Diffuse Radiation Calculation

Methods

by

Uday P. Singh

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved April 2016 by the
Graduate Supervisory Committee:

Nathan Johnson, Chair
Bradley Rogers
Govindasamy Tamizhmani

ARIZONA STATE UNIVERSITY

May 2016

ABSTRACT

With the recent rise in solar energy projects around the world there is an utmost need of proper estimation of solar energy. Significant error lands in estimation of energy production from the solar collectors due to the inaccurate assessment of solar energy. Substantial amount of error arises when the diffuse and direct part is separated from the global radiation using mathematical models. Diffuse radiation plays an important part in energy estimation from solar thermal and solar photovoltaic and is difficult to measure and in some parts of the developed world and in most parts of the developing world there is a scarcity of instruments. Diffuse radiation is estimated from global radiation by mathematical correlations computation, neural network and fuzzy logic. Present study validates which existing model works best in different geographical and sky conditions and also suggest a new method for diffuse radiation estimation. While most of the studies are focused on developing piecewise models for a particular country or particular location this study comes up with a global model i.e. continuous in nature and has been developed using seven US location data and four Global location data. Moreover, site specific continuous models are developed for ten locations. Results for the global and site specific models are better than the existing models in literature and also indicates that the models perform better in different sky conditions e.g. clear or cloudy sky. Study also shows that the continuous models perform equivalent or better than the piecewise models implemented. There are some intervals in which the existing models perform better. In those intervals, a best performing model is implemented while the remaining intervals e.g. 0.80 – 1.00 can still keep the newly obtained fit which will improve the overall performance of modeling techniques used in diffuse radiation estimation.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Nomenclature	Definition
G_{on}	Extraterrestrial radiation normal to the surface of the earth ($\frac{\mathrm{W}}{\mathrm{m}^2}$)
G_o	Extraterrestrial radiation on horizontal plane ($\frac{\mathrm{W}}{\mathrm{m}^2}$)
G_{sc}	Solar constant ($\frac{\mathrm{W}}{\mathrm{m}^2}$)
I	Global Horizontal Irradiance on a horizontal plane ($\frac{\mathrm{W}}{\mathrm{m}^2}$)
I_d	Diffuse radiation on a horizontal plane ($\frac{\mathrm{W}}{\mathrm{m}^2}$)
I_{dn}	Direct beam radiation on a horizontal plane ($\frac{\mathrm{W}}{\mathrm{m}^2}$)
I_d/I	Diffuse fraction
I_{dc}	Diffuse radiation calculated by models ($\frac{\mathrm{W}}{\mathrm{m}^2}$)
I_{dm}	Diffuse radiation measured ($\frac{\mathrm{W}}{\mathrm{m}^2}$)
k_t	Clearness index
b, B	Constant to be used in equation of time (°)
E_t	Equation of time (hours)
t_s	Local solar time (hours)
t_c	Local clock time (hours)
Z_c	Time zone
Λ	Longitude (°)
ϕ	Latitude (°)
δ	Solar declination angle (°)
$\omega, \omega_1, \omega_2$	Solar hour angles (°)
θ_z	Zenith angle (°)
N	Number of data points in model fit
$RMSE$	Root mean square error
RE	Relative error
R^2	R-squared value
T	Ambient air temperature (°C)
ρ	Relative humidity (%)
ρ_a	Absolute humidity ($\frac{\mathrm{gram}}{\mathrm{m}^3}$)

Chapter 1. INTRODUCTION

Energy generation in world is dominated by fossil fuels that resulted in a lot of research and development in the improvement of energy production methods from non – renewable energy sources. Improvement in conventional power generation methods does not stop the recent rise in pollution levels and global warming that makes renewable an alternative source of energy very attractive to most of the governments around the world e.g. Germany power generation from renewable stands around 26.2% of total power generation in 2014 with a potential of reaching 100% by 2050 (Szarka, N.). The total installed capacity in world from solar photovoltaic stands at 180 GW as on 2014 (Wirth and Schneider 2015) and continues to grow in future with China leading the installed capacity in Solar photovoltaic (Solar Power Europe 2015). The renewables in China will be cost competitive with fossil based generation by 2040 and will further increase the penetration of renewables in electric grid (Deloitte Report 2015). This will further drive research and development in area spanning from producing grid reliable equipment to the computational modeling techniques required for better estimation of energy from renewables and will plummet negatives of renewable electricity on electric grid. One of the major aspects in this domain is improvement in solar resource assessment to fully appreciate the use of solar power in off grid and on grid applications.

Diffuse radiation data for much of the world is computed using mathematical models e.g. China has 726 long term meteorological stations of which 98 measures global radiation and 19 measures diffuse radiation in China (Li et al. 2012). Moreover, most of the ground based data measurement is limited to the developed world and very scarce in developing world as the technology is still fledgling (Khalil & Shaffie 2013). Liu and Jordan (Liu &

Jordan 1960) have laid the foundation in computational modeling for assessment of diffuse radiation. Their work is further extended by Orgill and Hollands, Erbs et al. and other scholars around the globe.

Several statistical studies are conducted to explore different models in different world locations and various comparisons have been done to find best model that can be used for all location to improve the diffuse radiation estimation. This study is more complete in analyzing the models in different geographical conditions and different clearness index regions such as 0.30, 0.40 or 0.60 and regions such as 0.00 – 0.20, 0.40 – 0.60 etc. Moreover, annual comparison and daily comparisons are performed to look for the models' behavior in intermonth and intraday which was not performed before by any other scholar. A unique approach is adopted to improve the performance of the models and statistical comparison is done to find the better performing technique between continuous regression and piecewise regression. A regression analysis is done on eleven years of dataset obtained from World Radiation Data Center (WRDC 2016) to come up with one global model. Moreover, ten site specific models are proposed for the better estimation of a diffuse radiation. Present work helps to find the best method to calculate the diffuse radiation which will in turn improve the solar resource assessment hence the bankability of the solar thermal and solar photovoltaic system. This makes renewable energy system more attractive to the governments around the world and will lead to increase in the penetration of the renewables in electric market.

Chapter 2. REVIEW OF EXISTING DIFFUSE RADIATION MODELING TECHNIQUES

Pioneering work in the field of diffuse radiation calculation was done by the Liu and Jordan in 1960 when they explored different relations to estimate the diffuse radiation. A relationship was developed utilizing Hump Mountain, North Carolina data.

$$\frac{\overline{I_d}}{\overline{E_{tr}}} = 0.2710 - 0.2939 \times \frac{\overline{I_{dn}}}{\overline{E_{tr}}} \tag{2.1}$$

Where: $\overline{I_d}$ is a daily diffuse radiation on a horizontal plane. $\overline{I_{dn}}$ is a daily direct normal radiation on horizontal plane. $\overline{E_{tr}}$ is a daily extraterrestrial radiation on horizontal plane.

Work by Liu and Jordan was impressive and extensively used but the relation developed was based on a single site data and also did not give any hourly estimates. In addition, Ruth and chant (Ruth & Chant 1976) conducted their utilizing Canadian location and came up with a conclusion that the Liu and Jordan model significantly deviates from the measured values if location was changed.

In 1976, Orgill and Hollands (Orgill and Hollands 1971) proposed a new model for hourly estimation of diffuse radiation for a latitude between 43°N and 54°N. They obtained Toronto, Canada data of four years from period of Sept. 1967 – August 1971 and came up with a linear model. Other notable difference from Liu and Jordan was binning the data according to the clearness index which represented cloudy and uncloudy condition. Four years of data was binned in to three different intervals in which 32.4% data lies in $0 \leq k_t < 0.35$, 62% data lies in $0.35 \leq k_t \leq 0.75$ and 5.6% lies in $0.75 < k_t$. A linear model was fitted for the 32.4% and 62% data while the 5.6 % data was fitted with a constant. A

limited number of data points were available therefore it was not justified to use a complex model for this range at that time.

In 1981, Erbs et al. (Erbs et al. 1981) developed a new relationship between hourly diffuse fraction and the clearness index applying US data. Four US site was selected comprising of Fort Hood, Texas, Livermore, California, Raleigh, North Carolina, Maynard, Massachusetts and Albuquerque, New Mexico. The data for all the states were of different time period and interval e.g. some state data consisted of two year like Massachusetts and some of it was of 4 years like New Mexico. Erbs et al. also did the data binning according to the clearness index but implemented different clearness index bins for the regression modeling and also used a similar concept of fitting a constant in to a data in $0.8 < k_t$ as used by Orgill and Hollands. Erbs et al. not only utilized the clearness index for binning but also binned the data according to sunset hour angle which depends on the season. Models were analyzed implementing Mean bias error and Standard deviation to know how models behaved w.r.t. the measure values.

In 1982, Spencer (Spencer 1982) developed correlations for diffuse fraction which were dependent on the latitude of the place and the clearness index. The data constituted of five Australian sites of which the latitude varies from 20° S to 45° S. Absolute error was calculated and the correlation was compared with the Orgill and Hollands, Boes et al., Liu and Jordan and Bugler et al.

$$\frac{I_d}{I} = 0.94 + 0.0118 \times \phi - (1.185 + 0.0135 \times \phi) \times k_t,$$
$$0.35 < k_t < 0.75, \qquad 20° \text{ S} \leq \phi \leq 45° \text{ S} \tag{2.2}$$

In 1992, Reindl et al. (Reindl et al. 1992) further extended the work by taking data set from four European and two US locations, covering latitude from 28.4° N to 59.56° N. Some of the site data was of a single year and some was of two and three years. Twenty eight predictor variables were analyzed and stepwise regression was used to narrow it down to four. Those four predicator variables were temperature, relative humidity, solar altitude angle and clearness index. Different set of equations were developed using the same concept of binning the data according to the clearness index. Liu and Jordan and Orgill and Hollands developed linear relation while Erbs et al. and Reindl et al. developed the polynomial fits. Reindl et al. used composite residual sum square (CRSS). Reindl et al. correlation improved the fit by 14.4% over the Liu and Jordan fit.

In 1992, Al Riahi et al. (Al Riahi et al. 1992) also came up with correlations and collected data of two and a half years of Fudhaliyah, Iraq. Clearness index bins were used as was in the other studies and results were compared with Spencer, Erbs et al. and Orgill and Hollands. RMSE and Mean bias error was utilized for the comparison with the other models. Most of the studies done by 1992 implemented a common polynomial regression method and came uppolynomial piecewise models rather than exploring methods like continuous fit, rational fit or exponential fit.

In 1996, Janjai et al. (Janjai et al. 1996) developed a model for Bangkok, Thailand utilizing four locations: King Mongkut's Institute of Technology Thonbmi (KMllT) in the south, Silpakom University Snamchan Campus (SU) in the west and the Department of Meteorology (MET) in the southeast of Bangkok with a collection period of four, eight and seven years. They utilized the clearness index, temperature and relative humidity as a predictor variable to estimate diffuse radiation from global radiation. Error calculation was

done by RMSE and Mean bias error (MBE). Their model utilizing clearness index, temperature and relative humidity observed the better performance when compared with Erbs et al. and Liu and Jordan which only utilized clearness index for the estimation of diffuse radiation.

$$\frac{\overline{I_d}}{\bar{I}} = 0.913 - 0.146 \times \overline{k_t} - 0.014 \times \bar{\rho} + 0.0118 \times \bar{T} \tag{2.3}$$

Where: $\overline{I_d}$ is a monthly average daily diffuse radiation on a horizontal plane. $\bar{I}$ is a monthly average global horizontal radiation on horizontal plane. $\overline{k_t}$ is a monthly average daily clearness index. $\bar{\rho}$ is a monthly average daily relative humidity and $\bar{T}$ is a monthly average daily temperature.

In 2006, El-Sebbai et al. (El-Sebbai et al. 2006) came up with different regression models utilizing different predictor variables for estimation of diffuse radiation. Data of Jeddah, Saudi Arabia from 1996 – 2004 was analyzed and different fits were obtained utilizing different predictor variables such as clearness index, sunshine duration, temperature and relative humidity. El-Sebbai et al. also developed continuous models utilizing cloud coverage ratio as a predictor variable. The fits obtained were compared with each using MBE, RMSE and Mean percentage error (MPE). The models obtained were continuous and were linear.

$$\frac{I_d}{I} = -1.92 + 2.60 \times \left(\frac{S}{S_o}\right) + 0.06 \times T \tag{2.4}$$

$$\frac{I_d}{I} = -1.62 + 2.24 \times \left(\frac{S}{S_o}\right) + 0.332 \times \rho \tag{2.5}$$

$$\frac{I_d}{I} = 0.139 - 0.003 \times T + 0.896 \times \rho \tag{2.6}$$

Where: s is monthly average of daily bright sunshine hours (h), s_o is monthly average of maximum possible number of sunshine hours (h).

The MBE and RMSE is increased when relative humidity (Eq. 2.5) was used in place of Temperature in Eq. 2.4. Moreover, the RMSE and the MBE values were same for Eq. 2.4 and Eq. 2.6 indicating that the sunshine data could be replaced by relative humidity.

In 2008, Bolan et al. (Bolan et al. 2008) developed a rational model utilizing two Australian sites: Adelaide and Geelong, three European sites: Bracknell, Lisbon and Uccle and one Asian sites: Macau. A quadratic programming was also developed for removing the erroneous diffuse radiation values from data set. Absolute percentage error (APE) was implemented to check the model performance. Eq. 2.7 represents Bolan et al. model.

$$\frac{I_d}{I} = \frac{1.0}{1.0 + e^{-5.0+8.6\times k_t}} \tag{2.7}$$

In 2011, Li et al. (Li et al. 2011) developed continuous models utilizing thirty years of (1971 – 2000) monthly average daily Guangzhou data. They developed ten different models using clearness index, temperature, relative humidity, solar altitude angle and sunshine duration. Performance of the models was estimated by RMSE, MBE,R^2, Mean Absolute Percentage Error (MAPE) and Nashe – Sutcliffe equation (NSE).

$$\frac{I_d}{I} = 0.4461 + 0.4187 \times k_t - 0.8972 \times T + 0.0049 \times \rho + 0.3231 \times \sin(\alpha) \tag{2.8}$$

$$\frac{I_d}{I} = 0.5686 - 0.3724 \times (\frac{s}{s_o}) - 0.2991 \times log(\frac{s}{s_o}) + 0.0031 \times \rho + 0.2035 \times T \tag{2.9}$$

Where: $\sin(\alpha)$ is a solar altitude angle.

Their study found that the usage of solar altitude angle did not improve the performance of the diffuse radiation calculation, though the temperature and the relative humidity improved performance of models.

In 2016, Mohammadi et al. (Mohammadi et al. 2016) did an analysis to rank the usefulness of the predictor variables for the estimation of diffuse radiation. Ten parameters were selected e.g. sunshine duration, temperature, relative humidity, solar declination angle, water vapor pressure and clearness index etc. The dataset was obtained from city of Kerman, located in south central part of Iran. Adaptive neuro fuzzy inference system was applied to select the most influential parameter for the predication of diffuse radiation. RMSE, MBE, R^2, and Mean absolute bias error (MABE) were utilized for the performance measurement.

The findings observed by Mohammadi et al. indicated that the relative humidity is a least significant factor for the estimation of diffuse solar radiation for Kerman, Iran whereas the sunshine duration was considered as a most significant parameter for diffuse radiation estimation. Elminir et al. (Elminir et al. 2006) conducted a study for comparing the models generating using regression method with the models generated using artificial neural network technique (ANN). They found that the models generated by ANN technique for Egypt performed better than the models generated using regression techniques.

Chapter 3. DIFFUSE RADIATION CALCULATION METHODS

A paper to be submitted to *Applied Energy*

Uday P. Singh, Nathan G. Johnson

3.1 Introduction

In recent times renewable energy gained a lot of traction in different parts of the world. Most of the growth in renewables is driven by the government policies like providing subsidies for the renewable energy sources (Menanteau et al. 2003). Policies are structured to reduce the greenhouse gas emission (CO_2 emission increased by 52% globally from 1990 to 2012) (Deloitte 2015). The other driving factor for renewables growth is reducing the dependence on fossil fuels because of their limited availability and increasing cost of fossil fuels. This lead to a phenomenal growth in the installation of small and large scale renewable energy systems e.g. solar, wind etc. Initially development in solar energy is driven by the European nations with Germany leading the solar installation and generation till 2012 afterwards China, USA and Japan captured the majority of the market and currently driving the solar photovoltaic and thermal installation (Solar Power Europe 2015). On the other hand, growing economies like India set an ambitious target of 100 GW of solar installation by 2022 (Parkes 2016).

The potential of solar power can be further realized by analyzing the amount of solar energy received by the Earth. The total amount of incident solar power on Earth is 166,000 Terawatts (TW). Thirty percent is reflected back into space and approximately half (85,000 TW) is available for terrestrial collectors like solar thermal or solar photovoltaic systems (Abbott 2012). The world consumes 19.10 TWh (2012) of electricity per year; therefore, the total solar energy available is far more than the current electric energy needs (US EIA

2016). If 1 percent of the earth surface is reserved for solar power generating systems, and given 10% efficiency, then there will be sufficient electricity production for a population of 10 billion people with each person demanding 10 kW (Goswami et al. 2000). Recent estimates suggest that renewable energy capacity will be 3,930 GW by 2035 representing 31.2% of total power generation in which 690 GW will come from solar i.e. still a fraction of amount what Earth receives (Deloitte 2015). Still large-scale electricity generation from photovoltaic was limited because of high cost and long return on investment (Iyer 2015). Although, favorable conditions like easy to install, takes no time for start – up, no or very less moving part and machinery and its cost competitiveness to non–renewable sources of generation by 2040 in countries like China will further propel its deployment (Deloitte 2015).

Solar technology is new and developed lately compared to non – renewable generation e.g. terrestrial usage of solar arrays in US find its actual application in 1973/1974 after the oil shock (Goetzberger & Hoffmann 2005), therefore, many radiation collection laboratories are not equipped with instruments that measures all three component of radiation such as global horizontal radiation, diffuse horizontal radiation and direct normal radiation. Each component has its own usage like direct normal radiation finds its application in solar thermal (CSP) and concentrated photovoltaic technology (CPV) whereas solar photovoltaic relies on application of both. Majority of the countries relies on the mathematical models to compute the diffuse radiation values e.g. China has 726 long term meteorological stations of which 98 measures global radiation and 19 measures the diffuse radiation (Li et al. 2012). Moreover, the ground based data measurement is limited to the developed world and very scarce in developing world (Khalil & Shaffie 2013). Liu

and Jordan (Liu & Jordan 1960) has laid the foundation in computational modeling for an assessment of diffuse radiation. Their work is further extended by Orgill and Hollands (Orgill & Hollands 1976), Erbs et al. (Erbs et al. 1981), (Reindl et al. 1992) and other scholars around the globe.

Several statistical studies are conducted to develop models for a particular country but none came up with a model that can fit to different continents in world. Also, various comparisons have been done to find the best model that can be used for all location but none of the studies analyzed the performance of continuous and non – continuous i.e. piecewise models. This study not only utilizes statistical techniques to find the best performing model in different geographical location but also emphasizes on the models' behavior in different clearness index regions such as 0.00 – 0.20, 0.20 – 0.40, 0.40 – 0.60 or 0.80 – 1.00. Regression analysis is done on 11 years of dataset obtained from World Radiation Data Center to come up with one global model and 10 site specific models for the calculation of diffuse radiation. Present study is most complete in terms of validation of existing model such as performance in different clearness index conditions, yearly and daily evaluation, analyzing effects of temperature and relative on diffuse fraction, providing a new global model and new site specific model.

3.2 Background

3.2.1 *Classification of radiation and measurement techniques*

The radiations travelling through the space can be transmitted as it is or absorbed by the particles in the atmosphere or can be scattered by the particles like ozone, aerosol, water or dust in the atmosphere depending on wavelength. Based on the interaction of radiations

with the atmosphere it can be divided in to three components which are important to different technologies utilized for solar energy conversion.

Direct Normal (DNI) & Circumsolar Irradiance – It is the irradiance on a surface perpendicular to the vector from the observer to the center of the sun caused by radiation that did not interact with the atmosphere. This definition useful in atmospheric physics and radiative transfer models but in solar energy it is understood as the radiation received from a small solid angle centered on the sun's disk. The size of this "small solid angle" for DNI measurements is recommended to be 5×10^{-3} sr (corresponding to and approximate 2.5 degree half angle). Whereas circumsolar region closely surrounds solar disk and looks very bright, the radiation coming from this region is called circumsolar irradiance. DNI plays a vital role in concentrating solar power/photovoltaic. DNI is measured by a Pyrheliometer, the receiving surfaces of which is arranged to be normal to the solar direction (Sengupta et al. 2015).

Diffuse Horizontal Irradiance (DHI) – This is the scattered or reflected part of the DNI by the particles present in the atmosphere or the light reflected by the earth surface also termed as albedo is a part of DHI. Rayleigh, Mie and Young explained scattering of light that explained why sky looks blue and why sun looks red or yellow during the different time of the day (Kerker 1993). DHI is measured by the Pyranometer shaded with a shade ring.

Global Horizontal Irradiance (GHI) – Sum of DNI and DHI is termed as GHI. It is calculated using Eq. 3.1.

$$GHI = DNI \times cos(\theta_z) + DHI \tag{3.1}$$

Ground-based instruments widely used for collecting solar data like solar radiation intensity are Pyranometers and Pyrheliometers (Thekaekara 1976). World Radiation Data Center has a collection of solar data – e.g., global horizontal radiation, direct normal radiation, diffuse horizontal radiation – for most countries, while National Oceanic and Atmospheric Administration measures solar data for 7 sites in the United States at 1 minute resolution (NOAA 2016). Baseline Solar Radiation Network (WRMC–BSRN 2016), Fluxnet Network (ORNL DAAC 2015) and Swiss Institute of Meteorology also collects solar data. Moreover, there are models converting satellite images in to different radiation components and giving better estimation of radiation components compared to estimation done for a site using nearby ground station. A comparative study is done on Geomodel in Bratislava (SolarGis), Helioclim Soda (Heliostat 3v3), 3 Tier Company, University of Oldenburg (EnMetSol-Solis and EnMetSol-Dumortier) and IrSolAv by P. Ineichen in 2011 and confirmed that SolarGis and EnMetSol holds the better results for radiation estimation (Inchien 2011).

3.2.2 *Uncertainty in radiation measurements*

Importance of good solar data is realized when economic feasibility and system sizing for photovoltaic and solar thermal is done. Solar resource assessment directly affects the project cost and quality (Gueymard & Wilcox 2009). Also, the project financers are interested in a renewable energy project if they see higher returns in a shorter period of time with less uncertainty. Therefore, reducing the sources of energy uncertainty form

photovoltaic and solar thermal is important. These sources of energy uncertainty are enumerated by Marie Schnitzer et al. (Schnitzer et al. 2012).

- Annual Degradation (0.50 – 1.00%)
- Transposition to Plane of Array (0.50 – 2.00%)
- Energy Simulation & Plant Losses (3.00 – 5.00%)
- Solar resource uncertainty (5.00 – 17.00%)

Enormous emphasis is made on good data collection and can be seen in SOLRMAP (NREL Website) that consists of high quality solar data for particular locations which can be used by solar thermal projects. Furthermore, there are physical models that estimate radiation values based on atmospheric parameters like turbidity, and aerosol etc. and splits the diffuse and direct radiation value from the measured GHI. Models implemented for the separation of DHI and DNI from GHI are major sources of uncertainty (Gueymard 2009).

3.2.3 *Studies conducted for the calculation of DHI on horizontal plane*

Measuring DHI component of the radiation is a complex process. First methods requires Pyranometer with a small shading disc following the sun's motion. The technique is costly and requires a lot of maintenance. The second method uses a shadow ring/band. The ring/band is parallel to the sun path and hence blocks the DNI. This method not only blocks the DNI but also blocks the part of DHI reaching the receiver hence poor estimation (Gueymard & Myers 2009). In addition, there is a non – uniform temperature response, cosine error and thermal imbalance. Consequently, there is a need of mathematical models proposed by Drummond (1956), Steven (1984), Lebaron et al. (1990), Batles et al. (1995) and Muneer and Zhang (2002) to correct the DHI values (Sánchez et al. 2012). Considering

the complexities associated with the measurement scholars proposed alternate methods of estimation of DHI from GHI. This study compares the model which are widely prevalent in solar resource assessment and currently utilized in the photovoltaic simulation software like Homer etc. Also, a new model is generated and compared with these established models.

1) Orgill and Hollands,

$$\frac{I_d}{I} = 1.0 - 0.249 \times k_t \; for \; 0 \leq k_t < 0.35 \tag{3.2}$$

$$\frac{I_d}{I} = 1.577 - 1.84 \times k_t \; for \; 0.35 \leq k_t \leq 0.75 \tag{3.3}$$

$$\frac{I_d}{I} = 0.177 \; for \; 0.75 < k_t \tag{3.4}$$

2) Erbs et al.

$$\frac{I_d}{I} = \; 1.0 - 0.09 \times k_t \; for \; k_t \leq 0.22 \tag{3.5}$$

$$\frac{I_d}{I} = 0.9511 - 0.1604 \times k_t + 4.388 \times k_t^2 - 16.638 \times k_t^3 + \; 12.336 \times k_t^4 \tag{3.6}$$

$$for \; 0.22 < k_t \leq 0.80$$

$$\frac{I_d}{I} = \; 0.165 \times k_t \; for \; 0.80 < k_t \tag{3.7}$$

3) Reindl et al.

Constraint: $I_d/I \leq 1$

$$\frac{I_d}{I} = 1.020 - 0.248 \times k_t \; for \; 0 \leq k_t \leq 0.30 \tag{3.8}$$

$$\frac{I_d}{I} = 1.45 - 1.67 \times k_t for \; 0.3 < k_t < 0.78 \tag{3.9}$$

$$\frac{I_d}{I} = 0.147\ for\ 0.78 \leq k_t \tag{3.10}$$

4) Al-Riahi et al.

$$\frac{I_d}{I} = 0.932\ for\ k_t < 0.25 \tag{3.11}$$

$$\frac{I_d}{I} = 1.293 - 0.249 \times k_t\ for\ 0.25 \leq\ k_t \leq 0.70 \tag{3.12}$$

$$\frac{I_d}{I} = 0.151\ for\ 0.7 < k_t \tag{3.13}$$

These relations utilized regression analysis in which diffuse fraction (diffuse fraction is defined as a ratio of diffuse horizontal radiation to the global horizontal radiation) is a function of k_t (k_t is defined as the ratio of extraterrestrial radiation and global horizontal radiation). There are models proposed by Reindl et al., Li et al. (Lie et al. 2011) which considered parameter for example relative humidity and temperature for the estimation of diffuse component of light.

3.3 Methodology

3.3.1 *Solar resource data*

Global horizontal and diffuse radiation for Argentina, Australia, Germany, Japan and US are taken from the World Radiation Data Center (WRDC 2016). Temperature and Relative Humidity data for Germany is gathered from Weather Underground (WU 2016). A short python script has been developed and implemented to access data from Weather Underground. A data access key has been issued by Weather Underground to make hundred calls in a minute and five thousand calls in a day. Each location has its unique id that is required to access the data.

All data set is of hourly resolution. The radiation data is further filtered by replacing non-existent values with null value and then ignoring null values in regression analysis. Data points with $0 < I_d/I \leq 1$ is considered. This study covers the behavior of diffuse radiation models in four different continents and tries building a new model which can fit in all locations. Table 1 gives detail indicating location variability with annual average k_t. Year with most complete dataset has been selected. For example, the 2013 and 2014 data from Germany is incomplete resulting in a selection of 2012 for regression analysis. Negative sign on a latitude column indicates that latitude of location is in southern hemisphere whereas no sign is considered as a positive which indicates northern hemisphere. Similarly, the negative Longitude and the negative time zone indicates west of GMT while no sign considered as a positive that represents east of GMT.

Table 1. Site details and data from World Radiation Data Centre.

Location	**ϕ**	**Λ**	**Z_c**	**k_t**	**Year**
Hohenpeissenberg, Bavaria, Germany	47.80	11.00	1	0.44	2012
Wagga Wagga, New South Wales, Australia	-35.60	147.46	10	0.61	2014
Sapparo, Hokkaido, Japan	43.07	141.35	9	0.44	2014
Ushuaia, Tierra del Fuego, Argentina	-54.82	-68.33	-3	0.37	2014
Sioux Falls, South Dakota, USA	43.58	-96.75	-6	0.44	2014
Fort Peck, Montana, USA	48.31	-105.10	-7	0.58	2014
Bondville, Illinois, USA	40.72	-77.94	-5	0.60	2014
Boulder, Colorado, USA	40.13	-105.24	-7	0.59	2014
Desert Rock, Nevada, USA	36.62	-116.03	-8	0.68	2014
Goodwin Creek, Mississippi, USA	34.23	-89.87	-6	0.55	2014
Rock Spring, Pennsylvania, USA	40.72	-77.93	-5	0.49	2014

3.3.2 *Extraterrestrial radiation calculation*

Hourly extraterrestrial radiation data on horizontal plane is calculated for US sites and for other locations around the world utilized for the model development in this study. Mathematical procedures provided in Duffie and Beckman are employed for the calculations (Duffie & Beckman 1980).

$$\delta = 23.45 \times \frac{\sin\left(360 \times (284 + day)\right)}{365} \tag{3.14}$$

$$b = 2 \times 3.14 \times \frac{day}{365} \tag{3.15}$$

$$B = 360 \times \frac{(day - 1)}{365} \tag{3.16}$$

$$G_{on} = G_{sc} \times (1.00011 + 0.034221 \times \cos(b) + 0.001280 \times \sin(b) + 0.000719 \times \cos(2 \times b) + 0.000077 \times \sin(2 \times b)) \tag{3.17}$$

$$E_t = 3.82 \times (0.000075 + 0.001868 \times \cos(B) - 0.032077 \times \sin(B) - 0.014615 \times \cos(2 \times B) - 0.04089 \times \sin(2 \times B)) \tag{3.18}$$

$$t_s = t_c + \left(\frac{\lambda}{15}\right) - Z_c + E_t \tag{3.19}$$

$$\omega = (t_s - 12) \times 15 \tag{3.20}$$

$$E_{tr} = ((12/3.14) \times G \times ((\cos(\phi) \times \cos(\delta) \times (\sin(\omega_1) - \sin(\omega_2)) + (0.0174 \times (\omega_1 - \omega_2) \times \sin(\phi) \times \sin(\delta)))))) \tag{3.21}$$

3.3.3 *New model development*

Initially all the existing models are compared for identifying best among them and then a new model is generated doing a continuous and a piecewise regression. A global model is developed by doing a regression analysis on a data set of selected set of countries based

on their geographical locations. Data is divided in five regions based on k_t values 0.0 – 0.20, 0.20 – 0.40, 0.40 – 0.60, 0.60 – 0.80, 0.80 – 1.00 and a continuous fit is performed. This segmentation of data helped to determine where the existing and newly developed models are not performing well therefore a new fit can be applied in the regions of low R^2 values and high root mean square values.

$$I_{dc} = I \times (a \times k_t + b) \tag{3.22}$$

$$I_{dc} = I \times (a \times k_t^2 + b \times k_t + c) \tag{3.23}$$

$$I_{dc} = I \times (a \times k_t^3 + b \times k_t^2 + c \times k_t + d) \tag{3.24}$$

$$I_{dc} = I \times (a \times k_t^4 + b \times k_t^3 + c \times k_t^2 + d \times k_t + e) \tag{3.25}$$

Also, the k_t intervals existing in the studies of Erbs et al., Orgill and Hollande, Al Riahi et al and Reindl et al. are explored and new fits are applied in the existing intervals to determine which intervals are the best and why the these intervals are selected.

3.3.4 *Error calculation*

The measured value of the diffuse radiation is compared against the calculated value of the diffuse radiations using models. The error calculation is completed using RMSE, R^2 and RE values.

$$RMSE = \left(\sum_{i=1}^{N} \frac{\left(I_{dm,i} - I_{dc,i}\right)^2}{N_i} \right)^{0.5} \tag{3.26}$$

$$R^2 = 1 - \left(\sum_{i=1}^{N} \frac{\left(I_{dm,i} - I_{dc,i}\right)^2}{(I_{d,i} - mean(I_{dc})^2} \right) \tag{3.27}$$

$$RE = \frac{(I_{dm} - I_{dc})^2}{I_{dm}} \tag{3.28}$$

3.4 Results and Analysis

3.4.1 *Comparison of results from existing models for new locations*

An annual comparison is completed for the models to analyze which model fits best for all the locations or most of the locations and can be applied worldwide for diffuse radiation calculation on horizontal plane for solar photovoltaic and solar thermal power generation. RMSE and R^2 values are calculated for nine different locations for four different models which are mentioned in Table 2.

Table 2. Comparison of diffuse radiation models for different locations using RMSE (R^2).

Location	Orgill and Hollands	Erbs et al.	Reindl et al.	Al-Riahi et al.
Hohenpeissenberg, Bavaria, Germany	0.153 (0.793)	0.155 (0.775)	0.154 (0.785)	0.517 (0.563)
Wagga Wagga, New South Wales, Australia	0.175 (0.675)	0.173 (0.676)	0.169 (0.681)	0.360 (0.512)
Sapparo, Hokkaido, Japan	0.125 (0.846)	0.127 (0.828)	0.129 (0.829)	0.467 (0.542)
Ushuaia, Tierra Del Fuego, Argentina	0.261 (0.619)	0.270 (0.622)	0.262 (0.666)	0.450 (0.518)
Sioux Falls, South Dakota, USA	0.158 (0.727)	0.160 (0.700)	0.159 (0.704)	0.236 (0.473)
Fort Peck, Montana, USA	0.152 (0.768)	0.154 (0.746)	0.153 (0.754)	0.261 (0.512)
Bondville, Illinois, USA	0.133 (0.784)	0.136 (0.761)	0.135 (0.774)	0.402 (0.492)
Boulder, Colorado, USA	0.165 (0.659)	0.167 (0.640)	0.164 (0.653)	0.302 (0.442)
Desert Rock, Nevada, USA	0.143 (0.652)	0.139 (0.655)	0.136 (0.671)	0.166 (0.464)

Based on the values of RMSE and R^2 given in Table 2 the best model is Orgill and Hollands model that fits best for six locations. Orgill and Hollands model not only captures the variability in weather by performing well in different annual average k_t but also captures the geographical variability by better than others existing models in three

international locations: Germany, Japan, Argentina and three US locations: Illinois, Montana, South Dakota. Reindl et al. model is a second best that fits better than existing models for two US locations: Nevada, Colorado and one Australian location: New South Wales. Al Riahi model is least efficient compared to the other models. Results can be further confirmed by graphical analysis completed in MATLAB.

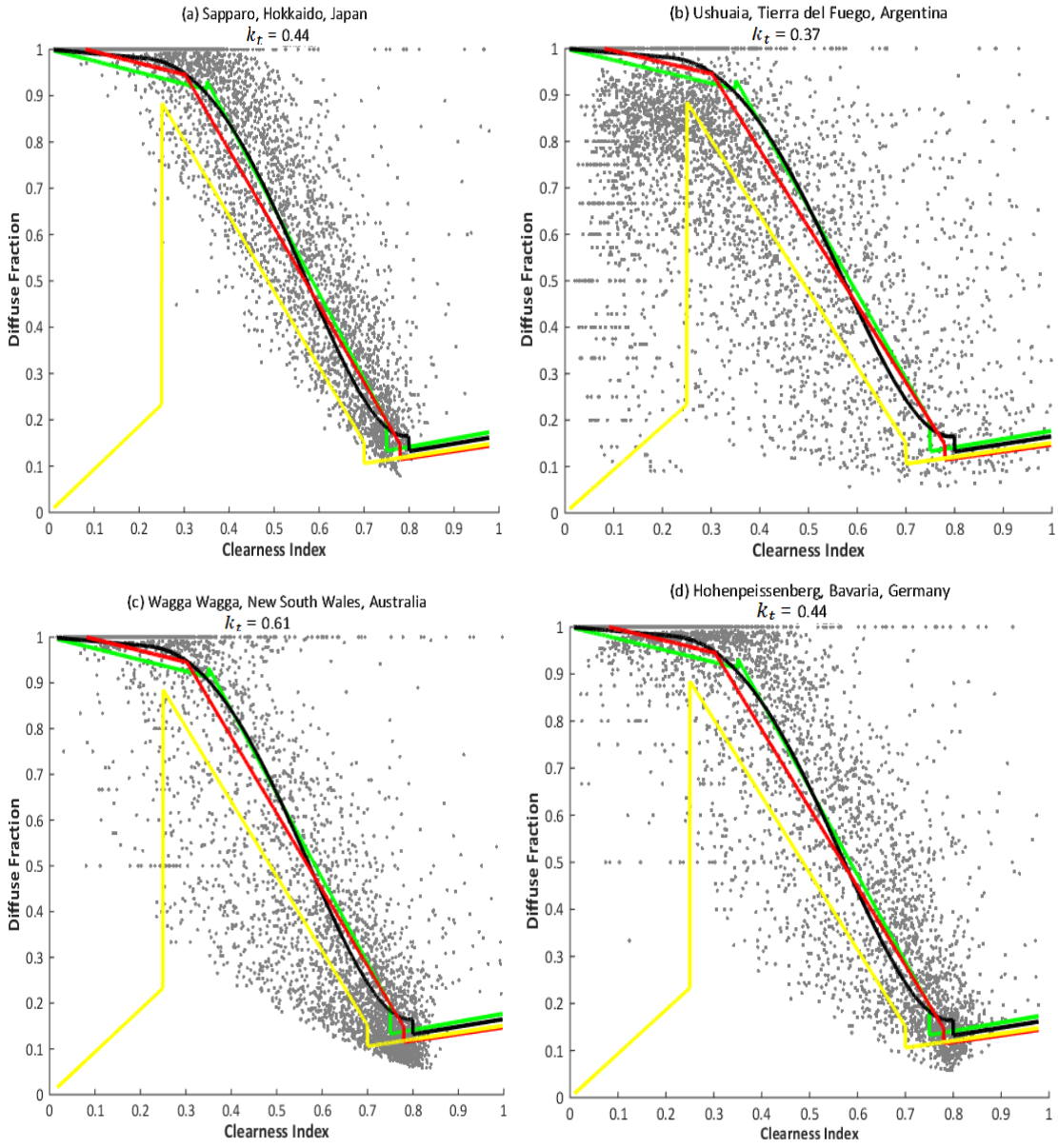

Figure 1. Model comparison for Global locations.

Figure 1 evaluates the performance of models in different climatic conditions around the globe that rules out the implementation of Al Riahi et al. model by visual inspection. Al Riahi et al. model does not fit well to global locations in an interval between 0.00 – 0.20 and can be seen in the graphs for all location, also, in the interval between 0.20 – 0.70, fit is not close to the other fits and lies far below from rest of the fit lines in that region. Calculated values of diffuse fraction by Al Riahi et al. model starts from 0.00 and then assumes a straight line at $k_t = 0.25$ which does not follow the k_t distribution with respect to diffuse fraction while the rest of the models follow the same pattern as diffuse measured values follows. Al Riahi et al. model has less R^2 value and large RMSE value compared to rest of the models and deviates far more from original values. Orgill and Hollands, Erbs et al. and Reindl et al. performs almost similar on the annual scale and the variability in their performance can only be observed by the RMSE and R^2 given in Table 2. Orgill and Hollands work best in three out of four global locations while Reindl et al. only fits best to one global location. Moreover, Figure 1(b) (Argentina) indicates a low annual k_t and values are scattered all over the plot which is difficult to capture by the models resulting in high RMSE and low R^2 for all the models compared to the other locations for which comparison has been done.

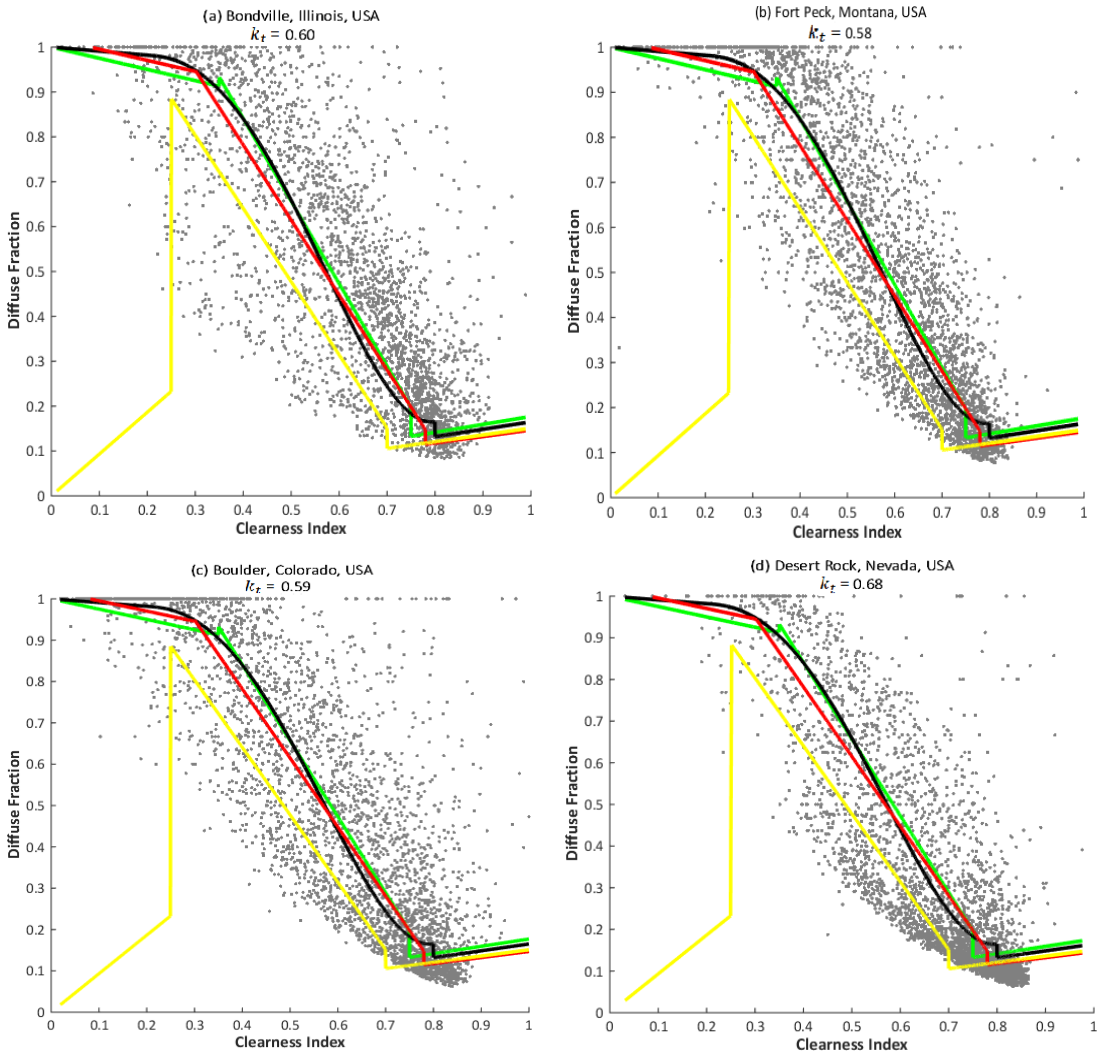

Figure 2. Model comparison for US locations.

Figure 2 also rules out the implementation of Al Riahi et al. model because it does not fit well to US locations in an interval between 0.0 – 0.20 and can be seen in the graphs for all location, also, in the interval between 0.20 – 0.70 fit is not close to the other fits and lies far below compared to the rest of the fits. The fit for US location repeats its behavior as observed in global locations. Performance by the models such as Orgill and Hollands, Erbs et al. and Reindl et al for US locations is similar to global locations. Orgill and Hollands

work best for three US locations while Reindl et al. work best for two US location. Significant deviation from measured diffuse fraction and the calculated diffuse fraction for all models lies in the region of k_t (0.80 – 1.00). The deviation for high values of k_t is further analyzed by doing a daily comparison for unique k_t values in below section.

3.4.2 *Comparison of clearness index on model results*

Figure 3. Model comparison for different locations with varying k_t values.

For this study, the region between 0.00 – 0.20 for k_t is considered to be a low k_t region, region between 0.20 – 0.50 considered to be a medium k_t region and 0.50 – 1.00 is considered to be a high k_t region. Selected days are based on k_t values to understand how the behavior of models are affected by the magnitude of k_t. Figure 3 clearly indicates that for high k_t values models are not performing well compared to medium and low k_t values. The relative error is high for Germany (k_t = 0.58), South Dakota (k_t = 0.55), Colorado (k_t = 0.53) and the lines are farther from x axis representing high magnitude in relative error. For low and medium k_t, the lines are particularly flat and are close to the x axis. This is the case for Pennsylvania (k_t = 0.26), Montana (k_t = 0.40), Illinois (k_t = 0.45) and Argentina (k_t = 0.24). Therefore, this high error region resulted due to higher value of k_t needs to be improved for the existing models. Findings are consolidated in Figure 4, where Bavaria, Germany is selected for a comparison and a particular time period is selected so that the position of the Sun in sky won't affect the duration and magnitude of extraterrestrial radiation on a horizontal plane received by earth hence performance of the models. Day 11 is a sunny day with high k_t and the performance of the models are worst compared to the Day 4 which has a low k_t value and Day 13 which has a medium k_t value.

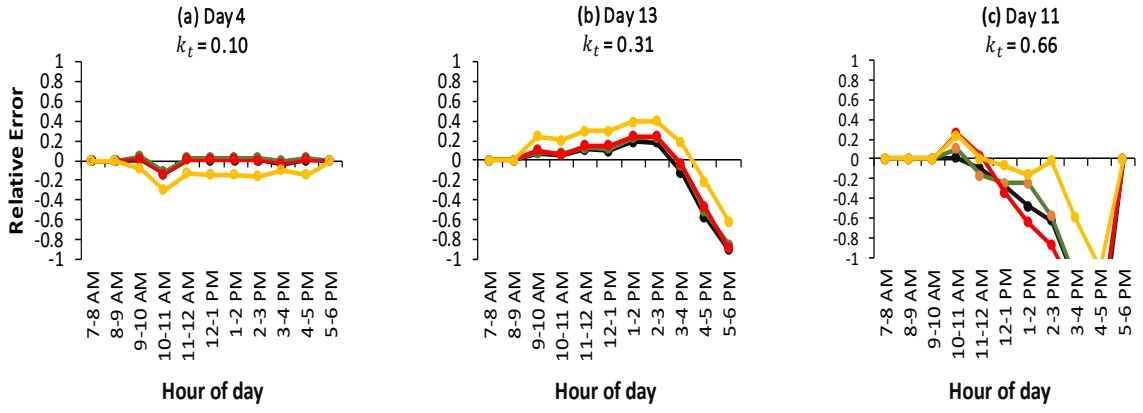

Figure 4. Model comparison for Bavaria, Germany for different k_t values.

3.4.3 *New models using continuous and piecewise fit*

Continuous fit

A continuous fit is implied utilizing a year's dataset of ten locations given in Table 1. A new model is obtained.

$$I_{dc} = I \times (8.307 \times k_t^4 - 11.240 \times k_t^3 + 2.729 \times k_t^2 - 0.123 \times k_t + 0.8846) \quad (3.29)$$

The newly developed model significantly improves the performance in high k_t region (0.80 – 1.00) for all locations around the world. Moreover, it improves the assessment in low k_t (0.00 – 0.02) region. This is illustrated in the below Figure 5.

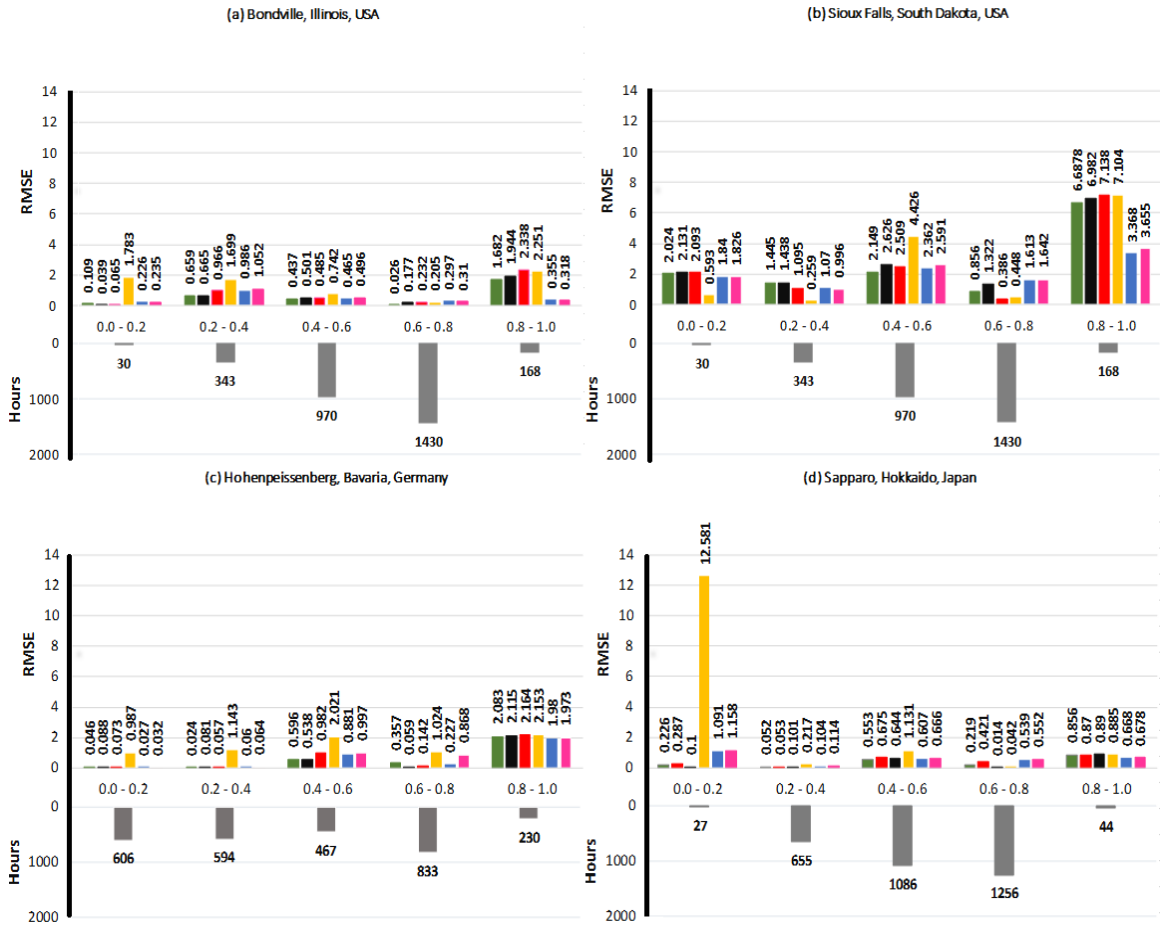

Figure 5. Comparison of model with Singh's US and Singh's global model for global locations.

In Figure 5(a), improvement in RMSE achieved by newly developed model (Singh Global) is 20.70% compared to the best model i.e. Orgill and Hollands in the region of 0.80 – 1.00 with 44 data points corresponding to 44 sun hours in a year. In Figure 5(b), for interval 0.80 – 1.00, improvement achieved by Singh's model is 5.20% compared to the best model i.e. Orgill and Hollande model. In Figure 6(c), for interval 0.80 – 1.00, improvement attained is 81% over the best model i.e. Orgill and Hollande for 230 sun hours in year. In Figure 6 (d), for interval 0.80 – 1.00, the assessment is improved by over 45% for 117 hours in a year. This is a very strong indication of using newly developed model

for calculation of diffuse radiation on a horizontal plane in interval of 0.80 – 1.00 and 0.00 – 0.20. Moreover, the new model improves fit for Germany in interval of 0.00 – 0.20 with improvement of 30.43% and for South Dakota with improvement of 9.78%.

Furthermore, a new site specific model is generated for Montana using a year's data (2014

) from World Radiation Data Center of Montana, USA and existing models in literature are compared with this newly developed continuous model. Four different regression fit is utilized: linear, quadratic, cubic and quartic. Quartic fit has been selected because of better RMSE and R^2 values. Figure 6 is a pictorial representation of comparison of the new model with the existing models.

$$I_{dc} = I \times (11.42 \times k_t^4 - 16.84 \times k_t^3 + 6.104 \times k_t^2 - 1.006 \times k_t + 1.026) \qquad (3.30)$$

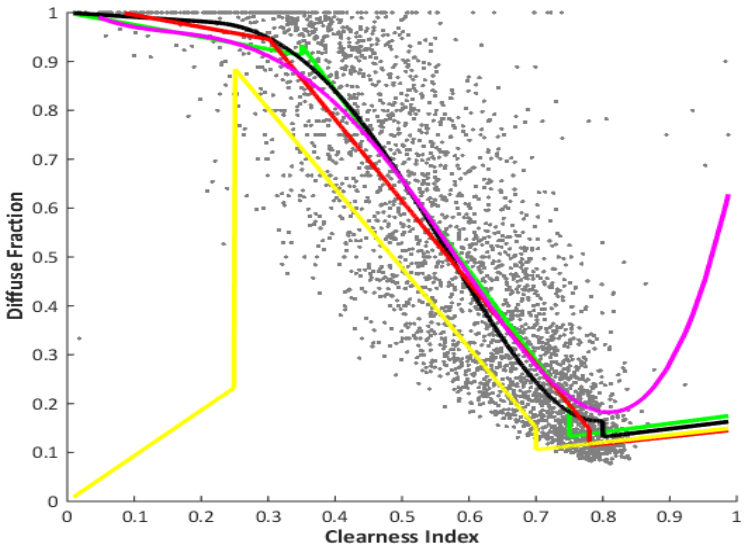

Figure 6. Model comparison for Montana, US.

An improvement of 3% in RMSE over the best performing model i.e. Orgill and Hollands is achieved giving an indication of developing site specific continuous model utilizing site specific data set rather than using a common piecewise model such as Erbs et

al. etc. This finding is consolidated by completing regression analysis for the ten locations that comprises of six US locations and four international locations. A continuous fit is generated for each site utilizing site specific data. The data is obtained from World Radiation Data Center. These site specific models are continuous in nature and performs better than established piecewise models. Table 3 and Table 4 gives the details about the RMSE and R^2 value.

Table 3. Fit results of diffuse radiation models for US locations using RMSE (R^2).

Fit type	Boulder, Colorado, USA	Bondville, Illinois, USA	Fort Peck, Montana, USA	Desert rock, Nevada, USA	Rock Spring, Pennsylvania, USA	Sioux Fall, South Dakota, USA
Linear	0.166 (0.715)	0.149 (0.713)	0.159 (0.713)	0.138 (0.704)	0.145 (0.785)	0.264 (0.165)
Quadratic	0.166 (0.716)	0.148 (0.716)	0.156 (0.723)	0.136 (0.712)	0.135 (0.813)	0.260 (0.189)
Cubic	0.161 (0.734)	0.143 (0.735)	0.152 (0.739)	0.133 (0.724)	0.131 (0.824)	0.260 (0.190)
Quartic	0.160 (0.739)	0.140 (0.745)	0.150 (0.744)	0.132 (0.725)	0.130 (0.828)	0.259 (0.120)

Table 3 provides the RMSE and R^2 for continuous models developed for US locations. For Colorado the best performing piecewise fit gives a RMSE of 0.164 (Table 2) while the newly developed continuous quartic fit gives 0.16 an overall improvement of 2.40%. Similarly, for Nevada an improvement of 2.90% is noted. Comparison is run on five US locations in which newly developed models perform better on three locations and piecewise models still do better on rest two locations i.e. Illinois and South Dakota.

Table 4. Fit results of diffuse radiation models for global locations using RMSE (R^2).

Fit type	Ushuaia, Tierra del Fuego, Argentina	Wagga Wagga, New South Wales, Australia	Sapparo, Hokkaido, Japan	Hohenpeissenberg, Bavaria, Germany
Linear	0.212 (0.349)	0.162 (0.680)	0.142 (0.783)	0.163 (0.769)
Quadratic	0.211 (0.392)	0.162 (0.680)	0.126 (0.827)	0.155 (0.790)
Cubic	0.205 (0.392)	0.156 (0.705)	0.122 (0.838)	0.150 (0.804)
Quartic	0.205 (0.392)	0.155 (0.708)	0.120 (0.845)	0.148 (0.809)

Table 4 provides the RMSE and R^2 for continuous models developed for global locations. Best performing piecewise model for Germany and Japan gives RMSE of 0.153 and 0.125 (Table 2) while the RMSE obtained by newly developed continuous models are 0.148 and 0.120, an improvement of 3.27% and 4.17%. Similarly, for Australia and Argentina, improvement of 8.28% and 21.45 % is noticed.

Piecewise fit

Furthermore, a comparative analysis is done between the piecewise models and continuous models. Three different locations are selected: Bavaria, Germany, South Dakota, USA and Illinois, USA. Intervals utilized are taken from the existing models but new fits such as constant, linear and quadratic are performed on the data set obtained from World Radiation Data Center. Table 5 tells that about gives the results of comparative analysis done between the piecewise and the continuous models for Bavaria, Germany.

Table 5. Piecewise fits for Bavaria, Germany with model results shown using RMSE (R^2).

Model	Interval	Constant	Linear	Quadratic	Hours
Erbs et al. Model Discontinuity	0.00 – 0.22	0.040 (-)	0.041 (0.010)	0.041 (0.013)	679
	0.22 – 0.80	0.323 (-)	0.157 (0.765)	0.153 (0.776)	1817
	0.80 – 1.00	0.168 (-)	0.155 (0.156)	0.154 (0.166)	229

Al Riahi et al. Model Discontinuity	0.00 – 0.25	0.044 (-)	0.043 (0.019)	0.043 (0.023)	792
	0.25 – 0.70	0.264 (-)	0.168 (0.595)	0.166 (0.606)	1168
	0.70 – 1.00	0.152 (-)	0.151 (0.012)	0.142 (0.135)	765
Orgill and Hollands Model Discontinuity	0.00 – 0.35	0.078 (-)	0.074 (0.096)	0.073 (0.117)	1083
	0.35 – 0.75	0.271 (-)	0.170 (0.604)	0.171 (0.604)	1105
	0.75 – 1.00	0.149 (-)	0.147 (0.029)	0.242 (0.102)	537
Reindl. et al. Model Discontinuity	0.00 – 0.30	0.064 (-)	0.062 (0.065)	0.061 (0.091)	943
	0.30 – 0.78	0.296 (-)	0.164 (0.694)	0.163 (0.696)	1414
	0.78 – 1.00	0.156 (-)	0.147 (0.110)	0.145 (0.138)	368

Table 5 indicates that the best value of RMSE and of R^2 are in interval of 0.22 – 0.80 for Erbs et al. model i.e. 0.153 and 0.776. Rest of the intervals such as 0.25 – 0.70, 0.35 – 0.75 and 0.30 – 0.78 do not provide a low RMSE value or high R^2 value compared to Erbs's region. Though, other existing models perform better in different intervals like 0.00 – 0.35 or 0.00 – 0.25. Orgill and Hollands model has a very low RMSE value in 0.00 – 0.25 interval i.e. 0.073. Additionally, lowest RMSE obtained from piecewise modeling is 0.153 for 1817 hours that lies in Erb's region on the other hand the quartic continuous model gives a RMSE of 0.148 which indicates an overall improvement of 3.27%.

Table 6 and Table 7 also confirms that continuous quartic models are comparable or even better than the piecewise linear or quadratic models and can be replaced by the continuous models. This is a very interesting finding and can be further looked upon by performing comparison in other locations like Illinois and South Dakota.

Table 6. Piecewise fits for Illinois, USA with model results shown using RMSE (R^2).

Model	Interval	Constant	Linear	Quadratic	Hours
Erbs et al. Model Discontinuity	0.00 – 0.22	0.065 (-)	0.064 (0.035)	0.065 (0.036)	43
	0.22 – 0.80	0.269 (-)	0.145 (0.712)	0.144 (0.715)	2726

	0.80 – 1.00	0.115 (-)	0.111 (0.069)	0.110 (0.092)	167
Al Riahi et al. Model Discontinuity	0.00 – 0.25	0.071 (-)	0.070 (0.043)	0.071 (0.043)	58
	0.25 – 0.70	0.236 (-)	0.156 (0.566)	0.156 (0.567)	1948
	0.70 – 1.00	0.120 (-)	0.113 (0.104)	0.113 (0.148)	930
Orgill and Hollands Model Discontinuity	0.00 – 0.35	0.098 (-)	0.094 (0.089)	0.093 (0.100)	231
	0.35 – 0.75	0.239 (-)	0.154 (0.588)	0.154 (0.588)	2189
	0.75 – 1.00	0.099 (-)	0.049 (0.483)	0.096 (0.075)	516
Reindl. et al. Model Discontinuity	0.00 – 0.30	0.076 (-)	0.074 (0.064)	0.074 (0.073)	111
	0.30 – 0.78	0.258 (-)	0.148 (0.670)	0.148 (0.671)	2530
	0.78 – 1.00	0.098 (-)	0.073 (0.094)	0.093 (0.092)	297

Table 6 also gives the similar results to Table 5. Quadratic fit in the Erbs's region of 0.22 – 0.80 has a minimum RMSE of 0.144 and maximum R^2 value of 0.715. Whereas the continuous quartic model is applied, it gives a RMSE of 0.140 and the R^2 value of 0.745. An improvement of 2.78% has been observed utilizing a continuous quadratic fit over piecewise quadratic fit. Improvement is further increased by using the quartic model. This signifies the importance of using a continuous quartic models rather than a piecewise quadratic or linear model as used in studies like Orgill and Hollands, Al Riahi et al. etc. Comparison is extended for one more location to confirm whether the findings are coherent or not.

Table 7. Piecewise fits for South Dakota, USA with model results shown using RMSE (R^2).

Model	Interval	Constant	Linear	Quadratic	Hours
Erbs et al. Model Discontinuity	0.00 – 0.22	0.254 (-)	0.234 (0.154)	0.232 (0.167)	297
	0.22 – 0.80	0.276 (-)	0.268 (0.057)	0.268 (0.058)	884
	0.80 – 1.00	0.253 (-)	0.252 (0.011)	0.248 (0.057)	116

Al Riahi et al. Model Discontinuity	0.00 – 0.25	0.259 (-)	0.241 (0.138)	0.238 (0.160)	339
	0.25 – 0.70	0.277 (-)	0.274 (0.024)	0.271 (0.022)	738
	0.70 – 1.00	0.240 (-)	0.239 (0.009)	0.237 (0.033)	219
Orgill and Hollands Model Discontinuity	0.00 – 0.35	0.273 (-)	0.251 (0.158)	0.250 (0.166)	492
	0.35 – 0.75	0.273 (-)	0.270 (0.023)	0.270 (0.023)	633
	0.75 – 1.00	0.247 (-)	0.245 (0.028)	0.245 (0.033)	172
Reindl. et al. Model Discontinuity	0.00 – 0.30	0.265 (-)	0.248 (0.130)	0.245 (0.152)	409
	0.30 – 0.78	0.273 (-)	0.268 (0.038)	0.267 (0.041)	749
	0.78 – 1.00	0.253 (-)	0.253 (0.007)	0.249 (0.041)	139

In Table 7 for South Dakota, Reindl discontinuity of 0.25 – 0.70 works best and gives a low error of 0.267 and R^2 value of 0.041. While the same continuous model i.e. quadratic when applied gives an error of 0.260 and R^2 value of 0.189 (Table 3). The continuo us model not only performs better than the piecewise model but also reduces the complexity associated with the piece wise models. An improvement of 2.62% in estimation of diffuse radiation is achieved using continuous quartic model over piecewise quadratic models. Continuous quartic models perform better than the piecewise quadratic models in all three locations which justifies usage of continuous quartic models and can implemented for the estimation of diffuse radiation calculation.

3.4.4 *Regressions using Relative Humidity, Absolute humidity and Ambient Air Temperature*

For improving diffuse fraction assessment in Germany some more parameters are explored. Same parameters are also explored in other studies such as Reindl et al. explored the elevation angle, temperature and relative humidity, Iqbal (Iqbal 1979) explored the

sunshine duration and Al Riahi et al. explored the sunshine duration and clearness index for improving diffuse radiation estimation. In this study, temperature, absolute humidity, relative humidity and clearness index have been explored and plotted with respect to diffuse fraction. Figure 7 is a distribution of relative humidity and temperature with diffuse fraction. A regression analysis was performed utilizing clearness index data only, clearness index, temperature and relative humidity data only and using clearness index and temperature only.

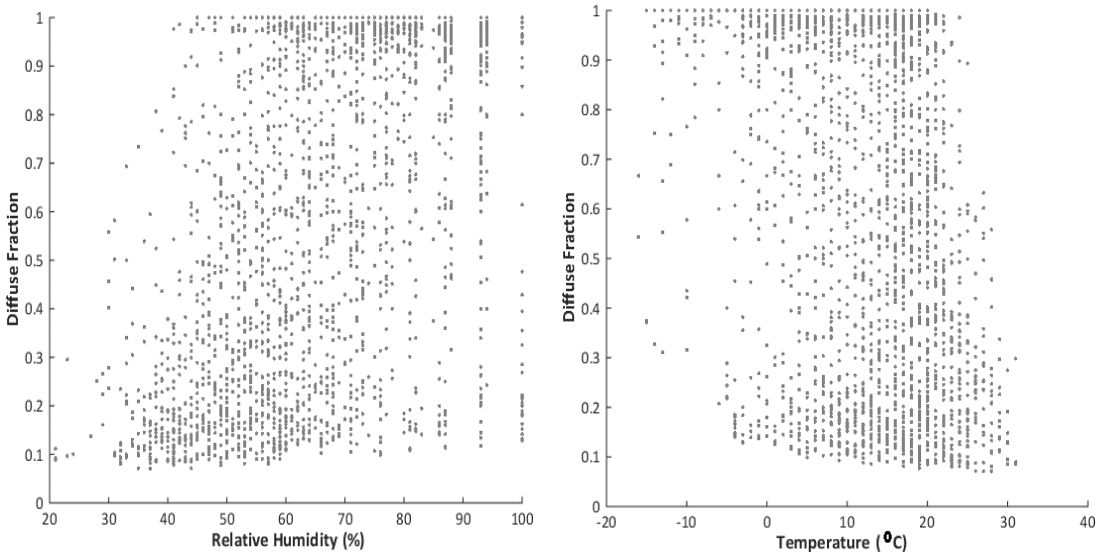

Figure 7. Effect of relative humidity and temperature on diffuse fraction.

It is observed in Table 8 that the RMSE has been improved by 6.10 % and R^2 value has been improved by 5.80 %. The RMSE and R^2 values remains same in linear fit even when relative humidity is not included in the regression analysis while RMSE and R^2 observe a fractional change in quadratic fit. A slight increase of 0.002 in R^2 that can be justified by the increase in number of variables and a slight decrease of 0.004 in $RMSE$ which can be justified by eliminating the parameter that is not required therefore reducing the RMSE.

Table 8. Bavaria, Germany continuous fit, RMSE on left and R^2 on right.

Predictor Variable	Linear	Quadratic
k_t	0.163 (0.769)	0.155 (0.790)
k_t , T , ρ	0.154 (0.803)	0.140 (0.836)
k_t , T	0.154 (0.803)	0.142 (0.832)

1) Using T, ρ and k_t as a predictor variable in linear model (Eq. 3.31) and quadratic model (Eq. 3.32).

$$I_d = I \times (1.391 - 1.1224 \times k_t + 0.00085 \times \rho - 0.00023629 \times T + 0.8846) \quad (3.31)$$

$$I_d = I \times (1.2761 - 0.61573 \times k_t - 0.00327 \times \rho - 0.010083 \times T + 0.0045948 \times k_t \times \rho - 0.000103 \times k_t \times T - 0.89936 \times k_t^2 + 7.36 \times 10^{-6} \times T^2 + 0.00011302 \times \rho^2) \quad (3.32)$$

2) Using T and k_t as a predictor variable in linear model (Eq. 3.33) and quadratic model (Eq. 3.34).

$$I_d = I \times (1.2201 - 1.1538 \times k_t - 0.0029763 \times T \quad (3.33)$$

$$I_d = I \times (1.0342 - 1.0371 \times k_t + 0.0006621 \times T - 0.0097391 \times k_t \times T - 1.041 \times k_t^2 + 5.317 \times 10^{-6} \times T^2) \quad (3.34)$$

3) Using k_t as a predictor variable in linear model (Eq. 3.35) and quadratic model (Eq. 3.36)

$$I_d = I \times (1.174 - 1.155 \times k_t) \quad (3.35)$$

$$I_d = I \times (1.045 - 0.2863 \times k_t - 0.9749 \times k_t^2) \quad (3.36)$$

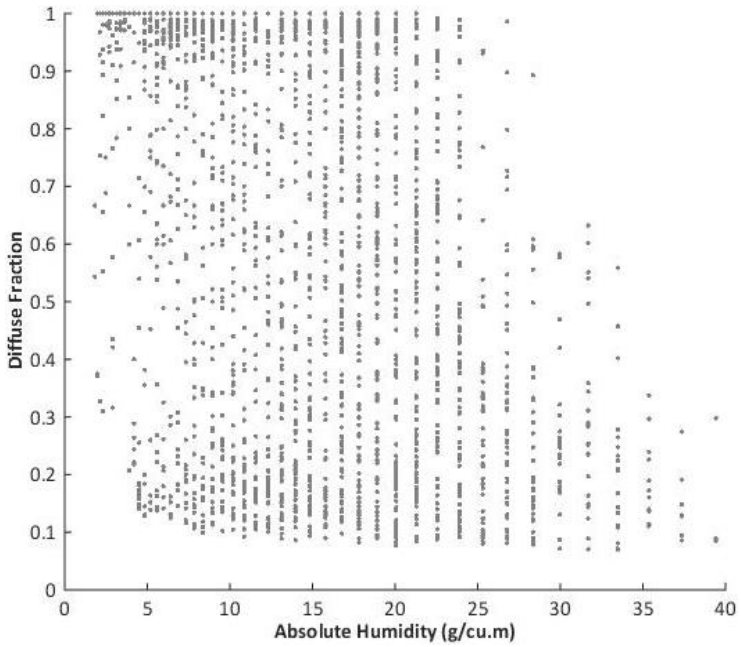

Figure 8. Effect of absolute humidity on diffuse fraction.

After analyzing three different parameters a fourth parameter that is absolute humidity is also studied and all possible combinations are analyzed. Table 9 is summary of the RMSE and R^2. The absolute humidity and temperature has the same values of RMSE and R^2 which can be explained by the fact that the absolute humidity is a function of temperature while the there is a slight improvement in RMSE and R^2 values when relative humidity is used. It is also observed that the either the relative humidity or temperature when used with the clearness index improves the fit. The clearness index is the most important variable after that relative humidity and temperature both produces the same RMSE and R^2 and ranked at the second place. Using relative humidity, temperature and clearness index together increases the complexity without improving the RMSE and R^2 values.

Table 9. Continuous fit with different predictor variables, RMSE on left and R^2 on right.

Predictor Variable	Linear	Quadratic
ρ_a	0.310 (0.200)	0.310 (0.200)
ρ	0.266 (0.408)	0.265 (0.413)
T	0.314 (0.180)	0.310 (0.200)
k_t , T	0.154 (0.803)	0.142 (0.832)
k_t , ρ_a	0.154 (0.802)	0.142 (0.831)
k_t , ρ	0.154 (0.801)	0.141 (0.833)
k_t , T , ρ	0.154 (0.803)	0.140 (0.836)
k_t , T , ρ_a	0.154 (0.803)	0.141 (0.834)

3.5 Conclusion and Future Work

Present study is conducted for four continents i.e. North America, South America, Australia and Asia Pacific all possessing different climatic conditions. The three most important conclusions obtained from the study are explained as: First, gives a best performing model based on the values of RMSE and R^2 values. An annual comparison is done among existing models and it has been found that Orgill and Hollands model worked best for six locations out of nine locations for which comparison has been run. The findings are in parallel with the findings in studies conducted by Dervishi and Mahadavi (Dervishi and Mahadavi 2012), Wong and Chow (Wong & Chow 2001), Eliminir (Eliminir 2007) and Jacovide et al. (Jacovide et al. 2006).

Second, exploits the models' vulnerability in low, medium and high k_t regions. It is observed that the existing models are prone to high relative error in regions of 0.50 – 1.00. New global model is developed to improve the fit in this region and the improvement is also realized in other regions like 0.00 – 0.20. The new global model performs better in low k_t region for 2 different sites when comparison is run for four different locations.

In third part a comparative analysis between the piecewise fittings and continuous fittings resulted in a conclusion that the continuous models work as good as or better than the piecewise models and can be implemented for the diffuse radiation estimation. Moreover, site specific models that are continuous in nature perform better than the global models such as Orgill and Hollands etc. If there is no data available for a particular site and hence no model can be generated for that site in that case a model which works best for most of the locations should be implemented with the improvements suggested. For example Orgill and Hollands model should be used where fit cannot be obtained because of data unavailability. Orgill and Hollands model must be complemented with the newly developed model in the region of 0.80 – 1.00 which will overall improves diffuse radiation estimation. For better estimation of diffuse radiation, site specific models generated in this study should be used compared to the existing models in literature. Also, study finds out the best working discontinuity region for the piecewise models. For Erbs et al. a high R^2 value and low RMSE is noted for discontinuity of 0.22 – 0.80 which is better than the rest of the discontinuities utilized in other models e.g. 025 – 0.70 etc. Therefore, if a piecewise fit is obtained then Erbs's region should be considered for better estimation of diffuse radiation.

Study is further narrowed down to Germany in which different predictor variables are explored. The effect of clearness index, relative humidity, absolute humidity and temperature is analyzed in improving the diffuse radiation calculation for Bavaria, Germany. The clearness index plays a major role in improving the diffuse radiation calculation after that temperature, relative humidity and absolute humidity all plays a similar role. A combination of clearness index and temperature is as significant as the

combination of clearness index and relative humidity in improving the calculation of diffuse radiation. Models are developed utilizing three, two and one predictor variable and a model with two predictor variable will be sufficient to calculate diffuse radiation for Bavaria, Germany.

The present work can be extended to build models for all locations and implementing those in the software utilized for the solar power estimation like HOMER, PVSYST, SAM, PVWATTS and PV SOL etc. This will be a cumbersome work but there are several studies already conducted in the world for the estimation of diffuse radiation for example Choudhary (Choudhary 1963) for India, Bolan et al (Bolan et al. 2008) for Australia, Srinivasan et al. (Srinivasan et al. 1986) for Saudi Arabia, Lam and Li (Lam & Li 1996) for Hong Kong, Muneer et al. (Muneer et al. 2007) for UK and Spain. These models can be gathered and can be implemented in the softwares as per the location. Moreover, present studies are concentrated on linear and nonlinear regression models for the estimation of diffuse radiations. This can be replaced by the rational models, exponential models or logarithmic models. For example Bolan et al. used the rational model for the estimation of diffuse radiation. Furthermore, Piri and kisi (Piri & Kisi 2015) used neural network for the estimation of diffuse radiation. Improvement in these methods will further result in better estimation of energy from solar photovoltaic or thermal.

References

Abbott, D. (2010). Keeping the energy debate clean: how do we supply the world's energy needs? *Proceedings of the IEEE, 98*(1), 42-66.

Al-Riahi, M., Al-Hamdani, N., & Tahir, K. (1992). An empirical method for estimation of hourly diffuse fraction of global radiation. *Renewable energy*,*2*(4), 451-456.

Boland, J., Ridley, B., & Brown, B. (2008). Models of diffuse solar radiation. *Renewable Energy*, *33*(4), 575-584.

Charles, N., L., Baseline Surface Radiation Network.

Choudhury, N. K. D. (1963). Solar radiation at New Delhi. *Solar Energy*, *7*(2), 44-52.

Deloitte. (2015) Future of global power sector: Preparing for emerging opportunities and threat.

Dervishi, S., & Mahdavi, A. (2012). Computing diffuse fraction of global horizontal solar radiation: A model comparison. *Solar energy*, *86*(6), 1796-1802.

Duffie, J. A., & Beckman, W. A. (1980). *Solar engineering of thermal processes* (Vol. 3). New York etc.: Wiley.

Elminir, H. K. (2007). Experimental and theoretical investigation of diffuse solar radiation: data and models quality tested for Egyptian sites. *Energy*,*32*(1), 73-82.

Elminir, H. K., Azzam, Y. A., & Younes, F. I. (2007). Prediction of hourly and daily diffuse fraction using neural network, as compared to linear regression models. *Energy*, *32*(8), 1513-1523.

El-Sebaii, A. A., Al-Hazmi, F. S., Al-Ghamdi, A. A., & Yaghmour, S. J. (2010). Global, direct and diffuse solar radiation on horizontal and tilted surfaces in Jeddah, Saudi Arabia. *Applied Energy*, *87*(2), 568-576.

Erbs, D. G., Klein, S. A., & Duffie, J. A. (1982). Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. *Solar energy*, *28*(4), 293-302.

Goetzberger, A., & Hoffmann, V. U. (2005). *Photovoltaic solar energy generation* (Vol. 112). Springer Science & Business Media.

Goswami, D. Y., Kreith, F., & Kreider, J. F. (2000). *Principles of solar engineering*. CRC Press

Gueymard, C. A. (2009). Direct and indirect uncertainties in the prediction of tilted irradiance for solar engineering applications. *Solar Energy*, *83*(3), 432-444.

Gueymard, C. A., & Myers, D. R. (2009). Evaluation of conventional and high-performance routine solar radiation measurements for improved solar resource, climatological trends, and radiative modeling. *Solar Energy*, *83*(2), 171-185.

Gueymard, C. A., & Wilcox, S. M. (2009, January). Spatial and temporal variability in the solar resource: Assessing the value of short-term measurements at potential solar power plant sites. In *Solar 2009 ASES Conf*.

Ineichen, P. (2011). Five satellite products deriving beam and global irradiance validation on data from 23 ground stations.

Iqbal, M. (1979). Correlation of average diffuse and beam radiation with hours of bright sunshine. *Solar Energy*, *23*(2), 169-173.

Iqbal, M. (2012). *An introduction to solar radiation*.

IYER, S. S. K. (2015, September). Solar Photovoltaic Energy Harnessing. In*Proc Indian Natn Sci Acad* (Vol. 81, No. 4, pp. 1001-1021).

Jacovides, C. P., Tymvios, F. S., Assimakopoulos, V. D., & Kaltsounides, N. A. (2006). Comparative study of various correlations in estimating hourly diffuse fraction of global solar radiation. *Renewable Energy*, *31*(15), 2492-2504.

Janjai, S., Praditwong, P., & Moonin, C. (1996). A new model for computing monthly average daily diffuse radiation for Bangkok. *Renewable energy*, *9*(1), 1283-1286.

Kerker, M. (2013). *The Scattering of Light and Other Electromagnetic Radiation: Physical Chemistry: A Series of Monographs* (Vol. 16). Academic press.

Khalil, S. A., & Shaffie, A. M. (2013). A comparative study of total, direct and diffuse solar irradiance by using different models on horizontal and inclined surfaces for Cairo, Egypt. *Renewable and Sustainable Energy Reviews*, *27*, 853-863.

Lam, J. C., & Li, D. H. (1996). Correlation between global solar radiation and its direct and diffuse components. *Building and environment*, *31*(6), 527-535.

Li, H., Bu, X., Long, Z., Zhao, L., & Ma, W. (2012). Calculating the diffuse solar radiation in regions without solar radiation measurements. *Energy*, *44*(1), 611-615.

Li, H., Ma, W., Wang, X., & Lian, Y. (2011). Estimating monthly average daily diffuse solar radiation with multiple predictors: a case study. *Renewable energy*, *36*(7), 1944-1948.

Liu, B. Y., & Jordan, R. C. (1960). The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. *Solar energy*, *4*(3), 1-19.

Menanteau, P., Finon, D., & Lamy, M. L. (2003). Prices versus quantities: choosing policies for promoting the development of renewable energy.*Energy policy*, *31*(8), 799-812.

Mohammadi, K., Shamshirband, S., Petković, D., & Khorasanizadeh, H. (2016). Determining the most important variables for diffuse solar radiation prediction using adaptive neuro-fuzzy methodology; case study: City of Kerman, Iran. *Renewable and Sustainable Energy Reviews*, *53*, 1570-1579.

Muneer, T., Younes, S., & Munawwar, S. (2007). Discourses on solar radiation modeling. *Renewable and Sustainable Energy Reviews*, *11*(4), 551-602.

Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC). 2015. FLUXNET Web Page. Available online [http://fluxnet.ornl.gov] from ORNL DAAC, Oak Ridge, Tennessee, U.S.A. Accessed May 5, 2015.

Orgill, J. F., & Hollands, K. G. T. (1977). Correlation equation for hourly diffuse radiation on a horizontal surface. *Solar energy*, *19*(4), 357-359.

Piri, J., & Kisi, O. (2015). Modelling solar radiation reached to the Earth using ANFIS, NN-ARX, and empirical models (Case studies: Zahedan and Bojnurd stations). *Journal of Atmospheric and Solar-Terrestrial Physics*, *123*, 39-47.

Rauschenbach, H. S. (2012). *Solar cell array design handbook: the principles and technology of photovoltaic energy conversion*. Springer Science & Business Media.

Reindl, D. T., Beckman, W. A., & Duffie, J. A. (1990). Diffuse fraction correlations. *Solar energy*, *45*(1), 1-7.

Ruth, D. W., & Chant, R. E. (1976). The relationship of diffuse radiation to total radiation in Canada. *Solar Energy*, *18*(2), 153-154.

Sánchez, G., Serrano, A., Cancillo, M. L., & García, J. A. (2012). Comparison of shadow-ring correction models for diffuse solar irradiance. *Journal of Geophysical Research: Atmospheres*, *117*(D9).

Schnitzer, M., Johnson, P., Thuman, C., & Freeman, J. (2012, June). Solar input data for photovoltaic performance modeling. In *Photovoltaic Specialists Conference (PVSC), 2012 38th IEEE* (pp. 003056-003060). IEEE.

Sengupta, M., Habte, A., Kurtz, S., Dobos, A., Wilbert, S., Lorenz, E., & Wilcox, S. (2015). *Best practices handbook for the collection and use of solar resource data for solar energy applications* (Doctoral dissertation, National Renewable Energy Laboratory).

Solar Power Europe. (2015) Global Market outlook.

Solar Radiation, National Centers for Environmental Information, National Oceanic and Atmospheric Administration. https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/solar-radiation

Spencer, J. W. (1982). A comparison of methods for estimating hourly diffuse solar radiation from global solar radiation. *Solar Energy*, *29*(1), 19-32.

Srinivasan, R., Bahel, V., & Bakhsh, H. (1986). Correlation for estimation of diffuse fraction of daily global radiation. *Energy*, *11*(7), 697-701.

Szarka, N., Eichhorn, M., Kittler, R., Bezama, A., & Thrän, D. (2016). Interpreting long-term energy scenarios and the role of bioenergy in Germany. *Renewable and Sustainable Energy Reviews*.

Thekaekara, M. P. (1976). Solar radiation measurement: techniques and instrumentation. *Solar Energy*, *18*(4), 309-325.

Tsvetkov, A., V., Voeikov Main Geophysical Observatory, World Radiation Data Centre. Retrieved from: http://wrdc.mgo.rssi.ru/

US Department of Energy, Energy efficiency and Renewable energy. History of Solar.

US energy information administration. International Energy Statistics. Retrieved from: https://www.eia.gov/cfapps/ipdbproject/IEDIndex3.cfm?tid=2&pid=2&aid=2

Weather history for Hohenpeissenberg, Bavaria, Germany (2012): Weather Underground.

Wilcox, S., National Renewable Energy Laboratory. Concentrating solar power research http://www.nrel.gov/csp/solrmap.html

Wirth, H., & Schneider, K. (2015). Recent facts about photovoltaics in Germany. *Report from Fraunhofer Institute for Solar Energy Systems, Germany*.

Wong, L. T., & Chow, W. K. (2001). Solar radiation model. *Applied Energy*, *69*(3), 191-224.

World Meteorological Organization. (1983). *Guide to meteorological instruments and methods of observation*. Secretariat of the World Meteorological Organization.

APPENDIX A

NEWLY DEVLOPED MODELS

Sioux Falls, South Dakota, USA Discontinuous Models

Erbs Model Discontinuity (Constant, Linear, Quadratic)

$$\frac{I_d}{I} = 0.6931$$

$$\frac{I_d}{I} = 0.45987$$

$$\frac{I_d}{I} = 0.37488$$

$$\frac{I_d}{I} = 0.88905 - 1.629 \times k_t$$

$$\frac{I_d}{I} = 0.6668 - 0.4208 \times k_t$$

$$\frac{I_d}{I} = -0.057515 + 0.48069 \times k_t$$

$$\frac{I_d}{I} = 0.96535 - 3.6802 \times k_t + 8.9358 \times k_t^2$$

$$\frac{I_d}{I} = 0.63382 - 0.27515 \times k_t - 0.14404 \times k_t^2$$

$$\frac{I_d}{I} = 14.924 - 33.104 \times k_t \ + 18.756 \times k_t^2$$

Reindl Model Discontinuity (Constant, Linear, Quadratic)

$$\frac{I_d}{I} = 0.66501$$

$$\frac{I_d}{I} = 0.44611$$

$$\frac{I_d}{I} = 0.37258$$

$$\frac{I_d}{I} = 0.83711 - 1.1447 \times k_t$$

$$\frac{I_d}{I} = 0.65342 - 0.4007 \times k_t$$

$$\frac{I_d}{I} = 0.07055 + 0.34601 \times k_t$$

$$\frac{I_d}{I} = 0.94183 - 3.0801 \times k_t + 6.3025 \times k_t^2$$

$$\frac{I_d}{I} = 0.40787 + 0.58705 \times k_t - 0.92966 \times k_t^2$$

$$\frac{I_d}{I} = 9.7926 - 21.848 \times k_t \ + 12.602 \times k_t^2$$

Orgill and Hollands Model Discontinuity (Constant, Linear, Quadratic)

$$\frac{I_d}{I} = 0.63324$$

$$\frac{I_d}{I} = 0.44968$$

$$\frac{I_d}{I} = 0.35115$$

$$\frac{I_d}{I} = 0.83107 - 1.085 \times k_t$$

$$\frac{I_d}{I} = 0.64663 - 0.37137 \times k_t$$

$$\frac{I_d}{I} = -0.14934 + 0.58698 \times k_t$$

$$\frac{I_d}{I} = 0.89723 - 2.1237 \times k_t + 2.8508 \times k_t^2$$

$$\frac{I_d}{I} = 0.57471 - 0.92048 \times k_t - 0.25941 \times k_t^2$$

$$\frac{I_d}{I} = 2.8344 - 6.3546 \times k_t \ + 4.009 \times k_t^2$$

Al Riahi Model Discontinuity (Constant, Linear, Quadratic)

$$\frac{I_d}{I} = 0.68709$$

$$\frac{I_d}{I} = 0.47106$$

$$\frac{I_d}{I} = 0.35301$$

$$\frac{I_d}{I} = 0.85938 - 1.3761 \times k_t$$

$$\frac{I_d}{I} = 0.6324 - 0.34308 \times k_t$$

$$\frac{I_d}{I} = -0.12014 + 0.2841 \times k_t$$

$$\frac{I_d}{I} = 0.96581 - 3.6828 \times k_t + 8.891 \times k_t^2$$

$$\frac{I_d}{I} = 0.7642 - 0.76962 \times k_t - 0.44994 \times k_t^2$$

$$\frac{I_d}{I} = 4.2173 - 9.563 \times k_t \ + 5.857 \times k_t^2$$

Bondville, Illinois, USA Piecewise model

Erbs Model Discontinuity (Constant, Linear, Quadratic)

$$\frac{I_d}{I} = 0.97382$$

$$\frac{I_d}{I} = 0.50283$$

$$\frac{I_d}{I} = 0.18382$$

$$\frac{I_d}{I} = 1.0041 - 0.1978 \times k_t$$

$$\frac{I_d}{I} = 1.4501 - 1.6072 \times k_t$$

$$\frac{I_d}{I} = -1.0388 + 1.4853 \times k_t$$

$$\frac{I_d}{I} = 1.0104 - 0.32941 \times k_t + 0.51106 \times k_t^2$$

$$\frac{I_d}{I} = 1.2422 - 0.80449 \times k_t - 0.7217 \times k_t^2$$

$$\frac{I_d}{I} = 22.927 - 56.023 \times k_t \ + 34.472 \times k_t^2$$

Reindl Model Discontinuity (Constant, Linear, Quadratic)

$$\frac{I_d}{I} = 0.9555$$

$$\frac{I_d}{I} = 0.50867$$

$$\frac{I_d}{I} = 0.17039$$

$$\frac{I_d}{I} = 1.0172 - 0.27616 \times k_t$$

$$\frac{I_d}{I} = 1.4668 - 1.6292 \times k_t$$

$$\frac{I_d}{I} = -0.76088 + 1.1517 \times k_t$$

$$\frac{I_d}{I} = 0.98553 + 0.16387 \times k_t - 1.217 \times k_t^2$$

$$\frac{I_d}{I} = 1.3276 - 1.1053 \times k_t - 0.46647 \times k_t^2$$

$$\frac{I_d}{I} = 12.715 - 31.703 \times k_t \ + 20.005 \times k_t^2$$

Orgill and Hollands Model Discontinuity (Constant, Linear, Quadratic)

$$\frac{I_d}{I} = 0.92819$$

$$\frac{I_d}{I} = 0.51759$$

$$\frac{I_d}{I} = 0.18578$$

$$\frac{I_d}{I} = 1.0415 - 0.40948 \times k_t$$

$$\frac{I_d}{I} = 1.4806 - 1.6479 \times k_t$$

$$\frac{I_d}{I} = 0.24706 - 0.077577 \times k_t$$

$$\frac{I_d}{I} = 0.97984 + 0.25969 \times k_t - 1.5133 \times k_t^2$$

$$\frac{I_d}{I} = 1.4383 - 1.4923 \times k_t - 0.1373 \times k_t^2$$

$$\frac{I_d}{I} = 17.285 - 42.627 \times k_t \ + 26.526 \times k_t^2$$

Al Riahi Model Discontinuity (Constant, Linear, Quadratic)

$$\frac{I_d}{I} = 0.9681$$

$$\frac{I_d}{I} = 0.60467$$

$$\frac{I_d}{I} = 0.22499$$

$$\frac{I_d}{I} = 1.008 - 0.22807 \times k_t$$

$$\frac{I_d}{I} = 1.42 - 1.5387 \times k_t$$

$$\frac{I_d}{I} = 0.97682 - 0.987 \times k_t$$

$$\frac{I_d}{I} = 1.0029 - 0.13608 \times k_t - 0.31944 \times k_t^2$$

$$\frac{I_d}{I} = 1.2839 - 0.96905 \times k_t - 0.56353 \times k_t^2$$

$$\frac{I_d}{I} = 9.1906 - 22.327 \times k_t \ + 13.823 \times k_t^2$$

Bavaria, Hohenpeissenberg, Germany

Erbs Model Discontinuity (Constant, Linear, Quadratic)

$$\frac{I_d}{I} = 0.98746$$

$$\frac{I_d}{I} = 0.57986$$

$$\frac{I_d}{I} = 0.21597$$

$$\frac{I_d}{I} = 0.91807 - 0.080938 \times k_t$$

$$\frac{I_d}{I} = 1.4073 - 1.5219 \times k_t$$

$$\frac{I_d}{I} = -1.5049 + 2.0712 \times k_t$$

$$\frac{I_d}{I} = 0.98769 - 0.11174 \times k_t - 0.75098 \times k_t^2$$

$$\frac{I_d}{I} = 1.1125 - 0.22283 \times k_t - 1.2469 \times k_t^2$$

$$\frac{I_d}{I} = 6.6391 - 16.826 \times k_t \ + 10.931 \times k_t^2$$

Reindl Model Discontinuity (Constant, Linear, Quadratic)

$$\frac{I_d}{I} = 0.97783$$

$$\frac{I_d}{I} = 0.55019$$

$$\frac{I_d}{I} = 0.1997$$

$$\frac{I_d}{I} = 1.015 - 0.223 \times k_t$$

$$\frac{I_d}{I} = 1.5023 - 1.6619 \times k_t$$

$$\frac{I_d}{I} = -1.111 + 1.6071 \times k_t$$

$$\frac{I_d}{I} = 0.96959 - 0.49426 \times k_t - 1.9943 \times k_t^2$$

$$\frac{I_d}{I} = 1.3065 - 0.89427 \times k_t - 0.69675 \times k_t^2$$

$$\frac{I_d}{I} = 9.0989 - 22.536 \times k_t \ + 14.231 \times k_t^2$$

Orgill and Hollande Model Discontinuity (Constant, Linear, Quadratic)

$$\frac{I_d}{I} = 0.97304$$

$$\frac{I_d}{I} = 0.55354$$

$$\frac{I_d}{I} = 0.20674$$

$$\frac{I_d}{I} = 1.0226 - 0.28012 \times k_t$$

$$\frac{I_d}{I} = 1.5588 - 1.7482 \times k_t$$

$$\frac{I_d}{I} = -0.3583 + 0.70638 \times k_t$$

$$\frac{I_d}{I} = 0.97693 + 0.31991 \times k_t - 1.5708 \times k_t^2$$

$$\frac{I_d}{I} = 1.5283 - 1.6335 \times k_t - 0.10251 \times k_t^2$$

$$\frac{I_d}{I} = 11.371 - 27.767 \times k_t \ + 17.23 \times k_t^2$$

Al Riahi Model Discontinuity (Constant, Linear, Quadratic)

$$\frac{I_d}{I} = 0.98542$$

$$\frac{I_d}{I} = 0.69587$$

$$\frac{I_d}{I} = 0.23161$$

$$\frac{I_d}{I} = 1.0004 - 0.10248 \times k_t$$

$$\frac{I_d}{I} = 1.4024 - 1.4756 \times k_t$$

$$\frac{I_d}{I} = -0.51441 - 0.36384 \times k_t$$

$$\frac{I_d}{I} = 0.98812 + 0.10291 \times k_t - 0.71436 \times k_t^2$$

$$\frac{I_d}{I} = 1.0288 + 0.23374 \times k_t - 1.17916 \times k_t^2$$

$$\frac{I_d}{I} = 10.197 - 24.932 \times k_t \ + 15.525 \times k_t^2$$

Colorado, Boulder, USA Continuous Models

$$\frac{I_d}{I} = 1.239 - 1.286 \times k_t$$

$$\frac{I_d}{I} = -0.01706 \times k_t^2 - 0.2711 \times k_t + 0.4901$$

$$\frac{I_d}{I} = 0.03535 \times k_t^3 + 0.03825 \times k_t^2 - 0.3281 \times k_t + 0.4649$$

$$\frac{I_d}{I} = 0.0162 \times k_t^4 + 0.06682 \times k_t^3 + 0.002466 \times k_t - 0.3674 \times k_t + 0.4796$$

Bondville, Illinois, USA Continuous Models

$$\frac{I_d}{I} = 1.388 - 1.498 \times k_t$$

$$\frac{I_d}{I} = -0.4506 \times k_t^2 - k_t + 1.263$$

$$\frac{I_d}{I} = 5.478 \times k_t^3 - 8.933 \times k_t^2 + 3.014 \times k_t + 0.7036$$

$$\frac{I_d}{I} = 14.06 \times k_t^4 - \ 22.8 \times k_t^3 + 10.64 \times k_t - 2.298 \times k_t + 1.143$$

Desert Rock, Nevada, USA Continuous Models

$$\frac{I_d}{I} = 0.3068 - 0.212 \times k_t$$

$$\frac{I_d}{I} = -0.4506 \times k_t^2 - k_t + 1.263$$

$$\frac{I_d}{I} = 0.01687 \times k_t^3 + 0.05804 \times k_t^2 - 0.2103 \times k_t + 0.2653$$

$$\frac{I_d}{I} = 0.002444 \times k_t^4 + 0.002595 \times k_t^3 + 0.05842 \times k_t - 0.2222 \times k_t + 0.2658$$

Rock spring, Pennsylvania USA Continuous Models

$$\frac{I_d}{I} = -0.6232 - 0.2766 \times k_t$$

$$\frac{I_d}{I} = -0.05388 \times k_t^2 - 0.2941 \times k_t + 0.6771$$

$$\frac{I_d}{I} = 0.03657 \times k_t^3 - 0.03162 \times k_t^2 - 0.3616 \times k_t + 0.6667$$

$$\frac{I_d}{I} = 0.01837 \times k_t^4 + 0.0511 \times k_t^3 - 0.08011 \times k_t - 0.3784 \times k_t + 0.6823$$

Sioux Falls, South Dakota, USA Continuous Models

$$\frac{I_d}{I} = 0.5059 - 0.1174 \times k_t$$

$$\frac{I_d}{I} = -0.04175 \times k_t^2 - 0.1245 \times k_t + 0.4642$$

$$\frac{I_d}{I} = -0.01034 \times k_t^3 + 0.04641 \times k_t^2 - 0.1032 \times k_t + 0.4613$$

$$\frac{I_d}{I} = 0.0271 \times k_t^4 - 0.02998 \times k_t^3 - 0.03389 \times k_t - 0.07659 \times k_t + 0.4867$$

Fort Peck, Montana, USA Continuous Models

$$\frac{I_d}{I} = 0.4956 + 0.2506 \times k_t$$

$$\frac{I_d}{I} = -0.02458 \times k_t^2 - 0.267 \times k_t + 0.5202$$

$$\frac{I_d}{I} = 0.02406 \times k_t^3 + 0.01709 \times k_t^2 - 0.3087 \times k_t + 0.4946$$

$$\frac{I_d}{I} = 0.009809 \times k_t^4 + 0.04836 \times k_t^3 - 0.004578 \times k_t - 0.3434 \times k_t + 0.5042$$

Ushusaia, Tierra Del Fuego, Argentina Continuous Models

$$\frac{I_d}{I} = 0.8976 - 0.6738 \times k_t$$

$$\frac{I_d}{I} = -0.3633 \times k_t^2 - 0.3508 \times k_t + 0.847$$

$$\frac{I_d}{I} = 3.647 \times k_t^3 - 5.458 \times k_t^2 + 1.588 \times k_t + 0.6695$$

$$\frac{I_d}{I} = -1.042 \times k_t^4 + 5.637 \times k_t^3 - 6.695 \times k_t + 1.867 \times k_t + 0.6523$$

Wagga Wagga, New South Wales, Australia Continuous Models

$$\frac{I_d}{I} = 1.145 - 1.269 \times k_t$$

$$\frac{I_d}{I} = 0.4874 \times k_t^2 - 1.321 \times k_t + 1.157$$

$$\frac{I_d}{I} = 5.656 \times k_t^3 - 0.8657 \times k_t^2 + 2.713 \times k_t + 0.2654$$

$$\frac{I_d}{I} = 7.789 \times k_t^4 - 10.47 \times k_t^3 + 2.851 \times k_t - 0.5285 \times k_t + 0.914$$

Sapparo, Hokkaido, Japan Continuous Models

$$\frac{I_d}{I} = 1.145 - 1.269 \times k_t$$

$$\frac{I_d}{I} = 0.4874 \times k_t^2 - 1.321 \times k_t + 1.157$$

$$\frac{I_d}{I} = 5.656 \times k_t^3 - 0.8657 \times k_t^2 + 2.713 \times k_t + 0.2654$$

$$\frac{I_d}{I} = 7.789 \times k_t^4 - 10.47 \times k_t^3 + 2.851 \times k_t - 0.5285 \times k_t + 0.914$$

Hohenpeissenberg, Bavaria, Germany

$$\frac{I_d}{I} = 1.174 - 1.155 \times k_t$$

$$\frac{I_d}{I} = -0.9749 \times k_t^2 - 0.2863 \times k_t + 1.045$$

$$\frac{I_d}{I} = 3.522 \times k_t^3 - 5.696 \times k_t^2 + 1.451 \times k_t + 0.8975$$

$$\frac{I_d}{I} = 8.125 \times k_t^4 - 11.37 \times k_t^3 + 3.269 \times k_t - 0.5006 \times k_t + 1.012$$

APPENDIX B

MATLAB PROGRAM - MODEL COMPARISON ON ANNUAL BASIS

```
function diffusedradiation (name,r_fname,lati,longi,Zc)
% Extraterrestrial radiation calculation
for day= 1:366
decA = 23.45*sind((360*(284+day))/366); % declination angle for a particular day.
Tc1= [0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24];
Tc2= [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 0];
b(day) = 2*3.14*day/366;
G = 1367*(1.00011 + 0.034221 * cos(b(day)) + 0.001280 * sin(b(day)) + 0.000719 *
    cos(2*b(day)) + 0.000077 * sin(2*b(day)));
B(day) = 360*(day-1)/366;
E(day) = 3.82*(0.000075+0.001868*cosd(B(day))-0.032077*sind(B(day))-
    0.014615*cosd(2*B(day))-0.04089*sind(2*B(day)));
for x= 1:24
Ts1(day,x)= Tc1(x) + (longi/15)-Zc + E(day); % Solar time corresponding to the local
    time tc1
Ts2(day,x)= Tc2(x) + (longi/15)-Zc + E(day); % Solar time corresponding to the local
    time tc2
w3(day,x) = (Ts1(day,x)-12)*15; % Hour angle corresponding to Ts1
w4(day,x) = (Ts2(day,x)-12)*15; % Hour angle corresponding to Ts2
Etr1(day,x) = ((12/(3.14*2.77))*G*((cosd(lati)*cosd(decA)*(sind(w4(day,x))-
    sind(w3(day,x)))+ (0.0174*(w4(day,x)-w3(day,x))*sind(lati)*sind(decA))));%
    Extraterrestrial radiation
end
end
ETR = transpose(Etr1);
xlswrite(r_fname,ETR,'Sheet4');
m1 =  xlsread(r_fname,'Sheet4');
an = m1(:);
xlswrite(r_fname,an,'Sheet3','B1');
m11 =  xlsread(r_fname,'Sheet1');
m22 =  xlsread(r_fname,'Sheet2');
an1 = m11(:);
am = m22(:);
xlswrite(r_fname,an1,'Sheet3','C1');
xlswrite(r_fname,am,'Sheet3','I1');
for x1= 1:numel(an1)
akt(x1,1) = an1(x1) / an(x1);
end
xlswrite(r_fname,akt,'Sheet3','D1');
m = xlsread(r_fname,'Sheet3');
I = m(:,2);
kt = m(:,3);
Id = m(:,8);
for i = 1:numel(I)
min = kt(i);
```

```
40 for j= i+1:numel(I)
41 if(min>kt(j))
        a.  min = kt(j);
        b.  idx = j;
        c.  temp = kt(i);
        d.  kt(i)= kt(idx);
        e.  kt(idx) = temp;
        f.  temp = I(i);
        g.  I(i) = I(idx);
        h.  I(idx) = temp;
        i.  temp = Id(i);
        j.  Id(i) = Id(idx);
        k.  Id(idx) = temp;
42 end
43 end
44 end
45 z = 1;
46 for i = 1:numel(kt)
47 if kt(i)>0 && kt(i)<= 1
48 x(z,1) = kt(i,1);
49 y(z,1) = I(i,1);
50 w(z,1) = Id(i,1);
51 z= z+1;
52 end
53 end
54 disp(z);
55 xlswrite(r_fname,y,'sheet5','B1')
56 xlswrite(r_fname,x,'sheet5','C1')
57 xlswrite(r_fname,w,'sheet5','I1')
58 N = xlsread(r_fname,'sheet5');
59 I = N(:,1);
60 kt = N(:,2);
61 Id = N(:,8);
62 Idf = zeros(z-2,1);
63 for i= 1:z-1
64 Idf(i,1) = Id(i,1)/I(i,1);
65 end
66 xlswrite(r_fname,Idf,'sheet5','H1');
67 % For taking values out of bound for Id/I form the data  set
68 o=1;
69 for i = 1:numel(kt)
70 if Idf(i)>0 && Idf(i) <= 1
71 p(o,1) = kt(i,1);
72 q(o,1) = Idf(i,1);
73 r(o,1) = I(i,1);
```

```
74 s(o,1) = Id(i,1);
75 o= o+1;
76 end
77 end
78 xlswrite(r_fname,p,'sheet5','J1');
        xlswrite(r_fname,q,'sheet5','K1');xlswrite(r_fname,r,'sheet5','u1');xlswrite(r_fname
        ,s,'sheet5','v1');
79 disp(numel(p));disp(numel(q));disp(numel(r));disp(numel(s));
80 %Orgills model diffuse radiation calculation
81 c1 =0 ; d1 =0; e1 =0;
82 z1 = 0;
83 z1 = numel(p);
84 for j = 1:z1
85 if(0 <p(j)) && (p(j) < 0.35)
86 c1 = c1+1;
87 elseif (0.35 <= p(j)) && (p(j) <= 0.75)
88 d1= d1+1;
89 elseif(0.75< p(j) && p(j) <1)
90 e1= e1+1;
91 else
92 end
93 end
94 for i=1:c1;
95 Idc(i,1) = r(i)*(1-0.249*p(i));
96 Idc1(i,1) = Idc(i,1)/r(i);
97 end
98 for i=c1+1:c1+d1
99 Idc(i,1) = r(i)*(1.577-1.84*p(i));
100 Idc1(i,1) = Idc(i,1)/r(i);
101 end
102 for i= c1+d1+1:c1+d1+e1;
103 Idc(i,1)= 0.177*(r(i)*p(i));
104 Idc1(i,1) = Idc(i,1)/r(i);
105 end
106 xlswrite(r_fname,Idc,'sheet5','E1');
107 Rsq1 = 1 - sum((s - Idc).^2)/sum((s - mean(Idc)).^2);
108 disp(Rsq1);
109 RMSE1 = sqrt(sum((q(:)- Idc1(:)).^2)/numel(q));
110 Idf1 = zeros(z1-1,1);
111 for i= 1:z1
112 Idf1(i,1) = Idc(i,1)/r(i,1);
113 end
114 % For taking values out of bound for Id/I form the data  set
115 o1=1;
116 for i = 1:numel(p)
```

```
117 if Idf1(i)>0 && Idf1(i) <= 1
118 p1(o1,1) = p(i,1);
119 q1(o1,1) = Idf1(i,1);
120 o1= o1+1;
121 end
122 end
123 xlswrite(r_fname,p1,'sheet5','M1');xlswrite(r_fname,q1,'sheet5','N1')
124 %Erbs Model model diffuse radiation calculation
125 c2 =0 ; d2 = 0; e2 =0;
126 for j = 1:z1
127 if(0 <p(j)) && (p(j) <= 0.22)
128 c2 = c2+1;
129 elseif (0.22 < p(j)) && (p(j) <= 0.80)
130 d2= d2+1;
131 elseif(0.80 < p(j) && p(j) <1)
132 e2= e2+1;
133 else
134 end
135 end
136 for i=1:c2;
137 Idc(i,1) = r(i)*(1-0.09*p(i));
138 Idc1(i,1) = Idc(i,1)/r(i);
139 end
140 for i=c2+1:c2+d2
141 Idc(i,1) = r(i)*(0.9511-0.1604*p(i)+4.388*(p(i)^2)-
        16.638*(p(i)^3)+12.336*(p(i)^4));
142 Idc1(i,1) = Idc(i,1)/r(i);
143 end
144 for i= c2+d2+1:c2+d2+e2;
145 Idc(i,1)= 0.165*(r(i)*p(i));
146 Idc1(i,1) = Idc(i,1)/r(i);
147 end
148 Rsq2 = 1 - sum((s(:) - Idc(:)).^2)/sum((s(:) - mean(Idc(:))).^2);
149 disp(Rsq2);
150 RMSE2 = sqrt(sum((q(:)- Idc1(:)).^2)/numel(q));
151 xlswrite(r_fname,Idc,'sheet5','D1');
152 Idf2 = zeros(z1-1,1);
153 for i= 1:z1
154 Idf2(i,1) = Idc(i,1)/r(i,1);
155 end
156 % For taking values out of bound for Id/I form the data  set  Reindl Model
157 o2=1;
158 for i = 1:numel(p)
159 if Idf2(i)>0 && Idf2(i) <= 1
160 p2(o2,1) = p(i,1);
```

```
q2(o2,1) = Idf2(i,1);
o2= o2+1;
end
end
xlswrite(r_fname,p2,'sheet5','O1');xlswrite(r_fname,q2,'sheet5','P1')
hold on;
c3 =0 ; d3 = 0; e3 =0;
for j = 1:z1
if(0<= p(j)) && (p(j) <=0.3)
c3 = c3+1;
elseif (0.3 < p(j)) && (p(j) < 0.78)
d3= d3+1;
elseif(0.78<= p(j) && p(j) <1)
e3= e3+1;
else
end
end
for i=1:c3;
Idc(i,1) = r(i)*(1.020-0.248*p(i));
Idc1(i,1) = Idc(i,1)/r(i);
end
for i=c3+1:c3+d3
Idc(i,1) = r(i)*(1.45-1.67*p(i));
Idc1(i,1) = Idc(i,1)/r(i);
end
for i= c3+d3+1:c3+d3+e3;
Idc(i,1)= 0.147*(r(i)*p(i));
Idc1(i,1) = Idc(i,1)/r(i);
end
Rsq3 = 1 - sum((s - Idc).^2)/sum((s - mean(Idc)).^2);
disp(Rsq3);
RMSE3 = sqrt(sum((q(:)- Idc1(:)).^2)/numel(q));
xlswrite(r_fname,Idc,'sheet5','F1');
for i= 1:z1
Idf3(i,1) = Idc(i,1)/r(i,1);
end
% For taking values out of bound for Id/I form the data  set
o3=1;
for i = 1:numel(p)
if Idf3(i)>0 && Idf3(i) <= 1
p3(o3,1) = p(i,1);
q3(o3,1) = Idf3(i,1);
o3= o3+1;
end
end
```

```
206 xlswrite(r_fname,p3,'sheet5','Q1');xlswrite(r_fname,q3,'sheet5','R1')
207 hold on;
208 %AL Riahi Model diffuse radiation calculation
209 c4 =0 ; d4 = 0; e4 =0;
210 for j = 1:z1
211 if(0 <= p(j)) && (p(j) < 0.25)
212 c4 = c4+1;
213 elseif (0.25 <= p(j)) && (p(j) <= 0.70)
214 d4= d4+1;
215 elseif(0.70 < p(j) && p(j) <1)
216 e4= e4+1;
217 else
218 end
219 end
220 for i=1:c4;
221 Idc(i,1) = r(i)*(0.932*p(i));
222 Idc1(i,1) = Idc(i,1)/r(i);
223 end
224 for i=c4+1:c4+d4
225 Idc(i,1) = r(i)*(1.293-1.631*p(i));
226 Idc1(i,1) = Idc(i,1)/r(i);
227 end
228 for i= c4+d4+1:c4+d4+e4;
229 Idc(i,1)= 0.151*(r(i)*p(i));
230 Idc1(i,1) = Idc(i,1)/r(i);
231 end
232 Rsq4 = 1 - sum((s - Idc).^2)/sum((s - mean(Idc)).^2);
233 disp(Rsq4);
234 RMSE4 = sqrt(sum((q(:)- Idc1(:)).^2)/numel(q));
235 xlswrite(r_fname,Idc,'sheet5','G1');
236 for i= 1:z1
237 Idf4(i,1) = Idc(i,1)/r(i,1);
238 end
239 % For taking values out of bound for Id/I form the data set
240 o4=1;
241 for i = 1:numel(p)
242 if Idf4(i)>0 && Idf4(i) <= 1
243 p4(o4,1) = p(i,1);
244 q4(o4,1) = Idf4(i,1);
245 o4= o4+1;
246 end
247 end
248 %New Model calculation
249 c5 =0 ; d5 = 0; e5 =0;
250 for j = 1:z1
```

```
if(0 <= p(j)) && (p(j) < 0.25)
c5 = c5+1;
elseif (0.25 <= p(j)) && (p(j) <= 0.70)
d5= d5+1;
elseif(0.70 < p(j) && p(j) <1)
e5= e5+1;
else
end
end
for i=1:c5;
Idc(i,1) = r(i)*(11.42*p(i)^4-16.84*p(i)^3+6.104*p(i)^2-1.006*p(i)+1.026);
Idc1(i,1) = Idc(i,1)/r(i);
end
for i=c5+1:c5+d5
Idc(i,1) = r(i)*(11.42*p(i)^4-16.84*p(i)^3+6.104*p(i)^2-1.006*p(i)+1.026);
Idc1(i,1) = Idc(i,1)/r(i);
end
for i= c4+d4+1:c4+d4+e4;
Idc(i,1) = r(i)*(11.42*p(i)^4-16.84*p(i)^3+6.104*p(i)^2-1.006*p(i)+1.026);
Idc1(i,1) = Idc(i,1)/r(i);
end
Rsq5 = 1 - sum((s - Idc).^2)/sum((s - mean(Idc)).^2);
disp(Rsq5);
RMSE5 = sqrt(sum((q(:)- Idc1(:)).^2)/numel(q));
for i= 1:z1
Idf5(i,1) = Idc(i,1)/r(i,1);
end
% For taking values out of bound for Id/I form the data  set
o5=1;
for i = 1:numel(p)
if Idf5(i)>0 && Idf5(i) <= 1
p5(o5,1) = p(i,1);
q5(o5,1) = Idf5(i,1);
o5= o5+1;
end
end
fprintf(1,'Rmse1 %5.3f\n Rmse2 %5.3f\n Rmse3 %5.3f\n Rmse4 %5.3f\n Rmse5
    %5.3f\n',RMSE1, RMSE2, RMSE3, RMSE4, RMSE5)
%fprintf(1,'Rmse1 %5.3f\n Rmse2 %5.3f\n Rmse3 %5.3f\n Rmse4 %5.3f\n',RMSE1,
    RMSE2, RMSE3, RMSE4);
xlswrite(r_fname,p4,'sheet5','S1');xlswrite(r_fname,q4,'sheet5','T1')
plot(p,q,'.','color',[0.5,0.5,0.5]);
hold on;
plot(p1,q1,'g',p2,q2,'k',p3,q3,'r',p4,q4,'y',p5,q5,'m','LineWidth',2,'LineWidth',2,'Line
    Width',2,'LineWidth',2,'LineWidth',2);
```

```
293 %plot(p1,q1,'g',p2,q2,'k',p3,q3,'r',p4,q4,'y','LineWidth',2,'LineWidth',2,'LineWidth',2,
        'LineWidth',2);
294 title(name);
295 xlabel('Clearness Index');
296 ylabel('Diffuse Fraction');
297 legend('kt - Grey','Orgills - Green','Erbs - Black','Reindl - Red','Al riahi - yellow');
298 end
```

APPENDIX C

MATLAB PROGRAM - MODEL PERFORMANCE ASSESSMENT IN DIFFERENT CLEARNESS INDEX REGIONS

```
1 Function piecewisef1()
2 m = xlsread('Germany TH data.xlsx','Sheet7');
3 I =  m(:,1);
4 Id = m(:,2);
5 kt=  m(:,3);
6 Idf = m(:,6);
7 p = 1; q1 =1; q2 =1; q3 =1; q4 =1; r =1 ; s = 1; t = 1; a= 1; w = 1 ; y = 1;
8 s1 = 1; s2 = 1 ;
9 for x1 = 1 : numel(Idf)
10 if 0 <= Idf(x1) && Idf(x1) <= 1
11 Idfn(a) =  Idf(x1);
12 kt(a) =  kt(x1);
13 I(a) =  I(x1);
14 Id(a) = Id(x1);
15 a =  a +1 ;
16 end
17 end
18 for x = 1 : a-1
19 if  kt(x) <= 0.2
20 Idc1(p,1) = I(x)*(1-0.249*kt(x)); % Orgills Model
21 Idc2(p,1) = I(x)*(1-0.09*kt(x));  % Erbs Model
22 Idc3(p,1) = I(x)*(1.020-0.248*kt(x)); % Reidnl Model
23 Idc4(p,1) = I(x)*(0.932*kt(x)); % Al Riahi Model
24 Idcn(p,1) = I(x)*(10.64*kt(x)^4-16.23*kt(x)^3+6.318*kt(x)^2 - 1.116*kt(x)+0.98); %
        Uday US model
25 Idcnn(p,1) = I(x)*(6.732*kt(x)^4-7.929*kt(x)^3+ 0.5146*kt(x)^2 -
        0.3875*kt(x)+0.8481); % Uday International Model
26 Idcng(p,1) = I(x)*(8.307*kt(x)^4-11.24*kt(x)^3+ 2.729*kt(x)^2 -
        0.1227*kt(x)+0.8846); % Uday Global Model
27 Id1(p) = Id(x);
28 p= p+1;
29 elseif 0.2 < kt(x) && kt(x) <= 0.4
                a. q1  = q1 + 1;
                b. if 0.2 < kt(x) && kt(x) <0.35            % Orgills Model
                c. Idc5(q1,1) = I(x)*(1-0.249*kt(x));
                d. Id2(q1,1) = Id(x);
                e. elseif 0.35 <= kt(x) && kt(x) <=0.4
                f. Idc5(q1,1) = I(x)*(1.577-1.84*kt(x));
                g. Id2(q1,1) = Id(x);
                h. end
                i. q2 = q2+1;
                j. if 0.2 < kt(x) && kt(x) <= 0.22           % Erbs Model
                k. Idc6(q2,1) = I(x)*(1-0.09*kt(x));
                l. Id3(q2,1) = Id(x);
                m. elseif 0.22 < kt(x) && kt(x) <=0.4
```

n. Idc6(q2,1) = I(x)*(0.9511-0.1604*kt(x)+4.388*(kt(x)^2)-16.638*(kt(x)^3)+12.336*(kt(x)^4));
o. Id3(q2,1) = Id(x);
p. end
q. q3 = q3+1;
r. if 0.2 < kt(x) && kt(x) <= 0.3 % Reidnl Model
s. Idc7(q3,1) = I(x)*(1.020-0.248*kt(x));
t. Id4(q3,1) = Id(x);
u. elseif 0.3 < kt(x) && kt(x) <= 0.4
v. Idc7(q3,1) = I(x)*(1.45-1.67*kt(x));
w. Id4(q3,1) = Id(x);
x. end
y. q4 = q4+1;
z. if 0.2 < kt(x) && kt(x) < 0.25 % Al Riahi Model
aa. Idc8(q4,1) = I(x)*(0.932*kt(x));
bb. Idcn1(q4,1) = I(x)*(10.64*kt(x)^4-16.23*kt(x)^3+6.318*kt(x)^2 - 1.116*kt(x)+0.98); % Uday US model
cc. Idcnn1(q4,1) = I(x)*(6.732*kt(x)^4-7.929*kt(x)^3+ 0.5146*kt(x)^2 - 0.3875*kt(x)+0.8481); % Uday International Model
dd. Idcng1(q4,1) = I(x)*(8.307*kt(x)^4-11.24*kt(x)^3+ 2.729*kt(x)^2 - 0.1227*kt(x)+0.8846); % Uday Global Model
ee. Id5(q4,1) = Id(x);
ff. elseif 0.25 <= kt(x) && kt(x) <=0.4
gg. Idc8(q4,1) = I(x)*(1.293-1.631*kt(x));
hh. Idcn1(q4,1) = I(x)*(10.64*kt(x)^4-16.23*kt(x)^3+6.318*kt(x)^2 - 1.116*kt(x)+0.98); % Uday US model
ii. Idcnn1(q4,1) = I(x)*(6.732*kt(x)^4-7.929*kt(x)^3+ 0.5146*kt(x)^2 - 0.3875*kt(x)+0.8481); % Uday International Model
jj. Idcng1(q4,1) = I(x)*(8.307*kt(x)^4-11.24*kt(x)^3+ 2.729*kt(x)^2 - 0.1227*kt(x)+0.8846); % Uday Global Model
kk. Id5(q4,1) = Id(x);
ll. end
mm. w = w +1;

30 elseif 0.4 < kt(x) && kt(x) <= 0.6
a. Idc9(r,1) = I(x)*(1.577-1.84*kt(x));
b. Idc10(r,1) = I(x)*(0.9511-0.1604*kt(x)+4.388*(kt(x)^2)-16.638*(kt(x)^3)+12.336*(kt(x)^4));
c. Idc11(r,1) = I(x)*(1.45-1.67*kt(x));
d. Idc12(r,1) = I(x)*(1.293-1.631*kt(x));
e. Idcn2(r,1) = I(x)*(10.64*kt(x)^4-16.23*kt(x)^3+6.318*kt(x)^2 - 1.116*kt(x)+0.98); % Uday US model
f. Idcnn2(r,1) = I(x)*(6.732*kt(x)^4-7.929*kt(x)^3+ 0.5146*kt(x)^2 - 0.3875*kt(x)+0.8481); % Uday International Model
g. Idcng2(r,1) = I(x)*(8.307*kt(x)^4-11.24*kt(x)^3+ 2.729*kt(x)^2 - 0.1227*kt(x)+0.8846); % Uday Global Model

h. Id6(r) = Id(x);
i. r = r +1;

31 elseif 0.6 < kt(x) && kt(x) <= 0.8

i. s = s +1 ;

b. if 0.6 <= kt(x) && kt(x) <= 0.75 % Orgills Model
   i. Idc13(s,1) = I(x)*(1.577-1.84*kt(x));
   ii. Idc14(s,1) = I(x)*(0.9511-0.1604*kt(x)+4.388*(kt(x)^2)-16.638*(kt(x)^3)+12.336*(kt(x)^4)); % Erbs Model
   iii. Id7(s,1) = Id(x);
c. elseif 0.75 < kt(x) && kt(x) <= 0.8
   i. Idc13(s,1)= 0.177*(I(x)*kt(x));
   ii. Idc14(s,1) = I(x)*(0.9511-0.1604*kt(x)+4.388*(kt(x)^2)-16.638*(kt(x)^3)+12.336*(kt(x)^4)); % Erbs Model
   iii. Id7(s,1) = Id(x);
d. end
   i. s1 = s1+1;
e. if 0.6 <= kt(x) && kt(x) < 0.78 % Reidnl Model
   i. Idc15(s1,1) = I(x)*(1.45-1.67*kt(x));
   ii. Id9(s1,1) = Id(x);
f. elseif 0.78 <= kt(x) && kt(x) <= 0.8
   i. Idc15(s1,1)= 0.147*(I(x)*kt(x));
   ii. Id9(s1,1) = Id(x);
g. end
   i. s2 = s2 +1;
h. if 0.6 <= kt(x) && kt(x) <= 0.70 % Al Riahi Model
   i. Idc16(s2,1) = I(x)*(1.293-1.631*kt(x));
   ii. Idcn3(s2,1) = I(x)*(10.64*kt(x)^4-16.23*kt(x)^3+6.318*kt(x)^2 - 1.116*kt(x)+0.98); % Uday US model
   iii. Idcnn3(s2,1) = I(x)*(6.732*kt(x)^4-7.929*kt(x)^3+ 0.5146*kt(x)^2 - 0.3875*kt(x)+0.8481); % Uday International Model
   iv. Idcng3(s2,1) = I(x)*(8.307*kt(x)^4-11.24*kt(x)^3+ 2.729*kt(x)^2 - 0.1227*kt(x)+0.8846); % Uday Global Model
   v. Idcngl3(s2,1) = I(x)*(143.9*kt(x)^4-358*kt(x)^3+329.1*kt(x)^2-133.9*kt(x)+20.94); % Uday Global cubic Model
   vi. Id10(s2,1) = Id(x);
i. elseif 0.7 < kt(x) && kt(x) <= 0.80
   i. Idc16(s2,1)= 0.151*(I(x)*kt(x));
   ii. Idcn3(s2,1) = I(x)*(10.64*kt(x)^4-16.23*kt(x)^3+6.318*kt(x)^2 - 1.116*kt(x)+0.98); % Uday US model

iii. Idcnn3(s2,1) = I(x)*(6.732*kt(x)^4-7.929*kt(x)^3+ 0.5146*kt(x)^2 - 0.3875*kt(x)+0.8481); % Uday International Model
iv. Idcng3(s2,1) = I(x)*(8.307*kt(x)^4-11.24*kt(x)^3+ 2.729*kt(x)^2 - 0.1227*kt(x)+0.8846); % Uday Global Model
v. Idcngl3(s2,1) = I(x)*(143.9*kt(x)^4-358*kt(x)^3+329.1*kt(x)^2-133.9*kt(x)+20.94); % Uday Global cubic Model
vi. Id10(s2,1) = Id(x);

j. end

i. y = y +1;

32 else

a. Idc17(t,1)= 0.177*(I(x)*kt(x));
b. Idc18(t,1)= 0.165*(I(x)*kt(x));
c. Idc19(t,1)= 0.147*(I(x)*kt(x));
d. Idc20(t,1)= 0.151*(I(x)*kt(x));
e. Idcn4(t,1) = I(x)*(10.64*kt(x)^4-16.23*kt(x)^3+6.318*kt(x)^2 - 1.116*kt(x)+0.98); % Uday US model
f. Idcnn4(t,1) = I(x)*(6.732*kt(x)^4-7.929*kt(x)^3+ 0.5146*kt(x)^2 - 0.3875*kt(x)+0.8481); % Uday International Model
g. Idcng4(t,1) = I(x)*(8.307*kt(x)^4-11.24*kt(x)^3+ 2.729*kt(x)^2 - 0.1227*kt(x)+0.8846); % Uday Global Model
h. Id11(t) = Id(x);
i. t = t+1;

33 end
34 end
35 for x2 = 1 : p-1
36 RMSE1 = sqrt(sum((Id1(x2)- Idc1(x2,1)).^2)/p);
37 RMSE2 = sqrt(sum((Id1(x2)- Idc2(x2,1)).^2)/p);
38 RMSE3 = sqrt(sum((Id1(x2)- Idc3(x2,1)).^2)/p);
39 RMSE4 = sqrt(sum((Id1(x2)- Idc4(x2,1)).^2)/p);
40 RMSEn = sqrt(sum((Id1(x2)- Idcn(x2,1)).^2)/p);
41 RMSEnn = sqrt(sum((Id1(x2)- Idcnn(x2,1)).^2)/p);
42 RMSEng = sqrt(sum((Id1(x2)- Idcng(x2,1)).^2)/p);
43 end
44 for x3 = 1 : w-1
45 RMSE5 = sqrt(sum((Id2(x3)- Idc5(x3,1)).^2)/w);
46 RMSE6 = sqrt(sum((Id3(x3)- Idc6(x3,1)).^2)/w);
47 RMSE7 = sqrt(sum((Id4(x3)- Idc7(x3,1)).^2)/w);
48 RMSE8 = sqrt(sum((Id5(x3)- Idc8(x3,1)).^2)/w);
49 RMSEn1 = sqrt(sum((Id5(x3)- Idcn1(x3,1)).^2)/w);
50 RMSEnn1 = sqrt(sum((Id5(x3)- Idcnn1(x3,1)).^2)/w);

```
51 RMSEng1 = sqrt(sum((Id5(x3)- Idcng1(x3,1)).^2)/w);
52 end
53 for x4 = 1 : r-1
54 RMSE9 = sqrt(sum((Id6(x4)- Idc9(x4,1)).^2)/r);
55 RMSE10 = sqrt(sum((Id6(x4)- Idc10(x4,1)).^2)/r);
56 RMSE11 = sqrt(sum((Id6(x4)- Idc11(x4,1)).^2)/r);
57 RMSE12 = sqrt(sum((Id6(x4)- Idc12(x4,1)).^2)/r);
58 RMSEn2 = sqrt(sum((Id6(x4)- Idcn2(x4,1)).^2)/r);
59 RMSEnn2 = sqrt(sum((Id6(x4)- Idcnn2(x4,1)).^2)/r);
60 RMSEng2 = sqrt(sum((Id6(x4)- Idcng2(x4,1)).^2)/r);
61 end
62 for x5 = 1 :y-2
63 RMSE13 = sqrt(sum((Id7(x5)- Idc13(x5,1)).^2)/y);
64 RMSE14 = sqrt(sum((Id7(x5)- Idc14(x5,1)).^2)/y);
65 RMSE15 = sqrt(sum((Id9(x5)- Idc15(x5,1)).^2)/y);
66 RMSE16= sqrt(sum((Id10(x5)- Idc16(x5,1)).^2)/y);
67 RMSEn3 = sqrt(sum((Id10(x5)- Idcn3(x5,1)).^2)/y);
68 RMSEnn3 = sqrt(sum((Id10(x5)- Idcnn3(x5,1)).^2)/y);
69 RMSEng3 = sqrt(sum((Id10(x5)- Idcng3(x5,1)).^2)/y);
70 RMSEngl3 = sqrt(sum((Id10(x5)- Idcngl3(x5,1)).^2)/y);
71 end
72 for x6 = 1 : t-1
73 RMSE17 = sqrt(sum((Id11(x6)- Idc17(x6,1)).^2)/t);
74 RMSE18 = sqrt(sum((Id11(x6)- Idc18(x6,1)).^2)/t);
75 RMSE19 = sqrt(sum((Id11(x6)- Idc19(x6,1)).^2)/t);
76 RMSE20 = sqrt(sum((Id11(x6)- Idc20(x6,1)).^2)/t);
77 RMSEn4 = sqrt(sum((Id11(x6)- Idcn4(x6,1)).^2)/t);
78 RMSEnn4 = sqrt(sum((Id11(x6)- Idcnn4(x6,1)).^2)/t);
79 RMSEng4 = sqrt(sum((Id11(x6)- Idcng4(x6,1)).^2)/t);
80 end
81 fprintf(1,'Interval1 %5.0f\n Interval2 %5.0f\n Interval3 %5.0f\n Interval4 %5.0f\n
       Interval5 %5.0f\n ',p,w,r,y,t);
82 fprintf(1,'Rmse1 %5.3f\n Rmse2 %5.3f\n Rmse3 %5.3f\n Rmse4 %5.3f\n ',RMSE1,
       RMSE2, RMSE3, RMSE4);
83 fprintf(1,'Rmse5 %5.3f\n Rmse6 %5.3f\n Rmse7 %5.3f\n Rmse8 %5.3f\n ',RMSE5,
       RMSE6, RMSE7, RMSE8);
84 fprintf(1,'Rmse9 %5.3f\n Rmse10 %5.3f\n Rmse11 %5.3f\n Rmse12 %5.3f\n
       ',RMSE9, RMSE10, RMSE11, RMSE12);
85 fprintf(1,'Rmse13 %5.3f\n Rmse14 %5.3f\n Rmse15 %5.3f\n Rmse16 %5.3f\n
       ',RMSE13, RMSE14, RMSE15, RMSE16);
86 fprintf(1,'Rmse17 %5.3f\n Rmse18 %5.3f\n Rmse19 %5.3f\n Rmse20 %5.3f\n
       ',RMSE17, RMSE18, RMSE19, RMSE20);
87 fprintf(1,'Rmse21 %5.3f\n Rmse22 %5.3f\n Rmse23 %5.3f\n Rmse24 %5.3f\n
       Rmse25 %5.3f\n ',RMSEn, RMSEn1, RMSEn2,RMSEn3,RMSEn4);
```

```
88 fprintf(1,'Rmse26 %5.3f\n Rmse27 %5.3f\n Rmse28 %5.3f\n Rmse29 %5.3f\n
       Rmse30 %5.3f\n ',RMSEnn, RMSEnn1, RMSEnn2,RMSEnn3,RMSEnn4);
89 fprintf(1,'Rmseng31 %5.3f\n Rmse32 %5.3f\n Rmse33 %5.3f\n Rmse34 %5.3f\n
       Rmse35 %5.3f\n Rmse36 %5.3f\n',RMSEng, RMSEng1,
       RMSEng2,RMSEng3,RMSEng4,RMSEngl3);
90 end
```