.3$,

$$k_d = 1.020 - 0.248\, k_t\ ,\ k_d \leq 1.0. \tag{2.28}$$

For $0.3 < k_t < 0.78$,

$$k_d = 1.45 - 1.67\, k_t. \tag{2.29}$$

For $k_t \geq 0.78$,

$$k_d = 0.147. \tag{2.30}$$

The choice of the ranges of the clearness index was based on finding the intermediate interval that will have minimum standard error of the final correlation. The new combined model was named Reindl*. The combination is taken from the two sets of equations listed above. Moreover, Helbig stated that the combined model was also built for clear sky conditions because very few scattered data can be found when k_t is higher than 0.8. For a partially cloudy sky interval, α is included, and for a cloudy interval it is assumed to have a minimal effect and is excluded in this interval. The set of equations for the Reindl* model is provided below.

For $0 \leq k_t \leq 0.3$,

$$k_d = 1.020 - 0.248\, k_t\ ,\ k_d \leq 1.0. \tag{2.31}$$

For $0.3 < k_t < 0.78$,

$$k_d = 1.400 - 1.749\, k_t + 0.177\, \sin(\alpha)\ ,\ 0.1\ \leq\ k_d \leq 0.97\ . \tag{2.32}$$

For $k_t \geq 0.78$,

$$k_d = 0.147. \tag{2.33}$$

One of the demerits of this model is the discontinuity of the diffuse fraction when $k_t = 0.3$ and $k_t = 0.78$ [16], [31].

2.5.2 DISC model

Many models have been highlighted as described by Maxwell [12] indicating that the relationship between k_n and k_t is a function of air mass, cloud cover, perceptible water vapor, atmospheric turbidity, and albedo. Nevertheless, getting an accurate estimate of DNI from GHI was uncertain. In 1987 at the National Renewable Energy Laboratory (NREL), Maxwell [12] proposed a quasi-physical model to estimate hourly direct normal irradiance values directly from hourly global horizontal irradiance values. Moreover, the diffuse fraction is not required for computing direct normal irradiance, but it can be used for computing diffuse horizontal irradiance. The development of this model was based on a computer program called the direct insolation simulation code (DISC). Maxwell used one year data set for Atlanta, Georgia, U.S., in the development of his model. He also validated the model by comparing estimated DNI with ground truth DNI using three U.S. locations. The term ”ground truth” is defined as “ information obtained by direct observation of a real system, as opposed to a model or simulation; a set of data that is considered to be accurate and reliable, and is used to calibrate a model, algorithm, procedure, etc.” [32] . The selected locations have different climatic conditions. These locations are Brownsville, Texas (Lat. 25.9º N, Long. 97.4º W, Elev. 6 m), Albuquerque, New Mexico (Lat. 35.05º N, Long. 106.6º W, Elev. 1619 m), and Bismarck, North Dakota (Lat. 46.7 º N, Long. 100.75 º W, Elev. 502 m). The model only requires hourly global horizontal irradiance values in order to obtain the hourly beam irradiance values. The clear-sky beam transmittance $k_{n,c}$ as a function of air mass is computed as:

$$k_{n,c} = 0.866 - 0.122AM + 0.0121AM^2\text{-}0.000653AM^3 + 0.000014AM^4. \tag{2.34}$$

The air mass is calculated using an expression developed by Kasten (1966). It is given as,

$$AM = \frac{1}{\cos(Z)+\frac{0.15}{(93.885-Z)^{1.253}}} \cdot \left(\frac{P}{P_0}\right), \tag{2.35}$$

where P is the pressure station and P_o is the standard pressure (1013.25 mbar). Another parameter used in this model which limits values of direct normal transmittance under clear-sky conditions is Δk_n. It is the reduction in the clear sky transmittance and it is computed as:

$$\Delta k_n = k_{n,c} - k_n, \tag{2.36}$$

where k_n is the hourly calculated direct beam transmittance.

Polynomial functions were implemented to calculate coefficients (a, b, & c) because the author observed a simple exponential relationship between AM and Δk_n. So, the computation of $k_{n,c}$ helps detects changes in k_n , which can be calculated as a function of k_t and AM .The relationship is illustrated by the following equation:

$$\Delta k_n = a + be^{c*AM}. \tag{2.37}$$

Next, the coefficients were calculated according to two ranges of k_t. They were grouped to two ranges when $k_t \leq 0.6$ (low interval) and $k_t > 0.6$ (high interval) depending on the observation made between Δk_n and AM.

For $k_t \leq 0.6$,

$$a = 0.512 - 1.56(k_t) + 2.286(k_t)^2 - 2.222(k_t)^3, \tag{2.38}$$

$$b = 0.370 + 0.962(k_t), \tag{2.39}$$

$$c = -0.280 + 0.932(k_t) - 2.048(k_t)^2 . \tag{2.40}$$

For $k_t > 0.6$,

$$a = -5.743 + 21.77(k_t) - 27.49(k_t)^2 + 11.56(k_t)^3, \tag{2.41}$$

$$b = 41.40 - 118.5(k_t) + 66.05(k_t)^2 + 31.90(k_t)^3, \tag{2.42}$$

$$c = -47.01 + 184.2(k_t) - 222.0(k_t)^2 + 73.81(k_t)^3. \tag{2.43}$$

Finally, the estimated DNI can be found by rearranging Equation (2.22) to calculate k_n and calculating the extraterrestrial radiation at the top of the atmosphere.

$$DNI_{DISC} = k_n * ETR, \tag{2.44}$$

$$k_n = k_{nc} - \Delta k_n. \tag{2.45}$$

2.5.3 Perez models

Perez *et al*. [33] evaluated the DISC model and determined that some improvements could be made to enhance its performance. Their work was sponsored by the American Society of Heating Refrigerating and Air Conditioning Engineers (ASHARE). As a result of their research, two models were developed. The first model was called DIRINT and is an improved version of the DISC model. The other model will be explained only briefly in this section since it is not used in this work. The aim of the Perez model is the same as that of the DISC model. Moreover, it converts the hourly global irradiance into hourly direct irradiance. The first two dimensions considered in the Perez model are the zenith angle (*Z*) and the clearness index ($k_{t'}$). A new expression of the clearness index ($k_{t'}$) has been developed by Perez *et al*. [33] and it is independent of the zenith angle (i.e. neglecting the sun's position). It is formulated as [34]:

$$k_{t'} = \frac{k_t}{\left(1.031 * \exp\left(\frac{-1.4}{0.9 + \frac{9.4}{AM}}\right) + 0.1\right)}. \tag{2.46}$$

Perez *et al.* [34] mentioned that the independence on the Z contributed to reduce the dispersion in the measured irradiance on the *x*-axis when plotted against modeled irradiance on the *y*-axis. Also, the new formula showed that higher values of k_t are possible near higher values of *Z* (around 80º). The formula for the air mass is the same as the one used in the DISC model. A new dimension is introduced, which is the stability index ($\Delta k_{t'}$) that varies with time and tells the stability condition of the clearness index. Moreover, low values of ($\Delta k_{t'}$) indicate stable conditions and high values of ($\Delta k_{t'}$) indicate unstable conditions. A turbidity factor of 1.4 was used in Equation (2.46). The stability index ($\Delta k_{t'}$) is expressed as:

$$\Delta k_{t'} = 0.5 * (|k_{t'} - k_{t'_{i+1}}| + (|k_{t'} - k_{t'_{i-1}}|), \tag{2.47}$$

where the subscripts (*i*,) (*i+1)*, and (*i-1)* refer to the current, next and former value of the hourly values, respectively. If the former or next hourly value is missing, then $\Delta k_{t'}$ is formulated as:

$$\Delta k_{t'} = \left| k_{t'} \pm k_{t'_{i+1}} \right|. \tag{2.48}$$

Dew point temperature data (Td) (ºC) was added as an independent dimension because relevant data are usually available, plus this dimension gives an estimation of the precipitable water content in the atmosphere. Moreover, the amount of moisture in the air absorbs solar radiation and enhances aerosols growth leading to the diffusion of the direct beam [33]. So, the atmospheric precipitable water (*W*) is considered to be a fourth dimension in the Perez model. The equation of the atmospheric precipitable water (*W*) is calculated as:

$$W = \exp(0.07 * Td - 0.075). \tag{2.49}$$

The dew temperature values are calculated as [35]:

$$Td = T - (\frac{100 - RH}{5}). \tag{2.50}$$

So, the net total is four dimensions that are used to correct the DNI modeled by the DISC model.

$$DNI_{DIRINT} = DNI_{DISC} \cdot X(k_{t'}, Z, W, \Delta k_{t'}), \tag{2.51}$$

where $X(k_{t'}, Z, W, \Delta k_{t'})$ is the correction matrix consisting of 6 × 6 × 5 × 7 matrix. The coefficients of the correction matrix are obtained by looking up the table listed below, which is divided into seven bins. The binning was used in order to treat each dimension independently. The matrix can be reduced depending on the number of the available dimensions (e.g. 3D in case *W* data are not available). The coefficient correction matrix is obtained from Dr. Richard Perez [36].

Table 2.11 Bins used for $X(k_{t'}, Z, W, \Delta k_{t'})$.

Bin #	$k_{t'}$	Z (°)	W (cm)	$\Delta k_{t'}$
1	0 - 0.24	0 - 25	0 - 1	0 – 0.015
2	0.24 – 0.4	25 – 40	1 – 2	0.015 – 0.035
3	0.4 – 0.56	40 – 55	2 – 3	0.035 – 0.07
4	0.56 – 0.7	55 – 70	3 - ∞	0.07 – 0.15
5	0.7 – 0.8	70 – 80	0 - ∞[1]	0.15 – 0.3
6	0.8 – 1[2]	80 – 90[3]		0.3 - 1[2]
7				0 - 1[4]

The other model developed by Perez *et al*. [33] uses simple linear relationships to model the direct normal beam.

$$DNI = \frac{ETR}{0.87291} \cdot kb' \cdot \exp\left(\frac{-1.4}{0.9+\frac{9.4}{AM}}\right), \tag{2.52}$$

where kb' is a direct transmittance that is independent of the sun`s position. It is given as:

For $kb' < 0.2$, $kb' = 0$, otherwise

$$kb' = a\left(k_{t'}, Z, W, \Delta k_{t'}\right) \cdot k_{t'} + b\left(k_{t'}, Z, W, \Delta k_{t'}\right), \tag{2.53}$$

where *a* and *b* can be obtained from a four dimensional table consisting of $8 \times 5 \times 4 \times 6$ matrix depending on the number of bins. The table for the second model is not shown since the light is shaded on the effect of the correction matrix on the DISC modeled values. Also, the DIRINT is commonly used in comparison to the second model. An outline of this table can be found in Perez's work [33].

[1] Values are used if W is not available (2D or 3D) model.
[2] Values of $k_{t'}$ exceeding 0.87 should be dealt with as suspect for locations below 1000 m altitude.
[3] The model was developed for $Z < 85^{o}$. Caution is required for values exceeding this level.
[4] Values are used if $\Delta k_{t'}$ is not available (2D or 3D) model.

2.5.4 BRL model

The BRL model was named after the researchers who developed it: Boland, Ridley, and Lauret [37]. One of the reasons for the development of this model is that models developed for northern hemisphere locations may not be effective for southern hemisphere locations [16]. The objective of this model is to compute the diffuse fraction first. Then DHI can be found knowing the GHI values. If DNI is needed, it can then be calculated by applying the equation that governs the three components. This model uses a decaying logistic function as it has been demonstrated to be a good method to predict DHI and is extensively employed in the field of ecology and species growth representation. Also, the logistic function has a curve of "S" shape and it behaves similarly as the diffuse fraction decreases when the clearness index increases and vice versa.

Lanini [31] indicated that the diffuse fraction estimated by the Reindl *et al*. [14] model varies with solar altitude in the middle interval of the clearness index ($0.3 < k_t < 0.78$). This variation decreases near the end intervals when $k_t < 0.3$ and $k_t > 078$ [16]. As previously mentioned, the Reindl *et al*. [14] model utilizes the use of solar altitude and hourly clearness index in modeling diffuse fraction. In 2001, Boland *et al*. [38] established a model for Australian cities using a logistic function model of the form, $\frac{1}{1+\exp(\beta_0+\beta_1 k_t)}$. Here, the numerator equals 1 because this value represents the maximum value of the *y-axis* (i.e. diffuse fraction). Seven years later, Boland *et al*.[39] implemented statistical techniques to assess the logistic function [28]. Moreover, the authors added more predictors in order to cover a wide spread of data near the two ends of the clearness index. The model was established based on hourly data obtained from Australian data for Adelaide and Darwin, located in the southern hemisphere, and other stations around the world located in the northern hemisphere (Bracknell, Maputo, Macau, Lisbon, and Uccle). The data for Adelaide and Darwin were obtained from the Australian Bureau of Meteorology. The authors claim that this model can be implemented globally in any location. Many predictors are included in this model: hourly clearness index (k_t), daily clearness index (K_t), apparent solar time (*AST*), persistence ($\boldsymbol{\psi}$), and solar elevation angle (α). The assumption made for the apparent solar time is that it is not symmetric about solar noon which might help explain the atmosphere conditions between morning and afternoon. The persistence is defined as the

average value of the previous hour and the next hour of the clearness index. It is calculated as:

$$\Psi = \begin{cases} \frac{[k_{t+1}+k_{t-1}]}{2} & sunrise \leq t \leq sunset \\ k_{t+1} & t = sunrise \\ k_{t-1} & t = sunset \end{cases}, \tag{2.54}$$

where *t* represents time (hour) of the day. The daily clearness index is computed as shown in the following equation:

$$K_t = \frac{\sum_{i=1}^{24} GHI_i}{\sum_{i=1}^{24} Ho_i}, \tag{2.55}$$

where GHI_i is the global radiation at hour *i* and Ho_i is the extraterrestrial radiation at an hour *i*. The authors included the daily clearness index because the entire day may feature a unique pattern. The coefficients introduced in the model were obtained by the Bayesian interference method [28]. Finally, the diffuse fraction can be notated as:

$$k_d = \frac{1}{1+\exp(-5.38+6.63k_t-0.06AST-0.007\alpha+1.725K_t+1.3\Psi)}. \tag{2.56}$$

From the above equation, it is evident that the daily clearness index (K_t) and persistence ($\boldsymbol{\psi}$) are strongly correlated to the hourly clearness index (k_t), especially in the interval ranging between (0.25 and 0.75). On the other hand, the model's output is less sensitive to the variation in the apparent solar time (*AST*) and solar elevation angle (α). Although the latter two variables have an insignificant effect on the model's outcome, the authors stated that it is still important to include them in the model.

The BRL model is easy to apply since it only requires one equation to estimate the diffuse fraction; however, using this model could be time-consuming if it is applied to long one-minute timestamp data.

2.5.5 Summary of models

Table 2.12 summarizes the models used in this study in terms of parameters required and the ability of each model to directly estimate DNI from the given input parameters.

Table 2.12 Summary of the parameters required for each model to estimate either DNI or DHI.

Model	Parameters required	Estimation of DNI	
		Immediate	Indirect
BRL	k_t, AST, α, K_t, Ψ		✓
Reindl	k_t, α		✓
Reindl*	k_t, α		✓
DISC	k_t, α, AM	✓	
Perez	k_t, α, AM, $X(k_{t'}, Z, W, \Delta k_{t'})$	✓	

Also, a summary of the advantages and disadvantages of each model is shown in Table 2.13.

Table 2.13 Advantages and disadvantages of each model used in this study

Model	Advantages	Disadvantages
Reindl*	Easy to implement	-The model is discontinuous at the end of each interval (i.e. k_t=0.3 and k_t=0.78) since it consists of three main intervals. -Assumes the diffuse fraction is constant for high clearness values (i.e. >0.78).
BRL	-Requires one equation to be implemented. -Can be applied globally as the authors claim.	-Many parameters are required to compute the diffuse fraction. -It is time consuming for large short term data sets. -Assumes the radiation is mostly direct for clear sky events.

DISC	-Shows a good fit to some extent for high clearness index values.	-As with other models, the DISC model behavior does not match the behavior of the measured or calculated data, especially for partial cloud intervals.
Perez	-Requires one input to be multiplied by the four-dimensional correction matrix.	-Requires weather data. If the data are not available, it can still be applied, but the model could have low accuracy.

The following chapter gave details about the type of data provided, and instruments used for each solar insulation stations. Also, a thorough description of the methodology used for the processing of solar data by applying different tests to examine the validity of the data is presented. Regarding recorded meteorological data, an explanation is provided of this type of data for the Ma'an station in Jordan.

Chapter 3 Methodology

In this chapter, a summary of each ground station is included in terms of type of data is collected, time resolution, range of data, and different approaches used to process the acquired data to be able to compare data from the various test sites. We also include a discussion section on the filtering process used to remove any erroneous data points before the implementation of the four models.

3.1 Selection of locations

Four locations in different countries were chosen for testing the models. The cities are Al-Hanakiyah, Western region, Saudi Arabia (Lat. 24.85577º N, Long. 40.536º E, El. 873 m), Ma'an, Southern region, Jordan (Lat. 30.17º N, Long. 35.82º E, El. 1012 m), Boulder, Colorado, U.S. (Lat. 40.05º N, Long. 105 º W, El. 1584 m), and Ottawa, Ontario, Canada (Lat. 45.42º N, Long. 75.68º W, El. 70 m). Note the geographic location of each city represents the station's coordinates, not the coordinates of the cities themselves. The selection of Boulder was based on the fact that that altitude of this location exceeds 1000 m. The increase in the location's elevation is associated with great potential of solar spectrum. The availability of high solar radiation means availability of high GHI, DNI, and low availability of DHI. Another reason is that, the DISC model was developed in Boulder which makes this location ideal to test this model in comparison to other models. The city of Ottawa was selected because it is one of the representative cities of the North America climate as well it has a ground station installed on the main campus at the University of Ottawa which is operated and administrated by the SUNLAB group members. It is known that the solar potential available in most Middle East countries is prominent. A project was planned to be executed in Taibah University located in Medina, Western region, Saudi Arabia (Lat. 24.5247° N, Long. 39.5692° E, El. 610 m) in 2014. The project aimed to install ground station to measure the three components. However, due to financial delay issues the station was not mounted in time for this research work. The Taibah University location build is delayed until 2017. As an alternative, Al-Hanakiyah station was selected since its geographical coordinates as well as weather pattern are similar to those of Medina. The Ma'an location was chosen for the same reason, too. Another cause of choosing Ma'an is to

test the models performance where two locations have similar geographical coordinates. Finally, the solar development in the Middle East region is growing rapidly and data for these locations were made available in the past two years with state-of-the-art equipment and excellent equipment upkeep and data acquisition capabilities. The locations vary in terms of weather as illustrated in Table 3.1.

Table 3.1 Type of climate in the four locations.

Station	Country	Climate
Al-Hanakiyah	Saudi Arabia	Hot/desert in summer and mild in winter
Boulder	United States	Hot/humid in summer and mild /bitter cold in winter
Ma'an	Jordan	Hot/desert in summer and mild/cold in winter
Ottawa	Canada	Temperate/humid in summer and cold/bitter cold in winter

3.2 Importance of high quality data

The measurement of GHI, DNI, and DHI requires a high level of accuracy for correct analysis. Error and uncertainty are usually associated with the measurement of solar radiation [40]. One of the objectives of this thesis is to understand the sources of uncertainty in the modeled data and ensure it is as close to the measured data when comparisons are possible between both. If measured data are applied to a model without having highly accurate data, this can lead to incorrect assessment of both measured and estimated data. Obviously, there is a direct relationship between accuracy of collected data and accuracy of the model. The following section will briefly describe possible errors and problems of uncertainty which result when obtaining inaccurate data and making an improper assessment. Figure 3-1 depicts the impact of incorrect data on the model's behavior. The example here shows a great

deviation between the measured data and the estimated data of DNI in Boulder (i.e. high uncertainty). The deviation is related to unexpected values of GHI in the entire month of June 2014. Since the models depend on the clearness index, calculated from the GHI values, incorrect data will produce higher uncertainty. As a result, if a comparison is made before and after elimination of these data, high improvement in the performance of a model could be achieved. These data are also critical for long-term evaluation for studying the bankability and possibility of installing solar systems in such a location.

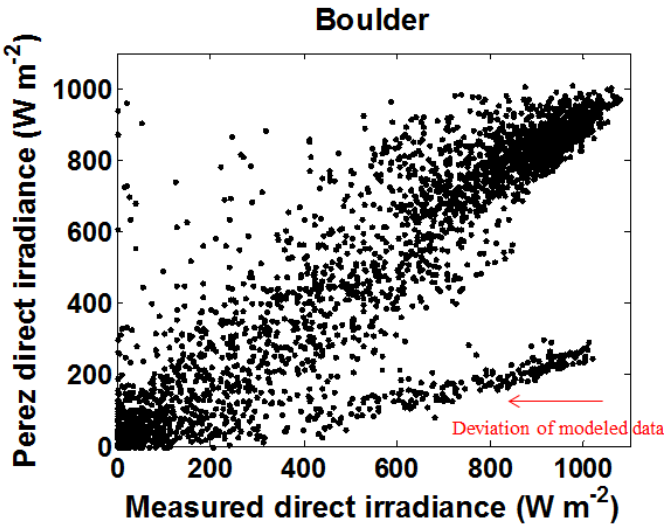


Figure 3-1 Observed deviation from the clustered data due to unreasonable GHI data**.**

3.2.1 Obtaining high quality data

Ensuring high quality data is challenging and depends on many factors. Some of them are addressed in this section.

A. Maintenance

Frequent maintenance of instruments is necessary for keeping measured values within an acceptable range. For example, in 2008, the World Meteorological Organization (WMO) suggested that three days should be the maximum period between checks, especially if the ground station is installed in or near locations where the air pollution level is high.

B. Selection of an appropriate location

Careful consideration of a location for installing ground stations is crucial in order to guarantee accurate data. One of the features of an optimal location is that it should not be surrounded by obstacles such as mountains, trees, building, and so on.

C. Instrumentation

The type of instrument used to measure any component plays another role in the quality of the measurement. There are issues related to equipment errors such as cosine response, non-linearity and, shade ring misalignment. Also, other issues that should be taken into consideration are light intensity attenuating factors (i.e. dust, snow, bird droppings, and so on), malfunctioning stations, inexperienced operators, and more. So, the accuracy of the solar measurement data requires extensive experience and knowledge to make sure everything is carried out properly.

However, inaccurate data may also be the result of cleaning the instruments. So, at the time of cleaning, the measurement of solar irradiance data should be flagged or replaced using filling techniques available in literature.

3.3 Description of stations

Four different stations for measuring solar and meteorological data were selected for testing various models against ground truth data. Data related to stations were handled, corrected, and, filled using techniques described in this section; there were almost no missing timestamps from Al-Hanakiyah and the Ma'an ground stations. However, for the two remaining stations (i.e. Boulder and Ottawa), the missing timestamps required filling in where it was suitable. Table 3.2 shows the coordinates and type of data measured by each station in addition to time used to record the data. It is important to point out that there are some types of data that were not included in Table 3.2 since they are irrelevant to this study.

Table 3.2 Types of data provided for each location.

Location				
Parameter	**Al-Hanakiyah**	**Boulder**	**Ma'an**	**Ottawa**
GHI	✓	✓	✓	✓
DNI	✓	✓	✓	✓
DHI	✓	✓	✓	✓
Zenith angle	✓	✓	✖	✓
Pressure	✓	✓	✓	✓
Temperature	✓	✓	✓	✓
Relative humidity	✓	✓	✓	✓
UTC	✖	✓	✖	✖
Missing data	✓	✓	✖	✓

3.4 Boulder station

The Boulder station (Lat. 40.05º N, Long. 105º W, Elev. 1584 m), known as Boulder Atmospheric Observatory (BAO), was built by the National Oceanic and Atmospheric Administration (NOAA) in Erie, Colorado. The measurements of solar radiation are registered at one-minute time steps according to coordinated universal time (UTC); however, data of solar irradiance were recorded at three- minute time steps before 1988. The data include GHI, DNI, DHI, and zenith angle in addition to other data not related to this study. Also, pressure station, temperature, and relative humidity data are recorded at this station. Data denoted by -999 are related to "non-existing" or "missing" values. The Precision spectral pyranometer was installed to take measurements of GHI. For the measurements of DNI, the Eppley normal incidence pyrheliometer was mounted on the dual axis automatic solar tracker, model SMT-3 to keep track of the sun path both horizontally (azimuth) and vertically (elevation), during the whole year. The movement of the tracker is done by a motor where a computer program calculates the solar position for the given time and location. The measurement of DHI is achieved by using a shadowband or ball to block the direct radiation falling on the pyranometer's sensor; manual adjustment of the shadowband is regularly required.

3.4.1 Processing and filling missing solar radiation data

Solar irradiance data were chosen from 2014/01/01 to 2015/02/28. The reason for choosing this range of data was that the ground station in Al-Hanakiyah was installed only recently, in 2013. So, the same range was selected to keep the datasets from all locations consistent in order to analyze the solar radiation data properly. Data were provided in Excel spreadsheets with each spreadsheet consisting of a full year of data related to GHI, DNI, DHI and, zenith angle, plus other data. The total number of measurements was approximately 525,600 for each year. Matlab software was implemented to calculate the hourly average values for both the solar irradiance as well as meteorological data. The values were calculated on an hourly basis for two reasons. The first reason was that the models included in this study were based on hourly data; secondly, Al-Hanakiyah data were provided as hourly averages. The total number of measurements available for each hour was 60 measurements; however, some hours had some missing values of the solar irradiance data.

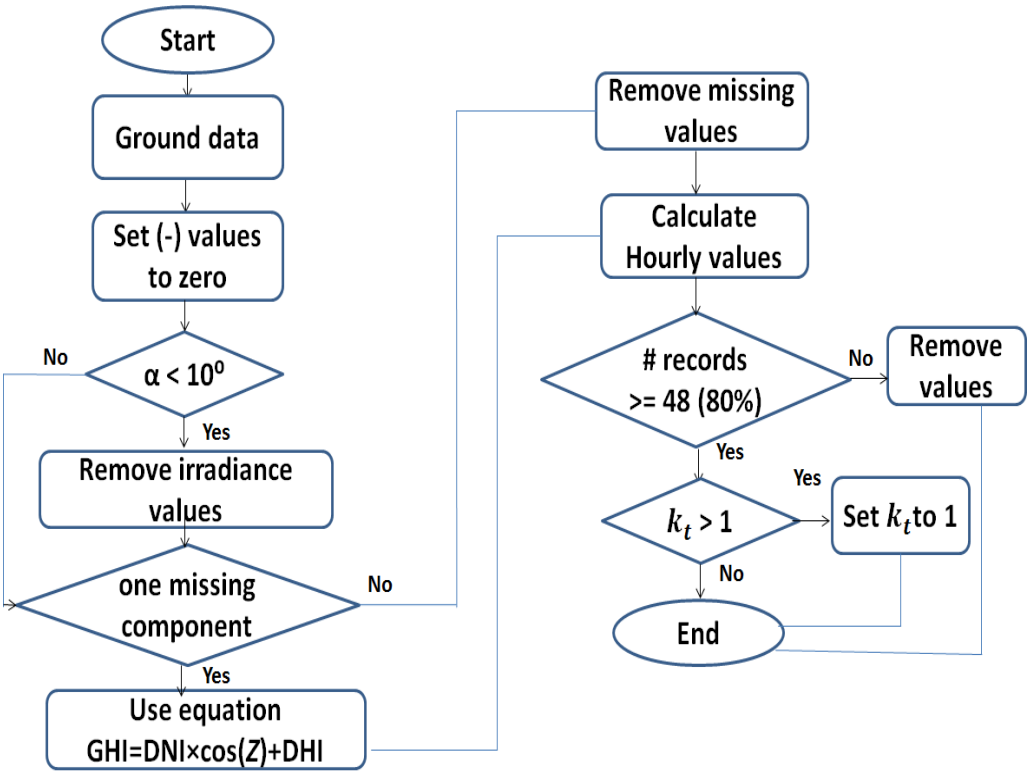


Figure 3-2 Methodology used for processing and filling missing data for Boulder location.

Missing values were eliminated or substituted depending on the approach used to fill the missing timestamps. Figure 3-2 illustrates the methodology used to process and fill in the missing data. First, the data were shifted by seven hours, since the data were originally recorded at UTC. After shifting the data, some values of irradiance were recorded as negative values. This was because of rapid change in the atmosphere such as overcast weather and rain in Boulder, so negative voltage signals were recorded due to frequent change in the radiation environment. As a consequence, the negative values were set to zero. The next filter implies that only values where the solar altitude is above 10º were selected. Values under 10º are usually small, so they were neglected. The third filter seeks to determine how many components (i.e. GHI, DNI, and DHI) were missing at a specific timestamp. Timestamps with two missing components or more were discarded. If only one component among the three is missing, then it can be calculated from Equation (2.14). After that, hourly averages are calculated and the values are stored at the beginning of the interval. The fourth filter checks if the available number of data within a specific hour is more than or equal to 80% (i.e. 48/60 measurements) of the original data. The primary number of data, including missing as well as actual data, is 60 per hour. So, the data were selected if only 48 or more measurements were available according to the approach followed for Al-Hanakiyah data [5]. Finally, the clearness index was calculated using Equation (2.20). If the value of the clearness index exceeds one, it is set to one because the clearness index may exceed one at higher temporal resolution. Moreover, GHI can sometimes exceed the extraterrestrial radiation due to the enhancement of the diffuse irradiance as the sun approaches the edge of a cloud [41]. After applying the above method, the total number of hours is 3526 daylight hours.

3.5 Al-Hanakiyah station

The program initiated by King Abdullah City for Atomic and Renewable Energy (K.A.CARE) is called the Renewable Resource Monitoring and Mapping (RRMM) program. It aims to benefit from renewable energy sources such as solar, wind, and geothermal sources available in the Kingdom [5]. Al-Hanakiyah station was one of the fifty stations deployed in the Kingdom of Saudi Arabia in 2013-2014. Figure 3-4 illustrates distributed ground-

installed and uninstalled stations in Saudi Arabia. The location of Al-Hanakiyah is located in a desert area, as shown in Figure 3-3.

Figure 3-3 Location of Al-Hanakiyah station [42].

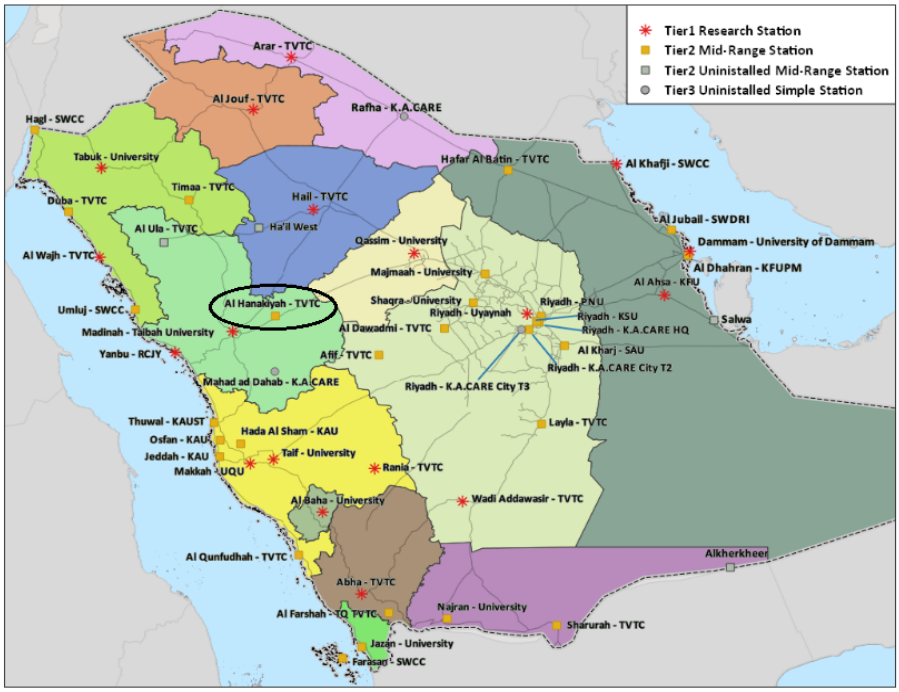


Figure 3-4 Installed ground station with different tiers across the Kingdom of Saudi Arabia [5].

As can be clearly seen in Figure 3-4, there are three different types of stations: Tier 1 (research stations), Tier 2 (mid-range stations), and Tier 3 (simple stations). Each station has specific requirements. Table 3.3 describes the features of each tier. GTI refers to the global tilted irradiance (i.e. global irradiance measured on an inclined plane).

Table 3.3 Description of different types of ground stations installed in Saudi Arabia.

Station type	Class	Description
Tier 1	A	Installed for research purposes where measurements of primary and secondary solar radiation data are registered plus measurements of basic meteorological data, dust deposition and horizontal visibility.
	B	Has the same features as A but without secondary data.
	C	Contains measurements of (GHI, GTI, DNI and DHI) and basic meteorological data.
Tier 2	-	RSRs are installed where GHI and DHI are recorded at one-minute time steps; however, DNI is calculated. Also, meteorological data are registered.
Tier 3	-	Only measurements of GHI and GTI plus temperature data are taken.

Al-Hanakiyah station (Lat. 24.85º N, Long. 40.53º E, Elev. 873 m) is classified as a Tier 2 station; the site name is Al-Hanakiyah Technical Institute. The data provided by this station are GHI, DNI, DHI, zenith angle, pressure, temperature, and relative humidity. The station is maintained and cleaned twice a week. The station started collecting data in August 2013, and continues to do so at the present. The hourly data provided to the user represents the end of interval. GHI and DHI are measured by the RSR model LICOR, which is manufactured by Irradiance Inc. The DNI is computed based on the algorithm used by the rotating shadowband instrument. The nominal uncertainty for the RSR is $\pm$ 5%. A secondary pyranometer is installed to measure GHI. The redundant measurements of GHI indicate if there is an issue with one of the sensors in case there is a discrepancy exceeding the default uncertainty [43]. Another advantage of redundant measurements is that they can be used to apply the three component tests detailed in the SERI quality control procedure [44]. Also, the station is equipped with a barometer, temp/RH probe, and an anemometer/wind vane to measure pressure, temperature, relative humidity, wind speed, and direction, respectively.

3.5.1 Methodology used by the K.A. CARE staff

The hourly data provided at Al-Hanakiyah station were already checked by the K.A.CARE staff. Only one-minute data as well as quality flags are available to the staff members. A Solar Energy Research Institute (SERI) quality control procedure was developed by the National Renewable Energy Laboratory (NREL) for checking collected solar radiation data [44]. The SERI procedure was implemented by the K.A.CARE staff members. During cleaning events, short data stops are filled in with interpolation techniques, or the missing component is calculated using the other two components. After filling in the data, any residual timestamps with an interval of more than 1.5 hours are flagged as missing data. Aggregated values are not calculated if more than 20% of the data is missing due to errors such as equipment malfunctioning. Finally, data can be accessed through an online request. Figure 3-5 illustrates the methodology used to fill in the missing or corrupted data and to ensure high quality solar radiation data. The data range selected for this study is from 2014/01/01 to 2015/02/28. Data are discarded where the solar elevation angle is below 10º. The total number of collected hourly data is 4104 daylight hours.

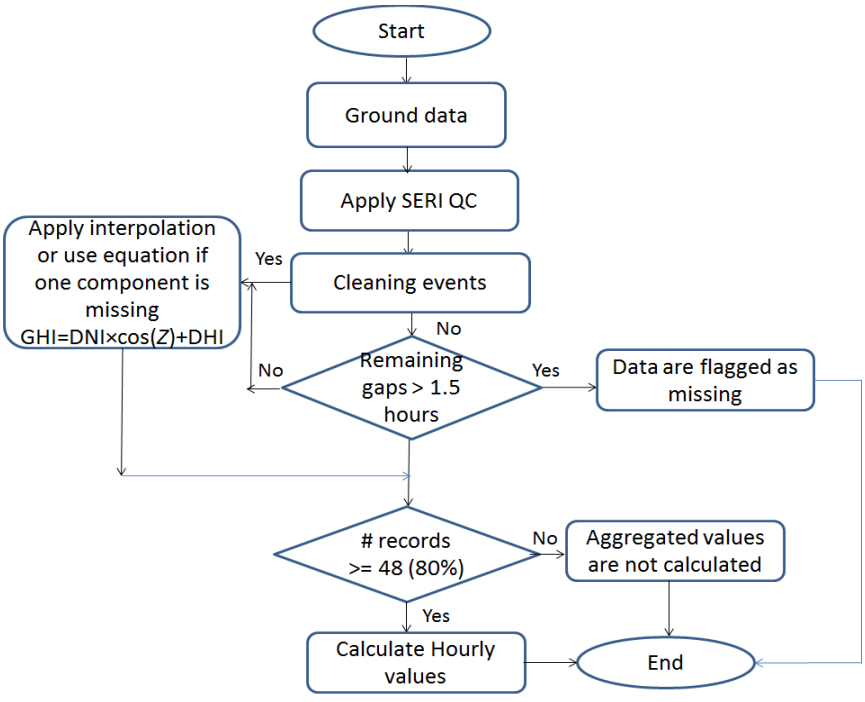


Figure 3-5 Methodology used for processing and filling missing data for Saudi Arabia.

3.6 Ma'an station

The enerMENA (Middle East and Northern Africa) project includes ten stations to measure on the ground local insulation conditions. These stations have been in operation since 2010. The quality control procedure for data screening and data flagging, known as the Meteorological Data Management System (MDMS), was created and implemented by Concentrating Solar Power (CSP) Services, DLR's Institute of Solar Research, and ARMINES [45], [46]. Five stations are currently installed in Morocco, two in Algeria, and one each for Tunisia, Jordan, and for Egypt. The Ma'an station (Lat. 30.17º N, Long. 35.82º E, Elev. 1012 m) started working on 2011/01/11. It is equipped with a CHP1 pyrheliometer for measuring DNI, and CMP21 pyranometers for measuring GHI and DHI. It is also equipped with instruments that take measurements of wind speed and wind direction at a 10 meter height, temperature, relative humidity, and pressure. The solar and meteorological data were obtained from Deutsches Zentrum für Luft and Raumfahrt (DLR) German Aerospace Center Institute of Solar Research [47]. The data are collected on a one-minute basis. They are then averaged to 10-minute automatically. The data selected for this study are from the period between 2014/01/01 and 2015/02/28. A total of 4569 daylight hours have been used in this work. The pyrheliometer is cleaned on weekdays only.

3.6.1 Methodology used to obtain more accurate data

Wolfertstetter stated that "Meteorological data has been measured at the meteorological station located at Ma'an Development Area in Jordan. The station was financed by enerMENA project funds provided by the German Federal Foreign Office" [48]. The data are collected at one-minute time steps and automatically averaged to 10-minute time steps. The gaps in data are either linearly interpolated, replaced by redundancy, or filled in with data from other stations. The data provided to the user include GHI, DNI, DHI, temperature, relative humidity, and pressure. The hourly values are stored at the end of the interval. The zenith angle values are calculated using the Solar Position Algorithm (SPA) [49]. Many quality control checks have been applied to the three components (i.e. GHI, DNI, and DHI). The strict limit for solar irradiance is 0 W/m^2. During night time, the limit is reduced to -5 W/m^2 for GHI and -1 W/m^2 for DNI. The values of the clearness index should be greater than

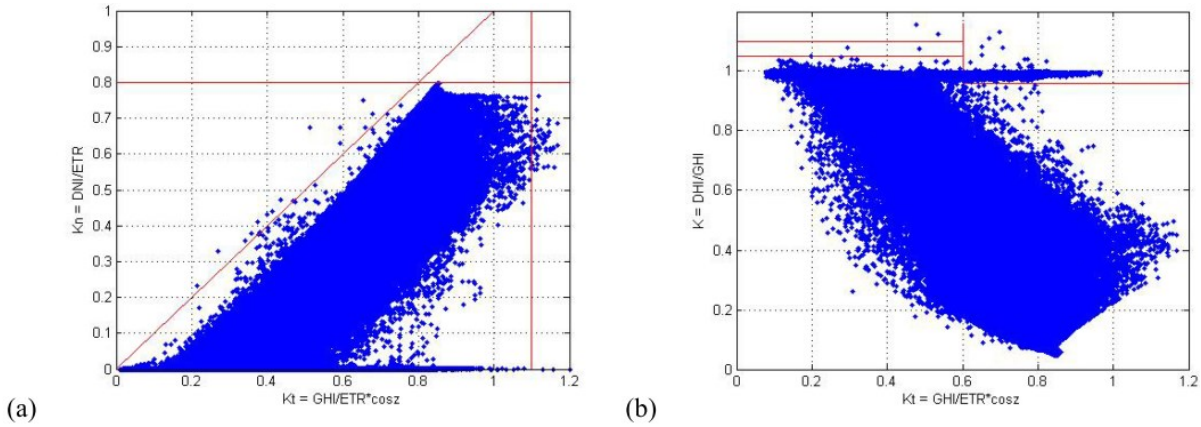
the values of the direct transmittance, and the upper limit of the clearness index should be 1.1 in the $k_t - k_n$ space. The conditions for the limits of the $k_t - k_d$ space are as follows [50]:

$$k_d > 1.05 \text{ for } Z < 75°,$$

and

$$k_d > 1.10 \text{ for } Z > 75°.$$

Figure 3-6 shows the limits (red lines) for $k_t - k_n$ space and $k_t - k_d$, respectively.



(a) (b)

Figure 3-6 Quality control checks for limits (red line) for a) DNI (k_t - k_n) and b) DHI (k_t – k_d) adapted from [45]. Low resolution of the figure carried forward from original source.

Next, a flag is set if the k_d is > 0.96 and k_t > 0.6 to indicate there is an issue in the function of the tracker/shadowband. Also, a flag is set when $k_n > k_t$ to detect cleaning events of the pyranometers. The following tests have been applied to ensure that the GHI, DNI and DHI lie between the ranges specified [51]:

$$0.03 \cdot ETR \cdot \cos(Z) < GHI < 1.2 \cdot ETR \cdot (\cos(Z))^{1.2} + 50 \text{ W/m}^2,$$

$$0 < DNI < 0.95 \cdot ETR \cdot (\cos(Z))^{0.2} + 10 \text{ W/m}^2,$$

$$0.03 \cdot ETR \cdot \cos(Z) < DHI < 0.75 \cdot ETR \cdot (\cos(Z))^{1.2} + 30 \text{ W/m}^2.$$

Consistency tests are performed to flag the data if they fulfill the conditions:

$$\frac{DHI}{GHI} > 1.05 \text{ for } Z > 75°,$$

and

$$\frac{DHI}{GHI} > 1.1 \text{ for } 93° > Z > 75°.$$

Additional consistency tests are applied but with less constraint compared to the above test.

$$\left|1 - \frac{GHI}{DHI+DNI \cdot \cos(Z)}\right| > 0.08 \text{ for Z} < 75°,$$

and

$$\left|1 - \frac{GHI}{DHI+DNI \cdot \cos(Z)}\right| > 0.15 \text{ for } 93° < Z < 75°.$$

After that, the change rate conditions are applied as [50]:

$$|GHI(t)/\cos(Z(t))/ETR - GHI(t-\Delta t)/cos(Z(t-\Delta t))/ETR|/\Delta t < 0.75 \text{ min}^{-1},$$

$$|DNI(t)/ETR - DNI(t-\Delta t|/ETR/\Delta t < 0.65 \text{ min}^{-1},$$

$$|DHI(t)/\cos(Z(t))/ETR - DHI(t-\Delta t)/cos(Z(t-\Delta t))/ETR|/\Delta t < 0.35 \text{ min}^{-1},$$

where Δt is the temporal resolution. This test can be applied for short timestamps (i.e.1 or 10 minutes). Higher change rates in DNI are detected by the implementation of the DNI coincidence (DNI_{coin}). It is expressed as [45]:

$$DNI_{coin} = DNI_{measured} - DNI_{calculated}. \quad (3.1)$$

Constant values of DNI coincidence are defined as the intervals having a length of 15 minutes and having a DNI coincidence of less than 4 W/m^2/min 2 min before and 2 min after the 15 min interval. During the 15 min interval, timestamps of the DNI coincidence exceeding 15 W/m^2/min are flagged. Figure 3-7 shows the use of the DNI coincidence test.

3.6.2 Specifying limits for ambient temperature data

The temperature range has been chosen to be between -10° C and +60° C; however, this range can be adapted to locations where the temperature reaches -30° C in winter. A coincidence check is applied to check the difference between individual values. An error flag is set when the difference between the individual values exceeds a certain limit (15 K for logger and temperature and 20 K for sensor and temperature). For a one-minute time

resolution data point, the alteration in temperature should not be more than ± 2 K/min. In the case of a 10-minute time resolution, the change in temperature should be less than or equal to ± 0.4 K/min.

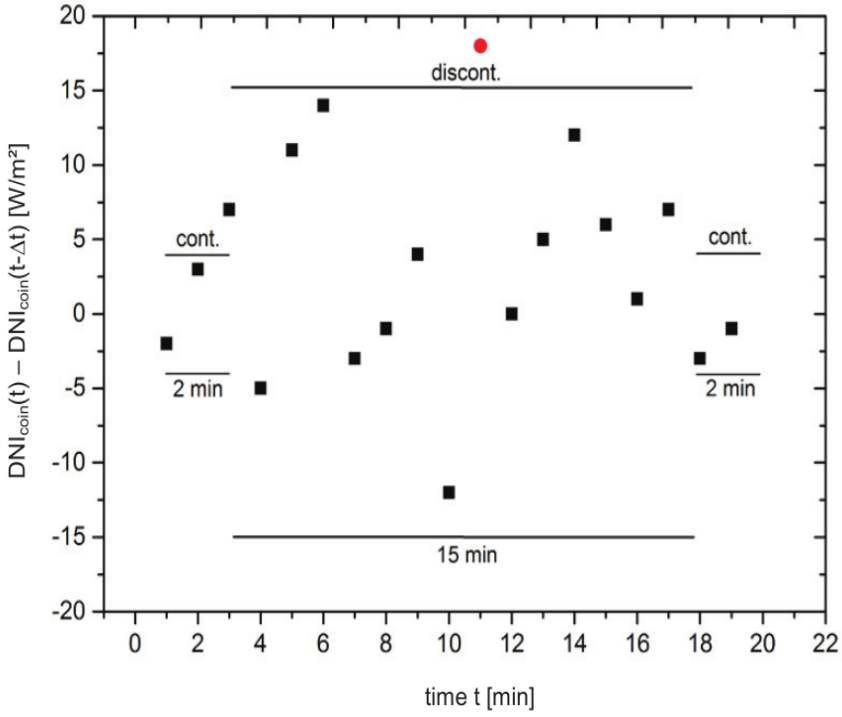


Figure 3-7 Process to check DNI values where a two-minute interval is checked before and after an interval of 15-minutes for DNI coincidence values of more than 4 W/m^2/min [45].

3.6.3 Specifying limits for pressure data

Measured pressure data should not exceed ±30 hPa from calculated pressure values according to the CSP Services. The equation to calculate the pressure knowing both the ambient temperature and site elevation is given as:

$$p = 1013.25 \cdot \left(1 - \frac{0.0065 \cdot \text{altitude}}{273.15 + T_{air}}\right)^{5.255}. \quad (3.2)$$

The measured pressure data show ±2 hPa and ±4 hPa for a one-minute and 10-minute time resolutions, respectively.

3.6.4 Specifying limits for relative humidity data and flagging irradiance data due to dew:

The range for the relative humidity data has been selected to be between 0% and 100%. The alteration rate should be less than 10%/min and 15%/10min for a one-minute and 10-minute

time resolution, respectively. The DNI data are flagged if the measured DNI is less in magnitude by 50 W/m^2 compared to its computed value, which is obtained by knowing the GHI, DHI, and the zenith/elevation angle.

3.7 Ottawa station

The Ottawa station (Lat. 45.42° N, Long. 75.68° W, Elev. 70 m) is located at the University of Ottawa where the SUNLAB research group is responsible for maintaining instruments and taking measurements of GHI, DNI, DHI, temperature, pressure, and relative humidity. One of the group's objectives is to study and analyze the performance of CPV systems, and to evaluate the atmospheric effects on the solar spectrum. The test site was installed in 2011 and the collection of data started on 2011/03/07. Secondary measurements of GHI are collected by the Lufft weather station; this helps to identify issues with the instruments by comparing primary and secondary measurements of GHI. The comparison provides good insight into detecting any issues or errors in data. The Eppley pyranometer and pyrheliometer take measurements of GHI and DNI, respectively. The data are collected at two-minute timestamps for the Eppley instruments while data are registered at one-minute timestamps for the Lufft weather station. The GHI data were corrected [52] by taking a polynomial fit curve of the cosine error for an Eppley PSP pyranometer provided by Reda [53].

3.7.1 Quality control procedure for obtaining high quality DNI data

A quality control procedure was implemented to ensure high quality DNI measurements; the uncertainty of DNI measurements at the test site is $\pm$ 2.3% [54]. Also, temperature dependence and tracker algorithms were developed to ensure accurate DNI data and to correct temperature data [54]. The Eppley pyrheliometer is fixed on the GoldenSUN dual-axis tracker. Data determining the position of the solar tracker are stored at one-minute time steps. The flow chart diagram in Figure 3-8 shows the steps followed to obtain accurate DNI data. Five categories have been created to group DNI data that have failed to fulfill the conditions in the quality control procedure. They are Instrument Error, No Tracker Data, Tracker under Maintenance, Tracker off Sun and Accurate DNI. Only data between sunrise and sunset are tested. The black arrow in Figure 3-8 indicates the data applied to test their accuracy and the red arrow means data have failed the test and have been removed. The first filer applied to the DNI measurements examines if the DNI records are below the

extraterrestrial radiation (*ETR*). If the condition is fulfilled, then the data pass to the second test. Otherwise, the data are removed and labeled as *Instrument Error*. The second condition implies that the tracker data timestamp should be matched with the timestamp of the DNI data. Sometimes, the tracker data are not available and are interpolated within a four-minute timestamp (i.e. two minutes before and after the missing data) given the condition that the DNI data are not missing. Otherwise, the data are removed and labeled as *No Tracker Data.* Next, the tracker algorithm will be applied to determine if the tracker is in closed loop condition under maintenance or installation by checking the difference between the azimuth and elevation values measured by potentiometers and the calculated values. If the difference exceeds 25°, then data are excluded from the DNI dataset and categorized as *Tracker under Maintenance.* The next filter examines the status of the tracker when it is in open loop condition (i.e. not enough sunlight). In overcast conditions, the DNI values are $\leq$ 160 W/m^2 (i.e. the sun sensor values are too small) [52]. If the tracker is in open loop condition and it is aligned, the data points may pass this test or else are removed and labeled as *Tracker off Sun.* The final filter compares the elevation and azimuth value differences in the case of low DNI values if they lie within the field of view of the pyrheliometer, which is about $\pm$ 2.9°. Consequently, the data are classified as *Accurate DNI* if they pass this test; otherwise, they are classified as *Tracker off Sun*.

3.7.2 Correction of DHI data

DHI measurements taken during the course of a day need to be corrected because the screening of the pyranometer's sensor blocks a portion of the anisotropic diffuse sky irradiance. Hence, a correction method is required to obtain accurate DHI values so the error between the estimated values and measured (corrected) values is minimized. To achieve this goal, two models by Ballet *et al*. were used [52]. The first model uses one correction factor for all sky conditions. It depends on multiple factors: the skylight brightness (Δ), the cloud condition (ε) and the zenith angle (Z). The equations for Δ and ε are provided below.

$$\Delta = \frac{DHI_{measur}}{(ETR \cdot \cos(\mathrm{Z}))}, \tag{3.3}$$

$$\varepsilon = \frac{(DNI_{measur} + DHI_{measur})}{DHI_{measur}}. \tag{3.4}$$

The second model divides the cloud conditions into four bins according to the values of ε: $\varepsilon \leq 3.5$, $3.5 \leq \varepsilon \leq 8$, $8 \leq \varepsilon \leq 11$, and $11 \leq \varepsilon$. The corrected diffuse irradiance is obtained by multiplying measured (uncorrected) diffuse irradiance by the algorithms detailed in [55].

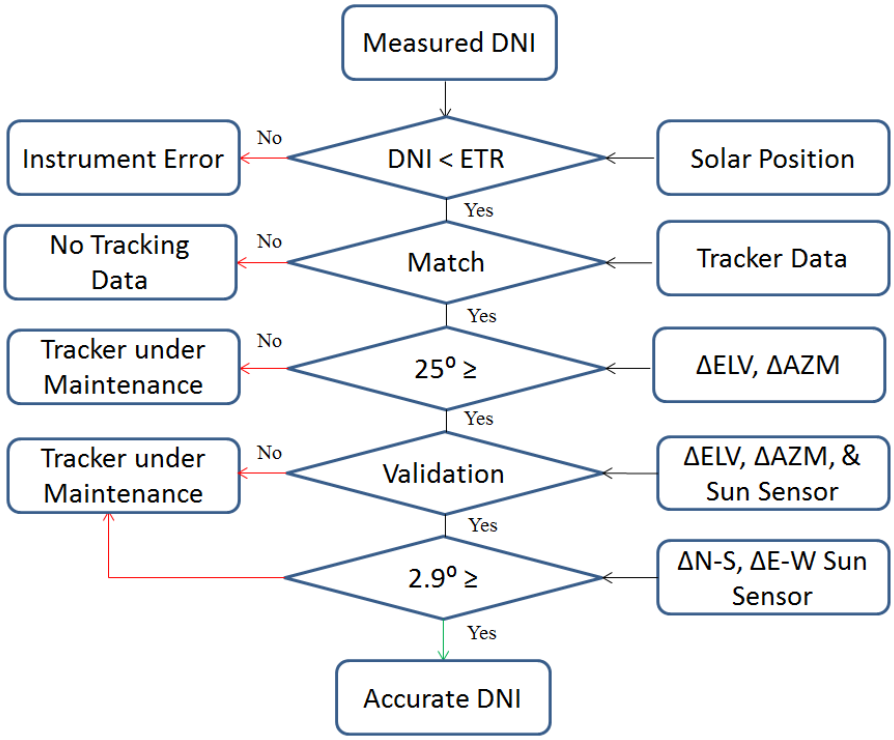


Figure 3-8 The flow chart diagram illustrates the procedure followed to obtain accurate DNI measurements at the Ottawa station. The black arrow indicates data are tested against different criterion while the red arrow represents DNI removed due to errors as shown in the flow chart. Reproduced with permission from [54].

3.7.3 Post processing of Ottawa data

The data set for solar radiation recorded at the SUNLAB test site contains very large data gaps for the three components for the period ranging from 2014/01/01 to 2015/02/28. This is due to many issues that occurred during this period such as vandalism which caused major damage to the pyranometer measuring GHI and DHI data, and tracker problems for measuring DNI data. The data gaps can be filled in various ways in order to have a complete data set. Some methods are linear interpolation for short gaps which can be from 3-5 hours (20%-25% of data) for high resolution data, redundancy of data, use of models to estimate DNI and DHI (which is not applicable in this case), or substituting the missing values from nearby stations. However, these filled data cannot be used as reference data since they are

not as accurate as the measurements taken from the radiometers to validate a model performance [56]. As previously mentioned, there are many missing timestamps for the three components. So, the approach followed here was to remove the large gaps for any components and keep only timestamps where 80% (24 out 30) of the data were available within an hour. Ottawa data also have been removed when the zenith angle is above 80º. Moreover, many values were missing for the three components for more than two months in total. Therefore, these values were removed and only values of more than or equal to 24 measurements within one hour were kept. After that, the hourly average was calculated after correcting DHI and ensuring high quality DNI values. The number of the hours was 2883 daylight hours. The timestamps of the meteorological data is matched with timestamps of the solar data.

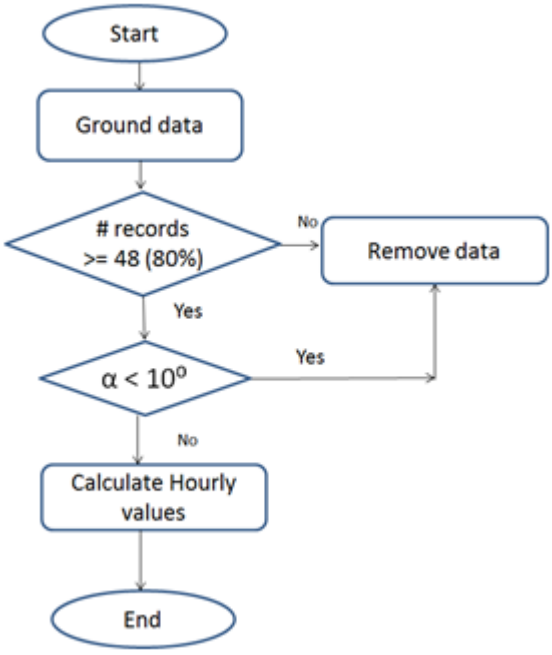


Figure 3-9 Methodology used for processing Ottawa data.

3.8 Comparison of modeled data with ground-truth data

The three solar radiation components are not always available at every ground-installed station around the world. One station may measure only GHI data. In this case, the DNI and DHI data can be found by looking for neighboring stations that measure DNI and DHI data, or models can be applied to measured GHI data to determine the missing components. The

need to employ models arises, especially when there is no other station installed near the location of interest. Some stations may measure two solar radiation components where one measurement is GHI and the other is DNI or DHI. Thus, the third solar radiation component can be calculated by applying the equation that governs the three components. The availability of all three components at the deployed station represents the best scenario since the model of the performance can be assessed given the condition that the ground (measured) data are highly accurate. This will lead to a detailed analysis of the models. So if there gaps in the measured data, the best model tested can be used to fill the missing timestamps. Figure 3-10 represents a flow chart of the relationship between the measured and the modeled irradiance components.

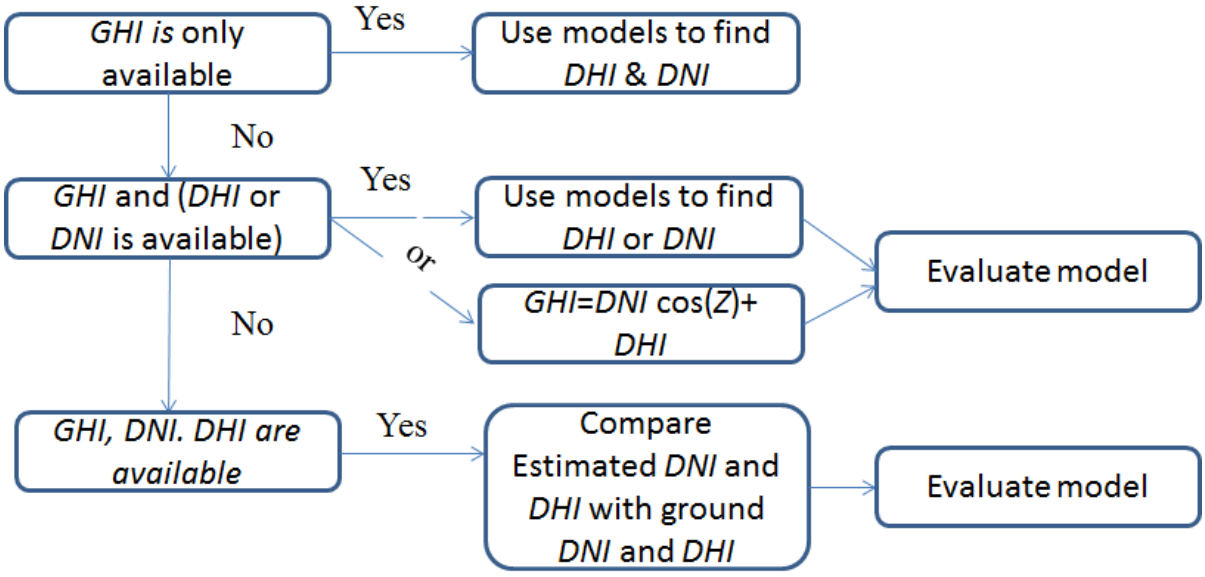


Figure 3-10 Flow chart of the relationship between the measured and the modeled irradiance components.

3.9 Evaluation of model's performance

Following the procedure to fill in missing data or filter out undesired data, statistical analyses are required in order to assess the performance of each model and compare the closeness of the estimated value to the actual value. The first tool applied is for determining the error between the two quantities (actual and estimated values). That is, the error (E) is given as:

$$E = Measured\ value - Modeled\ value. \tag{3.5}$$

After calculating the error, mean bias error (MBE) is calculated to see if the model is biased or unbiased. If the value of the mean bias error is zero, then the model is said to be unbiased. It is given as:

$$MBE = \frac{1}{N}\sum_{i=1}^{N} E, \tag{3.6}$$

where N is the number of data points. Mean absolute error (MAE) does not take into consideration the bias of a model. Instead, it computes the average absolute difference between the two quantities and is calculated as:

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |E|. \tag{3.7}$$

The root mean square error (RMSE) is a measure of the error between the two quantities. It is calculated as:

$$RMSE = \frac{1}{\sqrt{N}}\sqrt{\sum_{i=1}^{N} E^2}\ . \tag{3.8}$$

The standard deviation (σ) tells how close the values are to the average value.

$$\sigma = \frac{1}{\sqrt{N}}\sqrt{\sum_{i=1}^{N}(E - \bar{E})^2}. \tag{3.9}$$

The correlation coefficient (ρ) gives an indication of the degree of correlation between two quantities and it tests how close physically various parameters, considered by the four models, contribute to the correlation between independent variables (measured or calculated DNI and DHI) and dependent variables (modeled DHI and DNI). The higher the values of the (ρ) (positive correlation) means that the parameters described by the model have higher relevance to the independent variables. It is calculated as:

$$\rho = \frac{\sum_{i=1}^{N}(V_{predict} - \overline{V_{predict}}).(V_{measr} - \overline{V_{measr}})}{(\sigma_{predicted}).(\sigma_{measured})}, \tag{3.10}$$

where $V_{predict}$, $\overline{V_{predict}}$, V_{measr}, and $\overline{V_{measr}}$ represent predicted values, average predicted values, measured values, and average measured values, respectively. After applying various tests to filter data, the remaining data are compared with the modeled data, where the main input for the four models is the measured GHI. The other solar components are used as a reference when comparing the ground DNI and DHI with modeled DNI and DHI. The following chapter covers these aspects.

Chapter 4 Results

This chapter discusses the outcomes of this work and highlights the performance and ability of the models to predict DHI and DNI. Figures supported with numerical analyses are also included in this section Then, a comparison between ground data and modeled data is made in order to determine which model(s) are best representatives for any location.

4.1 Reindl* model

As previously mentioned, the Reindl* model tends to estimate the diffuse irradiance first. The direct irradiance can then be calculated knowing both GHI and estimated DHI values. The relationship between the clearness index and diffuse fraction can help identify different atmospheric changes. The diffuse fraction against the clearness index is plotted in Figure 4-1.

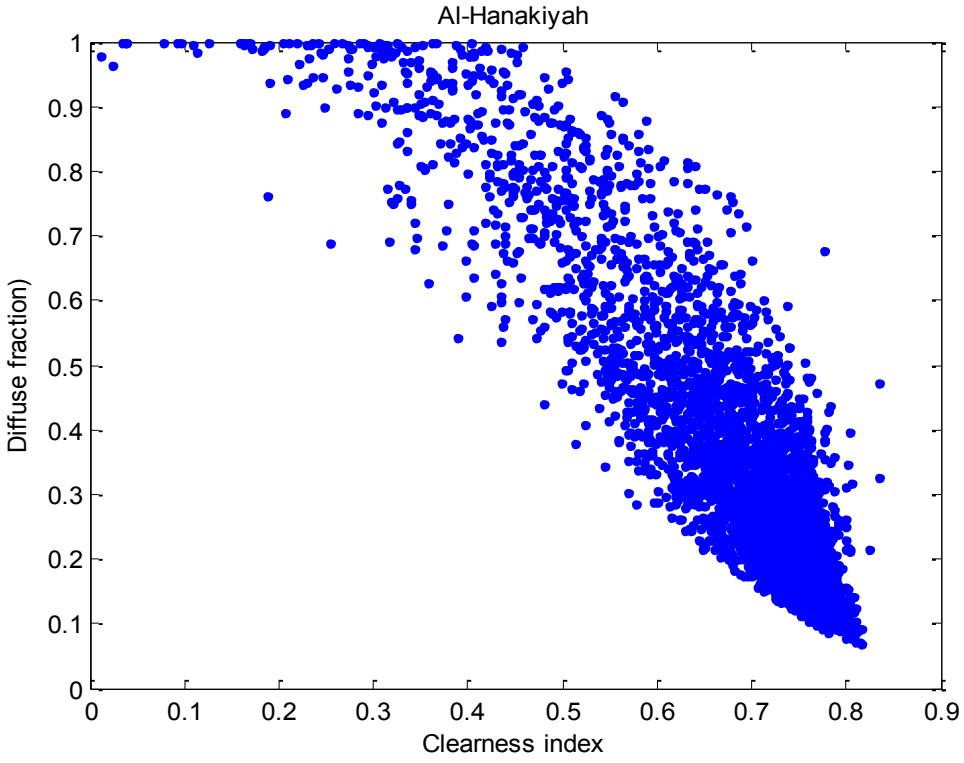


Figure 4-1 Scatter plot of the diffuse fraction against the clearness index for Al-Hanakiyah.

As Figure 4-1 shows, majority of the data points of the clearness index lie between 0.6 and 0.8. This means that overcast weather events are not common for Al-Hanakiyah. Very little scattering of data points is observed since SERI quality control ensures that outlier data

points are eliminated. The clear sky events are clustered when the clearness index ranges between 0.6 and 0.8. In this range, low values of the diffuse fraction are expected. Only a few higher diffuse fraction points are seen when the clearness index is above 0.8 which can be explained by the existence of dust particles in the atmosphere. Al-Hanakiyah station is located in a desert surrounded by mountains, as depicted in Figure 3-3, so scattering of the sunlight by dust particles is expected. Before evaluating the performance of the Reindl* model for the estimation of DNI, it is better to compare the modeled diffuse fraction against the calculated diffuse fraction (i.e. extracted from both measured GHI and DHI) obtained from Equation (2.21). The comparison is to detect the model's problems and to assess the capability of the model in generating modeled values that are close to the actual values. The values computed from measured GHI and measured or modeled DHI against the clearness index are illustrated in Figure 4-2. The red points represent the modeled diffuse fraction against the clearness index, and the blue points are the calculated diffuse fraction against the clearness index.

The Reindl* model assumes that the diffuse fraction cannot exceed one for low intervals of the clearness index (overcast events). This could be true for some locations; however, for other locations with a rapid occurrence of overcast events, the diffuse fraction can exceed one. The model uses a decreasing linear function in the low clearness index interval; however, for the range of the clearness index between 0 and 0.3, it is found that the model assumption is not correct since some data points of the calculated diffuse fraction equal unity while the modeled diffuse fraction keeps decreasing. The same issue is observed for the intermediate interval of the clearness index (0.3 to 0.78). The diffuse fraction decreases as a function of solar altitude and clearness index in the two intervals. This means that the model should include two or more parameters in order to reduce the error between the estimated and measured values. So, in low and intermediate clearness intervals, the Reindl* model tends to underestimate the diffuse fraction as well as the diffuse irradiance. For the higher clearness index range, the model assumes that the diffuse fraction is low and constant. This assumption is not always true. For example, Al-Hanakiyah station is located in a desert region which means that part of the sunlight will be scattered; the scattering depends on the amount and size of dust particles in the atmosphere. Also, the existence of clouds contributes

to the diffusion process, and this process cannot always be constant. It is obvious that the majority of the data points lie in the middle interval, as depicted in Table 4.1 and Figure 4-2.

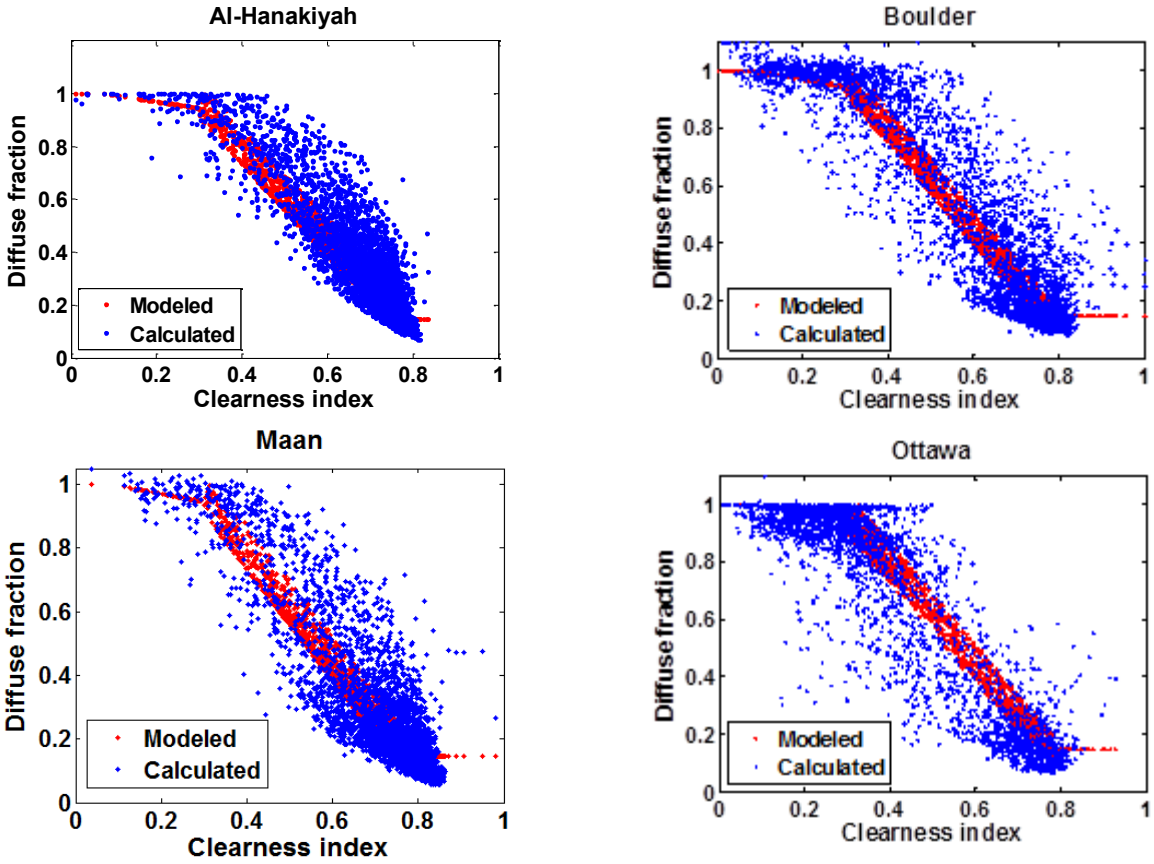


Figure 4-2 Scatter plot of the calculated and estimated diffuse fraction by the Reindl* model against the clearness index for the four locations.

Table 4.1 Number of available hourly data for each location according to the interval of the clearness index.

Location	Interval		
	$k_t \leq 0.3$ (Low interval)	$0.3 < k_t < 0.78$ (Intermediate interval)	$k_t \geq 0.78$ (High interval)
Al-Hanakiyah	64	3768	272
Boulder	699	2363	464
Ma'an	115	2809	1645
Ottawa	1056	1482	295

It can be seen that the model behavior is the same for all locations; nevertheless, the behavior of the measured data points varies for each location. It can be seen that similar behavior is observed for Boulder and Ottawa data. Since huge amount of data were removed from Ottawa dataset, the Reindl* model overstimates DHI compared to other locations. On ther other hand, similar behavior can also be seen for Al-Hanakiyah and Ma'an. The next step is to plot the scatter plot between the modeled diffuse irradiance and measured diffuse irradiance for each location to see how close the estimated values are to the actual values. Figure 4-3 depicts the estimated DHI versus the measured DHI. The cyan line represents the identity line and the green color is the fit line between the modeled and ground data, representing the best straight line between the independent and dependent variables. So, that means there is a poistive correlation betwwen measured/calculated and modeled quantities.

The scatter plot for all locations consists of three regions of the clearness index: blue (low interval), red (intermediate interval), and black (high interval). It is clearly seen from the identity line that model underestimates the DHI in Al-Hanakiyah, Boulder, and Ma'an. It is expected that the model may perform better in Boulder. One of the reasons is that the development of the model included two locations from the U.S.: Cape Canaveral and Albany. Since the majority of the data points lie in the intermediate interval, the model clearly underestimates DHI, as previously discussed for all locations. Also, by looking at the maximum values of both measured and estimated DHI, it is found the model cannot estimate values above 400 W/m^2 while the measured DHI can reach values higher than 700 W/m^2.

On the other hand, the model overestimates DNI for Al-Hanakiyah and Ma'an as depicted in Figure 4-4. The model performs better in Boulder since it shows a better correlation between measured and estimated DHI values. So, it is expected that the Reindl* model may estimate DNI with lower bias error as the identity and fitting lines are close to each other.The overestimation of DNI in the three locations could be due to the underestimation of DHI values mostly in the intermediate values. In higher clear sky events, some points are overestimated due to the condition set in this interval (i.e. $k_d = 0.147$). Moreover, many of the data points of DHI are underestimated so that many of the DNI values are overestimated

as a consequence except for Ottawa where many DNI values are understimated as depicted in Figure 4-4.

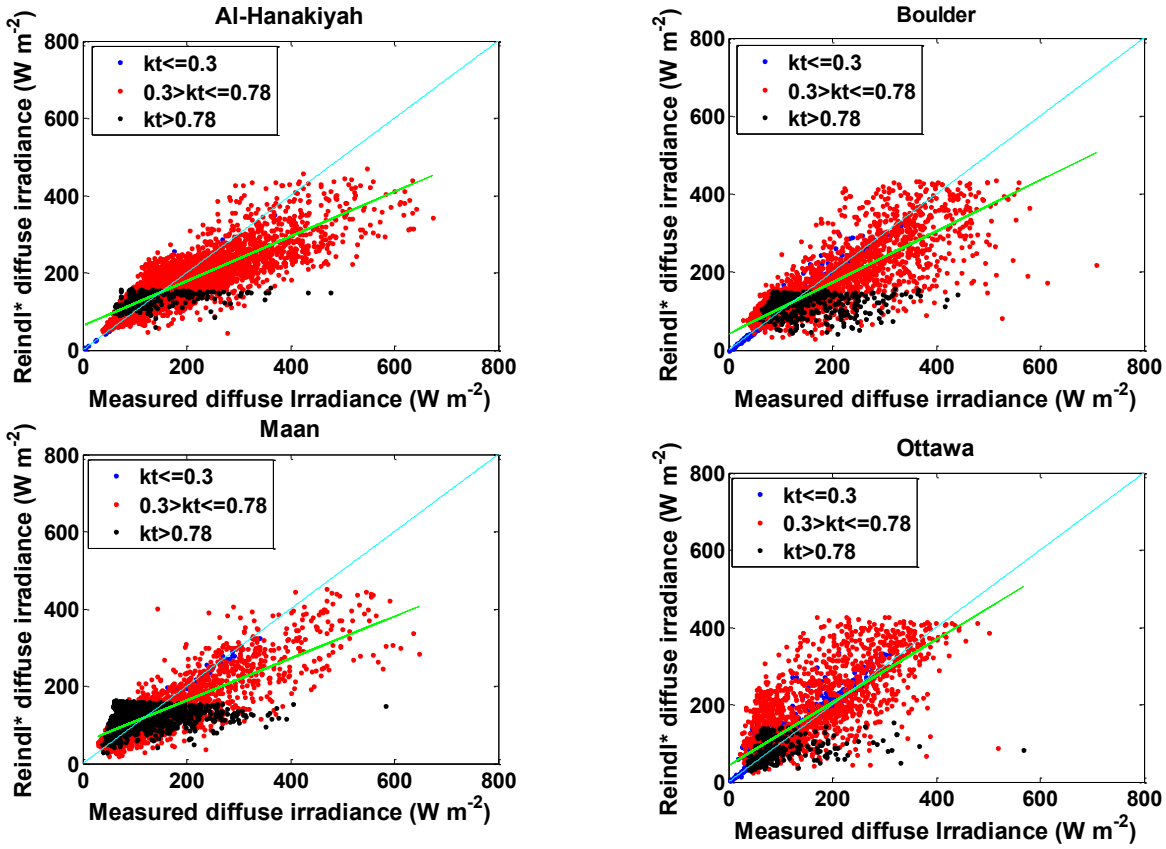


Figure 4-3 Scatter plot of the estimated DHI by the Reindl* model against the ground diffuse data. The cyan line is the identity line and the green line is the fit line between the estimated and measured or calculated data.

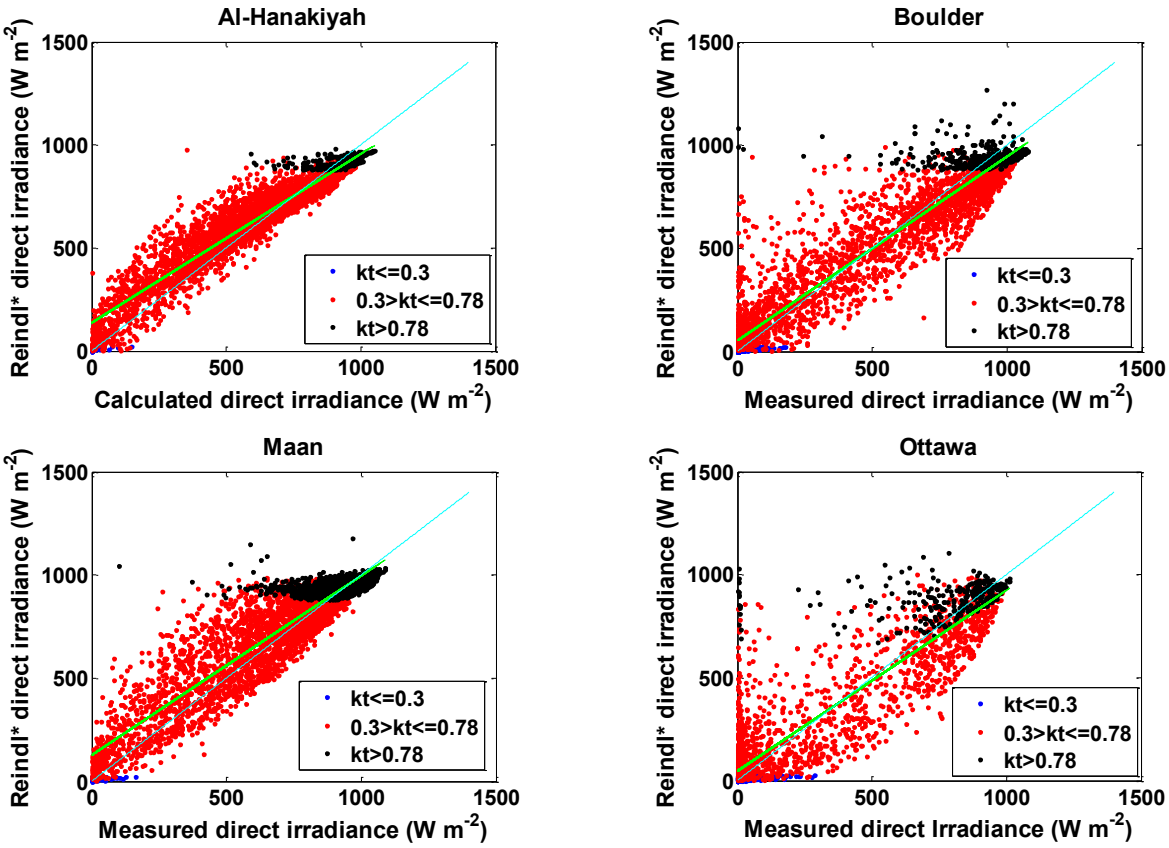


Figure 4-4 Scatter plot of the estimated DNI by the Reindl* model against the ground direct data. The cyan line is the identity line and the green line is the fit line between the estimated and measured or calculated data.

4.2 BRL model

Like the Reindl* model, the BRL model estimates diffuse irradiance so the direct irradiance can be calculated. The model depends on a logistic function; the behavior of the modeled data is the same for the four locations. The range of the estimated diffuse fraction ranges from 0 to 1 except for Al-Hanakiyah since the maximum value of the clearness index is 0.84, as shown in Figure 4-5. The modeled diffuse fraction shows a reasonable fit with the measured diffuse fraction for Boulder; however, it shows a poor fit for Al-Hanakiyah, Ma'an and Ottawa data. For a high clear sky interval, the model does not perform well, especially for Boulder. Moreover, the modeled diffuse fraction greatly underestimates the calculated diffuse fraction. So, as with the Reindl* model, it assumes the atmosphere is very clear with low diffusivity of sunlight, which is not always the case. Also, the same issue observed with the Reindl* model in the low clearness index interval is apparent here. The BRL model

assumes that diffuse irradiance is dominant in this region. Reflecting the behavior of the modeled diffuse fraction on the scatter plot between the measured and the modeled DHI data, it is found that the scattering of data points is very common in overcast events (high DHI values) since the model underestimates DHI in this area. Less scattering is observed in high clearness index values for Al-Hanakiyah since the maximum clearness value does not exceed 0.84. Figure 4-6 illustrates the scatter plot between measured DHI and modeled DHI. The green line is the fit line while the cyan line represents the identity line between the two quantities.

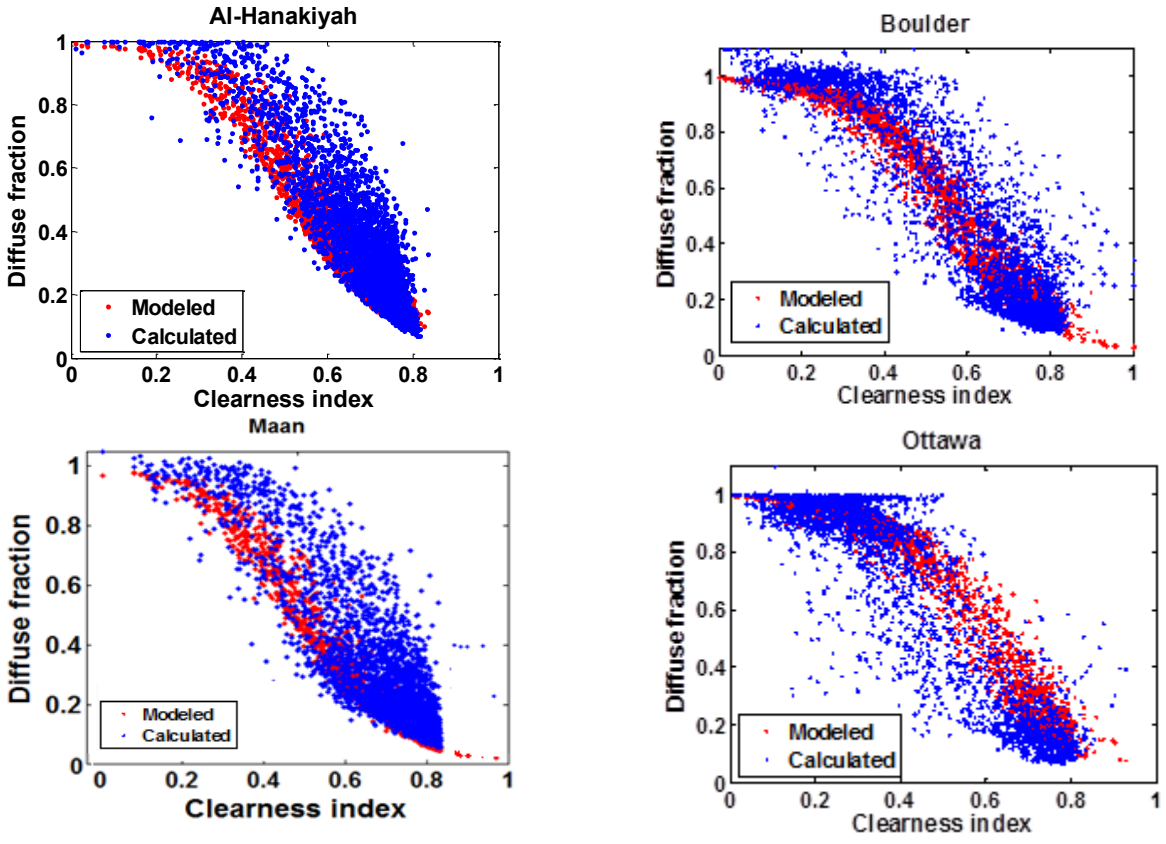


Figure 4-5 Scatter plot of the calculated and estimated diffuse fraction by the BRL model against the clearness index for the four locations.

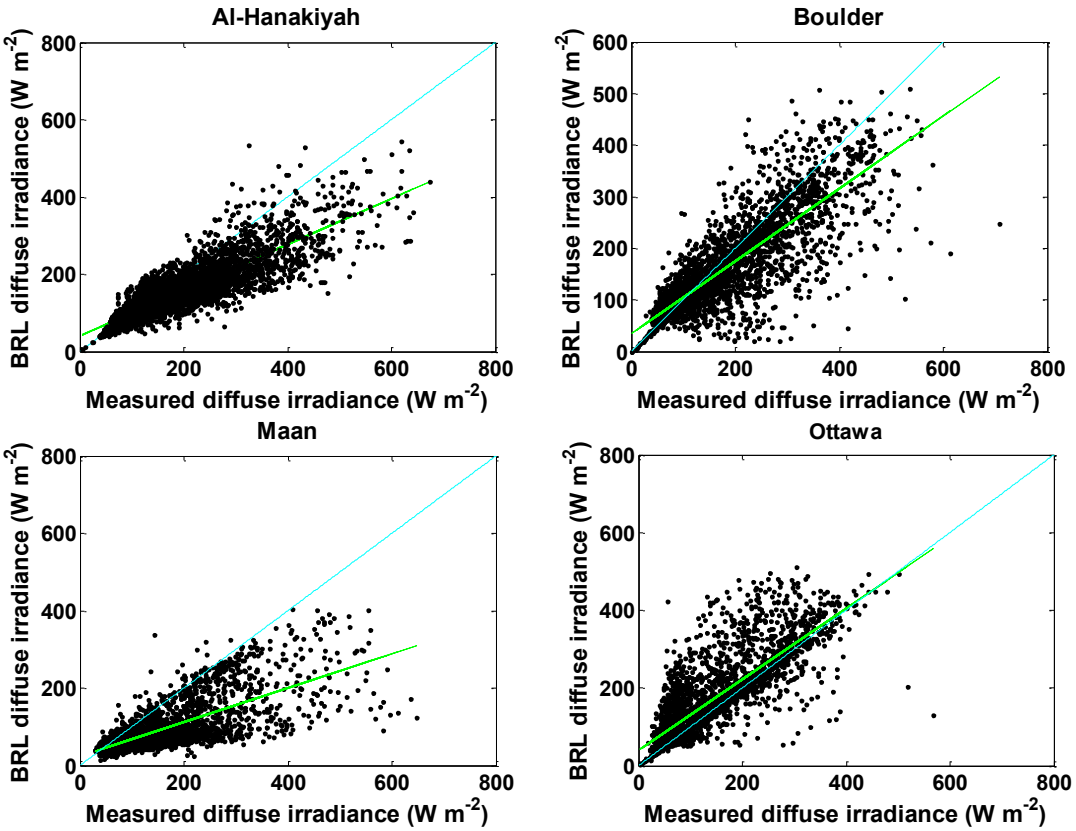


Figure 4-6 Scatter plot of the estimated DHI by the BRL model against the ground diffuse data. The cyan line is the identity line and the green line is the fit line between estimated and measured or calculated data.

The modeling of DNI for Boulder and Ottawa data shows less bias compared to the modeling of DNI for Al-Hanakiyah and Ma'an. Many scattering points are observed where some of them exceed 1000 W/m^2 for Ma'an in high clear sky intervals, as shown in Figure 4-7. This effect is also seen in Boulder but with fewer data points. This is because the number of modeled diffuse fraction points in Ma'an is higher compared to Boulder. Both Reindl* and BRL models perform poorly in the higher interval of the clearness index. The issue with these models represented in this work is that most of them have been developed according to stations located in Europe or North America. New locations, especially in the Middle East region, should be included in models in order to develop a broad model that may fit for any location.

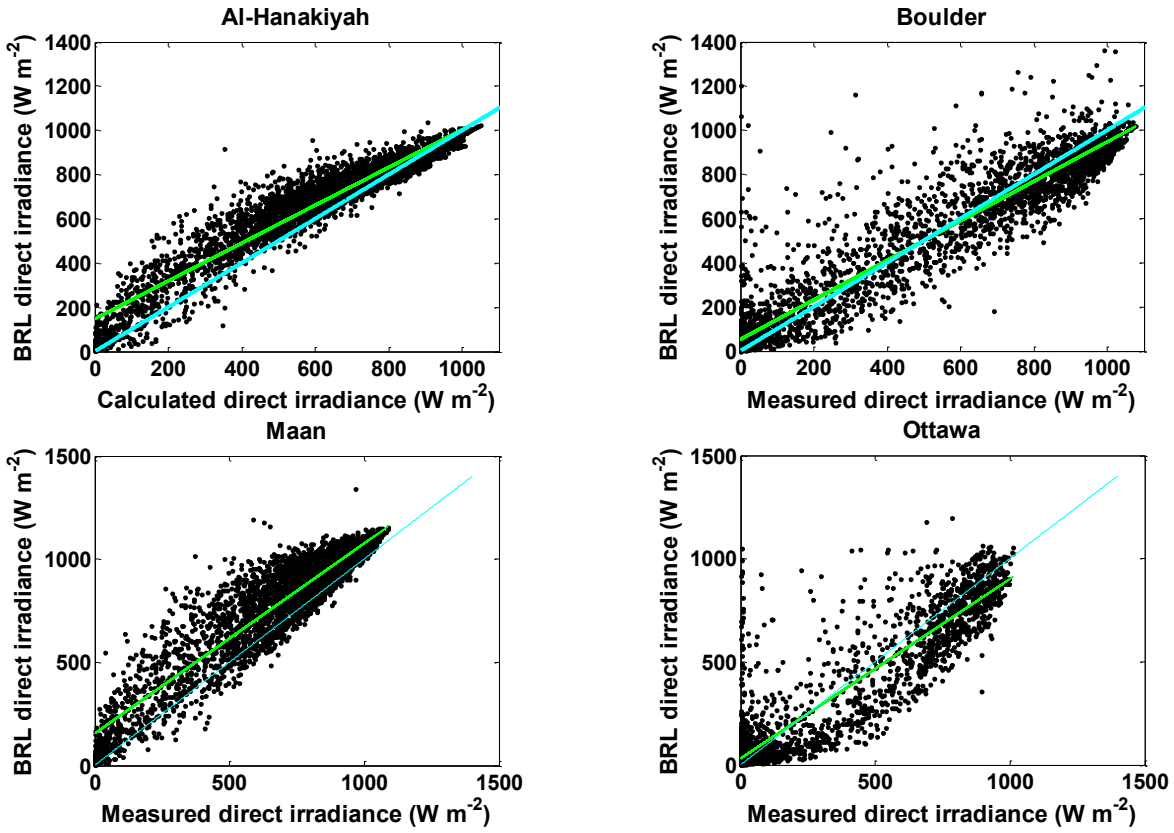


Figure 4-7 Scatter plot of the estimated DNI by the BRL model against the ground direct data. The cyan line is the identity line and the green line is the fit line between the estimated and measured or calculated data.

4.3 DISC model

The DISC model estimates DNI directly compared to the Reindl* and BRL models. The modeled diffuse fraction shows a good agreement compared to the other models. This is clearly seen in Boulder graph, as depicted in Figure 4-8. So, it is expected that the modeling of DHI and DNI will not reach higher irradiance values when the condition of the sky is clear. However, part of the modeled diffuse fraction does not merge well with the calculated diffuse fraction, as with Al-Hanakiyah data. This can be attributed to the amount of dust in the atmosphere, which increases the diffusivity of the sunlight. Scattering of high DHI values is observed for all locations, as illustrated in Figure 4-9.

The modeled DNI values do not exceed 1000 W/m^2 in high clearness index values, as shown in Figure 4-10. As previously mentioned, the BRL model assumes decaying change in the

diffusivity of the sunlight, and the Reindl* model assumes constant change in the diffused light.

4.4 Perez model

Figure 4-13 shows the modeled DNI using the Perez model against measured or calculated DNI. The behavior of the scatter plots are similar to the one shown in Figure 4-10 except that some values are corrected using the four-dimensional matrix. The Perez model shows moderately good fit for high clearness index values. Nonetheless, the modeled DHI data for Ottawa show a high increase in the high clearness index interval compared to calculated diffuse fraction, as shown in Figure 4-11. The model also shows a limitation for the intermediate clearness index interval, which is a common issue among all models.

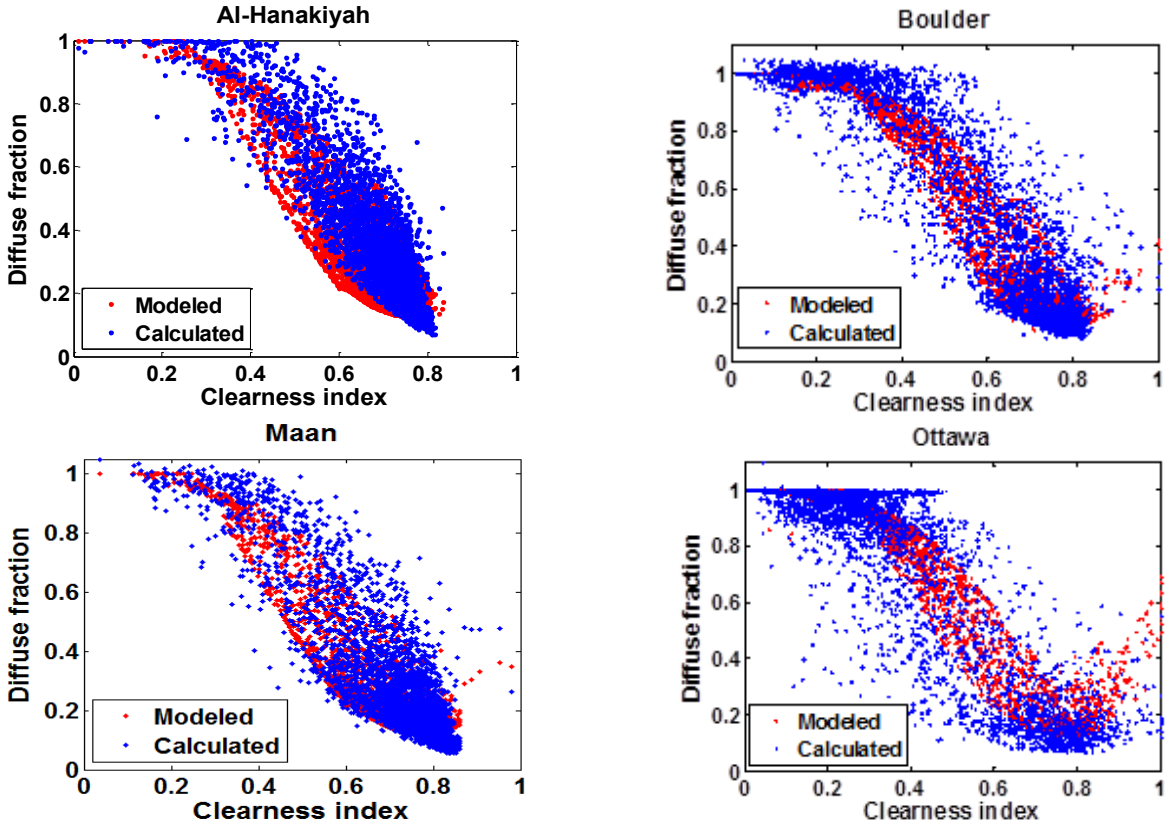


Figure 4-8 Scatter plot of the calculated and estimated diffuse fraction by the DISC model against the clearness index for the four locations.

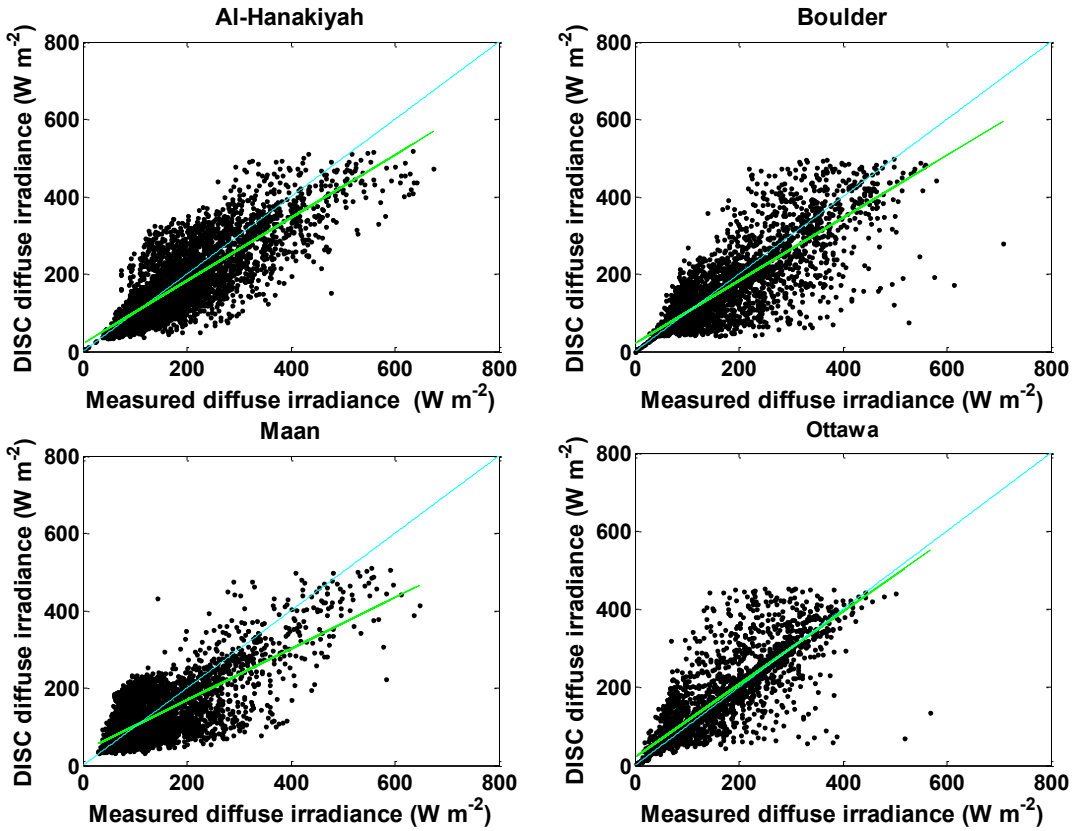


Figure 4-9 Scatter plot of the estimated DHI by the DISC model against the ground diffuse data. The cyan line is the identity line and the green line is the fit line between estimated and measured or calculated data.

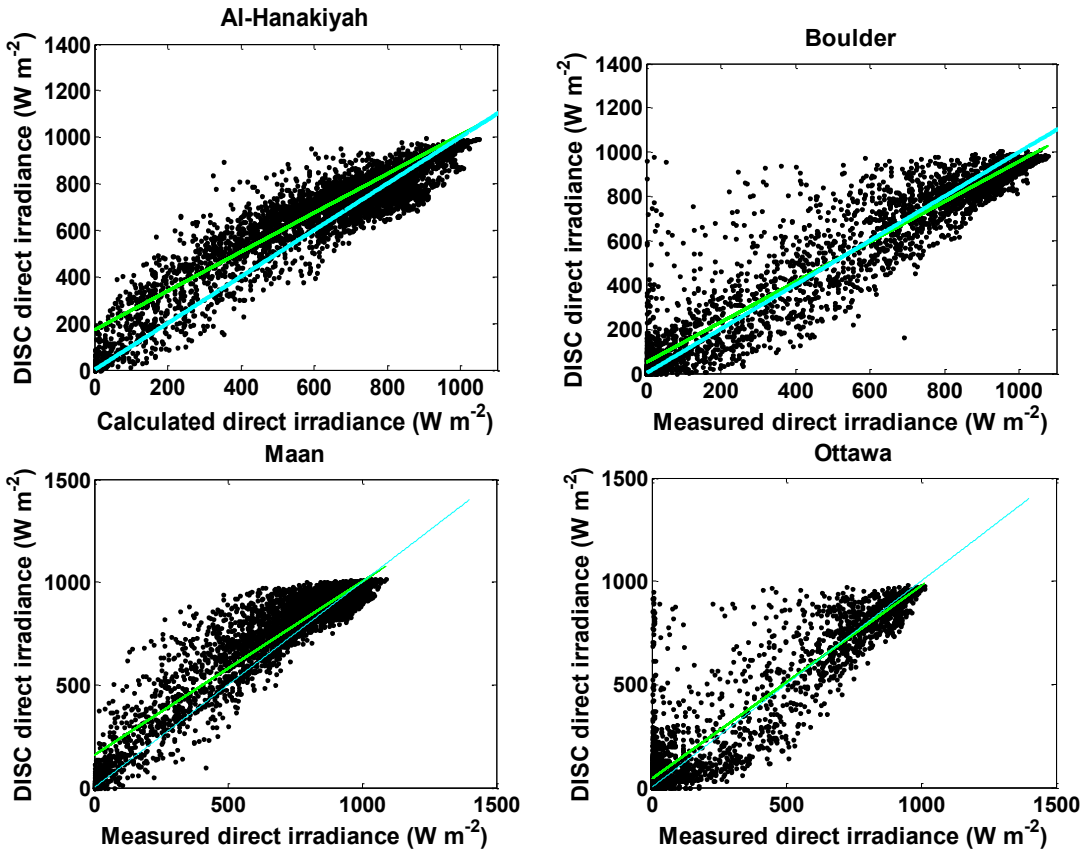
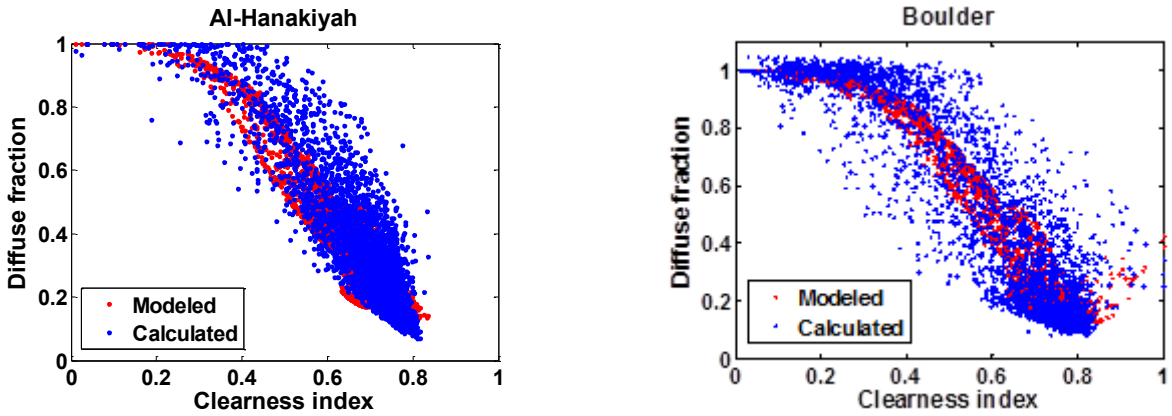


Figure 4-10 Scatter plot of the estimated DNI by the DISC model against the ground direct data. The cyan line is the identity line and the green line is the fit line between the estimated and measured or calculated data.



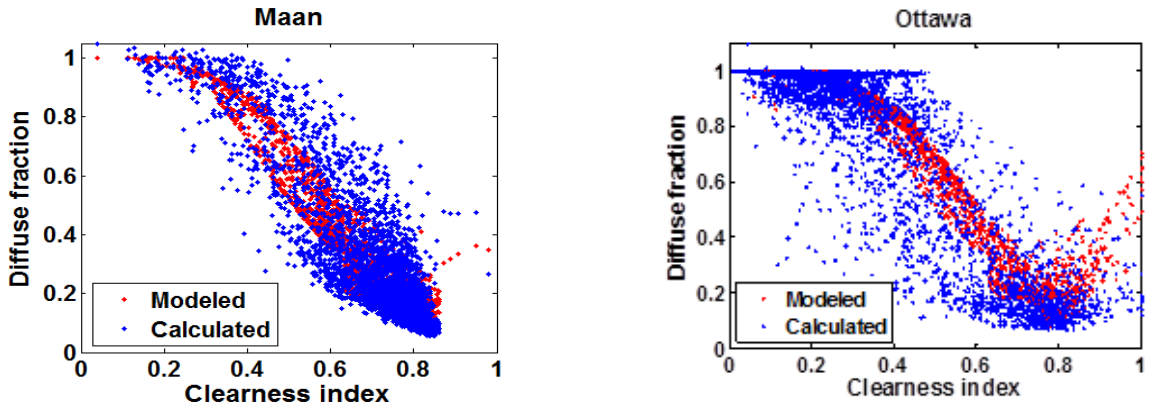


Figure 4-11 Scatter plot of the calculated and estimated diffuse fraction by the Perez model against the clearness index for the four locations.

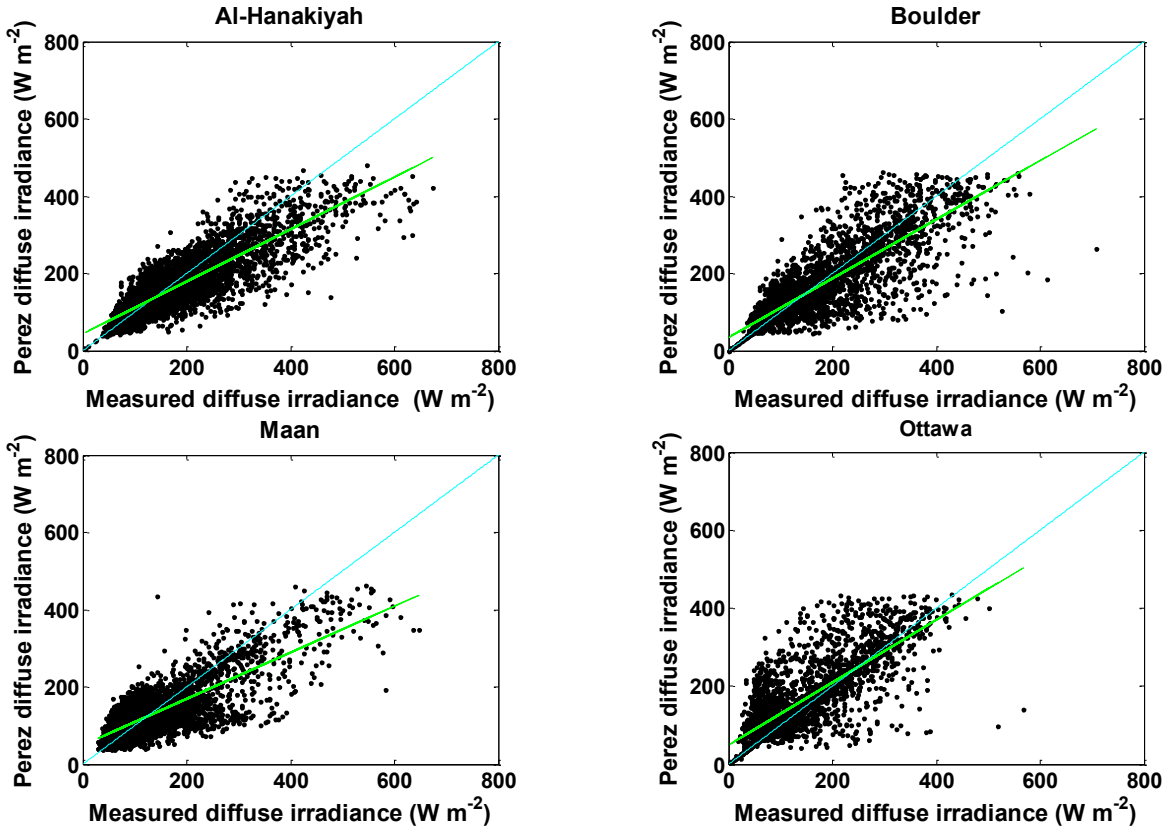


Figure 4-12 Scatter plot of the estimated DHI by the Perez model against the ground diffuse data. The cyan line is the identity line and the green line is the fit line between the estimated and measured or calculated data.

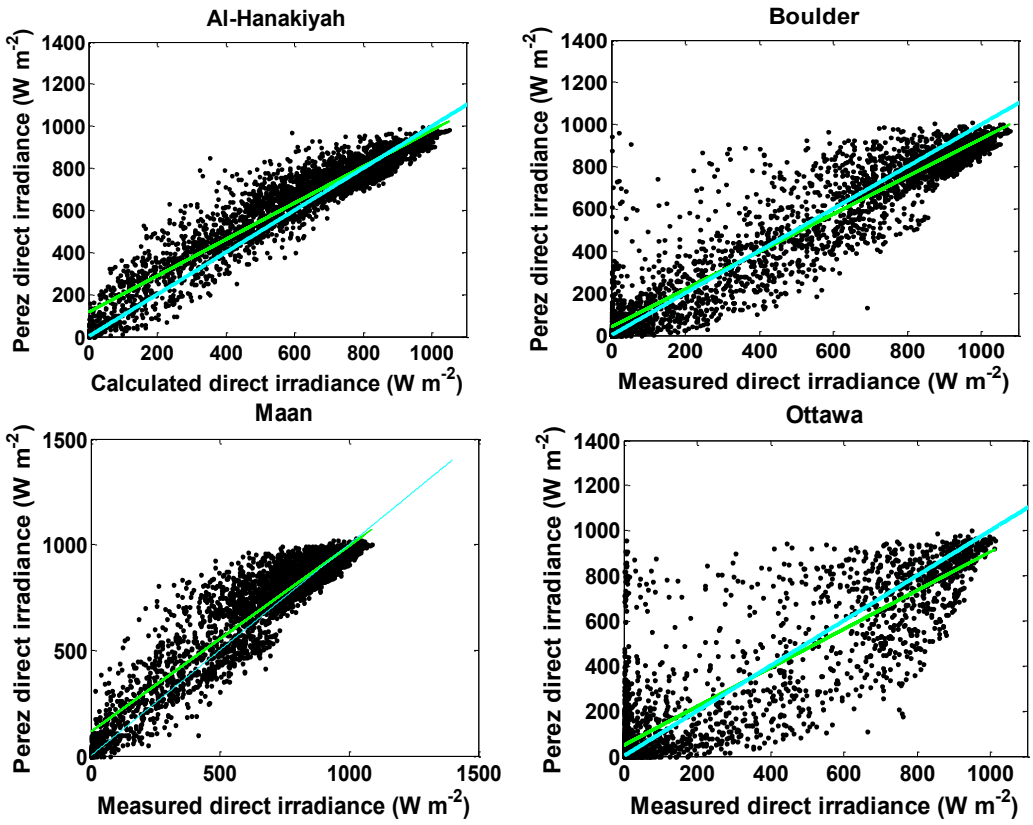


Figure 4-13 Scatter plot of the estimated DNI by the Perez model against the ground direct data. The cyan line is the identity line and the green line is the fit line between the estimated and measured or calculated data.

Table 4.2 Error analysis of each model for Al-Hanakiyah. The mean value of the calculated DNI is 618 W/m^2. The mean value of the measured DHI is 188 W/m^2.

Al-Hanakiyah (DHI)										
	(W/m^2)						Relative (%)			
Model	**Mean**	**MBE**	**MAE**	**RMSE**	**σ**	**ρ**	**MBE**	**MAE**	**RMSE**	**σ**
Reindl*	172	16	40	59	57	0.85	8.5	21.3	31.4	30.3
BRL	152	36	44	65	54	0.86	19.1	23.4	34.6	28.7
DISC	173	15	47	62	61	0.82	8.0	25	33.0	32.4
Perez	171	17	39	56	54	0.85	9.0	20.7	29.8	28.7

Al-Hanakiyah (DNI)										
	(W/m^2)						Relative (%)			
Model	**Mean**	**MBE**	**MAE**	**RMSE**	**σ**	**ρ**	**MBE**	**MAE**	**RMSE**	**σ**
Reindl*	646	-28	63	86	82	0.95	-4.5	10.2	13.9	13.3
BRL	677	-59	71	97	77	0.96	-9.6	11.5	15.7	12.5
DISC	662	-44	81	106	97	0.93	-7.1	13.1	17.2	15.7
Perez	651	-33	64	87	80	0.95	-5.3	10.4	14.1	13.0

Table 4.3 Error analysis of each model for Boulder. The mean value of the measured DNI is 471 W/m^2. The mean value of the measured DHI is 163 W/m^2.

Boulder (DHI)										
	(W/m^2)						Relative (%)			
Model	**Mean**	**MBE**	**MAE**	**RMSE**	**σ**	**ρ**	**MBE**	**MAE**	**RMSE**	**σ**
Reindl*	149	14	40	63	61	0.83	8.6	24.5	38.7	37.4
BRL	149	14	38	61	**59**	0.84	8.6	23.3	37.4	36.2
DISC	153	10	39	63	63	0.84	6.1	23.9	38.7	38.7
Perez	159	4	38	59	59	0.85	2.5	23.3	36.2	36.2

Boulder (DNI)										
	(W/m^2)						Relative (%)			
Model	**Mean**	**MBE**	**MAE**	**RMSE**	**σ**	**ρ**	**MBE**	**MAE**	**RMSE**	**σ**
Reindl*	474	-3	84	126	126	0.95	-0.6	17.8	26.8	26.8
BRL	475	-4	78	121	121	0.95	-0.9	16.6	25.7	25.7
DISC	478	-7	75	119	119	0.95	-1.5	16.0	25.3	25.3
Perez	460	11	78	118	118	0.95	2.3	16.6	25.1	25.1

Table 4.4 Error analysis of each model for Ma'an. The mean value of the measured DNI is 709 W/m^2. The mean value of the measured DHI is 133 W/m^2.

Ma'an (DHI)										
	(W/m^2)						Relative (%)			
Model	**Mean**	**MBE**	**MAE**	**RMSE**	**σ**	**ρ**	**MBE**	**MAE**	**RMSE**	**σ**
Reindl*	127	6	38	54	**54**	0.8	4.5	28.6	40.6	40.6
BRL	83	50	52	78	60	0.74	37.6	39.1	58.6	45.1
DISC	125	8	42	59	59	0.76	6.0	31.6	44.4	44.4
Perez	129	4	38	54	54	0.79	3	28.6	40.6	40.6

Ma'an (DNI)										
	(W/m^2)						Relative (%)			
Model	**Mean**	**MBE**	**MAE**	**RMSE**	**σ**	**ρ**	**MBE**	**MAE**	**RMSE**	**σ**
Reindl*	741	-32	74	105	100	0.93	-4.5	10.4	14.8	14.1
BRL	809	-100	107	140	98	0.94	-14.1	15.1	19.7	13.8
DISC	757	-48	80	114	103	0.93	-6.8	11.3	16.0	14.5
Perez	738	-29	70	101	96	0.94	-4.1	9.9	14.2	13.5

Table 4.5 Error analysis of each model for Ottawa. The mean value of the measured DNI is 266 W/m^2. The mean value of the measured DHI is 146 W/m^2.

Ottawa (DHI)										
	(W/m^2)						Relative (%)			
Model	**Mean**	**MBE**	**MAE**	**RMSE**	**σ**	**ρ**	**MBE**	**MAE**	**RMSE**	**σ**
Reindl*	163	-17	34	56	54	0.82	-11.6	23.3	38.4	37.0
BRL	174	-27	40	67	62	0.79	-18.5	27.4	45.9	42.5
DISC	163	-17	42	67	65	0.75	-11.6	28.8	45.9	44.5
Perez	167	-21	40	64	61	0.78	-14.4	27.4	43.8	41.8

Ottawa (DNI)										
	(W/m^2)						Relative (%)			
Model	**Mean**	**MBE**	**MAE**	**RMSE**	**σ**	**ρ**	**MBE**	**MAE**	**RMSE**	**σ**
Reindl*	279	-13	74	130	129	0.93	-4.9	27.8	48.9	48.5
BRL	262	4	78	134	134	0.92	1.5	29.3	50.4	50.4
DISC	289	-23	98	172	170	0.88	-8.6	36.8	64.7	63.9
Perez	277	-11	93	165	164	0.88	-4.1	34.9	62.0	61.7

It is important to highlight some points related to the removal of data points from Boulder and Ottawa datasets. First, the June data for Boulder were removed because of the unexpected behavior depicted in Figure 3-1. Also, many timestamps for Ottawa were removed since data were not available. This influenced the values presented in Table 4.3 and Table 4.5. For example, the removal of many data points affected the mean values of DHI and DNI for Boulder and Ottawa. This suggests that the weather conditions, especially in Ottawa are mostly cloudy which is not always true year round since the radiation in the spring and summer seasons have higher irradiance values. The second point is that large amount of data were removed from the Ottawa dataset. This will lead to matching between solar radiation data and weather data. As a result, part of the meteorological data is eliminated as well.

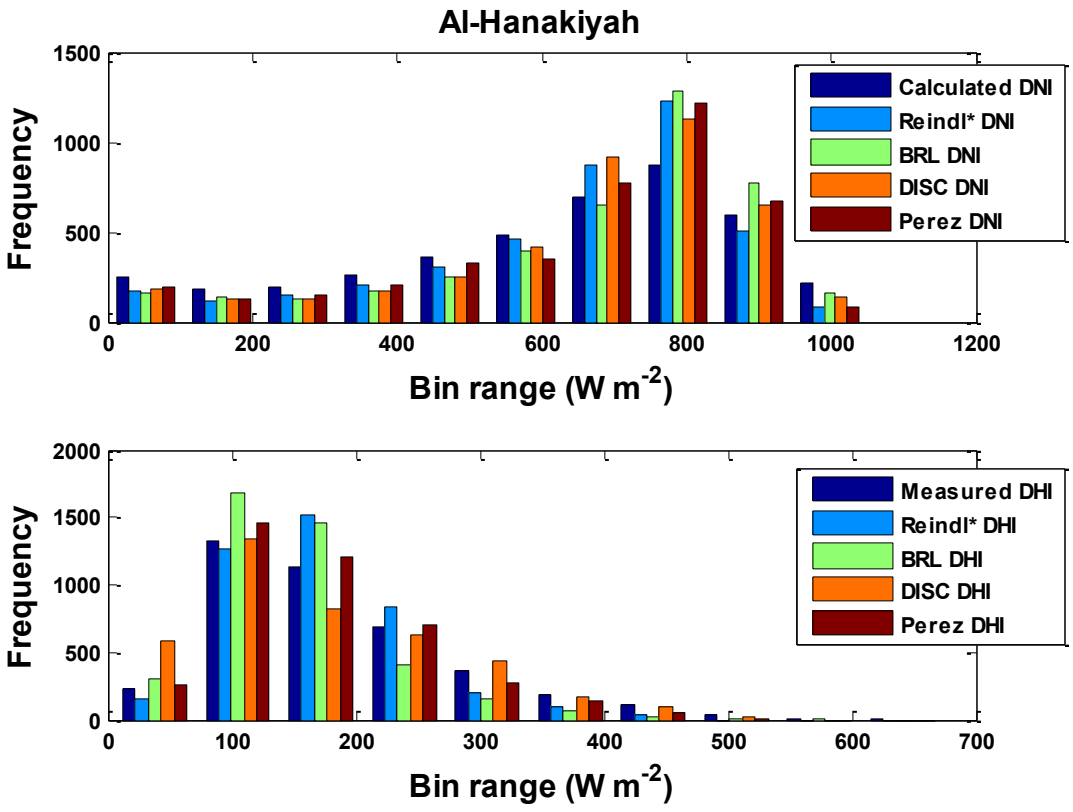
Having applied the metrics to the four locations, it is evident that the four models underestimate the measured DHI for Al-Hanakiyah. The BRL model has a higher bias of 36 W/m^2 compared to the other models. Since the modeled DHI values are lower than the real values, it is expected that the modeled DNI values are high compared to the measured values. The closest modeled mean value is obtained by the DISC (173 W/m^2) and the Reindl* (172 W/m^2) models for DHI. The Reindl* model has a mean value of 646 W/m^2 for DNI, as seen in Table 4.2. The metrics computed for the Reindl* model (i.e. MAE, RMSE) have values for DHI of 40 W/m^2 (21.3 %) and 59 W/m^2 (31.4 %), respectively. The Reindl* and Perez models show similar metrics values for modeled DHI and DNI. For Boulder, it is expected that the DISC and Perez models will outperform other models since the DISC was developed at Boulder using data from various U.S. cities, and the Perez model is a corrected version of the DISC model. The results show that the Perez model has the lowest bias error of 4 W/m^2 (2.5%), RMSE and σ of 59 W/m^2 (36.2%) when modeling DHI. The Reindl* and DISC models exhibit similar metrics error in terms of MAE and RMSE for DHI. The Perez and BRL models have similar performance in comparison to the other models for estimating DHI values. For modeled DNI data, the DISC and Perez models have better estimation abilities compared to the BRL and Reindl* models regarding RMSE and σ. Unlike the Perez model, the other three models tend to overestimate the DNI. For Ma'an, the BRL model shows higher error outcomes for DHI and DNI modeled data, as shown in Table 4.4. The same behavior is observed for Al-Hanakiyah. This could be interpreted due to the variation

between the actual or modeled data of the diffuse fraction against the clearness index, as shown in Figure 4-5. The Riendl* and Perez models outperform the other models when estimating DHI. Also, the metrics computed for the Reindl* and Perez models are similar in terms of the computation of the DNI values. The two models perform better in Al-Hanakiyah, which suggests that the two models work better in locations where the availability of high solar potential is present. Ottawa is a special case since large loss of data has produced significant deviation between measured and modeled data. For this location, the Reindl* model produces better metrics values compared to the other three models. For modeled DHI and DNI data, the Reindl* model shows an MAE of 34 and 74 W/m^2 respectively, an RMSE of 56 and 130 W/m^2 respectively, and σ of 54 and 129 W/m^2 respectively. The other models may perform better than this model but, the huge loss of data greatly affected the metrics used to calculate the error between the real and modeled values.

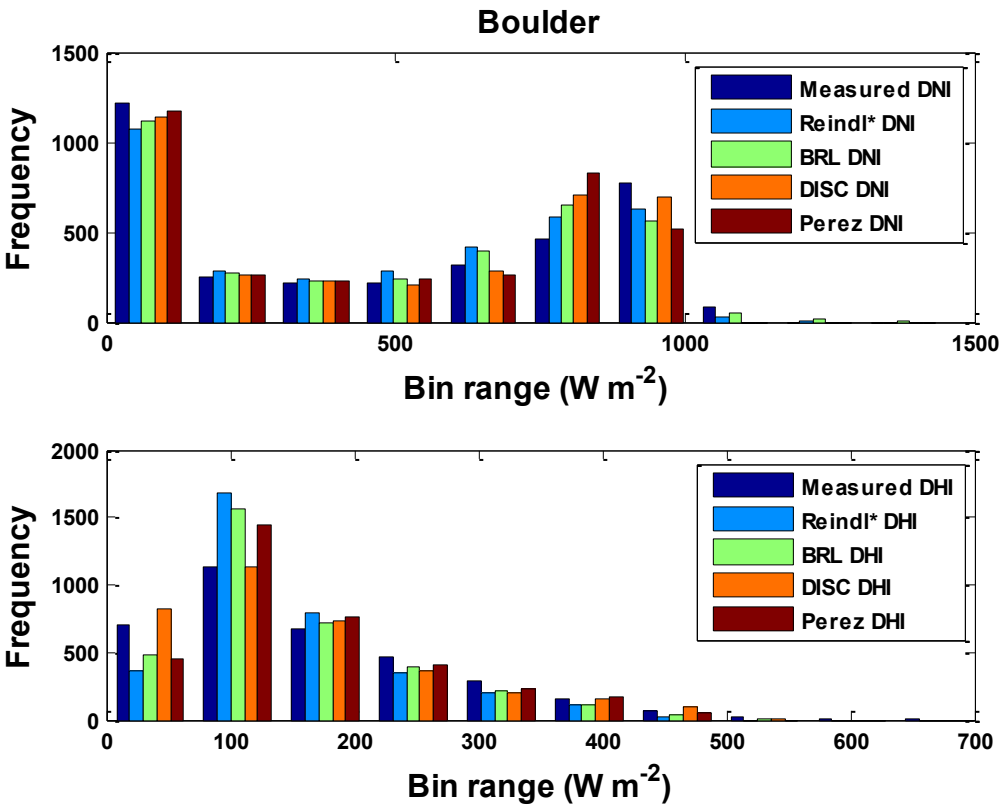
The histogram shown in Figure 4-14 depicts the ability of each model to estimate the desired component and examines how close the estimated value is from the real value depending on the number of bins. The *x*-axis represents solar radiation values which are divided into multiple bin ranges and the *y*-axis represents the number of times of occurrence of both modeled and real values of DNI and DHI for the four locations. As expected for Al-Hanakiyah location, the solar radiation potential is high. Moreover, the availability of high DNI values above 630 W/m^2 exceeds 694 hours which is high compared to the DNI values below 600 W/m^2. As the solar radiation exceeds 800 W/m^2 and approaches 1000 W/m^2, the occurrence of high DNI values decreases. The four models tend to produce lower values than the actual ones when the solar radiation is below 630 W/m^2. The BRL and Reindl* models deviate more when the sky condition has a low amount of cloud cover in comparison to the DISC and Perez models. The histogram of the DHI values illustrates that the weather conditions are usually clear although the station is located in a desert area. This could be due to the frequent cleaning events which help minimize the soiling effects on the instruments' sensors. In general, the Perez model seems to provide somewhat better estimates of DHI values between 200 W/m^2 and 339 W/m^2.

After the removal of the June data from the Boulder dataset, overcast events are common for the values of DNI below 145 W/m^2 as well as for high clear sky events with a period

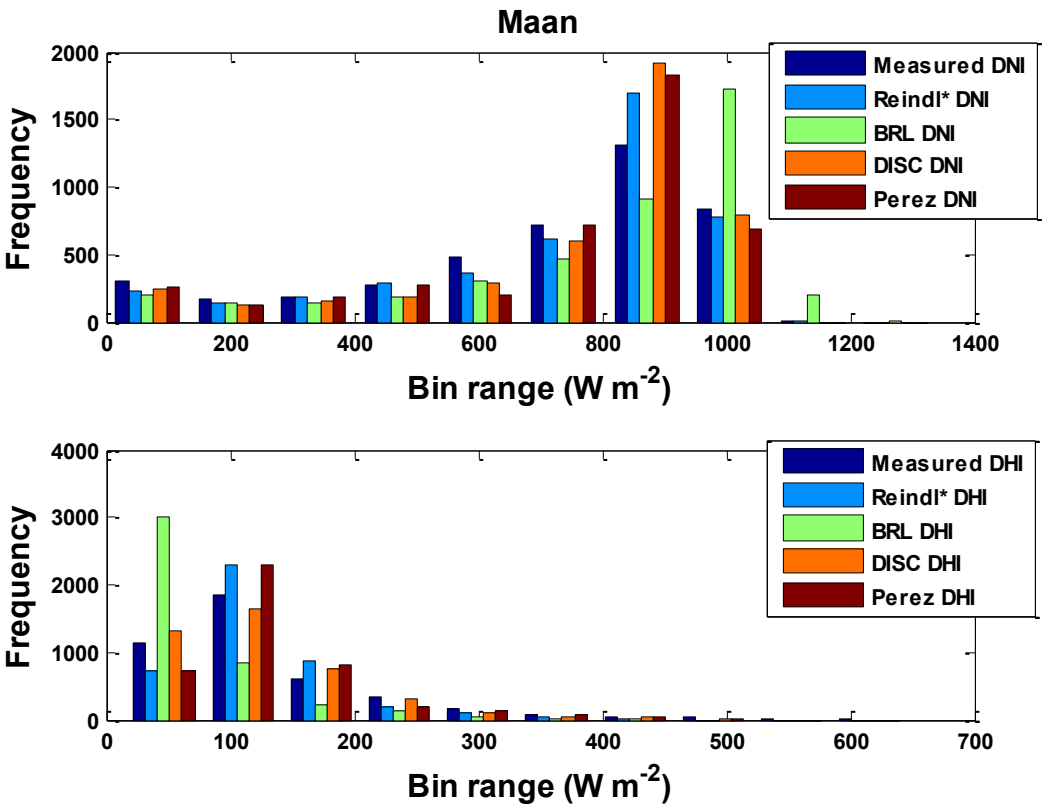
exceeding 760 hours. All models seem to provide close estimates of DNI values below 580 W/m^2. The DISC and Perez models are better for high clearness index values. However, for DNI values above 1000 W/m^2, the DISC and Perez models do not provide good estimates. In other words, the two models do not have the ability to estimate DNI values above 1000 W/m^2 in comparison to the BRL and Reindl* models. Nonetheless, the DISC model is the best model among the four models to provide close values for most of the ranges of DHI. By looking at Figure 4-14 (a) and (c), it can be seen that both Al-Hanakiyah and Ma'an have a higher availability of high solar radiation. The Reindl* model performs better in the range of 268 to 670 W/m^2. As solar irradiance increases from 938 to 1070 W/m^2, the Reindl*, DISC, and Perez models provide good and close estimates of DNI values, while the BRL model greatly overestimates DNI. The deviation of modeled DHI from the ground-truth values, as calculated by the DISC model, is not as large compared to the BRL and Reindl* models for DHI values less than 521 W/m^2. The elimination of Ottawa data due to vandalism causes a loss of valuable high solar radiation data. When DNI values are below 359 W/m^2, the BRL model exhibits less error compared to other models. Most of the high solar potential hours do not exceed 300 hours. The Reindl* model is better for partly cloudy conditions in estimating DNI values. All models tend to overestimate DHI values above 341 W/m^2. On the other hand, the models underestimate DHI below 59.5 W/m^2. Otherwise, the DISC model has lower estimation error compared to the Perez model.


Al-Hanakiyah
Frequency
Bin range (W m-2)
Calculated DNI
Reindl* DNI
BRL DNI
DISC DNI
Perez DNI
Measured DHI
Reindl* DHI
BRL DHI
DISC DHI
Perez DHI

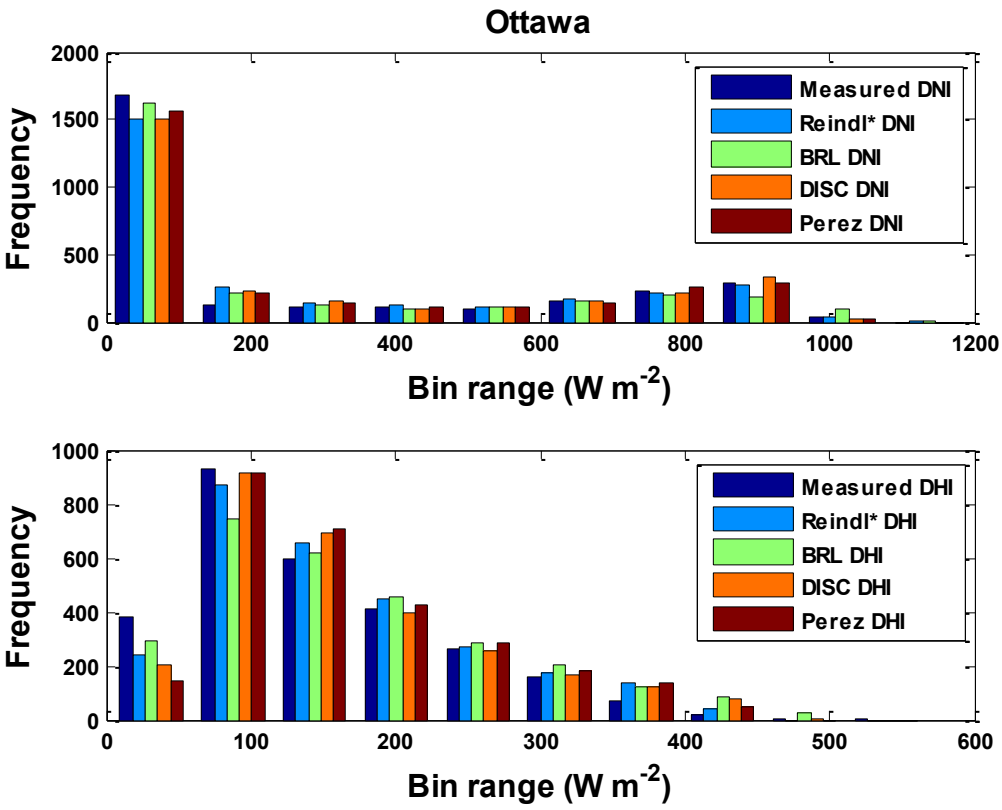
(a)


Boulder
Frequency
Bin range (W m-2)
Measured DNI
Reindl* DNI
BRL DNI
DISC DNI
Perez DNI
0
500
1000
1500
Frequency
Bin range (W m-2)
Measured DHI
Reindl* DHI
BRL DHI
DISC DHI
Perez DHI
0
100
200
300
400
500
600
700
2000

(b)


Maan
Frequency
Bin range (W m-2)
Measured DNI
Reindl* DNI
BRL DNI
DISC DNI
Perez DNI
Measured DHI
Reindl* DHI
BRL DHI
DISC DHI
Perez DHI

(c)



(d)

Figure 4-14 Comparison of the occurrence of calculated or measured values and modeled values for (a) Al-Hanakiyah, (b) Boulder, (c) Ma'an and (d) Ottawa. The *x*-axis represents the range of bins used to divide the solar irradiance values and the *y*-axis shows the frequency of values. The top part shows the comparison for DNI values and the bottom part represents the comparison of DHI values.

4.5 Importance of potential solar insolation

The generation of power is necessary for producing sufficient electricity to meet market and household needs. As previously highlighted, a solar resource assessment for the deployment of solar plants is required. CPV systems can make use of the high DNI values for each month in Ma'an so that the incoming radiation can be more concentrated to enhance the efficiency. The data were taken from Deutsches Zentrum für Luft and Raumfahrt (DLR) German Aerospace Center Institute of Solar Research [47]. Also, it is observed that PV systems can produce good power reaching up to 268 kWh/m^2 since the DHI represents the lowest portion with a maximum insolation of 64 kWh/m^2 compared to GHI and DNI, as depicted in Figure 4-15. For the whole period of (2014/01/01 to 2014/12/31), DNI represents the largest portion of the GHI, so the situation is good for PV and CPV systems. High availability of DNI with a maximum insolation of up to 316 kWh/m^2 can be generated when multiple CPV systems are installed; special attention must be given to the high cost of these systems. This reinforces the statement that the solar potential available in Ma'an is high and promising. Hence, having highly accurate data accompanied by highly standardized PV and CPV systems is optimal for the generation of desired electrical power.

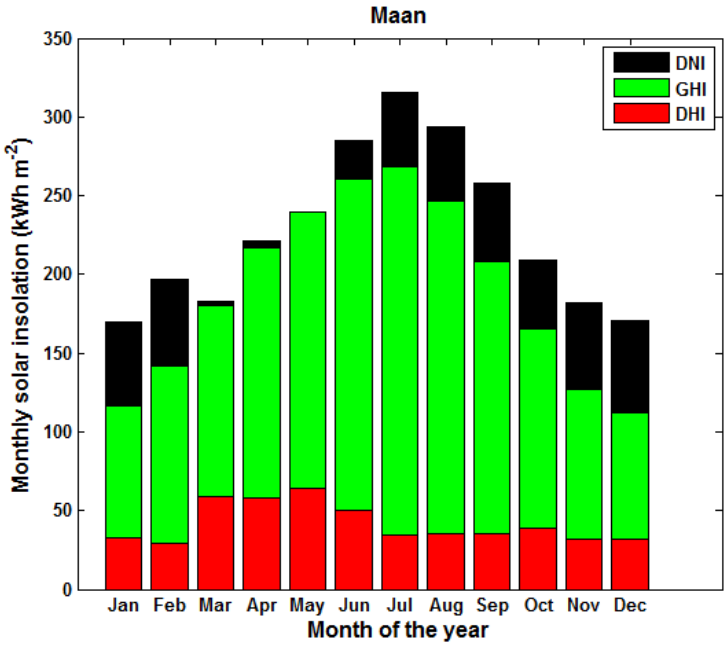


Figure 4-15 Monthly solar insolation in Ma'an from 2014/01/01 to 2014/12/31.

As previously described, each model assumes different pattern or behavior during the three interval of the clearness index. The Reindl* model assumes that the diffuse fraction is constant during the high clearness interval, and this model is a function of the clearness index and solar elevation only. The BRL model is a function of many parameters where the diffuse fraction is gradually decreasing when moving from low clearness index interval to high clearness index interval, where the diffuse fraction becomes almost zero. The DISC is a function of the air mass and the clearness index, and it cannot estimate values of the diffuse fraction in the high clearness index interval that should be close to the real values since the term high clearness index contradicts with high diffuse fraction. The Perez model is a correction of the DISC model but it still produces error similar to the DISC model. Therefore, a model could be developed by including parameters that have a great effect in the model's outcomes. The parameter may include solar elevation since the solar radiation changes with the change of the solar elevation angle. The albedo effect is an another important factor since part of the scattered radiation hits the ground and is reflected back towards a solar panel, which changes the amount of the modeled DHI if it is considered in the model. The aerosols available in the atmosphere affect the amount of the solar radiation reaching the Earth's surface. These measurements are not affordable for each installed-ground station, which is the case that is applied for the four locations in spite of its importance. These models were developed using hourly irradiance data, so short term data (i.e. one-minute -10 minute) to hourly data can be used in order to observe the variation during these intervals, where many years of data are recommended because few years of data could contain a year with a non-average amount of solar radiation data behavior. Although processing solar and meteorological data is important, spectral data are important, too since the change in the solar spectrum can affect the solar technology being used. Therefore, analysis of the effect of the solar spectrum should be considered in the future.

Chapter 5 Conclusion

Solar energy is promising as a replacement for other more conventional energy sources. The knowledge of how to analyze the solar spectrum is required to make the right decision about deploying ground solar plants and to assess the bankability of issuing solar projects. Measurements of solar components require special instruments. However, these instruments are not fully affordable for some countries where ground stations installed. To provide a full solar radiation dataset, one of the methods applied is the use of models to estimate either DHI or DNI. Then, the other component can be computed from known values of GHI and modeled DHI or DNI values. Thus, testing various models will help determine the strong and weak points of each model. The models are the Reindl*, BRL, DISC, and DIRINT.

Chapter 2 gives a brief background about the solar geometry concepts required to help understand the content of the atmosphere that hinders the sunlight from reaching the earth's surface. Also, these geometry concepts help calculate some parameters such as the solar zenith angle and the clearness index, which are two important parameters in each model. Also, a brief description and specifications are included for instruments measuring the three solar components. A detailed explanation of the four models along with their equations is also given in chapter 2.

In chapter 3, a description of four ground stations measuring solar components is given. The methodology to eliminate, filter, and correct solar data is explained thoroughly for the four locations: Al-Hanakiyah, Saudi Arabia, Boulder, U.S., Ma'an, Jordan, and Ottawa, Canada. This chapter also includes the processing of meteorological data for Ma'an.

Chapter 4 discusses the results of the evaluation of each model. The period of the data studied is from 2014/01/01 to 2015/02/28 for the four locations. Scatter plots of modeled data and real values are provided for assessing the correlation between the two datasets. Any data outside of the boundaries of the correlation between the two datasets helps identify the model problems. For instance, the Reindl* and BRL models overestimate DNI due to the underestimation of the DHI data when the clearness index is high. Error analyses have been applied to the four models. The metrics show that the Reindl* model produces the lowest error for modeling DHI and DNI for Ottawa. The Perez and BRL models show similar

metrics values in terms of MAE, RMSE, and σ when computing modeled DHI for Boulder. The DISC and Perez models have similar metrics values when modeling DNI for Boulder. For Al-Hanakiyah and Ma'an, the Reindl* and Perez models exhibit similar performance when DHI and DNI are modeled. Although the BRL model was meant to be applied globally, it shows high error results compared to the other models, especially for the Ma'an location.

The goal of this work was to find a common model that can be applied for modeling data for any location. The evaluation of solar data as well as is very important step for future installation of solar plants at a location of interest. It is recommended that a new model be developed in an attempt to produce less error between the modeled values and the actual values. Some recommendations that can help achieve this goal are listed below.

-Solar and meteorological data should be collected from many locations with different weather conditions around the world.

- Ground stations should be well maintained and periodically inspected to minimize the undesired loss of data and to ensure high quality data.

-After the collection of data, a global model should be developed as a means to estimate DHI or DNI, and the results should be tested against geographically diverse reference data.

-More solar resource model development is needed since most models fail to accurately estimate DHI and DNI during clear sky events.

-To make statistically significant conclusion, at least five years of data must be included. The reason is that one year could be misleading as the weather pattern observed at a location could exhibit a different behavior in comparison to multi-year data.

This work has not considered the spectral effect on different solar technologies but it deals with correction of solar data and evaluation of models. So, spectral information is a crucial step that can help improve accuracy and analyze their effect on solar technologies in the future. Also, computation of the global tilt irradiance is important since PV systems are usually oriented at a specific angle. Many models were developed to compute GTI. So, testing these models can be carried out in the future studies.

References

[1] "Fossil Fuels have made up at least 80% of U.S. fuel mix since 1900. Obtained from http://www.eia.gov/todayinenergy/detail.cfm?id=21912."

[2] A. Weaver, *Generation us: The challenge of global warming*. Orca Book Publishers., 2011.

[3] "Nave, R. (2012). Abundances of the Elements in the Earth's Crust. Georgia State University."

[4] "Honsberg, C., & Bowden, S. (2014). PV Education. ORG. http://pveducation. org/pvcdrom/properties-of-sunlight/Atmospheric Effects."

[5] "K.A.CARE, 2016. Renewable Resource Atlas, King Abdullah City for Atomic and Renewable Energy (K.A.CARE), Saudi Arabia, http://rratlas.kacare.gov.sa."

[6] B. Y. H. Liu and R. C. Jordan, "The interrelationship and characteristic distribution of direct, diffuse and total solar radiation," *Solar Energy*, vol. 4, no. 3, pp. 1–19, 1960.

[7] J. F. Orgill and K. G. T. Hollands, "Correlation equation for hourly diffuse radiation on a horizontal surface," *Solar Energy*, vol. 19, no. 4, pp. 357–359, 1977.

[8] D. G. Erbs, S. A. Klein, and J. A. Duffie, "Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation," *Solar Energy*, vol. 28, no. 4, pp. 293–302, 1982.

[9] M. Iqbal, "Prediction of hourly diffuse solar radiation from measured hourly global radiation on a horizontal surface," *Solar Energy*, vol. 24, no. 5, pp. 491–503, 1980.

[10] A. Skartveit and J. A. Olseth, "A model for the diffuse fraction of hourly global radiation," *Solar Energy*, vol. 38, no. 4, pp. 271–274, 1987.

[11] A. Skartveit, J. A. Olseth, and M. E. Tuft, "An hourly diffuse fraction model with correction for variability and surface albedo," *Solar Energy*, vol. 63, no. 3, pp. 173–183, 1998.

[12] E. Maxwell, "A quasi-physical model for converting hourly global horizontal to direct normal insolation," *No. SERI/TR-215-3087. Solar Energy Research Institute*. Golden, CO (USA), 1987.

[13] P. Ineichen, R. R. Perez, R. D. Seal, E. L. Maxwell, and A. Zalenka, "Article Dynamic global-to-direct irradiance conversion models," *Ashrae Trans.*, vol. 98, no. 1, pp. 354–369, 1992.

[14] D. T. Reindl, W. A. Beckman, and J. A. Duffie, "Diffuse fraction correlations," *Solar Energy*, vol. 45, no. 1, pp. 1–7, 1991.

[15] N. Helbig, "Application of the Radiosity Approach to the Radiation Balance in Complex Terrain," p. 209, 2009.

[16] J. Boland, J. Huang, and B. Ridley, “Decomposing global solar radiation into its direct and diffuse components,” *Renewable and Sustainable Energy Rev.*, vol. 28, no. November 2015, pp. 749–756, 2013.

[17] “National Renewable Energy Laboratory (NREL). Efficiency Chart. Accessed (February,28,2017) from https://www.nrel.gov/pv/assets/images/efficiency-chart.png.”

[18] C. J. Chen, *Physics of Solar Energy*. John Wiley & Sons, 2011.

[19] “Honsberg, C., & Bowden, S. (2014). PV Education. ORG. http://pveducation. org/pvcdrom/properties-of-sunlight/Blackbody Radiation.”

[20] M. Iqbal, *Introduction to Solar Radiation*. Elsevier, 2012.

[21] *ScienceBlogs. Accessed (February, 20,2017) from http://scienceblogs.com/startswithabang/files/2011/01/seasons.jpeg*. .

[22] M. Sengupta, A. Habte, S. Kurtz, A. Dobos, S. Wilbert, E. Lorenz, T. Stoffel, D. Renné, C. Gueymard, D. Myers, S. Wilcox, P. Blanc, and R. Perez, “Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications,” no. February, 2015.

[23] V. Tatsiankou, “Instrumentation Development for Site-Specific Prediction of Spectral Effects on Concentrated Photovoltaic System Performance Viktar Tatsiankou,” 2014.

[24] T. Vignola, F., Michalsky, J., & Stoffel, *Solar and Infrared Radiation Measurements*. CRC Press., 2012.

[25] T. Muneer, *Solar Radiation and Daylight Models*. Routledge, 2007.

[26] H. P. Garg, *Solar Energy:fundamentals and applications*. Tata McGraw-Hill Education., 2000.

[27] *Diagram of air mass. Accessed (February,20,2017) from http://assets.newport.com/web600w-EN/images/1085.gif*. .

[28] P. Lauret, J. Boland, and B. Ridley, “Derivation of a Solar Diffuse Fraction Model in a Bayesian Framework,” *Case Stud. business, Ind. Gov. Stat.*, vol. 3, no. 2, pp. 108–122, 2010.

[29] *Rotating Shadowband Radiometer. Irradiance Inc. Obtained from http://irradiance.com/sites/default/files/RSR_Brochure_080122.pdf*. .

[30] S. Wilbert, “Determination of Circumsolar Radiation and its Effect on Concentrating Solar Power (Doctoral dissertation),” Lehr-und Forschungsgebiet Solartechnik (DLR)., 2014.

[31] F. Lanini, “Division of global radiation into direct radiation and diffuse radiation (Master’s thesis),” University of Bern, 2010.

[32] *Ground truth. 2017. In https://en.oxforddictionaries.com. Retrieved (March, 02, 2017, from https://en.oxforddictionaries.com/definition/ground_truth*. .

[33] R. Perez and P. Ineichen, "Dynamic Global-To-Direct Irradiance Conversion Model (Ashrae)," *ASHRAE Trans Res*, vol. 98. pp. 354–69, 1992.

[34] R. Perez, P. Ineichen, R. Seals, and A. Zelenka, "Making full use of the clearness index for parameterizing hourly insolation conditions," *Solar Energy*, vol. 45, no. 2, pp. 111–114, 1990.

[35] A. S. Conversion, "The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air," no. July 2004, pp. 225–233, 2005.

[36] *R. Perez (2016). Senior Reaserach Associate. Atmospheric Sciences Research Center. State University of New York at Albany.* .

[37] B. Ridley, J. Boland, and P. Lauret, "Modelling of diffuse solar fraction with multiple predictors," *Renew. Energy*, vol. 35, no. 2, pp. 478–483, 2010.

[38] J. Boland, L. Scott, and M. Luther, "Modelling the diffuse fraction of global solar radiation on a horizontal surface," *Environmetrics*, vol. 12, no. 2, pp. 103–116, 2001.

[39] J. Boland, B. Ridley, and B. Brown, "Models of diffuse solar radiation," *Renew. Energy*, vol. 33, no. 4, pp. 575–584, 2008.

[40] S. Younes, R. Claywell, and T. Muneer, "Quality control of solar radiation data: Present status and proposed new approaches," *Energy*, vol. 30, no. 9, pp. 1533–1549, 2005.

[41] C. Riordan, E. Maxwell, T. Stoffel, M. Rymes, and S. Wilcox, "SERI Solar Radiation Resource Assessment Project: Fiscal Year 1990 Annual progress Report" No. SERI/TP-262-4118. Solar Energy Research Inst., Golden, CO (United States), 1991.

[42] *Google Maps. Hanakiyah Road. Accessed (February,28,2017) from https://www.google.ca/maps/place/24%C2%B051'20.8%22N+40%C2%B032'09.6%22E/@24.8598238,40.5418146,1020m/data=!3m1!1e3!4m5!3m4!1s0x0:0x0!8m2!3d24.85577!4d40.536*. .

[43] E. Zell, S. Gasim, S. Wilcox, S. Katamoura, T. Stoffel, H. Shibli, J. Engel-cox, and M. Al, "ScienceDirect Assessment of solar radiation resources in Saudi Arabia," *Solar Energy*, vol. 119, pp. 422–438, 2015.

[44] E. Maxwell, S. Wilcox, and M. Rymes, "User's Manual for SERI QC Software. Assessing the Quality of Solar Radiation Data," *Ratio*, no. December, 1993.

[45] N. Geuder, F. Wolfertstetter, S. Wilbert, D. Schüler, R. Affolter, B. Kraas, E. Lüpfert, and B. Espinar, "Screening and Flagging of Solar Irradiation and Ancillary Meteorological Data," *Energy Procedia*, vol. 69, pp. 1989–1998, 2015.

[46] D. Schüler, S. Wilbert, N. Geuder, R. Affolter, F. Wolfertstetter, C. Prahl, M. Röger,

G. Abdellatif, A. A. Guizani, M. Balghouthi, A. Khalil, A. Mezrhab, N. Yassaa, F. Chellali, D. Draou, P. Blanc, J. Dubranna, and O. M. K. Sabry, "The enerMENA Meteorological Network – Solar Radiation Measurements in the MENA Region," 2015.

[47] "Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), 2016. German Aerospace Center Institute of Solar Research , Paseo de Almería, 73-2°, 04001 Almería, Spain."

[48] *F. Wolfertstetter (2016). DLR German Aerospace Center, Institute of Solar Research, Plataforma Solar de Almeria, 04200 Tabernas, Spain.* .

[49] I. Reda and A. Andreas, "Solar Position Algorithm for Solar Radiation Applications (Revised)," *NREL/Tp-560-34302*, no. January, pp. 1–56, 2008.

[50] M. Journée and C. Bertrand, "Quality control of solar radiation data within the RMIB solar measurements network," *Solar Energy*, vol. 85, pp. 72–86, 2011.

[51] T. Espinar, B.; Wald, L.; Blanc, P.; Hoyer-Klick, C.; Schroedter-Homscheidt, M. & Wanderer, "Report on the harmonization and qualification of meteorological data Project ENDORSE, Energy Downstream Service Providing Energy Components for GMES," 2011.

[52] J. E. Haysom, L. De Salle, K. Hinzer, and H. Schriemer, "Multi-year Ground-based Irradiance Dataset in a Northern Urban Climate," pp. 1–6, 2016.

[53] I. Reda, "Method to Calculate Uncertainties in Measuring Shortwave Solar Irradiance Using Thermopile and Semiconductor Solar Radiometers," *Contract*, *303*, 275-3000. July, 2011.

[54] "P. White (Master student).Direct Solar Irradiance: Quality control methodologies for accurate ground measurement Internal SUNLAB report. SUNLAB group. University of Ottawa, Ottawa," 2016.

[55] F. J. Batlles, F. J. Olmo, and L. Alados-Arboledas, "Shadowband correction methods for diffuse irradiance measurements," *Solar energy*, vol. 54, no. 2, pp. 105–114, 1995.

[56] J. Polo, J., Wilbert, S., Ruiz-Arias, J. A., Meyer, R., Gueymard, C., Suri, M., ... & Boland, "Integration of ground measurements to model-derived data: A report of IEA SHC Task 46 Solar Resource Assessment and Forecasting.," 2015.