4
Solar Radiation Data Quality Control

4.1 INTRODUCTION

Any decision on an active or passive solar energy application needs a database with at least an hourly average radiation values. Moreover, for setting up a solar thermal power plant, high-frequency database (1 or 10-min interval) of minimum one-year duration is required, consisting of two or three radiation components with other meteorological parameters. But as discussed in the previous chapter, measured solar radiation always has some percentage of errors in it. Discussion and identification of various climatic days (for Jodhpur location) are given in the previous chapter. Now errors due to operational reasons are to be identified using quality control conditions and guidelines. Therefore basic, advanced and additional quality control guidelines are used for detailed analysis of available radiation database. Generated results are given in terms of quality flags, which finally show various types of errors present in it. For better understanding (of mixed error present at any measurement interval), a theoretical study of different flags, their combinations and relevance to specific radiation components and their error identification is performed. All possible conditions which can cause an error in any measurement value are discussed here. Finally, to verify outcome of this study, some sample days are selected and their climate condition with different error present in each day is identified.

The raw radiation database (measurement frequency of 10-minute) of a ground-based solar radiation measurement station (IMD-Jodhpur), is used for analysis. Now randomly one-week data, containing no missing intervals is selected. For each of these days climatic condition is identified using transmittance plots. Detailed quality control guidelines are applied and the errors identified for these days are cross-verified using theoretical conclusions discussed in this chapter. Now, this process (cloud study and quality control guidelines) is applied to the available radiation database and their results are analyzed. Some days are filtered from the selected database, which has various combinations of flags. These days help us in creating a quality baseline for the whole years cloud impact and radiation potential receiving at that location.

4.2 QUALITY CONTROL GUIDELINES

For the available radiation database, cloud analysis procedure is discussed in the previous chapter. By using (BSRN and CIE) quality control guidelines (algorithms discussed in Chapter 2), basic quality control test is performed. For individual radiation component verification, measured radiation database is evaluated with clear sky radiation model relations. Station local time, site coordinates, elevation, solar constants and Sun-Earth eccentricity factor (location specific) are required and theoretical solar angles are calculated. Similarly, solar time is calculated at corresponding solar noon. Plotting solar time-step measured radiation value with solar elevation angle; one can check the quality of measurements [Ineichen, 2013]. The guidelines implemented for checking the physical limits is provided by Muneer, (2004) and all relevant literature is applied to the database. Another approach used for checking the measured data is by using a quartile analysis approach (discussed in chapter 2). Also, advanced and currently used guidelines for data quality control are discussed here in this chapter.

Current quality control guidelines, applied to Indian measurement stations, is suggested by "GIZ" in cooperation with "MNRE". Here the three radiation components are checked and errors are identified by a suggested flagging system. These tests are inter-related to each other and flag number is attached. Where "Flag 1" means data passes the quality test and no error is present in the radiation component at that interval. Remaining all other flags show some error in the measured values. The guidelines provided by C-WET (GIZ), stores data in "ASCII" format with quality control results and suitable flags are assigned for radiation component-specific error identification.

4.2.1 Steps in Solar Radiation Data Quality Control Guidelines

Available solar radiation databases provided for analysis are first checked with uniform time step analysis. Now if for some intervals, data is found missing, then they are considered as data gaps. The quality control guidelines are applied to this database and each radiation component is typically analyzed. First test checks the maximum/minimum possible boundary of measured radiation values. After that test related to correlation equation is considered and at last individual radiation component tests are discussed. For each quality control test, different flag number is assigned and all data need to pass through all these tests. For better understanding, a supporting step-by-step radiation data test is also provided here.

(a) Missing Value Test

During the time-step checking of measured radiation dataset, if the measured value is not present in the database (during the sunshine durations), then this case is considered as a data missing condition. For this interval, data is represented as "Flag 0". Here, this data missing problem may be occurring due to the data actually missing from the database, heavy cloud conditions, an instrument related, etc. Hence after a proper investigation of the real cause of this missing data is needed and further processing is suggested.

(b) Minimum Physical Limit

Equations Eq. (4.5 and 4.9) given by (Meyer *et al.,* (2009) and Long and Dutton, (2002)), are used for calculating Extra-terrestrial values. Where "ε" (Earth's eccentricity factor) is used for calculating Extra-terrestrial radiation[SPA, 2017]. Also, some constants and cosine of zenith angle factor are used to compensate for the effect of sky conditions (which radiation has to pass through while traveling towards earth's surface). The conditional relations for physically possible and extremely rare minimum limits are provided below in Eq. (4.5 and 4.9). Here the minimum limits are usually taken as -4 (W/m^2) and -2 (W/m^2), due to radiative cooling of thermal sensors [Stoffel and Reda, 2013]. Values lower than these values show offset error in the measuring instruments, hence "Flag 2" is assigned.

(c) Maximum Physical Limit

The maximum radiation values for three radiation components are provided in Eq. (4.1 to 4.4) and Eq. (4.6 to 4.8). Any measured value, if identified greater then these values, then "Flag 3" is assigned. These conditions are rare and only seen during heavy cloud traveling in front of the sun or during actual instrument failure.

$$\text{Maximum Physical limits, } GHI = DNI_o \times 1.5(\cos z)^{1.2} + 100 \quad (W/m^2) \tag{4.1}$$

$$DNI = DNI_o \quad (W/m^2) \tag{4.2}$$

$$DHI = DNI_o \times 0.95(\cos z)^{1.2} + 50 \quad (W/m^2) \tag{4.3}$$

$$DNI_o = I_E \times \varepsilon \quad (W/m^2) \tag{4.4}$$

$$\text{Minimum Physical limit, } GHI, DHI \ \& \ DNI = -4 \quad (W/m^2) \tag{4.5}$$

$$\text{Extremely Rare Maximum limits, } GHI = DNI_o \times 1.2(\cos z)^{1.2} + 50 \quad (W/m^2) \tag{4.6}$$

$$DNI = DNI_o \times 0.95(\cos z)^{0.2} + 10 \quad (W/m^2) \tag{4.7}$$

$$DHI = DNI_o \times 0.75(\cos z)^{1.2} + 30 \quad (W/m^2) \tag{4.8}$$

$$\text{Extremely Rare Minimum limit, } GHI, DHI \ \& \ DNI = -2 \quad (W/m^2) \tag{4.9}$$

(d) Clear sky limits

In the clear-sky day condition, the atmosphere is clear with no cloud presence (< 10% cloud is acceptable). Here this test checks the maximum measured limit for GHI & DNI and minimum limit for DHI limit, possible at the ground location. Suggested theoretical clear sky models can be used to determine these values and if values go above these limits, they are given "Flag 4". However, in our analysis "ASHRAE" clear sky equation is used for calculation of radiation values, because of its simplicity (see *Annexure* D).

(e) Coherence Test

The theoretical co-relation between measured solar radiation components (GHI, DHI & DNI) is checked here [Long *et al.,* 2008]. Here this test checks the basic ratio of measured and calculated GHI value. The range is selected based on inherent error present in the instrument and an unavoidable error that occurs at the optical side. These guidelines provided in [Maxwell *et al.,* 1993] are still used as standard conditions for solar radiation data quality control. Where if "solar zenith angle" is less than 75 degrees, then "± 8 %" error is allowed, for greater than 75 degrees till 93 degrees, the range is increased up to "±15 %" (high error occurring at sunrise and sunset time). If the value is falling outside this range, they are given "Flag 5" (see Eq. (4.10) and Eq. (4.11)).

$$\left| 1 - \frac{GHI}{DHI + DNI(\cos(z))} \right| > 0.08, \text{ for } z < 75^o \tag{4.10}$$

$$\left| 1 - \frac{GHI}{DHI + DNI(\cos(z))} \right| > 0.15, \text{ for } 93^o > z > 75^o \tag{4.11}$$

(f) Diffuse Error

For identifying an error in diffuse radiation component, the ratio of diffuse to global radiation components is checked and limits are provided in the Eq.(4.12) and Eq. (4.13). The data failing in this test is given "Flag 6". Where the occurrence of this flag is usually seen under heavy cloud conditions day or cleaning issue at the instrument side.

$$DHI/GHI > 1.05 \quad , \text{for } z<75^o \tag{4.12}$$

$$DHI/GHI > 1.1 \quad , \text{for } 93^o > z > 75^o \tag{4.13}$$

(g) Rayleigh Limit

It is the theoretical limit provided for only diffuse radiation component, where the atmosphere is clear and dry sky conditions are followed [Long *et al.,* 2008]. Here the calculation is done for determining molecular scattering due to the presence of aerosol in the atmospheric column. Hence, by using various scattering values, a relation is proposed [Younkin *et al.,* 2004], where the diffuse limit is calculated using Eq.(4.14). But the constants used here are derived using only US location stations, hence not used in this thesis.

$$R_L = d(\cos(z)) + e(\cos(z))^2 + f(\cos(z))^3 + g(\cos(z))^4 + h(\cos(z))^5 + i(\cos(z))\,\mathrm{Pr}s \quad (4.14)$$

Where:
d=209.3, e=-708.3, f=1128.7, g=-911.2, h=287.85, i=0.046725 (Coefficients)
Prs =station surface pressure in milli-bars

(h) Tracking error

This error appears when DHI and DNI instruments are not correctly aligned towards the sun. Hence to verify this errors, two independent conditions (Eq.(4.15 and 4.16)) are given [Long *et al.,* 2008] and both are compared with their maximum possible component at that interval. If for both these conditions, the ratio is greater than 0.85, then recorded data observation is considered to be having tracking error (Flag 8 is assigned).

$$\frac{DHI + DNI(\cos(z))}{GHI_{clear}} < 0.85 \quad (4.15)$$

$$\frac{DHI}{DHI + DNI(\cos(z))} < 0.85 \quad (4.16)$$

Every dataset needs to pass through all these discussed tests, and for details see provided flowchart in Table 4.1. Where only data missing test (Flag 0) is given flag directly, remaining all data has to undergo this process. Now, these calculated QC results are stored in a standard format. For any location, the raw database provided is not disturbed at all, but one additional column of QC test results is included in it. While analysis, if these identified errors and missing values are less than "5%" of the whole database, then one may ignore them. Otherwise, one needs filling them using relevant gap filling approaches, which are discussed in the next chapter.

4.2.2 Additional Data Quality Control Tests

Other than the tests discussed above, some other special tests are listed here, which can be applied according to location-specific conditions. These tests are applicable for detecting special errors in measured radiation database values, like step test (Snow and shadow presence), persistence, spatial and sunshine duration tests.

(a) Snow Detection Test

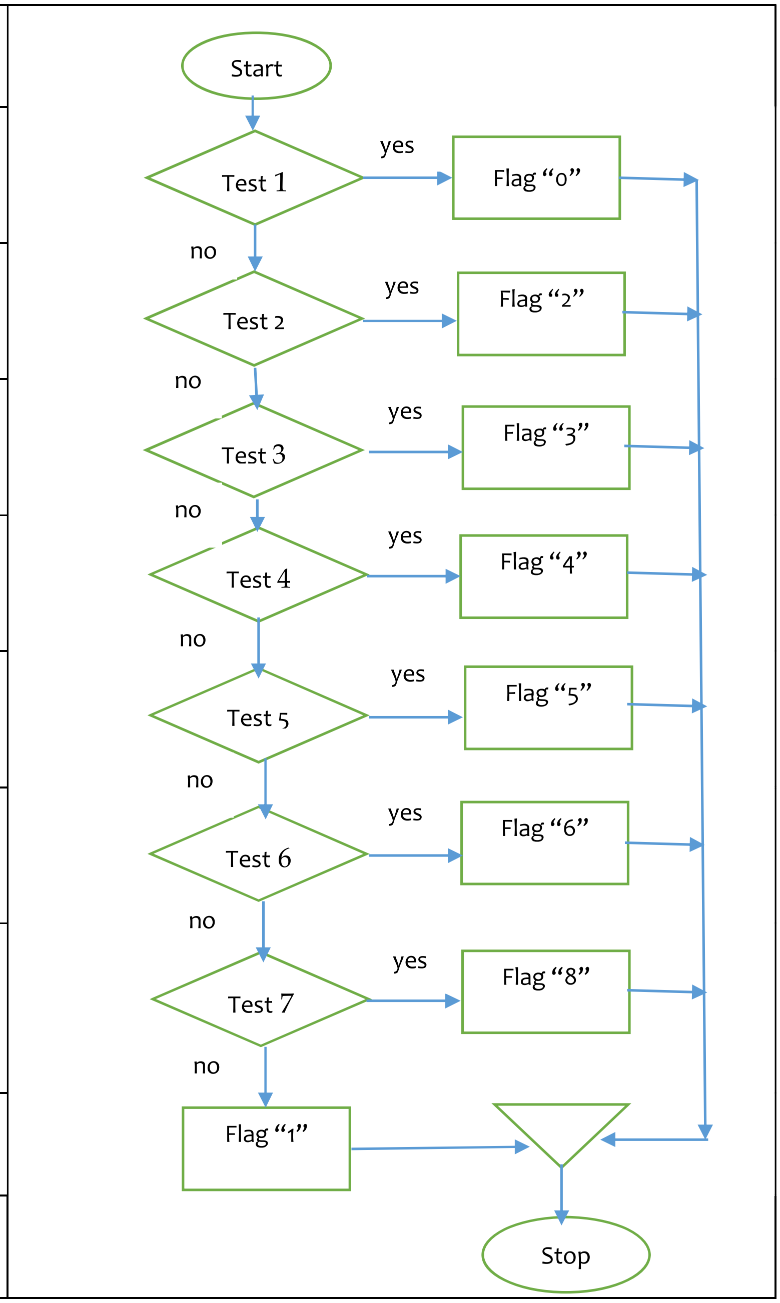
Under snow conditions, due to the white reflective background, the majority of radiation gets reflected back to space, with little/no energy absorption. Hence if the ratio of beam radiation component and global radiation component is greater than unity and along with that reflected radiation component with global radiation component is less than 0.7, then snow presence is confirmed (see Eq.(4.17) and (4.18)).

$$(DNI / GHI) \geq 1 \quad (4.17)$$

$$(RDNI / GHI) \geq 0.7 \quad (4.18)$$

Table 4.1: Flowchart for Basic Quality Control Guidelines

Start	
Missing Data GHI or DHI or DNI=0	
Minimum Physical Limit GHI or DHI or DNI < -4 (W/m^2)	
Maximum Physical Limit GHI or DHI or DNI > Max limit	
Clear Sky Limits GHI or DHI or DNI > Clear sky	
Coherence Test GHI=DNI*cos(z)+ DHI	
Diffuse Error DHI < GHI	
Tracking Error GHI_C / GHI_{clear} < 0.85 DHI / GHI_C < 0.85	
Merge	
Stop	



(b) Spatial Consistency Test

Radiation data of nearby measurement stations is correlated here. Due to the variation of cloud cover over these stations, measured data similarity is not possible at high measurement frequency rates (1, 10 or 30-minute interval). But the daily total of global solar irradiance data of near stations can be compared with each other [Terzenbach, 1995]. The selected station for data analysis is compared with the nearby station using spatial interpolation relation. Here the distance between stations is converted into an inverse weighted function, including in the calculation. Now if the factor is greater than 50%, then station needs further data analysis (see Eq. (4.19) to Eq. (4.21)).

$$\hat{G}(x_o) = \frac{1}{\sum_{z=1}^{N} w_z} \sum_{z=1}^{N} w_i G(x_z) \tag{4.19}$$

$$w_z = \left(\frac{1 - d_z/R}{d_z/R} \right)^2 \tag{4.20}$$

$$Bais = \left| \hat{G}(x_o) - G(x_o) \right| \tag{4.21}$$

Where,
x_o = Station location
x_z= Next station location
N = No. of stations
d_z = Station in between distance (x_o with other stations)
w_z = Interpolation coefficient
R = Maximum distance between stations
$\hat{G}(x_o)$ = Sum of all station in spatial consistency test
$G(x_o)$ = Individual station value in spatial consistency test

(c) Step Test

This test is applied to understand the reason behind the variation of measured radiation values. It is suggested in the literature, that there should be a maximum threshold limit between the variations of two successive time-steps. Where for a 1-minute interval radiation database, this limit is 800 (W/m^2), due to the presence of moving clouds [Shafer *et al.*, 2000; WMO, 2007]. But for 10-minute interval data due to high variations, normalizing the measured value is done by dividing each one of them with Extra-terrestrial values. Limits are provided for all three radiation components (see Eq. (4.22) to Eq. (4.24)), where solar elevation angle should always be greater than 2 degrees.

$$\left| \frac{G(t)}{E(t)} - \frac{G(t-1)}{E(t-1)} \right| < 0.75, \text{ if } h(t) > 2^o \tag{4.22}$$

$$\left| \frac{B(t)}{E(t)} - \frac{B(t-1)}{E(t-1)} \right| < 0.65, \text{ if } h(t) > 2^o \tag{4.23}$$

$$\left| \frac{D(t)}{E(t)} - \frac{D(t-1)}{E(t-1)} \right| < 0.35, \text{ if } h(t) > 2^o \tag{4.24}$$

(d) Shadow Detection Test

This test is similar to the step test, but here if the absence of solar radiation measured is for a longer duration, then its identification is possible. Hence this error occurs due to the blocking of the sensor (presence of shadow or any other solid objects). The conditions used in this test is given below (see Eq. (4.25) to (4.28)).

$$h(t) > 2^o \tag{4.25}$$

$$B(t) \geq G(t) \text{ and } B(t-1) < G(t-1) \tag{4.26}$$

$$\left|\frac{G(t)}{E(t)}-\frac{G(t-1)}{E(t-1)}\right|>0.1 \qquad (4.27)$$

$$\frac{G(t)-G(t-1)}{B(t)-B(t-1)}>3 \text{ or } <-1 \qquad (4.28)$$

(e) Persistence Test

This test checks, the measuring ability of solar radiation measuring instruments, which depends on the quality and health of measuring system and sensors. Failure of this test shows failure of the sensor, so quick actions are needed. Otherwise if not checked, it will keep on measuring low or high and consistently wrong measured values (even daily mean and standard deviation estimation may be wrong). Hence a limit is provided on these values (daily mean and standard deviation) in relation to measured radiation values. The data exceeding the standard deviation value limits need to be flagged (see Eq. (4.29) to Eq. (4.32)).

$$\frac{1}{8}\mu\frac{G}{E}\leq\sigma\left(\frac{G}{E}\right)\leq 0.35 \qquad (4.29)$$

$$\frac{1}{6}\mu\frac{D}{E}\leq\sigma\left(\frac{D}{E}\right)\leq 0.2 \qquad (4.30)$$

$$\lambda\leq\sigma\left(\frac{B}{E}\right)\leq 0.3 \qquad (4.31)$$

$$\lambda=\begin{cases} 0 & if,\mu\left(\frac{B}{E}\right)\leq 0.01 \\ \frac{1}{2}\mu\left(\frac{B}{E}\right) & if,0.01<\mu\left(\frac{B}{E}\right)\leq 0.2 \\ 0.1+\frac{1}{20}\left(\mu\left(\frac{B}{E}\right)-0.2\right) & otherwise \end{cases} \qquad (4.32)$$

(f) Sunshine Duration Test

Sunshine duration falling on the surface also helps in determining the solar radiation quality. Only Direct radiation component is considered here for analysis, according to [WMO, 2007]. For the overcast day (sunshine duration is relatively zero), maximum daily averaged direct radiation component can't exceed 120 (W/m^2) when calculated only during the period of sunshine duration. For other days, the ratio of actual and maximum possible sunshine duration averaged direct radiation component should be greater than the lower bound value (120 (W/m^2)) (see Eq. (4.33 to 4.35)).

$$0\leq SD\leq SD_{\max} \qquad (4.33)$$

$$SD=0\ ,DNI<120\ (W/m^2) \quad \text{[Over-casted Day]} \qquad (4.34)$$

$$SD>0\ ,(SD/SD_{\max})\geq 120\ \ (W/m^2) \qquad (4.35)$$

4.3 MODIFIED QUALITY CONTROL GUIDELINES

Guidelines discussed above for solar radiation quality control can give the overall measurement quality. But actual estimation of error in individual radiation components is not discussed anywhere. To better understand the measurement process and how various errors influence it, a theoretical study is required. Hence, in this study of errors in measured radiation database, one assumes possible failure conditions of radiation components and correspondingly the reason behind this condition is analyzed. Finally, this needs verification from quality flag results and combination derived from actual quality control results. This can help in proposing new guidelines.

4.3.1 Erroneous Radiation Components- Flags Determination

In this section, the possible cases of flag results, which result in radiation data analysis of any frequency (quality control flags) are analyzed. All possible variations applied to radiation measurement process are considered, to determine the valid reason for the error in radiation components. Also, it is assumed that all data has to pass maximum and minimum physical limit test. If not, consider them as data missing case and assign "Flag 0". Moreover, the "Rayleigh limit" test is not applied here, as the equation available in the literature is not found eligible for the selected location. For a clear sky day, diffuse limit is checked.

Quality flag outcomes from each radiation database interval are analyzed with various quality control tests conditions as shown here. Then the possible occurrence of error flag combination is discussed.

(a) All Radiation Components Passing the Quality Control Tests

(1) Correctly Measured Radiation Database (Flag 1)

According to the literature, if all values are measured correctly with proper operation and maintenance schedule, then all data passes this quality test. All QC tests will get passed and no error is found in the radiation database.

(2) Tracking Error is Present, but no Coherence Error Present (Flag 8, 1)

Data failing in tracking limit test but passes the coherence test. This concludes the presence of radiation averaging issue (due to any reason) or there is some dust presence on all sensors.

(b) All Radiation Components Failing in the Quality Control Tests

Data fails in all quality control tests, which implies the urgent maintenance requirement of the station. Either the instruments become faulty or there can be some serious operation problem.

(1) Station Maintenance poor (Flag 4, 5, 6, 8)

Sensor cleaned less frequently, due to this glass window is blocked and all radiation measured values are less than actual. Another reason for radiation blockage may be due to the bird's presence at the measurement site.

(2) Instruments Tracking and Alignment (Flag 4, 5, 6, 8)

Either no or fewer values are measured from the sensors. Some measured data is greater than clear sky values, which shows the malfunctioning of sensors.

(3) Data Communication Error (Flag 4, 5, 6, 8)

Data is measured but at the server level, the values are not able to get stored, hence all values become zeros.

(c) Anyone Radiation Component is Failing in the Quality Control Tests

(1) Only GHI component is Found Erroneous

- Instrument not connected for some period, a default high value (or no value) is recorded by the sensor for the whole period. (Flag 4, 5, 8)
- Due to the motion of heavy clouds under the influence of sun rays, a concentration effect occurs and values greater than clear sky values are recorded for that period. (Flag 4, 5, 8)
- Due to dust presence on the sensor, fewer values are measured. Manual or artificial blockage of GHI sensor by birds or shading effects (Flag 4, 5, 8)

(2) Only DNI Component is Found Erroneous

- Both coherence and tracking error is present but after analysis, only the DNI component is corrected. (Flag 4, 5, 8)
- The instrument by mistake interchanged with GHI sensor (Flag 5, 8) and with DHI sensor (Flag 6, 8)
- Solar tracker not working properly. Only tracking error is coming and DHI component is found acceptable (Flag 6, 8)
- Manually or artificially blockage of DNI sensor by birds or shading effects (Flag 4, 5, 8)

(3) Only DHI Component is Found Erroneous

- Instrument not connected (Flag 5, 6, 8)
- Shading disc missing and tracking error (Flag 5, 6, 8)
- Solar tracker not working properly. Only tracking error is coming and DNI component is found acceptable (Flag 6, 8)
- Only Diffuse and Tracking error (Flag 6, 8)
- Manual or artificial blockage of DHI sensor by birds (Flag 5, 6)
- Only Diffuse and Coherence error (Flag 6, 8)

(d) Any Two Radiation Components Failing the Quality Control Tests

(1) GHI and DHI Component are Found Erroneous

- On heavy cloud conditions during the rainy period. Due to cloud concentration, GHI is greater than the clear sky value and due to water cooling of sensors logically less DHI values are recorded. (Flag 5, 6)
- Tracking error but DNI component is measured correctly, shade ring missing error can be present. (Flag 5, 6)
- Sensor cleaning issues, under heavy-medium cloud condition days (Flag 5, 6)
- In snow area, due to the blocking of sensors (Flag 5, 6)
- Only Diffuse and Coherence error (Flag 6, 8)

(2) GHI and DNI Component is Found Erroneous

- On cloudy condition days and dust present on the sensors (Flag 4, 5, 8)
- Tracking error but DHI component is under acceptable range (Flag 4, 5, 8)
- Instruments not working or connected to the system (Flag 4, 5, 8)
- Only Tracking and Coherence error (Flag 5, 8)

(3) DHI and DNI Component is Found Erroneous

Usually, if no error is found in GHI values, then the correction of DHI and DNI is not done. But if identified with dust presence, all three components need to be corrected.

- On cloudy condition days and dust present on the sensors (Flag 4, 8)
- Tracking or alignment error and GHI measured is good (Flag 8)
- Instruments not working or connected to the system (Flag 4, 8)

No discussion is available, regarding theoretical combination of measurement errors and the possible reasons for which these might occur. Here a conceptual discussion is suggested, by analyzing various information available in the literature. For actual verification of these findings, they are applied to ground-based measured radiation data analysis. Now just for quick testing of these conclusions, one-week (randomly selected), and 10-minutes interval (uniformly) measured radiation data from IMD-Jodhpur is selected. Detailed quality control analysis is done for each day and corresponding plots and error flag information is discussed.

4.4 DETAILED QUALITY CONTROL GUIDELINES APPLIED TO RANDOMLY SELECTED RADIATION DATASET

IMD-Jodhpur database from 14-20 March 2016, 10-minute uniformly averaged radiation dataset is analyzed here. Now the first step applied to this database is its time-stamp analysis and checking the availability of other measured radiation and meteorological databases. All basic and advanced quality control guidelines are applied to the dataset, as discussed above. After that generated quality control results are in-depth analyzed and information regarding cloud conditions, dust presence, tracking error, etc. is concluded from it.

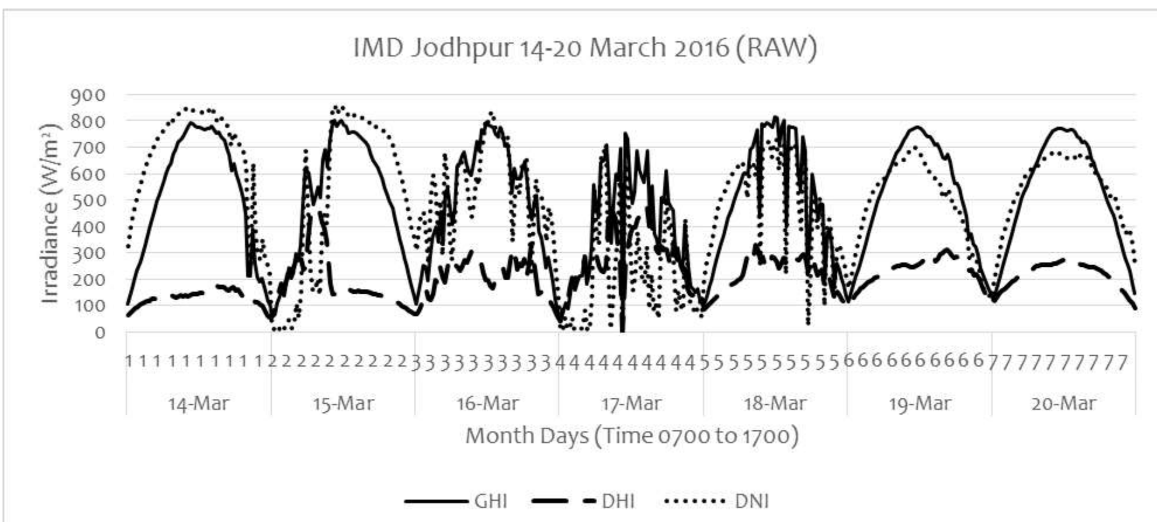


Figure 4.1: Basic Radiation Plot of three radiation Components (14 to 20 March 2016, IMD-Jodhpur)

4.4.1 Basic Visual Data Analysis

All days selected for analysis do not have any missing interval. Also, sunshine duration (0700 to 1700 hours) is selected, as it simplifies our dataset analysis procedure. Now after this modification, this database is plotted under the axis of irradiance and time duration (see Fig. 4.2). Basic visual analysis of this plot shows, daily climatic variation at the location. Now for further analysis, each day "k_t–k_n", "k_t–k" and "measured-calculated GHI" plots are produced (see Fig. 4.2 (a, b and c) to Figure 4.8 (a, b and c)) and comparison is done with historically available plot provided by Maxwell *et al.*, (1993).

Now by simply looking at this radiation dataset, one concludes that GHI measurement values are uncorrelated (analyzed using measured-calculated GHI plot), with respect to other radiation component values. Now each day is analyzed and final cloud type findings are summarized in Table 4.2. For 15 and 16-March 2016, after analyzing all plots shown in Fig 4.2 (a, b and c) and Fig 4.3 (a, b and c), both days are classified as mixture of clear-sky and passing

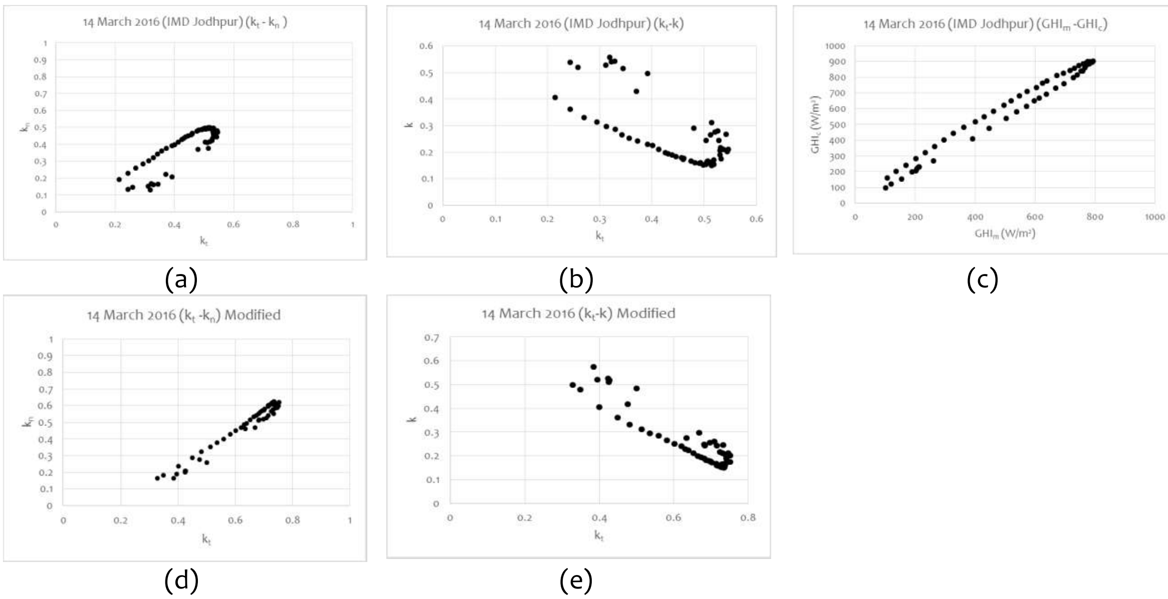


Figure 4.2: (a)k_t-k_n; (b) k_t-k; (c)GHI (Measured-Calculated); (d) k_t-k_n (Modified) and (e) k_t-k (Modified) (14 March 2016, IMD-Jodhpur)

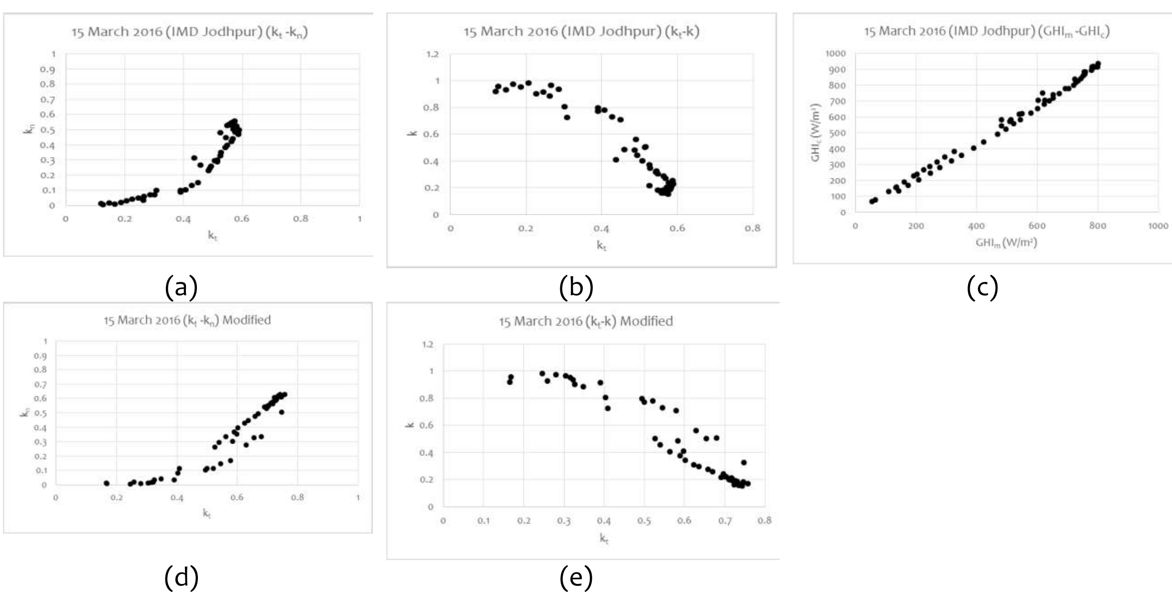


Figure 4.3: (a)k_t-k_n; (b) k_t-k; (c)GHI (Measured-Calculated); (d) k_t-k_n (Modified) and (e) k_t-k (Modified) (15 March 2016, IMD-Jodhpur)

cloud. However, the plotted data has wide spread, which shows measurement error present in radiation components. Now for 16 and 18-March 2016, presence of moving medium-size cloud verifies cloud condition is partly clouded. Due to these type of clouds, drastic change in radiation values is seen, which makes transmittance plot widely spread. Where on 17-March 2016, the day is identified with heavy cloud presence and partly clouded sky cloud condition. Here relatively low DNI values and similar GHI and DHI values are measured and the corresponding transmittance plot is also shown. Similarly, on 19 and 20-March 2016 a uniform thickness of cloud presence makes the day identified as haze cloud condition day.

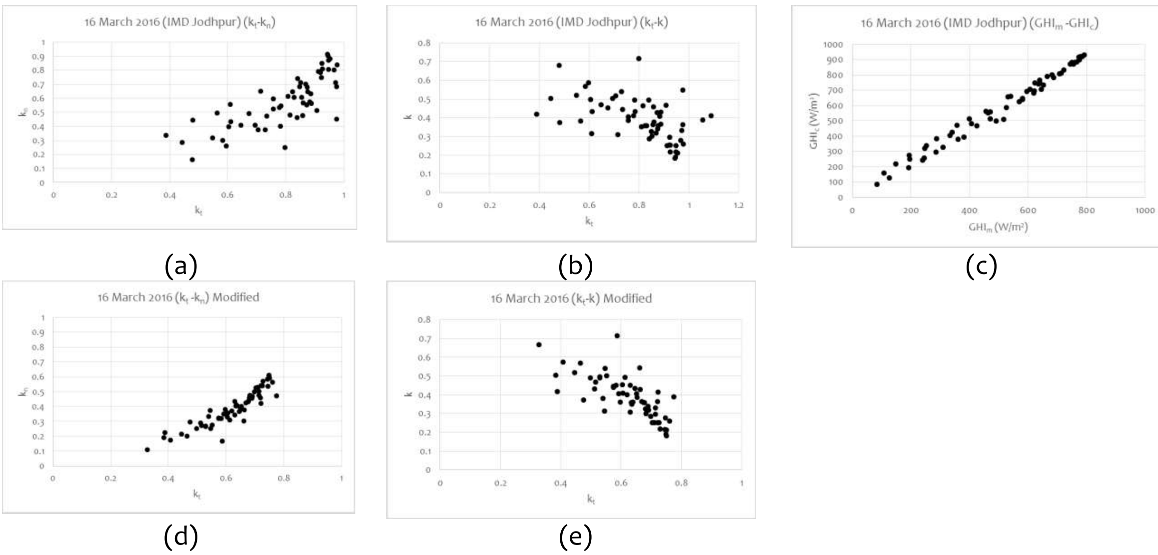


Figure 4.4: (a)k_t-k_n; (b) k_t-k; (c)GHI (Measured-Calculated); (d) k_t-k_n (Modified) and (e) k_t-k (Modified) (16 March 2016, IMD-Jodhpur)

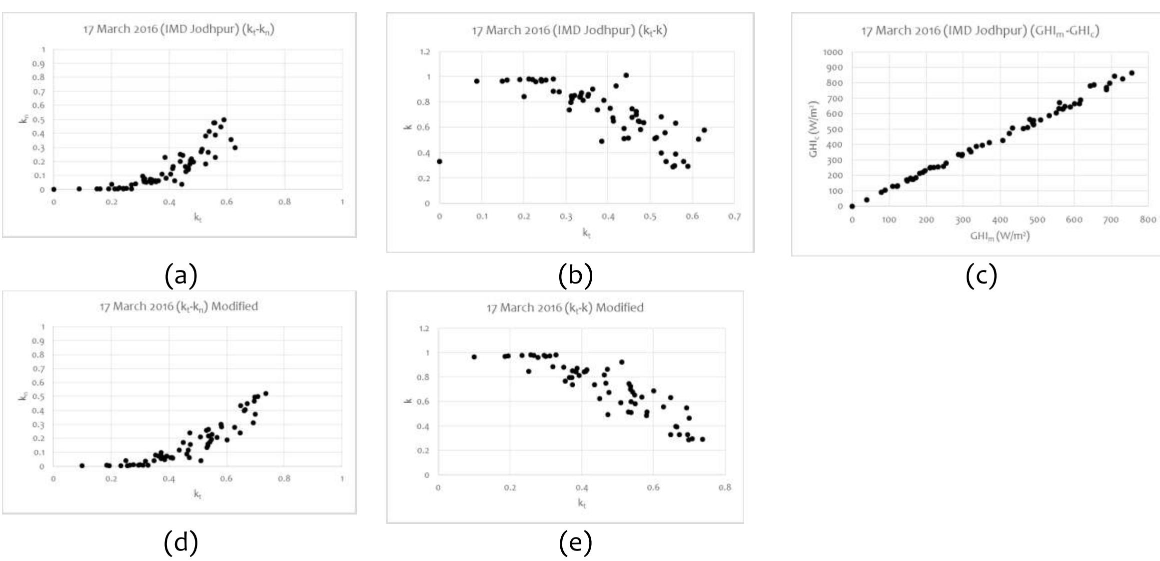


Figure 4.5: (a)k_t-k_n; (b) k_t-k; (c)GHI (Measured-Calculated); (d) k_t-k_n (Modified) and (e) k_t-k (Modified) (17 March 2016, IMD-Jodhpur)

But viewing only the plotted results (shown above), the exact error in the individual radiation component can't be identified. For example, 14-March and 16-March 2016 (k_t–k_n) in scatter plot some data points are found lying on or above the diagonal axis. Generally, data lying outside the proposed boundary region is termed as erroneous and hence removed or not considered during potential conclusion process. But if these values are corrected by any guideline, they can help in actual solar potential estimation. Hence, some additional tests and correlating conditions are applied to this database (detailed quality control guidelines). Now the actual reason for measurement error and its presence on individual radiation component is identified.

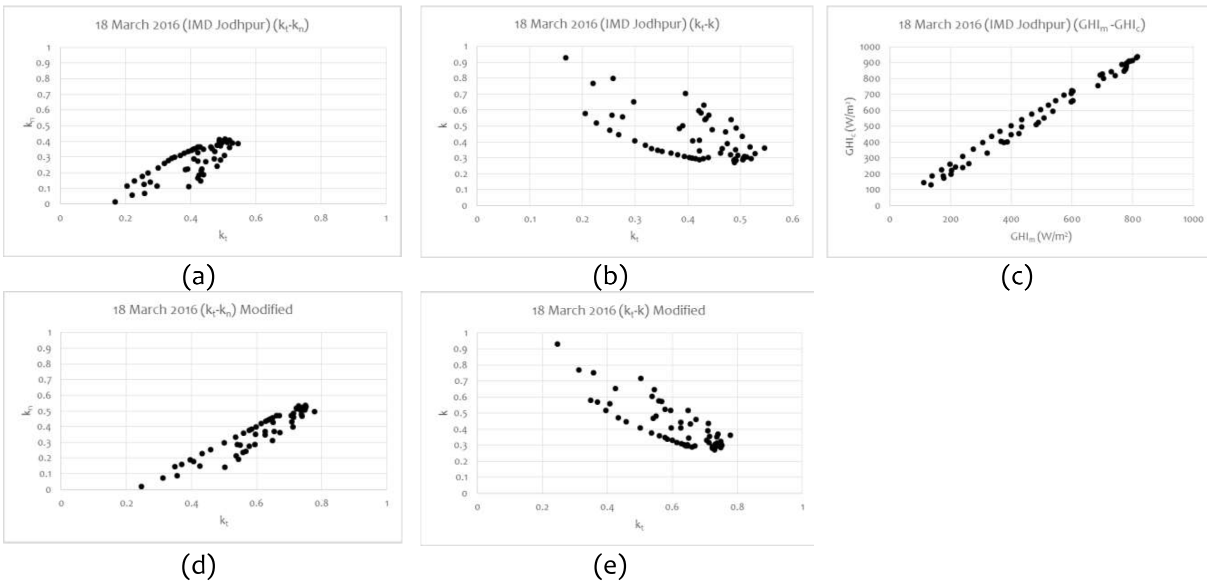


Figure 4.6: (a)k_t-k_n; (b) k_t-k; (c)GHI (Measured-Calculated); (d) k_t-k_n (Modified) and (e) k_t-k (Modified) (18 March 2016, IMD-Jodhpur)

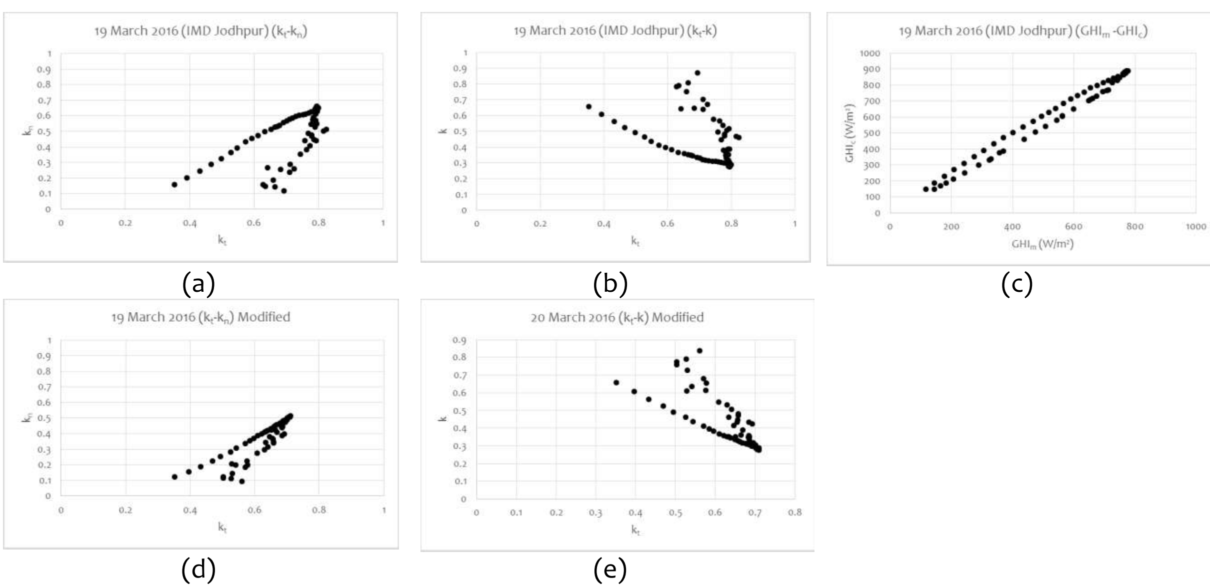


Figure 4.7: (a)k_t-k_n; (b) k_t-k; (c)GHI (Measured-Calculated); (d) k_t-k_n (Modified) and (e) k_t-k (Modified) (19 March 2016, IMD-Jodhpur)

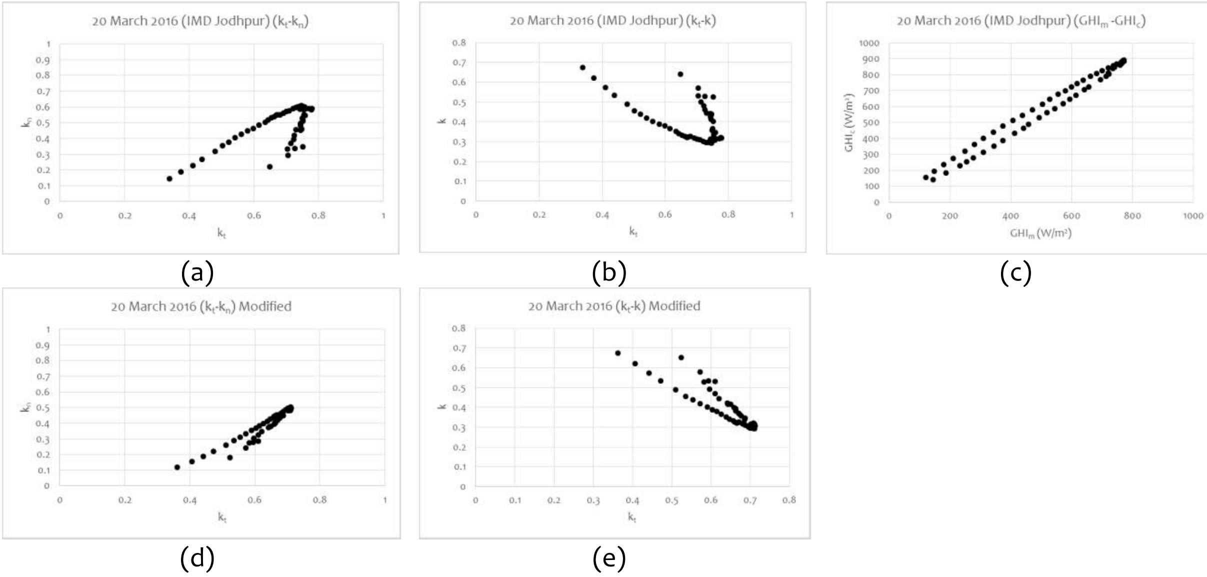


Figure 4.8: (a)k_t-k_n; (b) k_t-k; (c)GHI (Measured-Calculated); (d) k_t-k_n (Modified) and (e) k_t-k (Modified) (20 March 2016, IMD-Jodhpur)

Table 4.2: Climate Prediction Results (14-20 March 2016, IMD- Jodhpur)

Day	*Climate Prediction*
14 March 2016	Clear Sky with Some Cloud Presence in Evening
15 March 2016	Heavy Cloud Presence in Morning and then Clear Sky
16 March 2016	Whole Day Medium Moving Cloud Presence
17 March 2016	Whole Day Heavy Moving Cloud Presence
18 March 2016	Whole Day Medium Moving Cloud Presence
19 March 2016	Whole Daylight Cloud-like Haze Present
20 March 2016	Whole Daylight Cloud-like Haze Present

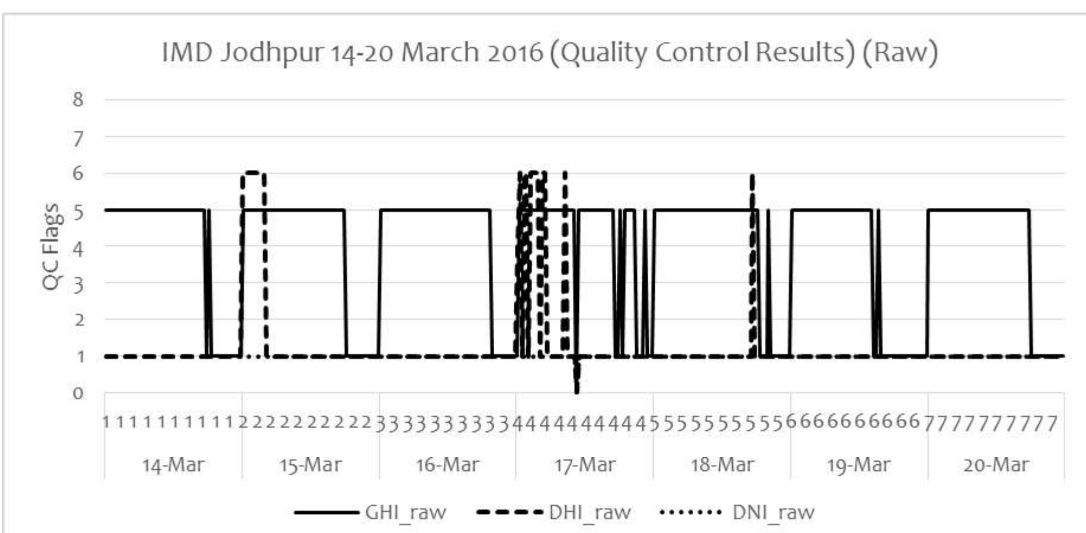


Figure 4.9: Quality Control Results (14-20 March 2016, IMD-Jodhpur)

4.4.2 Advance Data Quality Control Procedure

After identification of different cloud conditions present in a selected radiation database, quality control guidelines and other relevant tests are used to determine radiation database quality. Hence, using the flowchart in Table 4.2, measured radiation data sets are checked. They need to pass all tests and results generated are in terms of quality flags. Now Fig. 4.9, shows the final flag results plotted after applying detailed quality control test on the available database. It is identified that maximum data fails in a coherence test with some duration of tracking error present. Now coherence error comes when all three radiation components don't follow the physical radiation relationship between them, whereas tracking error comes when radiation sensors are not aligned with sunlight or heavy cloud presence. Moreover, for advanced understanding, step test is performed to each radiation component and if data are measured properly, the result should be in provided range (see Eq. (4.22) to Eq. (4.24)).

Step test results for each radiation component are shown in Fig. 4.11 (a, b, c). Where during the cloud conditions some variations are identified, but after that all values are in their suggested step range. Hence, one can conclude that the measuring instrument is working correctly, but the step change is low. Now only reason by which these errors can occur is due to instrument cleaning issues. The presence of dust on GHI measuring instrument, affects the radiation receiving at the sensor (similarly the DNI and DHI component). Hence during data correction due to cleaning issues, the first correction is applied for DNI radiation component (identified during the tracking error) and after that correction in GHI radiation component.

Now using any suggested gap-filling approach, one can correct these values (techniques yet to be discussed in the coming chapter). Now just for verification of our corrections for this database, [Kumar *et al.,* 2014] a relevant gap filling guideline is applied. Here if out of three radiation components, only one or two components need correction, then gap is filled by using physical radiation relationship (see Eq. 5.5) or by using interpolation approaches. Now after correcting the identified erroneous measurement values from the database, all data pass the quality control test (see Fig. 4.10). Moreover, the identified change in measured and corrected values for each radiation component is provided in Fig. 4.12 (a, b and c) for each radiation component. Modified databases calculated transmittance plot is shown in Fig. 4.2 (d and e) to Fig. 4.8 (d and e). After analyzing the results look reliable as compared to available raw database, and after considering cloud condition for that day, sites real radiation pattern can be identified.

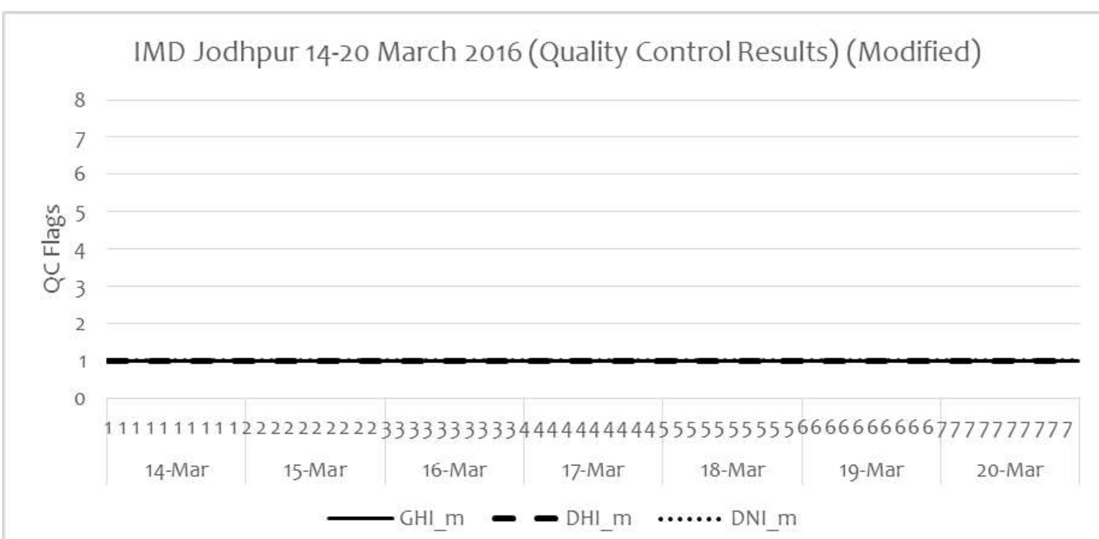
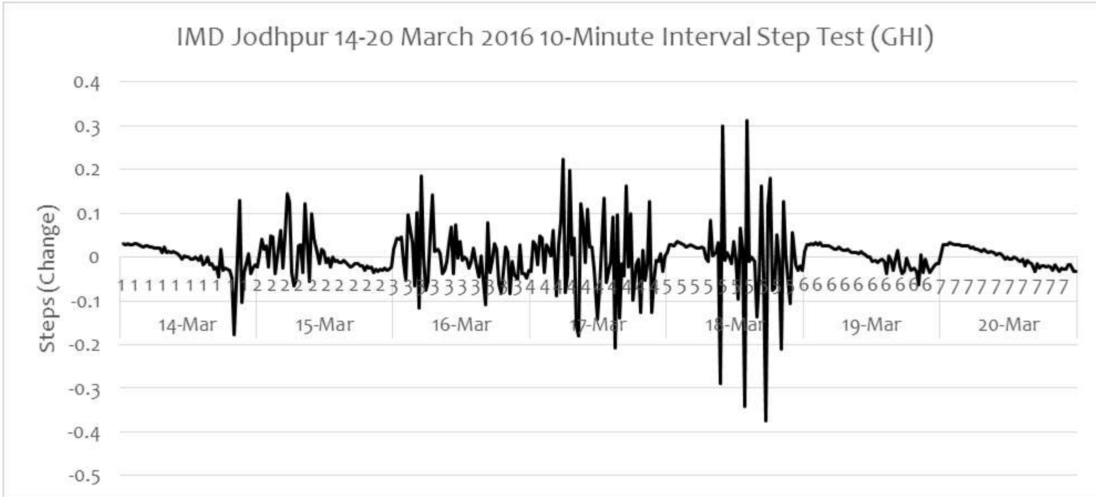
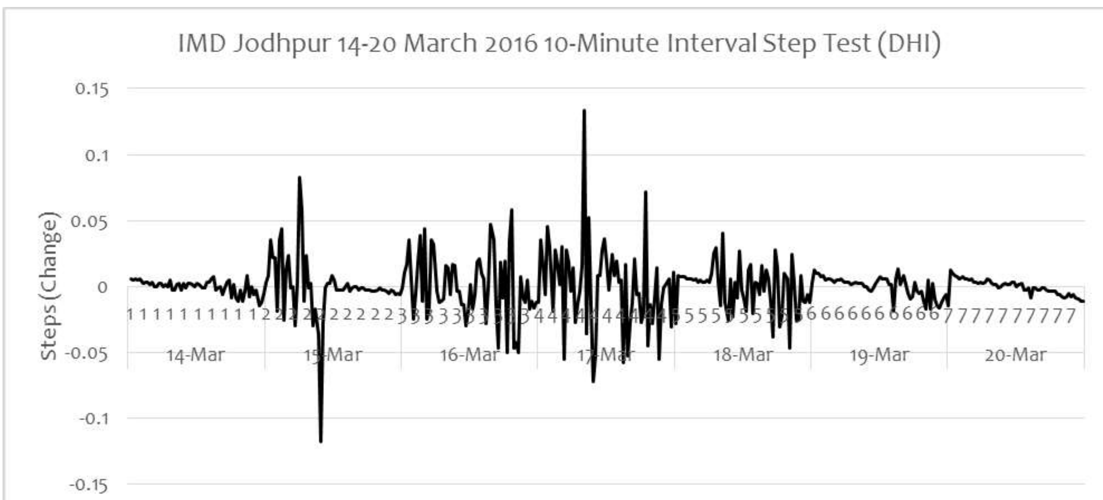


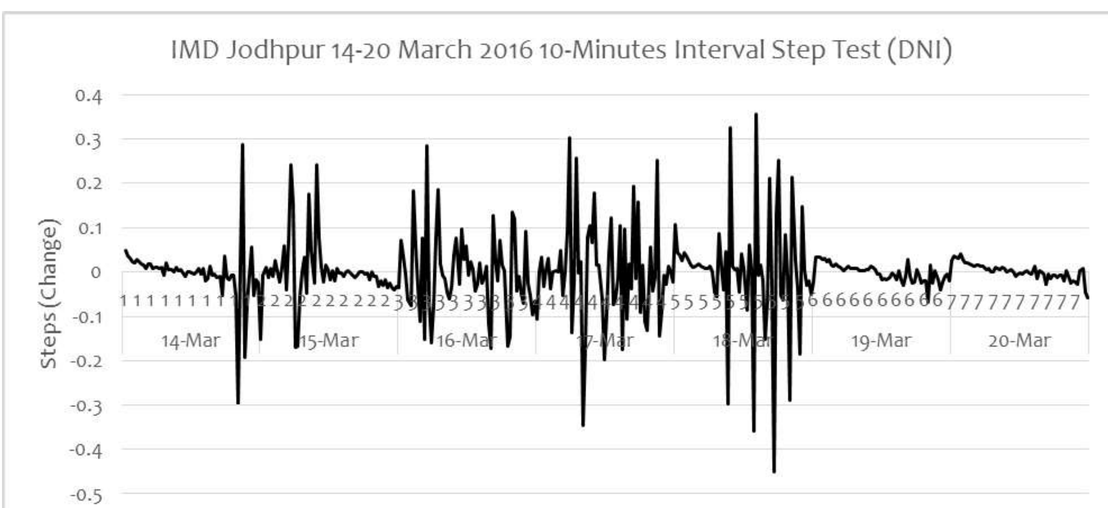
Figure 4.10: Corrected Database Quality Control Results (14-20 March 2016, IMD-Jodhpur)



(a)

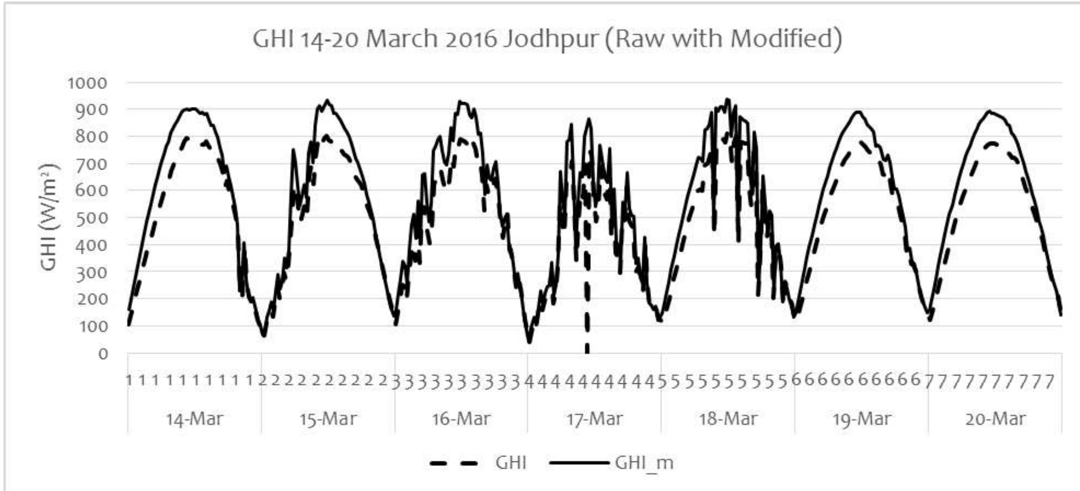


(b)

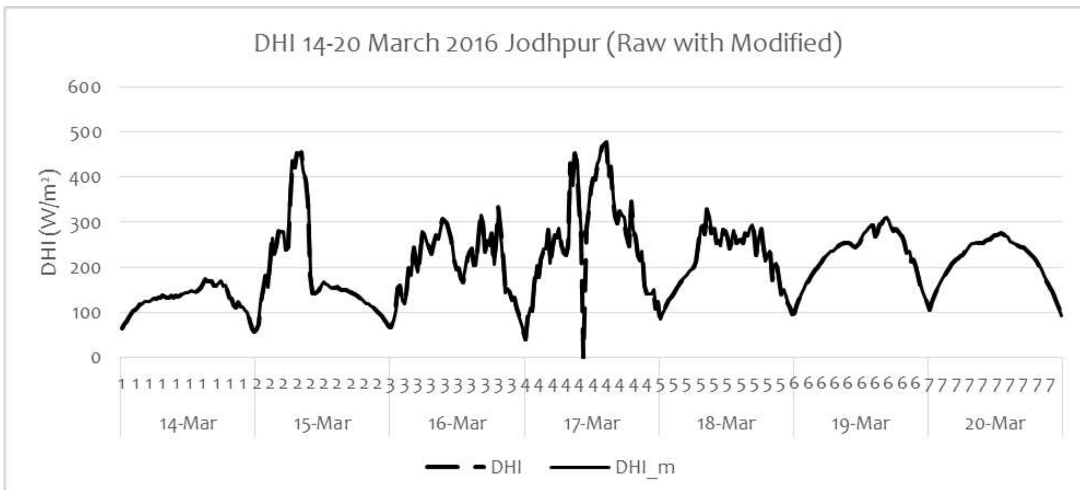


(c)

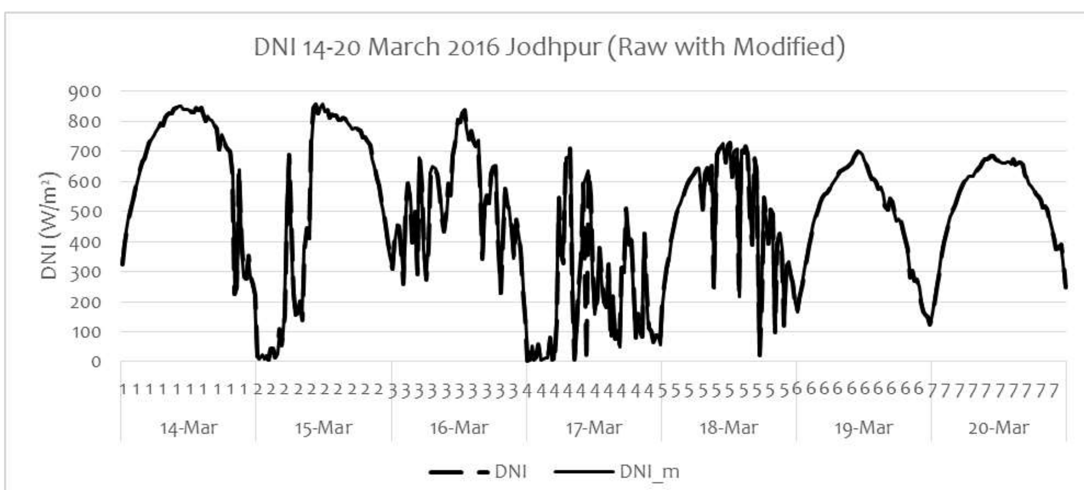
Figure 4.11: Step Test Results (a. GHI, b. DHI and c. DNI) (14-20 March 2016, IMD-Jodhpur)



(a)



(b)



(c)

Figure 4.12: Difference in Measured and Corrected Radiation Database (a. GHI, b. DHI and c. DNI) (14-20 March 2016, IMD-Jodhpur)

4.5 MIXED ERRORS IDENTIFICATION FROM THE DATABASE

The database discussed above deals with a simple case of operational (cleaning and maintenance) issue, which is common in most stations. Moreover, in reality, all days are not similar, due to different errors present at each interval, makes the data less reliable for analysis. Now from the available database (2015 IMD- Jodhpur), some unique days are selected on the basis of the pattern of measured radiation values: Days having same errors in different components. The local condition analysis and detailed quality control procedures are applied. The results obtained by this analysis are classified based on various quality flags, which are then used in terms of various errors present in the database. Data when compared with other radiation parameters and its clear sky values, results in the proper identification of individual errors in the measurement.

Now considering all feasible errors present in the measurement system, for each day, errors due to operational and instrument errors are identified. For determining error in each radiation component data is plotted. As seen in Table 4.3 and Fig. 4.13, the day is identified with a recording of low DNI value (one occurrence of haze type cloud is also present) for that duration, and other radiation components (DHI and GHI) are similar. Hence, heavy clouded sky conditions presence is verified, derived from transmittance plots and measured–calculated GHI plot. In quality control flags analysis, during the heavy cloud conditions the presence of tracking error in the measurement process is common but diffuse and coherence error at same interval shows, possible error due to cleaning issues. Hence cleaning of all measuring instruments is advised and data correction of GHI is suggested. Furthermore, in Table 4.4 and Fig. 4.14, for all day "haze like cloud" condition, is identified, with similar GHI and

Table 4.3: Detailed Climate, Quality Control and Error Information (1-January 2015)

Climate	: Heavily clouded conditions
Quality Control Errors	: Diffuse, Coherence and Tracking Error
Operational Errors	: Dust presence on DNI and GHI sensor
Instrumental Errors	: No

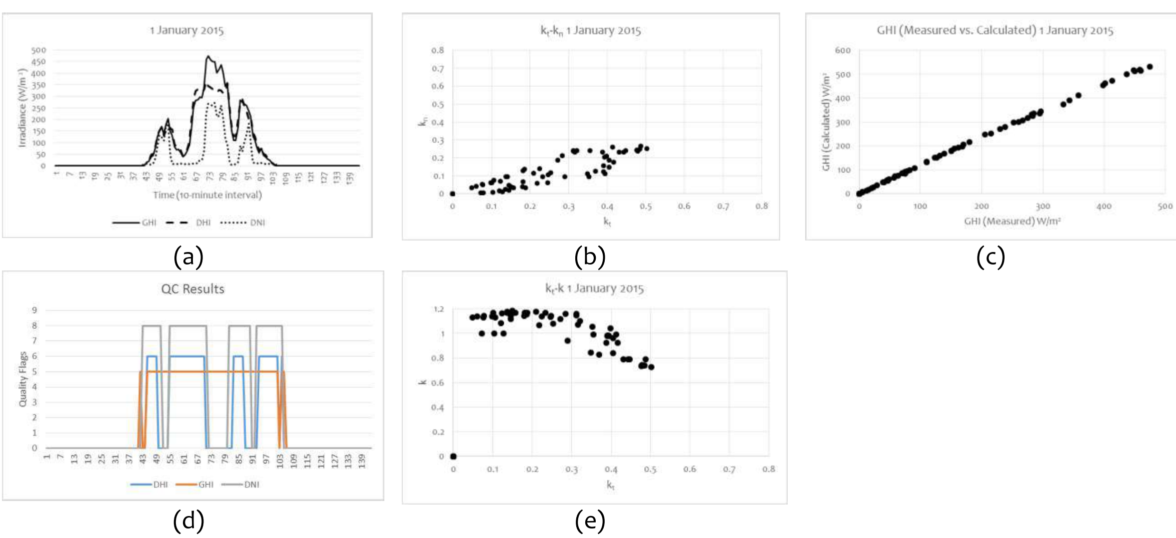


Figure 4.13: (1-January 2015) (a) Radiation Components Plot, (b) k_t–k_n Plot, (c) GHI (Measured vs. Calculated) Plot, (d) Quality Control Results and (e) k_t –k Plot

DNI values with comparable DHI radiation values. In quality control flag report, during morning and evening event time, error due to diffuse, coherence and tracking error is seen and in-between period only coherence error is seen. Here all cleaning of all instruments and data correction are suggested.

Table 4.4: Detailed Climate, Quality Control and Error Information (2-January 2015)

Climate	: Light clouded (Haze) conditions
Quality Control Errors	: Diffuse, Coherence and Tracking Error
Operational Errors	: Dust presence on GHI sensor
Instrumental Errors	: No

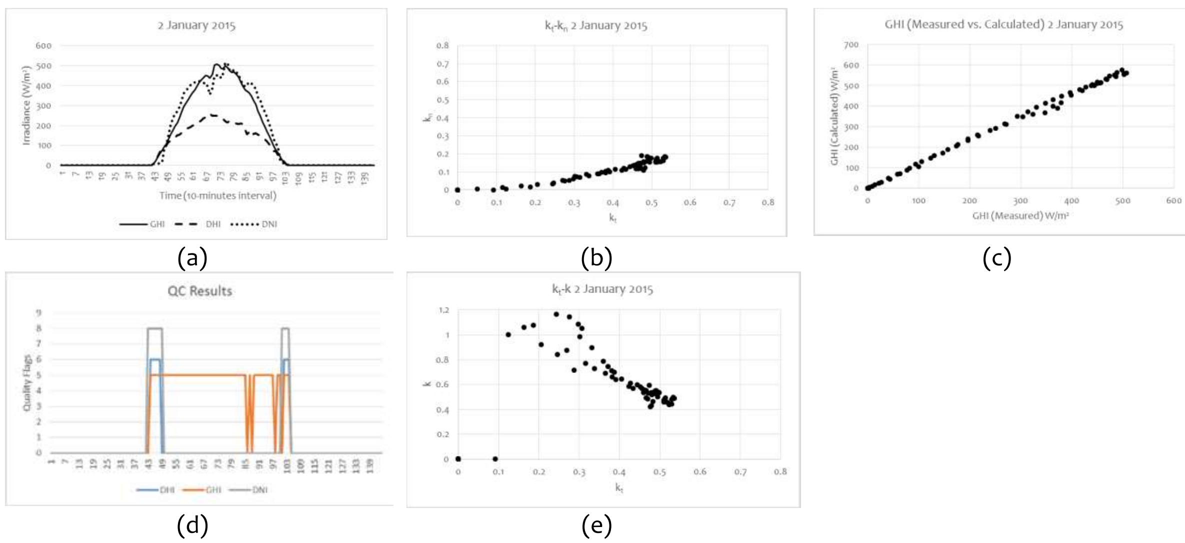


Figure 4.14: (2-January 2015) (a) Radiation Components Plot, (b) k_t–k_n Plot, (c) GHI (Measured vs. Calculated) Plot, (d) Quality Control Results and (e) k_t –k Plot

In Table 4.5 and Fig. 4.15, GHI and DHI are nearly same with good DNI measurement. Here day is identified as clear-sky with lightly clouded conditions, according to the transmittance plot study. The quality control guideline shows, data failure due to tracking and coherence conditions. Hence, DHI measuring instrument is found unreliable with installation and maintenance issues. Now another day with unique measurement condition is studied in Table 4.6 and Fig. 4.16. Here GHI and DNI measured values are found similar to high DHI values recorded. Theoretically, this plot is not possible, and the results are inconclusive. Here the "k_t-k_n" plot shows clear-sky day, whereas "k_t-k" and measured-calculated GHI plot shows out of the bound limit values. Hence, this case seems to be an error due to tracking, but analysis shows that the instrument is not correctly connected to the desired node.

Table 4.5: Detailed Climate, Quality Control and Error Information (5-February 2015)

Climate	: Lightly clouded to Clear sky conditions
Quality Control Errors	: Coherence and Tracking Error
Operational Errors	: Dust presence on GHI sensor
Instrumental Errors	: Tracking/Sensitivity error in DHI

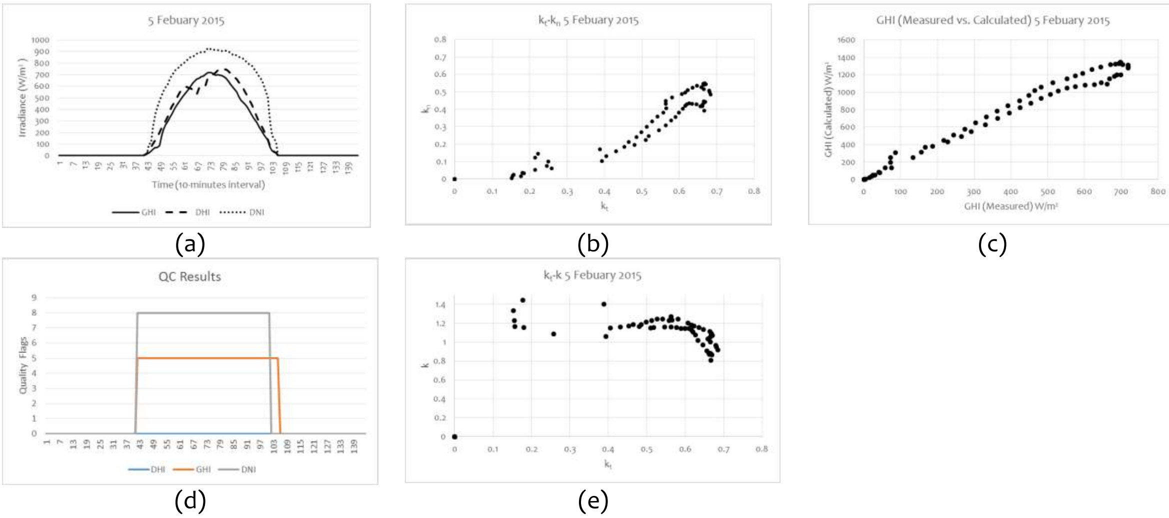


Figure 4.15: (5-February 2015) (a) Radiation Components Plot, (b) k_t–k_n Plot, (c) GHI (Measured vs. Calculated) Plot, (d) Quality Control Results and (e) k_t –k Plot

Table 4.6: Detailed Climate, Quality Control and Error Information (9-February 2015)

Climate	: Lightly clouded to Clear sky conditions
Quality Control Errors	: Coherence and Tracking Error
Operational Errors	: Dust presence on GHI and DNI sensor
Instrumental Errors	: Tracking/Sensitivity error in DHI

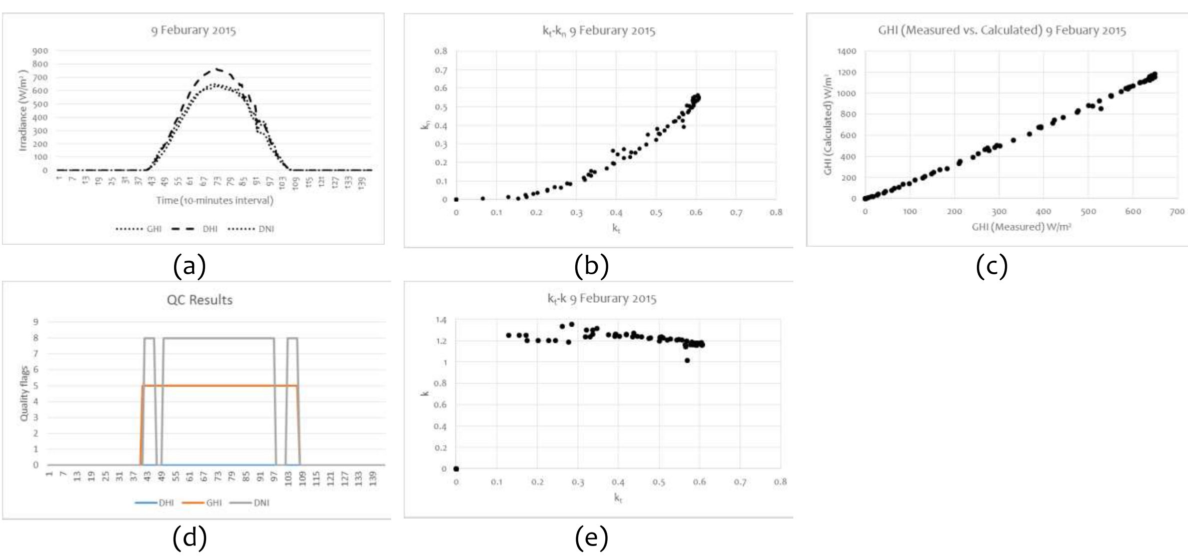


Figure 4.16: (9-February 2015) (a) Radiation Components Plot, (b) k_t–k_n Plot, (c) GHI (Measured vs. Calculated) Plot, (d) Quality Control Results and (e) k_t –k Plot

Similarly in Table 4.7 and Fig. 4.17, due to medium to heavy cloud presence, the DNI value is less than expected, with high GHI and DHI measured value (verified using "k_t-k_n" and "k_t-k" plot). The presence of tracking and coherence errors is identified for the majority of day due to cleaning issues. Hence cleaning of all measurement sensors is recommended. Similarly in Table 4.8 and Fig. 4.18, the day is identified having "passing and scattered" cloud day, having similar DNI and GHI measured values with high DHI values. Now, here all data failed in

Table 4.7: Detailed Climate, Quality Control and Error Information (28-April 2015)

Climate	: Medium clouded to Heavy sky conditions
Quality Control Errors	: Coherence and Tracking Error
Operational Errors	: Dust presence on GHI and DNI sensor
Instrumental Errors	: No

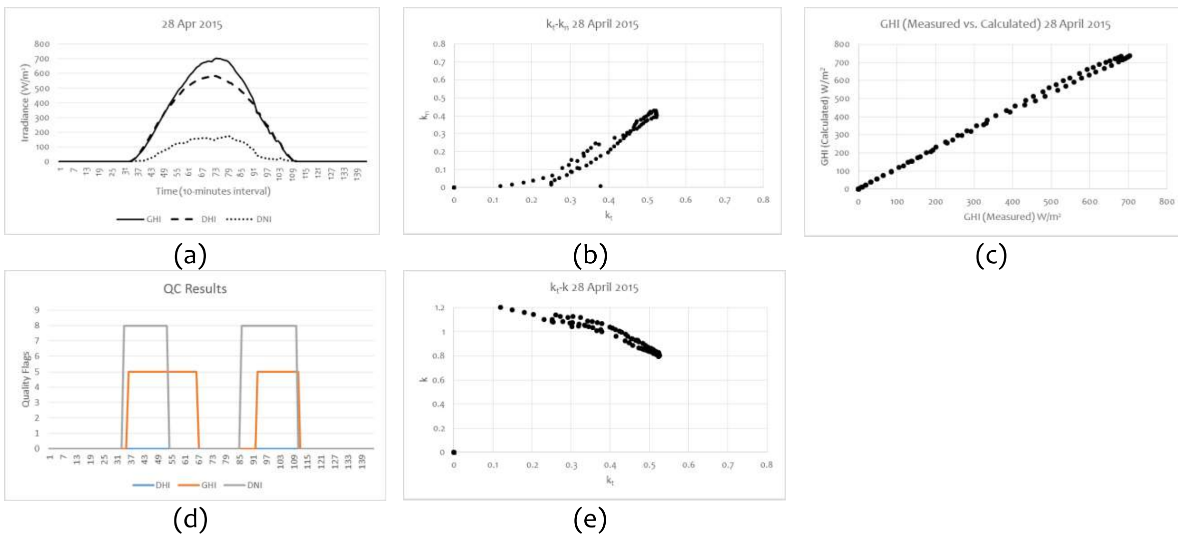


Figure 4.17: (28-April 2015) (a) Radiation Components Plot, (b) k_t–k_n Plot, (c) GHI (Measured vs. Calculated) Plot, (d) Quality Control Results and (e) k_t–k Plot

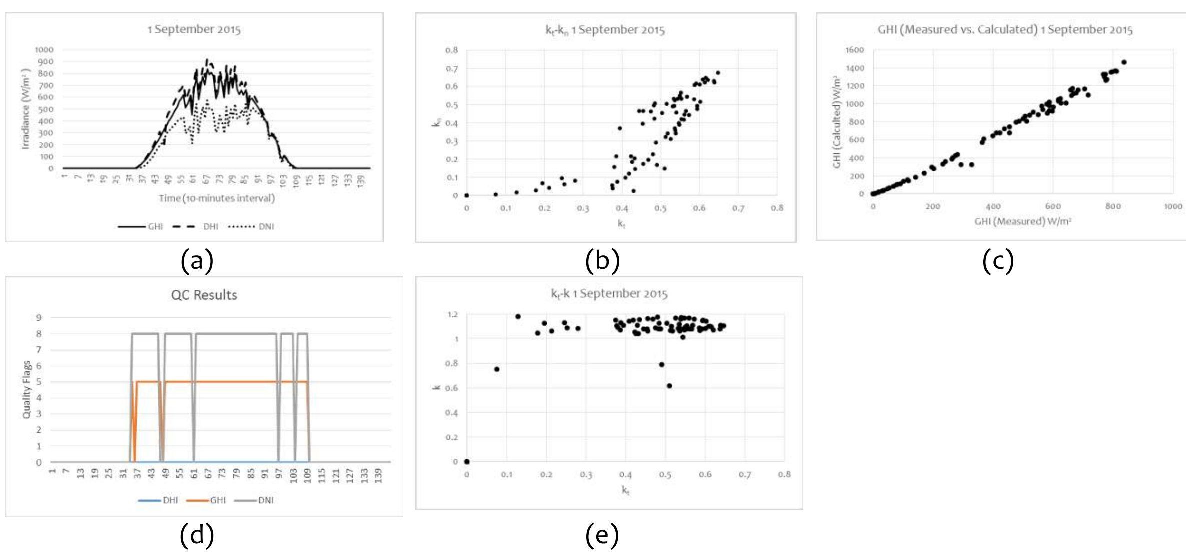


Figure 4.18: (1-September 2015) (a) Radiation Components Plot, (b) k_t–k_n Plot, (c) GHI (Measured vs. Calculated) Plot, (d) Quality Control Results and (e) k_t–k Plot

“coherence” and “tracking” error test, hence sensor cleaning and correction of DHI component is suggested.

In Table 4.9 and Fig. 4.19, the climate shows partly clouded sky, with the presence of moving medium sized clouds. Where for the first half of this day, no data are measured by the system, data filling required. Now in the next half, all data is getting failed in “coherence”, “diffuse” and “tracking” error tests. Hence whole day data is not reliable and instruments are having tracking problems as well as instrument cleaning issues. Moving forward in Table 4.10 and Fig. 4.20, the present climate is identified as haze cloud condition day with as expected GHI and DHI measured values, but no DNI values are stored for this period. Therefore, all plots are are wrong and gap filling of DNI values is suggested. Finally in Table 4.11 and Fig. 4.21, haze cloud condition day with the presence of some cloud is identified. But some abnormal behavior occurs during DNI and DHI measurements. Hence, both components need data correction, as suggested by quality control flags results.

Table 4.8: Detailed Climate, Quality Control and Error Information (1-September 2015)

Climate	: Light clouded sky conditions
Quality Control Errors	: Coherence and Tracking Error
Operational Errors	: Dust presence on GHI and DNI sensor
Instrumental Errors	: Tracking/Sensitivity error in DHI

Table 4.9: Detailed Climate, Quality Control and Error Information (24-July 2015)

Climate	: Medium clouded sky conditions
Quality Control Errors	: Diffuse, Coherence and Tracking Error
Operational Errors	: Dust presence, instrument error possible
Instrumental Errors	: Tracking error and sensor unstable.

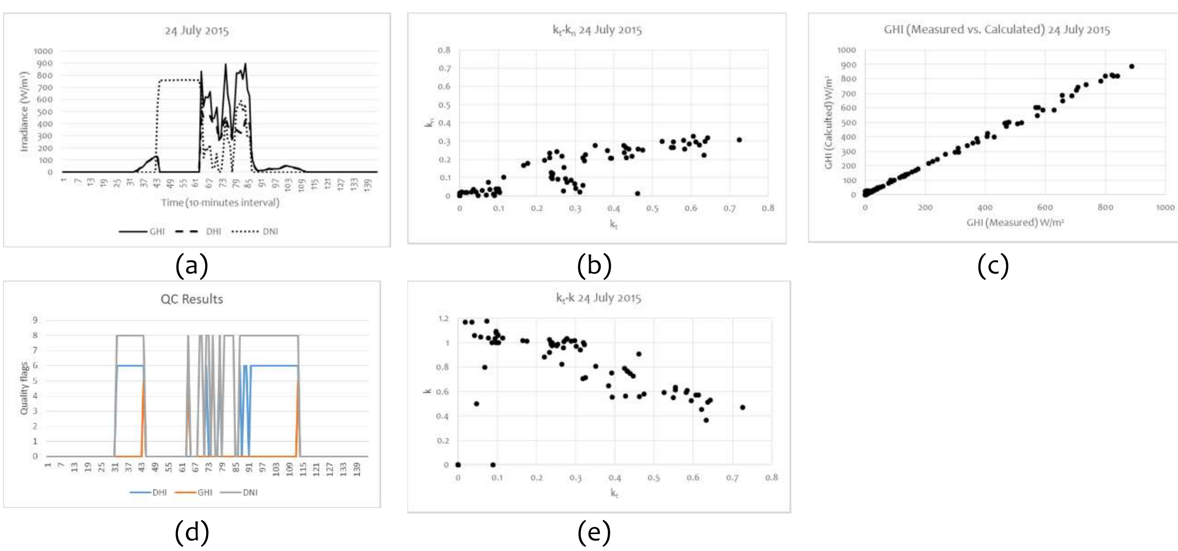


Figure 4.19: (24-July 2015) (a) Radiation Components Plot, (b) k_t –k_n Plot, (c) GHI (Measured vs. Calculated) Plot, (d) Quality Control Results and (e) k_t –k Plot

Table 4.10: Detailed Climate, Quality Control and Error Information (12-September 2015)

Climate	: Light clouded sky conditions with Haze
Quality Control Errors	: Coherence and Tracking Error
Operational Errors	: Dust presence GHI
Instrumental Errors	: Tracking/Sensitivity error in DNI

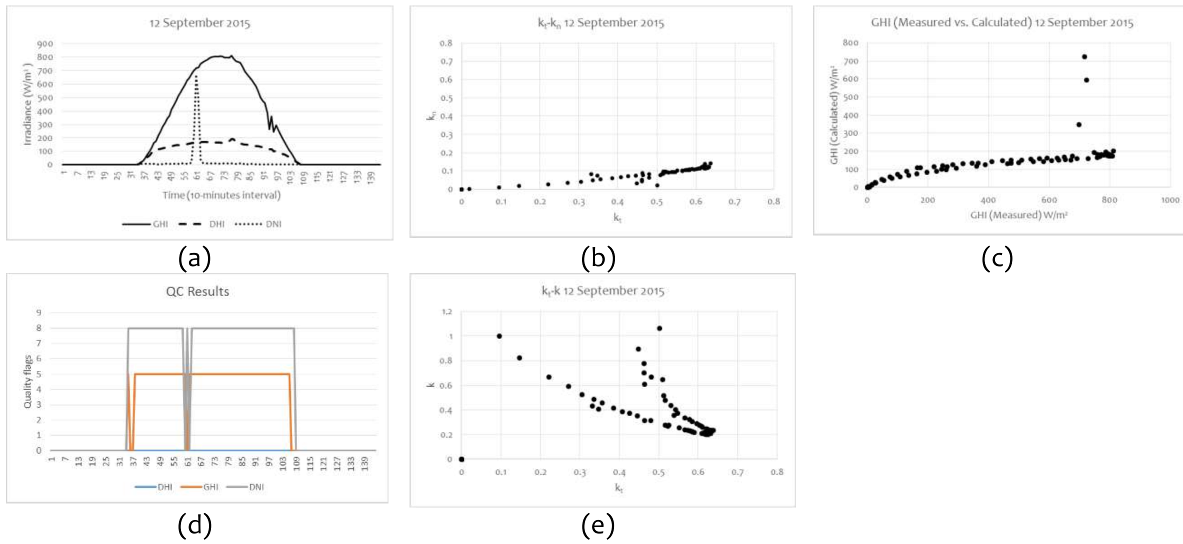


Figure 4.20: (12-September 2015) (a) Radiation Components Plot, (b) k_t–k_n Plot, (c) GHI (Measured vs. Calculated) Plot, (d) Quality Control Results and (e) k_t–k Plot

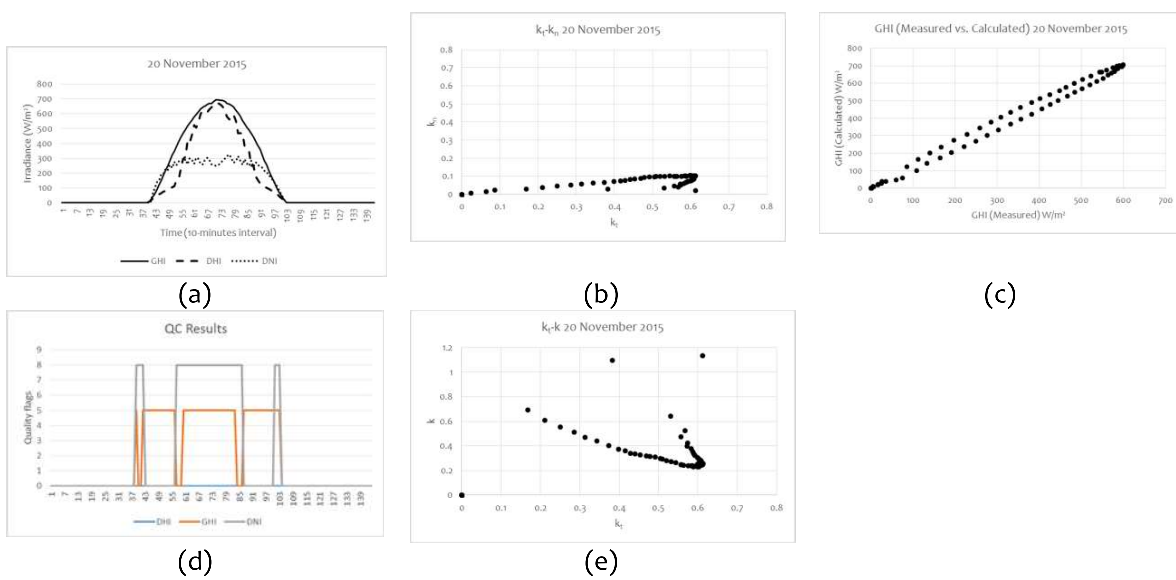


Figure 4.21: (20-November 2015) (a) Radiation Components Plot, (b) k_t–k_n Plot, (c) GHI (Measured vs. Calculated) Plot, (d) Quality Control Results and (e) k_t–k Plot

Here all days discussed above are selected on the identification of different measurement responses from the available instruments. The final gap identification decision is taken on the basis of results discussed with near day's climate conditions. According to which final decisions are provided for either ignoring (due to cloud presence) or application of gap filling approaches for each radiation component values.

Table 4.11: Detailed Climate, Quality Control and Error Information (20-November 2015)

Climate	: Light clouded sky conditions with Haze
Quality Control Errors	: Coherence and Tracking Error
Operational Errors	: Dust presence GHI, DHI and DNI
Instrumental Errors	: Tracking error DNI and DHI

4.6 SUMMARY

Detailed data quality control guidelines are explained in this chapter using the radiation database of Jodhpur. First time-step analysis is done which identifies the measuring frequency of radiation database and identifies missing intervals in it. Now in basic data analysis, suggested guidelines (BSRN and CIE) check the feasibility of the database for applications. This includes comparing the available database with standard clear sky model and plotting their transmittance ratios. In advance database analysis, individual tests for each radiation component error analysis are applied and results are shown in terms of quality flags. These check missing values, physical limits, clear sky limits, diffuse limit, coherence error and tracking error between radiation components. Some additional location-specific tests (step, persistence, spatial and sunshine duration) are discussed, which help in detailed error estimation. Now for accurate analysis, a theoretical study of possible flag combinations to specific radiation component error identification is done.

For data verification and to illustrate all the procedures discussed above, one-week dataset (a randomly selected) is used. The climate identification is done by plotting transmittance values and calculated-measured GHI values. By using quality control guidelines and step test, each measured radiation component is analyzed. After comparison of all generated plots and flag report, the error is shown to arise only due to maintenance issues. In addition, to study different responses from instruments, some unique days are selected from the available dataset and quality control process is applied to it. Where for each day, generated plots and flag report are used to identify root causes (either maintenance or instrument related). Moreover, the discussed cases can be used as a benchmark for identifying various errors for any location data analysis. Now the next step is to fill these gaps with appropriate algorithms.

…..