2
Review of Literature

2.1 INTRODUCTION

Solar resource assessment is important in selecting suitable sites for solar power generation. For this purpose one requires a good quality measured solar radiation database with additional (co-related) meteorological parameters. Recent and historical radiation databases are available from various sources. These databases contain in international (or location specific) time zone and sometimes solar angles as well. However, the solar radiation incident at any part of the *Earth's* surface is not always the same, therefore the importance of these solar angles. Two primary reasons which affect the solar radiation receiving on *Earth's* outer surface are the *Earth's* elliptical rotation around the sun and its own eccentric tilted axis rotation [Sukhatme and Nayak, 2008]. Also according to the solar radiation energy budget plot [Duffie and Beckman, 2013], the radiation incident on earth has to pass through various gaseous layers, before reaching the *Earth's* surface. Therefore the direct interaction of sunlight with (clouds, ozone, carbon-di-oxide and other) elements present in between, affects the actual radiation falling on the specific location surface. Different climatic conditions present at the chosen location are often identified, as this process helps in the solar radiation assessment. In addition, several companies have also shown interest in solar thermal power plant installation during India's National Solar Mission, but due to incorrect solar resource potential estimation, plants are unable to work at designed capacity [Bhusan *et al.*, 2015]. Therefore, a detailed resource assessment with a reliable database, is needed for further progress in solar power generation in India.

For any solar PV plant installation, one needs only hourly or daily averaged global irradiation component. However in solar thermal power plants, both direct and global components are required, and use of high-frequency database (1-min to 10-min) is suggested. For accurate yield forecast of large-scale solar plants at a particular location, the need of sub-hourly radiation dataset (Ex. 1 or 10-minutes) is currently in high demand. India holds the world second largest PV plant located at one specific location of 648 MW capacity and is spread over 2,500 acres or 10 km^2[PTI, 2016]. In addition to solar resource database selection, another important question identified while discussing the feasibility of the power plant is the land availability. Table 2.1 provides a brief overview of different power plant technologies and their specific land required [Nicholson, 2013]. From this one can conclude that conventional thermal power plant needs less area than any solar energy based plant. However in solar thermal systems, due to use of similar power block as in coal power plant, higher efficiency may be achieved [Parsons, 2014]. Therefore, the conversion/augmentation of coal based power station to a solar-based system may be possible.

Table 2.1: Various Electricity Generation Technologies, Efficiencies and their Land per MW Requirements

Technology	*Efficiency (%)*	*Land per MW (Acres)*
Mono Crystalline Silicon	18-24	3-4
Poly/Multi Crystalline Silicon	14-18	4-5
Thin Film	6-14	7.5-9
Solar Thermal Power Plant	9-31 [Darwish et *al.*, 2015]	7.5-10
Conventional Thermal Power Plant	33-48	1.42

Solar radiation databases available for any site analysis are categorized under satellite-derived, ground-measured and data generated using relations (derived from standard correlations and local data availability). In satellite-derived radiation databases, the solar radiation data is calculated from (Earth's sky and ground) images captured by Geostationary Metrological Satellite (GMS). Here the information provided by satellite images is identified by using its spatial (km) and temporal (Deg.) resolution. Also, by using various frequency filters at satellite measuring station, one can derive extra information like cloud type, aerosol presence, etc. Currently spatial resolution of 2.5 km and temporal resolution of half-hour database is available from satellite images; hence it makes easy comparison with ground-measured values. These historical radiation databases are created using these satellite images and some of them are also openly available to download.

There are several online radiation sources and historical databases from which location-specific data can be downloaded. Some possible radiation sources are "NREL", "NASA", "IMD" and "C-WET", and database of various frequencies are available there. These databases are produced by comparing and merging data from various solar radiation stations at ground level with corresponding satellite images. The historical dataset for various locations is present in hourly or daily averaged obtained from measurements at high frequency (1 to 10-minutes interval). The hourly database is freely accessible but sub-hourly needs to be purchased from various sources.

The quality of radiation values measured at any ground location depends upon instrument type used, stations' surroundings, data conversion approach, etc. Ground-based solar radiation measuring instrument stations, various sensors are installed at ground and data is measured at a high frequency. Radiation components and some meteorological parameters are typically measured at these stations. However the presence of errors in measured values is common, therefore good standards of operation and maintenance schedules are to be followed for accurate measurements. All measuring instruments used here are selected on the basis of their precision and build quality, suggested by World Meteorological Office [WMO (1), 2011] and Commission Internationale de l'Eclairage (CIE) guidelines [Kendrick *et al.*, 1994]. After that for accurate data, a standard measurement guideline for these instruments is provided in International Organization for Standardization [ISO-9060, 2008], which is universally followed for radiation measuring instrument classification. The instruments' inherent properties like tolerance, error, class, etc. are printed on the instruments, and according to the measured solar radiation data requirements, most promising instrument (based on measurement quality and accuracy) is selected. Also, before purchasing any solar radiation measuring instrument from the manufacturer, its sensor response and build properties need to be properly checked.

For determining theoretical solar radiation data, various mathematical models are provided in the literature. These models need some supporting parameters (averaged radiation and associated meteorological values) for radiation data estimation. Basic radiation models [Angstrom, 1924; Page, 1961; Lof *et al.*, 1966; Modi and Sukhatme, 1979; Gopinathan and Soler, 1995; ASHRAE, 1972] depend upon station's solar geometry angles with meteorological measured values. Using this information, location specific regression equations are derived by analyzing station's long-term radiation database and their relation with present meteorological parameters. Related sky condition and climatic properties are also collected from the meteorological station's historical database. All relations are discussed in the literature [Duffie and Beckman, 2013; Sukhatme and Nayak, 2008]; by utilizing them one can calculate daily and hourly averaged values. Some supporting climate models are also available, which calculate radiation data by using satellite images and climate parameters. They also predict various cloud condition radiation values and also support in the identification and correction of measured radiation values.

In ground-based measurement stations, if the recorded measurements are not following the standard radiation co-relation, then that data is considered as erroneous. The reason behind its occurrence could be according to instrument or operation related [Younes *et al.*, 2005]. Instrument based error is identified, if radiation incident on the sensor is unable to fall directly on the sensors. If the error is arising out of stations infrequent maintenance and effects due to local climate disturbances, then it confirms the presence of operation based errors. All these errors are briefly described below and for their identification, some conditions are also discussed.

For the identification of various errors in measured solar radiation databases, quality control tests are available in the literature. But before any detailed data analysis, the primary objective is to check the measured solar data time step and identify the missing intervals of data (if any). Now if measurement data is found missing in the database, then search for relevant meteorological parameters (Ex. Climate condition) and the missing interval is recorded for that day. After that quality control conditions (discussed below) can be used to check the measurements (using various boundary conditions) and their inter-correlation ability with different radiation components [Long and Dutton, 2002]. For each quality test condition failing, a type of error flag is identified. Now, this corresponding assigned error flag can help analyze the overall station radiation database. Gap filling approaches suggested in literature, are applied for completing the provided database. Some techniques used for filling the gaps are discussed, based on the frequency of missed data.

Finally, after successfully applying suggested data quality control and gap filling techniques, some statistical indicators or tests can be used to check the station's radiation potential. These suggested tests calculate the ratios between measured values with predicted values. Now the results generated by each approach are compared with other databases and some guidelines are suggested for "Hot and Dry" climate type locations.

2.2 CREATING RELIABLE RADIATION DATABASES

An identified database having the same time-stamp of radiation and meteorological component is the only raw material required for solar resource assessment of any site. The sources identified for getting the radiation datasets are provided in Table 2.2. These listed databases sources (NASA, 2017; NREL, 2017; Solemi, 2017; SOLARGIS, 2017, etc.) collect data from other satellite-based radiation databases (often verified from ground-based measured values). Here each historical database is recorded at daily or hourly interval duration. The ground-based measurement stations [IMD (3), 2017; C-WET (1)2017, etc.] measure at the sub-hourly interval and are considered to be the most accurate. But from these stations, radiation databases need to be purchased. Their measurement frequency is in 1-Second, 1-Minute to 10-Minute interval. However, in satellite or relation based radiation databases, sub-hourly frequency is unavailable. Hence in this research work, available ground-based radiation and related meteorological measurement data are selected for analysis. Meteorological data are also used for cross checking radiation measurements. Such comparisons and conclusions help in understanding climatic conditions at the site.

2.3 SOLAR RADIATION MEASUREMENT AND ERRORS IDENTIFICATION

In a solar radiation measurement station, for measuring (GHI, DHI and DNI) radiation components, two pyranometers and one pyrheliometer with a 2-axis tracking mechanism are used. Pyranometers with glass opening at the top allow both direct and diffuse radiation to interact with the sensor. For measuring diffuse radiation component, another shaded disc pyranometer with glass opening covered from the top is used. By using a solar radiation

Table 2.2: Solar Radiation Database Available Data for India [CCS, 2017]

Sources	*Details*
NASA *(eosweb.larc.nasa.gov/sse/)*	-Uncertainty 6%-12% -Monthly Averaged- 22 years of Satellite Data with the 30km Grid (resolution) -Hourly Data, derived using the Averaged Day of the Month (frequency)
Solemi *(www.solemi.de)*	-Data Derived using Meteosat-5 and Meteosat-7 Satellite -Spatial Resolution of 2.5 km and Temporal Resolution of Half Hour -Data Derived using Satellite Image, Visual Spectrum Analysis -RMSE Hourly values (GHI-15%) (DNI-35%)
Meteonorm *(www.meteonorm.com)*	-Data Collected from Ground Weather Stations and Supplemented Satellite Data -one Min. data, Hourly Values (TMY) Calculated using a Stochastic Model and monthly values. -Resolution 1/8 degrees.
NREL *(http://www.nrel.gov/rredc)*	-Hourly Database Collected from Meteosat-5 and Meteosat-7 Satellites -10-km Resolution Resource Map Available for India, using SUNNY Satellite Data - Uncertainty 5%-10% (RSI instrument)
Solar radiation handbook *[IMD(3), 2008]*	-Mean Monthly Global and Diffuse Solar Radiant Exposure, for 17 Indian Locations -Air Temperature & Mean Sunshine Hours -Instrument Accuracy (5%-7%) - (2km x 2km), 33 minute interval images
3Tier *(http://www.3tier.com/en/products/solar/)*	-Hourly Value of Radiation Values -Horizontal Resolution of 3 km (2 arc minutes) -Long term Mean (5% GHI) (20%DNI)
ISHRAE *(www.ishrae.in)*	-Hourly Average DNI Data for 58 Indian locations, using IMD Database -Used only for Building Energy Performance
IMD *(http://www.imdaws.com/viewradiationdata.aspx)*	-Satellite and ground-based measurement source (accuracy 15%) (10kmx10km maps) -Historical database (many years, satellite 30-min values) available for Indian locations. -Historical and live radiation (ground-based) database, of 1-min or 10-min interval
C-WET (SRRA) *(http://niwe.res.in/department_srra_SDSAP.php)*	-121 High quality basic and advanced ground-based radiation measurement stations -MNRE project supported by “SolMap” provides the best data quality radiation database (10kmx10 km) -Raw database available of 1-min interval solar and other meteorological parameters
SolarGIS *(http://solargis.com/products)*	-Data Collected from various Satellite Database with validation studies (lowest uncertainty) -Data available 10/15/30 minute time-step with 250mx250m spatial resolution -TMY created using sub-hourly time series.

blocking disc or strip, only diffuse radiation is allowed to interact with the sensor. For measuring direct radiation component, the instrument utilized is pyrheliometer in which the sensor is placed at the long narrow hinged/baffled tube end. This allows radiation to pass through instruments' glass opening of 5 degrees. Therefore precise tracking angles are involved for accurate DNI measurement. When the radiation component interacts with the sensor, the radiation value is sensed in electron volts [Kippzonen, 2017], which is then accordingly converted and recorded in solar radiation value in appropriate units.

2.3.1 Solar Radiation Components and their Inter-correlation

For measuring different radiation components, the types of instruments used and their inter-relationship is discussed here. The GHI radiation component is the total radiation (direct and reflected part) coming from the sun and for its measurement, a pyranometer is installed in the horizontal plane. When the instrument is aligned at some predefined (panel or location specific inclination) angle, then Global Tilted Irradiance (GTI) is calculated. Similarly, on the same horizontal plane, when the instrument is shaded with a tracking ball or disc (which stops direct radiation falling on the sensor), then radiation component measured is Diffuse Horizontal Irradiance (DHI). Finally for measuring direct radiation (DNI) component (here sun ray is coming directly to cylinder shape instrument with continuous sun tracking mechanism and) instrument used is "pyrheliometer". Now in their inter-correlation, the total GHI radiation is evaluated using DNI with cosine zenith angle, DHI and scattered radiation component (see Eq. (2.1)). In computation of GTI, components included are DNI, DHI, scattered and reflected radiation components. Here total extraterrestrial radiation falling on a location is calculated using Eq. (2.2).

$$I_G = I_D + I_B \tag{2.1}$$

$$I_E = I_G + I_{reflected} + I_{absorbed} \tag{2.2}$$

where,
I_E = Extraterrestrial Radiation
I_G = Global Horizontal Irradiation
$I_{reflected}$ = Reflected Irradiance (Back to Space)
$I_{absorbed}$ = Absorbed Radiation by Atmosphere
I_D = Diffused Horizontal Irradiation
I_B = Beam or Direct Horizontal Irradiation

2.3.2 Solar Radiation Measurement Standards

Instruments used for measuring solar radiation (pyranometer and pyrheliometer), have to follow some standard [ISO-9060, 2008; WMO (2), 2008] guidelines. These instruments are classified on the basis of its working class, sensitivity, working spectrum and response rate. In the working class, instruments are classified based on type of sensor used. The Primary Standard (Absolute Cavity Radiometer (*ACR*)) is the most advanced, accurate and reliable source (also known as standard instrument and is directly traceable to World Radiometric Reference Scale (*WRR*) standards). These instruments are highly sensitive to a slight change in local environmental conditions and provide relatively error-free data. These instruments are exclusively used for calibration purpose (not in field operations). Many countries have regional calibration center and it should include at least three (each radiation component) instruments. These instruments are slightly inferior to primary standards but are found stable at continuous

Table 2.3: Instrument Error when compared with Reference Instruments [Coulson, 1975]

Class	*Pyranometer*	*Pyrheliometer*
First class	±10%	±4%
Second class	±25%	±8%
Third Class	±32%	---%

use (termed as secondary Standard). Moreover, with the help of these sensors, all radiation sensors installed at different locations of a country can be calibrated. First, second and third measuring instruments are classified according to uncertainty and errors present in the instruments. See Table 2.3, for further details of these instruments when compared with reference instruments [Coulson, 1975].

For solar radiation and related meteorological data measurement practices, WMO (1), [2011] guidelines are used. Here different instrument limits related to the sunshine, relative humidity, pressure, temperature, etc. are provided. Using these guidelines, the standard solar radiation data measurement and data quality control guidelines are provided by BSRN [Long and Dutton, 2002]. The standard guidelines (discussed below) provide maximum and minimum possible radiation limits and other radiation components inter-related equations. These are significant for any solar radiation data analysis, as all solar radiation measurement stations installed in the world, have different time-zones with varying climatic and atmospheric conditions. Here measuring instruments type, radiation sensing procedure and measurement frequency can also be different. However, all stations may not measure all listed components. The non-uniformity in instruments and sampling interval makes a comparison of available measured databases time-consuming.

2.3.3 Factors Affecting Solar Radiation Data Measurement

The sun spreads radiation in all directions and radiation received at the outer level of the Earth's surface is termed as Extra-terrestrial radiation. However, this radiation has to pass through various types of density filters, after that radiation is received at the Earth's surface. In a broad scale, one can arrange these factors affecting radiation as [Iqbal, 1983]:

(a) Location (Space and Time)

Location coordinates (Latitude and Longitude) and duration of the day mostly influence the actual radiation falling on that location. Utilizing this information, solar angles are calculated, and their ratio with solar extraterrestrial radiation directly gives the theoretical radiation availability at that location.

(b) Cloud (Droplets and Ice)

The natural water cycle is constantly working to accurately maintain the life cycle of the Earth. The formation of various clouds and their movement is the next most influential component, which affects the incoming solar radiation.

(c) Total Precipitable Water

Water present in the atmosphere which settles down at night causes a high-density region at the location. It acts as a reflector for incoming radiation.

(d) Aerosol and Dust

The aerosol remains in the sky, as a collection of fine solid particles or liquid present in the gaseous form. Similarly, dust as small solid particles also floats in our atmosphere. They both reduce the intensity of solar radiation passing through them.

(e) Surface Albedo

It is a surface property and measures the brightness of the receiving surface. Therefore, if radiation is falling on any black surface, it gets completely absorbed and when the surface is white, it gets reflected without transferring its energy to the ground. In high mountain regions (ex. Ladakh mountains) snow reflects all radiation back to the sky, that's why in spite of getting high radiation, the climate is still cold. In addition, as no energy is absorbed by the surface, snow doesn't get melted easily for the whole year.

(f) Ozone

It is an important gas layer for our survival. Ultra-violet rays coming from the sun, gets filtered by ozone. Due to the release of gases (CFC, SOx, NOx, etc.) from the ground level, some variations are seen in the incoming solar radiation.

(g) Other Gases

The increase of various gases (CO_2, CH_4, N_2, etc.) and their mixture, also increases the temperature of the earth's surface. Their density cloud also reflects radiation back to space.

2.3.4 Errors in Solar Radiation Measurement

For identification of errors, some tests are suggested by Maxwell *et al.*, (1993); Journée and Bertrand, (2011); Geuder *et al.*, (2013) and Kumar *et al.*, (2014). But the most common way of identification of errors is by using transmittance ratios and their mutual plots [Maxwell *et al.*, 1993]. These ratios are calculated by using measured value with theoretical maximum radiation values. Different transmittance ratios calculated for global, direct and diffuse radiation components are provided in Table 2.4. Where for clearness index (or global horizontal transmittance), the actual measured value is divided with theoretical maximum clear sky radiation value at any horizontal surface. Similarly, for calculating direct beam transmittance and diffuse horizontal transmittance, the measured value is divided with Extra-terrestrial and measured global radiation value. Also, these ratios are unit-less components, hence used universally for checking the quality of measured radiation components at any location. Now by plotting two transmittance ratios with respect to each other, various climatic conditions and instrument measurement quality is identified [Maxwell *et al.*, 1993]. In Figure 2.1, one can see the "k_t-k_n" plot, which provides information about different climate conditions present at any day. Now for understanding "k_n" in this plot, one sees that as we go downwards the density of cloud increases. Therefore, the high one is clear sky zone, then comes in the partly clouded zone and at the end cloudy sky condition. Similarly, the "k_t" is also divided into zones, the irregular trapezoid having a base at (0.0 to 0.4) on "k_t" shows normal days plot region. Then triangle having a base in "k_t" shows heavy cloud condition with some measurement errors and after that space shows the region of light reflected or scattered due to clouds. However, the transmittance ratio range varies with location to location, but the maximum possible limit for both "k_t" and "k_n" is 1.0 and 1.0. Here, the data present in the upper triangle are not valid, due to data ratio unity.

Table 2.4: Various Transmittance Factors used in Data Analysis [Geuder *et al.*, 2015]

K_t = Clearness Index, or Global Horizontal Transmittance	k_n = Effective Direct Beam Transmittance	k= Effective Diffuse Horizontal Transmittance
$k_t = I_G/(I_E \cos(z))$	$k_n = I_B / I_E$	$k = I_D / I_G$

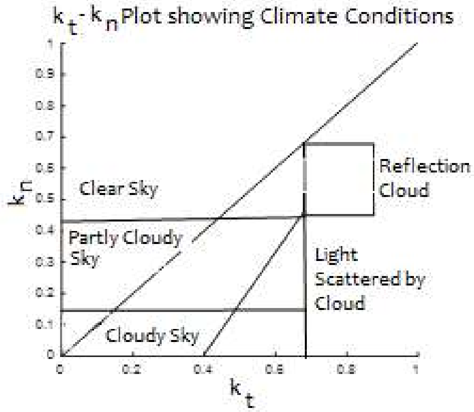

Figure 2.1: Transmittance Ratio Plot (k_t-k_n) for Different Atmospheric Conditions

The solar radiation which has reached at the ground, suffers from various filters during its measurement. The representation of the scale of errors for various irradiance sensors is provided in Table 2.3 [Stoffel *et al.*, 2001]. If instruments are incapable to measure solar radiation correctly, due to their own construction and sensing technique, then the errors identified are referred to as equipment errors. It is the duty of the manufacturer of the measuring instrument, to provide the best quality instruments and arrange periodic calibration sessions, to get the best results from them. Another source of errors is due to the operation of instruments, if any disturbance is identified related to the measurement process. Presence of dust, the shadow on the instrument, alignment issues, animal or birds presence, open magnetic and electric flux, etc. are some additional reasons behind these errors.

2.4 MATHEMATICAL MODELS FOR ESTIMATING SOLAR RADIATION

If radiation database for any location is not present, but its basic site analysis is required, then mathematical models are used. Here radiation components are calculated by implementing standard radiation, clear-sky and different cloud condition models. Relevant approaches are discussed below with a discussion of their input parameters. Here most parameters are dependent on other variables and need additional meteorological (transmittance) values from other sources.

2.4.1 Solar Radiation Prediction Models

If the station's solar radiation database is not available or missing from measurement database, but meteorological dataset is present, then by using radiation model equations, one can predict stations averaged radiation values. Hence, each such applicable approach is briefly discussed below:

a) Meteorological Radiation Model

Equations used here [Muneer and Gul, 2000; Muneer *et al.*, 1998; Gul *et al.*, 1998] calculate uniform resolution averaged radiation values (hourly diffuse horizontal irradiance (I_D), beam horizontal irradiance (I_B) and beam clearness index (k_t)). The input data required by model equations is daily averaged atmospheric pressure (P_{atm}) [Spokas and Forcella, 2007], dry bulb temperature (DBT) & wet bulb temperature (WBT) [Hargreaves and Samani, 1982; Annandale *et al.*, 2002] and the sunshine ratio (n/N) [Angstrom, 1924]. Additional accurate radiation model [Kambezidis and Psiloglou, 2008] includes other meteorological measured parameters. If all values are measured correctly, then approximate estimations can be made. Here different transmittance values (calculated for the different medium) are used for determination of radiation average values. Some remaining parameters are location specific constants and historical database average values. But the availability of various scattering values at a similar location is difficult; hence averaged values are used (see Eq. (2.3 and 2.4) for calculating I_D from I_B values).

$$\text{Diffuse beam ratio, DBR} = I_D/I_B = a(k_n)^b \tag{2.3}$$

$$k_n = I_B/I_E \tag{2.4}$$

b) Cloud Radiation Model

In meteorological radiation modeling, mathematical model uses meteorological parameters and different transmittance values to estimate radiation values. But still, the majority of people use cloud cover data for radiation calculations, because of simplicity in collecting images from satellite or ground stations. Therefore by following this model, global and diffuse horizontal hourly averaged irradiance relations are calculated. Now for deriving these empirical equations [Gul *et al.*,1998; Muneer and Gul, 2000; Kasten, F., and Czeplak, G.,1980] 10 years (1964-1973) continuous hourly radiation dataset is taken from Hamburg, Germany. The results derived after analysis are in terms of relations (using local coefficients) as provided in Eq.(2.5 to 2.7), where A_1 , B_1 , C_1 and D_1 are locally derived coefficients and "O" cloud cover on a scale of 8. However, for every location, these coefficients need to be calculated using a well-maintained radiation database.

$$I_{G,C} = A_1 \sin\alpha - B_1 \tag{2.5}$$

$$I_G = I_{G,C}\,(1-C_1\,(O/8)^{D_1}) \tag{2.6}$$

$$I_D = I_G\,(0.3 + 0.7\,(O/8)^2) \tag{2.7}$$

2.4.2 Clear Sky Radiation Models

For any station's radiation data analysis, comparing the measured radiation value with corresponding theoretical clear sky day is important for a data feasibility study. Here various spectral and broadband models [ASHRAE, 1972; Page, 1961; Yang *et al.*, 2001; Gueymard, 2008] are discussed and any suitable model can be applied for clear sky radiation data calculation.

a) Clear-sky Page Radiation Model

Meteorological Radiation Model (MRM) is a broadband horizontal irradiance model and it got evolved while developing European Solar Radiation Atlas (ESRA) [Page and Lebens, 1986]. Therefore, a cloudless model is suggested, which takes input parameters as solar altitude and Linke turbidity (derived using selected air mass value) (see Eq. (2.8)). For calculation of cloudless-sky beam and diffuse components (Eqs. (2.8-2.11)) can be used.

$$I_B = I_E K_D \exp(-0.8662\, m T_{LK}\, \delta_r(m)) \sin(\alpha) \tag{2.8}$$

Where "I_E" is the Extra-terrestrial irradiance, "K_D" the Earth-Sun correction factor, "m" air mass, "T_{LK}" the Linke turbidity, "δ_r" the Rayleigh's optical depth and "α" as solar altitude. Here for accurate Rayleigh optical depth calculation [Kasten, 1993] for Indian conditions, Eq. (2.8) requires some modification. In Eqs. (2.9) to (2.11) air mass is an input parameter and its correlation is provided here.

$$\delta_r(m) = [6.6296 + 1.7513\, m - 0.1202\, m^2 + 0.0065\, m^3 - 0.00013\, m^4]^{-1} \tag{2.9}$$

$$T_{rd} = -21.657 + 41.752\, T_{LK} + 0.5190\, {}^{2}_{LK} \tag{2.10}$$

$$I_d = K_d T_{rd}(n)\sin(\alpha) \tag{2.11}$$

For calculating (T_{rd}) theoretical diffuse component, Linke turbidity at air mass '2' (sunrise and sunset interval) and solar altitude is required. Moreover, page quality control guidelines [Grief and Scharmer, 2000; Guide, 2002] are used for determining beam and diffuse components.

b) Clear-sky Meteorological Radiation Model

Here beam component is derived using above relations and diffuse component by the clear-sky algorithm. These equations are obtained from Dave, 1949; Bird and Hulstorm, 1979; Pisimanis *et al.*, 1987. The input data required for analysis are air-mass, various transmittance ratios, ground albedo and cloudless sky albedo values (for detailed equations see Muneer, 2004). Another similar model is "Clear-sky Yang Meteorological Radiation" Model [Yang *et al.*, 2002], which is based on the product of different atmospheric transmittances values.

c) Clear-sky Bird Radiation Model

Using atmospheric scattering values and considering various atmospheric factors constant throughout the year, the total irradiance is calculated [Bird *et al.*, 1981]. The input parameters involved in this model are solar constant, zenith angle, surface atmospheric pressure, ground albedo, perceptible water vapor, and total ozone, turbidity at the 0.5-μm and 0.38-μm wavelength and aerosol forward scattering ratio. This model is widely accepted, as standard clear sky model and is designed by SERI, NREL [Maxwell *et al.*, 1993].

d) Clear-sky Mc-Clear Model

In addition to other empirical models, Copernicus clear-sky solar radiation model uses ground-based measured values of the aerosol and water vapor [Lefevre *et al.*, 2013]. But here database and various equations available are found to be applicable for only Europe and African locations. These relations need modification for Jodhpur location.

e) Clear-sky ASHRAE Model

This model determines hourly radiation values for cloudless skies, receiving on a horizontal surface. These equations are based on exponential decay and depend on the "angle of incidence" and "location specific" constants. Where these constants are determined on a monthly basis, and variations are observed in accordance with seasonal changes, dust presence, vapor content present in the atmosphere and Earth-Sun distance. Here the constants (A, B, C) used are derived for Indian locations by Iqbal (1983). For related equations and constants, see Appendix (D.1).

f) Gueymard REST2 Model

Consistent progress in solar radiation data analysis is examined by [Gueymard, 2008] and he also provided (Reference Evaluation of Solar Transmittance Model) REST2 model. The model takes input parameters as extraterrestrial irradiance with the Rayleigh, ozone, uniformly mixed gases, water vapor, aerosol and NO_2 atmospheric transmittance values. The relation used for beam horizontal transmittance calculation is provided in Eq. (2.12). Here various transmittance ratios are calculated, dependent on the location specific parameters (solar geometry, the vertical ozone column, perceptible water, site atmospheric pressure, air-mass and turbidity coefficients).

$$I_B = I_E \sin(\alpha)\tau_r \tau_o \tau_g \tau_w \tau_a \tau_{NO_2} \tag{2.12}$$

Any station selected for analysis can be compared with the above clear-sky model. But the availability of the uniform interval measured radiation values and all transmittance values given above are not guaranteed. Therefore the simplest and often used clear-sky model in this work is Clear-sky "ASHRAE" model. This model is selected because, for the selected location, different aerosol values are unavailable and it requires minimal data inputs.

2.5 IDENTIFICATION OF VARIOUS SKY CONDITIONS

In addition to the standard clear sky model, study of cloudy days is also needed, as correct radiation potential can't be calculated without it. During radiation data analysis, correct identification of day having clear sky is difficult. Various elements present in the sky (cloud, aerosol, etc.) create turbidity and affects the coming radiation. To study these components at ground level, various instruments and sky imagers are used. According to the literature [CIE 22, 1973], sky conditions (based on cloud presence or sky illuminance) can be classified into three groups: clear sky, intermediate cloud and heavy cloud (overcast) condition. A modification is available [CIE 110, 1994] for understanding the spatial distribution of daylight. Where for identifying clear sky day (quasi-clear-sky) Rayleigh limit equation is used with no clouds and gases present in the sky. Finally, by using available sky images and mathematical relations, sky type is identified [Long *et al.*, 2008].

Sky classification using satellite images and cloud cover values is popular. Here cloud cover values are either measured manually or calculated from image analysis tools in terms of "oktas". Now the visible upper sky is here divided into 8 parts (like pie) and percentage of cloud present in the sky, is decided on the scale of 1/8 to 8/8. Literature [Barker, 1992] shows that by using these "oktas" values and by doing image analysis, cloud percentage is identified. Where after the image analysis, if the numerator value is "1" that means clear sky day and if "8", then it is identified as overcast day [Lam and Li, 2001]. Another approach used for cloud classification is done by using available daily sunshine average values [Littlefair, 1988; Muneer, 2004]. This ratio is calculated by dividing measured sunshine hours with theoretical sunshine hours possible at that location. If the result is near to unity (less than 0.9), it means clear sky day and if zero then overcast clouds conditions. For further analysis, both of these approaches are combined [Page and Lebens, 1986; CIBSE, 2002].

In another approach, if the measured radiation database is available, then sky conditions are identified by plotting different transmittance ratios. Measurement quality information can be gathered by using these plots. Here for each radiation component, transmittance factor is calculated and when it is compared with another transmittance factor, the trend of plot reveals various errors in the database. Instrument error in measurement, dust effects and cloud classifications are identified here. However plotting of clear sky index (k_t) and the diffuse ratio (k) is popular, which is used for sky condition identification [Maxwell *et al.*, 1993]. Literature also indicates that while using these ratios, standard ranges (discussed below) and limits are proposed for classifying climatic conditions into groups.

For clear sky day, the clearness index varies between "0.7-0.9" [Iqbal, 1983], which is considered an important conclusion for GHI limits estimation. Here the analysis is further refined by [Perez *et al.*, 1992]. They introduced alternate sky clearness index, which is accurate but requires "air mass" values for computation. But in radiation inter-component studies, one has to analyze all three transmittance ratios by determining relations between the plots. According to Thevenard and Brunger, (2001) for diffuse transmission ratio (for their selected location) the upper and lower limits are 0.4 and 0.2 respectively. Similarly, Greif and Scharmer, (2000); Lam and Li, (2001) calculated upper and lower limits for "k" and "k_t", by using various climatic locations and correlated it with available cloud cover and sunshine fraction. Cloud classification using monthly averaged atmospheric turbidity are also discussed. Here for each station cumulative frequency distribution and plots of cloud cover, clearness index and diffuse ratio are studied. Constants are also derived, for example, for Chennai, "k" maximum limit is 0.21 and "k_t" minimum limit is 0.78. However, analysis shows every month transmission ratio value varies [Lanetz *et al.*, 2005] and this value is identified by using the correct cloud identification approach. Nowadays in place of zenith angle, elevation angle is used in various types of plots [Ineichen, 2006].

2.6 SOLAR RADIATION DATA QUALITY CONTROL TECHNIQUES

To obtain a reliable solar radiation database for the selected location, standard quality control guidelines are needed. These tests are important due to various measurement errors present in the database. But the identification of individual error parameter at each measuring interval is difficult. Solar radiation data quality guidelines and procedures proposed in the literature are described here.

Here the initial efforts in the solar radiation measured data quality analysis are suggested by Maxwell *et al.*, 1993 (NREL SERI-QC) and Long *et al.*, 2008. These standard guidelines are selected from BSRN [Long and Dutton, 2002] and location-specific improvements are included in these equations. Four quality control guidelines are proposed, where the first three conditions check the physical limits of measured values and last checks luminance values (used for checking sunshine duration measurement). These being the most basic approaches and applied to all historical datasets. Moreover in further analysis, by calculating the transmittance ratio for GHI, DHI and DNI components, a quartile envelope is created [Younes *et al.*, 2005]. For calculations of various transmittance ratios of solar radiation components Table 2.4 can be used. In this approach, the selected database is divided into a number of ranges and for each corresponding range mean and standard deviation are calculated. Now by joining the values, a quality envelope is created, which covers good quality data. The visual inspection of these ratios and plots provides information to identify outliers based on a local model for different weather conditions. Literature shows use of the approach for various climate regions, where results are in the suggested range.

Another relevant approach for data quality check is web-based data quality checking [Molineaux and Ineichen, 1994; Geiger *et al.*, 2002; Maxwell *et al.*, 1993]. Here as discussed above standard guidelines are combined and a web-based facility is created for data analysis. The maximum limits for radiation components are calculated with the help of standard models [Page and Lebens, 1986; Bird and Hulstorm, 1979]. The data to be analyzed are uploaded in a specified format with detailed quality control limits applied to it. The end results are obtained in the form of either text or detailed plots for daily and monthly analysis. Further developments occurred in checking the time-step of irradiation values [Ineichen, 2013], with respect to the solar noon. This step supports checking the recording error in the radiation database. In addition, Journee and Bertrand, (2011) provide detailed quality control analysis guidelines derived from high-frequency data (10-min or 30-min average) from Belgium network. Here tests included are physical limits, step, persistence, quality envelop, spatial consistency, sunshine duration, and other climatic analysis tests. This modified new guideline (discussed below) shows various limits which check the database in-depth and identify the condition in which measured data is failing. Moreover, some unique conditions are given which are not included in this procedure, but can be used for individual component error testing.

2.6.1 Basic Quality Control Test Procedure

These conditions are modified extensions of BSRN [Long and Dutton, 2002] guidelines and their development is done by using other radiation and meteorological databases. These conditions include checking of physical theoretical limits and errors present in individual radiation components. From the database different transmittance ratios are calculated and then they are compared with available standard transmittance plots. Hence in this analysis, the results and plots can predict different cloud conditions and instrument errors. The condition used here is described below:

a) First Basic Quality Control Test

The duration for which sunshine is available is considered here for analysis. Hence all measured values whose solar altitude angle is less than 7^o to 10^o are ignored from the analysis. As during this period (sunrise or sunset), the radiation level is low and high air mass causes a scattering in the atmosphere. Now dataset is passed to the next basic quality control test.

b) Second Basic Quality Control Test

The received dataset is now modified by calculating their respective clearness index and diffuse ratio values. Here "k_t, k_n and k" (clearness index, beam transmittance and diffuse transmittance) are used to standardize/normalize the radiation database. Hence now in the same space, one can perform detailed analysis, but the equation used here allows broad estimates (see Eq. (2.13) to (2.16)). As one goes in for additional detailed site analysis, the range of these ratios is further modified (Geuder *et al.*, (2015)). The value which falls outside this limit is considered as erroneous and needs further analysis.

$$0 < k_t < 1 \text{ and } 0 < k < 1 \tag{2.13}$$

$$k_n > k_t \text{ , } k_n < 0.8 \text{ and } k_t = 1.0 \tag{2.14}$$

$$k > 1.05 \qquad (z < 75^o) \tag{2.15}$$

$$k > 1.10 \qquad (z > 75^o) \tag{2.16}$$

c) Third Basic Quality Control Test

In this test, measured values in the database are compared with the standard clear-sky model (Page-model upper and lower boundaries). No measured value can go above these modeled values. Here equations are shown with their conditions [Muneer and Fairooz, 2002; Greif and Scharmer, 2000 and CIBSE, 2002] (see Eq. (2.17 to 2.19)).

$$I_G \le I_{G,C} \tag{2.17}$$

$$I_{G,C} = I_{B,C} + I_{D,C} \tag{2.18}$$

$$I_{D,OC} \le I_D \le I_{D,C} \tag{2.19}$$

2.6.2 Quartile Analysis Test

This is the next step after basic quality control analysis. As discussed above, by using a combination of "k_t-k_n" and "k_t-k" plots, the whole dataset can be analyzed in one space [Muneer and Fairooz, 2002]. Now for calculating transmittance ratios maximum and minimum limits, all databases are grouped on the basis of data location identified in the transmittance plot. Where for each identified range, the corresponding mean and standard deviation are calculated by using the outlier analysis. In this analysis, odd values are removed from the database. Also in addition, according to the available experience (statistical), the near-outlier factor is taken as "1.5" and for far-outliers, as "3" (see Eq. (2.20) and (2.21)) and the dataset is now arranged into four quartile limits. Where first, second, third and fourth quartile means, correspondingly 25%, 50%, 75% and 100% data presence near the whole database mean value.

Also by using near-outlier factor and quartile ratios, mathematically lower and upper outlier limits can be calculated.

Lower outlier limit = 1st quartile - 1.5 (3rd quartile – 1st quartile) (2.20)

Upper outlier limit = 3rd quartile + 1.5 (3rd quartile – 1st quartile) (2.21)

By using derived "k_t" and "k_n" mean value, corresponding weighted means and standard deviation values are identified. For an accurate estimation of limits, "k_t" and "k_n" values are divided into 10 equal bands of their means and standard deviations. Where each band, max $\left(\bar{k}_d + 2\sigma_{k_d}\right)$and min $\left(\bar{k}_d - 2\sigma_{k_d}\right)$points are calculated and accordingly they are marked on the plot itself. Now after identification of points, they are manually fitted by joining these points with any second-degree polynomial equation. Hence after joining, the curve traced is represented as $\left(\bar{k}_d, \bar{k}_d + 2\sigma_{k_d}\right), \left(\bar{k}_d, \bar{k}_d - 2\sigma_{k_d}\right)$ and boundaries are given in Eq. (2.22 and 2.23).

$$A(k_t) = Max(1, a_1 k_t^2 + b_1 k_t + c_1) \tag{2.22}$$

$$B(k_t) = Min(1, a_2 k_t^2 + b_2 k_t + c_2) \tag{2.23}$$

Now analyzing these results, one can understand the location's measurement response. But these generated plots can be affected by any type of error and climate disturbances. Both conditions tell us additional information about climate, measuring instruments, etc. Hence these conclusions are added with the historical database and suggested for further analysis. Some advanced guidelines and conditions will be discussed in detail in the following chapters. Combining cloud information results while doing quality control analysis, improves the accuracy of data.

2.7 SOLAR RADIATION GAP FILLING APPROACHES

After application of solar radiation data quality analysis, the data failing in the tests are considered as data gaps. These gap occurrences are twofold: (i) the data are not measured or stored for that period and (ii) data are found erroneous during quality control analysis. In both cases, data correction is suggested by using gap filling approaches. Now during the gap filling process, data missing intervals and their pattern are needed to kept in mind. As for each condition, different gap-filling approach is used. However, the final decision of data correction is to be expected from the competent authority. But if overlooked it may affect the site's actual average radiation calculation.

The most common and well established approach for filling the gaps, uses calculation of mean values or by data imputation (statistical models, skewness and kurtosis) techniques [Schneider, 2001; Pappas *et al.*, 2014; Meyer *et al.*, 2008; Hay, 1993; Annandale *et al.*, 2002]. In this thesis, all available gap filling approaches and averaging techniques studied are applied, to correctly predict the missing values. Another standard approach uses historical averaged data values, where the comparison is done between historical and available database values and accordingly missing values are filled. Now by involving the distance factor for an adjacent location or another meteorological measurement station, missing data is corrected. Furthermore, accurate approach uses computer simulation packages [Lundstrom, 2012] and automatic gap filling approaches [Long and Dutton, 2002; Kumar *et al.*, 2014; Tejera *et al.*, 2015] provided in the literature. All approaches and techniques discussed above are included in one platform. The selection of final gap-filling approach is based on the operator's decision. The following gap filling approaches can be used:

i) Data filling using interpolation (linear or polynomial) and averaging
ii) Data filling using models (Clear sky, Bird, Linke turbidity with standard parameters and averages)
iii) Data manually changed (offset included, cleaning event, instrument errors, and unavoidable reasons)

Each approach is unique and by applying any one, missing gaps in radiation can be filled. But in literature, no description of standard guideline is available for filling the data gaps and the results obtained by each approach vary. This requires some discussion.

2.8 STATISTICAL INDICATORS

These statistical ratios are used to compare the measured radiation database with theoretically calculated radiation values. Software packages like Excel, FORTRAN, SPSS and MATLAB have an in-built program for calculating these ratios. Therefore, one can carefully check the database by calculating these error values. Here both qualitative and quantitative procedures are used for measured value comparison, and results are plotted (scatter or histograms). Standard ratios/metrics used for calculating the possible errors in the measured database [Bland and Altman, 1996], are now listed here (see Eq.(2.21-2.25)).

(a) Mean ($\bar{x}$)

Sum of all values divided by number of samples

$$\bar{x} = (\sum_{i=1}^{n} x_i)/n \qquad (2.21)$$

(b) Standard Deviation (σ_n)

Measure to quantify the amount of variation or dispersion of a set of data values

$$\sigma_n = \sqrt{\frac{1}{n-1}\sum(\bar{x}_i - x_i)^2} \qquad (2.22)$$

(c) The Coefficient of Determination (R^2)

It is the ratio of the difference of predicted variation with a total variation of any time series (equation calculating R^2 is given in Eq.(2.23)). The acceptable range of this ratio is between zero to one, where zero means accurate prediction and one as not correlated at all. This parameter is not considered as a perfect technique to judge the provided dataset, hence it is verified with other error checking ratios for getting to any final decision.

$$R^2 = \frac{\sum(\bar{x}_i - \bar{x})}{\sum(x_i - \bar{x})} \qquad (2.23)$$

(d) Mean Bias Error (MBE)

It finds the trend in the measured database, by averaging all differences calculated at each measured interval. The results can be positive or negative according to location and time of year selected and can be expressed in terms of percentage or relative value. Here error is calculated by summing all errors present in the database, so it normalizes all database errors in one final error value. After estimation, if the total error is close to zero, then predicted value is considered accurate.

$$MBE = \frac{\sum(\bar{x}_i - x_i)}{n} \tag{2.24}$$

(e) Root Mean Square Error (RMSE)

It provides the degree or level of scatterness of the given model or dataset. Also checks the repeatability

$$RMSE = \sqrt{\frac{1}{n}\sum(\bar{x}_i - x_i)^2} \tag{2.25}$$

Where, x_i = Measurement value

$\bar{x}_i$= gap filling model output

2.9 SUMMARY

This chapter discussed a basic literature review available for identifying errors in solar radiation data. For detailed solar radiation potential estimation, ground station or other data analysis is needed. This includes local climatic study and analysis of solar radiation received at that location. Hence use of historical radiation databases (satellite, ground and relation-based) and their availability for any specific location is studied. High-frequency ground-based radiation database is preferred for detailed plant location study. Also, all instruments (installed at ground stations) used for measuring solar radiation components should follow standard guidelines, and selection is done on the basis of their measurement quality. Hence to achieve correct measured values, instruments need frequent calibration and their measurement system is standardized. For better understanding, all factors which affect the radiation received and quality of the measuring instruments are also discussed. This exercise helps in examining the theoretical possible radiation receiving at any location by using clear sky models. Now for identifying various errors present in the radiation database, detailed quality control guidelines and inter-correlations equations are suggested for each radiation component. The results derived from this analysis, help in understanding: location conditions, climate present there, instrument potential and stations current maintenance practices. Here the data failed during these quality control test, are considered as gaps and basic gap filling approaches is discussed. Finally, both original and modified database is compared using different statistical indicators, and solar plant's potential is evaluated. In the following chapters, one can see extensive discussion on each part given above and final objective of getting reliable radiation dataset for our selected location will be achieved.

…