SOLAR MODEL ADAPTATION REPORT - PAKISTAN

March 2017

This report was prepared by Solargis, under contract to the World Bank.

It is a deliverable from the activity "Global Solar Atlas" [Project ID: P161130] and therefore relates to the solar resource database available via the Global Solar Atlas. This activity was funded and managed by the Energy Sector Management Assistance Program (ESMAP), a multi-donor trust fund administered by the World Bank, under a global initiative on Renewable Energy Resource Mapping. Further details on the initiative can be obtained from the ESMAP website.

The content of this document is the sole responsibility of the consultant authors.

Attribution

Please cite the work as follows: ESMAP. 2019. Solar Model Adaptation Report - Pakistan. Washington, DC: World Bank.

World Bank Group, Global ESMAP Initiative

Renewable Energy Resource Mapping and Geospatial Planning – Pakistan

Contract No. 8005624

Regional Adaptation of Solargis Model for Reduced Uncertainty of Solar Resource Data and Maps – Pakistan

March 2017

This report has been prepared by Tomas Cebecauer, Marcel Suri, Branislav Schnierer, Nada Suriova, Juraj Betak, Artur Skoczek and Veronika Madlenakova from Solargis

All maps in this report are prepared by Solargis

Solargis s.r.o., Mytna 48, 811 07 Bratislava, Slovakia

Reference No. (Solargis): 170-05/2017

http://solargis.com

Solargis is ISO 9001:2008 certified company for quality management

Table of contents

Acronyms

AOD	Aerosol Optical Depth
CFSR	Climate Forecast System Reanalysis. The meteorological model operated by the US service NOAA (National Oceanic and Atmospheric Administration)
CFS v2	Climate Forecast System Version 2 CFSv2 model is the operational extension of the CFSR (NOAA, NCEP)
DIF	Diffuse Horizontal Irradiation, if integrated solar energy is assumed. Diffuse Horizontal Irradiance, if solar power values are discussed
DNI	Direct Normal Irradiation, if integrated solar energy is assumed. Direct Normal Irradiance, if solar power values are discussed.
GFS	Global Forecast System. The meteorological model operated by the US service NOAA (National Oceanic and Atmospheric Administration)
GHI	Global Horizontal Irradiation, if integrated solar energy is assumed. Global Horizontal Irradiance, if solar power values are discussed.
GTI	Global Tilted (in-plane) Irradiation, if integrated solar energy is assumed. Global Tilted Irradiance, if solar power values are discussed.
MACC	Monitoring Atmospheric Composition and Climate – meteorological model operated by the European service ECMWF (European Centre for Medium-Range Weather Forecasts)
MERRA-2	Modern-Era Retrospective analysis for Research and Applications, Version 2
Meteosat IODC	Meteosat First Generation satellite operated by EUMETSAT organization. For this report the data from the Meteosat IODC are used. The satellite is positioned over the Indian Ocean.

Glossary

Aerosols
Small solid or liquid particles suspended in air, for example desert sand or soil particles, sea salts, burning biomass, pollen, industrial and traffic pollution.

All-sky irradiance
The amount of solar radiation reaching the Earth's surface is mainly determined by Earth-Sun geometry (the position of a point on the Earth's surface relative to the Sun which is determined by latitude, the time of year and the time of day) and the atmospheric conditions (the level of cloud cover and the optical transparency of atmosphere). All-sky irradiance is computed with all factors taken into account

Bias
Represents systematic deviation (over- or underestimation) and it is determined by systematic or seasonal issues in cloud identification algorithms, coarse resolution and regional imperfections of atmospheric data (aerosols, water vapour), terrain, sun position, satellite viewing angle, microclimate effects, high mountains, etc.

Clear-sky irradiance
The clear sky irradiance is calculated similarly to all-sky irradiance, but without considering the impact of cloud cover.

Long-term average
Average value of selected parameter (GHI, DNI, etc.) based on multiyear historical time series. Long-term averages provide a basic overview of solar resource availability and its seasonal variability.

P50 value
Best estimate or median value represents 50% probability of exceedance. For annual and monthly solar irradiation summaries it is close to average, since multiyear distribution of solar radiation resembles normal distribution.

P90 value
Conservative estimate, assuming 90% probability of exceedance (with the 90% probability the value should be exceeded). When assuming normal distribution, the P90 value is also a lower boundary of the 80% probability of occurrence. P90 value can be calculated by subtracting uncertainty from the P50 value. In this report, we apply a simplified assumption of normal distribution of yearly values.

Root Mean Square Deviation (RMSD)
Represents spread of deviations given by random discrepancies between measured and modelled data and is calculated according to this formula:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n}\left(X_{i,measured} - X_{i,predicted}\right)^2}{n}}$$

On the modelling side, this could be low accuracy of cloud estimate (e.g. intermediate clouds), under/over estimation of atmospheric input data, terrain, microclimate and other effects, which are not captured by the model. Part of this discrepancy is natural - as satellite monitors large area (of approx. 3.2 x 3.3 km for the satellite pixel), while sensor sees only micro area of approx. 1 sq. centimetre. On the measurement side, the discrepancy may be determined by accuracy/quality and errors of the instrument, pollution of the detector, misalignment, data loggers, insufficient quality control, etc.

Solar irradiance
Solar power (instantaneous energy) falling on a unit area per unit time [W/m^2]. Solar resource or solar radiation is used when considering both irradiance and irradiation.

Solar irradiation
Amount of solar energy falling on a unit area over a stated time interval [Wh/m^2 or kWh/m^2].

Uncertainty of estimate, U_{est}	Is a parameter characterizing the possible dispersion of the values attributed to an estimated irradiance/irradiation values. In this report, uncertainty assessment of the solar resource model estimate is based on a detailed understanding of the achievable accuracy of the solar radiation model and its data inputs (satellite, atmospheric and other data), which is confronted by an extensive data validation experience. The second source of uncertainty is ground measurements. Their quality depends on accuracy of instruments, their maintenance and data quality control. Third contribution to the uncertainty is from the site adaptation method where ground-measured and satellite-based data are correlated.

Executive summary

This report presents results of the solar resource assessment and mapping activity undertaken by The World Bank in Pakistan, as a part of a broader technical assistance project covering biomass, solar and wind mapping funded by the Energy Sector Management Assistance Program (ESMAP).

This report describes accuracy enhancement of Solargis solar resource data layers for Pakistan and maps based on the ground measurements collected at nine solar meteorological stations across the country. The solar meteorological stations were installed and operated by CSP Services (Germany) with their partner PITCO (Pakistan), and commissioned by the World Bank over the years 2014 to 2017 under the same activity.

The data layers show long-term yearly and monthly averages of Direct Normal Irradiation (DNI) and Global Horizontal Irradiation (GHI), and they represent a period of the last 18 years, from 1999 to 2016. The data is calculated by aggregation of sub-hourly map-based time series calculated for the territory of Pakistan with 1-km spatial resolution. This data is delivered in the format that is compatible with Geographical Information Systems (GIS). Additionally, printable maps are available in digital format and ready-to-use for large-format printing.

The accuracy of the data layers is enhanced by regional adaptation of the Solargis model with use of ground measurements acquired at nine high-standard solar meteorological stations located in Pakistan. The measurements helped to reduce systematic deviation of the data inputs to the Solargis model, namely aerosol data, which is the driving factor for accuracy of the satellite-modelled data in arid and semiarid climates. Individual improvements of the Solargis model at the stations in Pakistan are shown in the table below. As a result of the regional adaptation of the Solargis model, the calculated GHI and DNI data layers show more accurate long-term yearly and monthly values with reduced uncertainty.

Table 0.1: Position of solar measuring stations in Pakistan:

Site	Province (district)	Latitude [°]	Longitude [°]	Elevation [metres a.s.l]
Peshawar	Khyber Pakhtunkhwa	34.00170	71.48540	367
Islamabad	Islamabad Federal Territory	33.64191	72.98380	558
Lahore	Punjab	31.69458	74.24410	207
Quetta	Balochistan	30.27080	66.93980	1586
Multan	Punjab	30.16540	71.49780	123
Bahawalpur	Punjab	29.32542	71.81877	123
Khuzdar	Balochistan	27.81780	66.62940	1254
Hyderabad	Sindh	25.41340	68.25950	63
Karachi	Sindh	24.93340	67.11160	45

Table 0.2: Results of regional adaptation for solar measuring stations in Pakistan:

Site	DNI original [kWh/m²]	DNI adapted [kWh/m²]	Difference to original [%]	GHI original [kWh/m²]	GHI adapted [kWh/m²]	Difference to original [%]
Peshawar	1244	1320	6.1%	1650	1695	2.7%
Islamabad	1389	1468	5.7%	1722	1758	2.1%
Lahore	1205	1181	-2.0%	1698	1700	0.1%
Quetta	2301	2277	-1.0%	2183	2191	0.4%
Multan	1309	1369	4.6%	1822	1858	2.0%
Bahawalpur	1455	1569	7.8%	1909	1959	2.6%
Khuzdar	2133	2239	5.0%	2183	2213	1.4%
Hyderabad	1618	1758	8.7%	2069	2108	1.9%
Karachi	1576	1522	-3.4%	1962	1949	-0.7%

** All values in this table show results of the regional adaptation of Solargis time series, long-term average yearly values. GHI – Global Horizontal Irradiance, DNI – Direct Normal Irradiance. The data is derived from the GIS layers after disaggregation.*

Table 0.3: Uncertainty of yearly value for original and regionally-adapted Solargis model:

	Direct Normal Irradiation			Global Horizontal Irradiation		
Data uncertainty	Low	Medium	High	Low	Medium	High
Original data	Up to ±10%	Up to ±16%	±18% and more	Up to ±6%	Up to ±8%	±8% and more
After adaptation	±6% to ±8%	Up to ±14%	±18% and more	±4% to ±4.5%	Up to ±6%	±8% and more
Best-achievable*	±4.0%	-	-	±2.5%	-	-

** Only achievable by site-specific model adaptation (values shown as a reference)*

1 Overview of GIS data layers

Solargis is a high-resolution global database that includes solar resource and meteorological parameters. This database is developed and operated by the company Solargis and is computed and updated on a daily basis by models using satellite, atmospheric and meteorological data inputs. The geographical extent of the database covers the land area between latitudes 60º North and 45º South. The innovative features of the delivered GIS data and maps for Pakistan include:

- Harmonized and accuracy-enhanced solar resource data: yearly and monthly long-term averages of GHI and DNI
- Historically representing the last 18 years (1999 to 2016) at high spatial and temporal resolution
- Solargis database is extensively validated and accuracy enhanced
- The time series data for any location can be accessed online (http://solargis.com).

Table 1.1: Description of GIS data layers that were accuracy enhanced by regional model adaptation

Acronym	Full name	Unit	Type of use	Type of data layers
GHI	Global Horizontal Irradiation	kWh/m²	Reference information for the assessment of flat-plate photovoltaic (PV) and solar heating technologies (e.g. hot water)	Long-term annual and monthly averages
DNI	Direct Normal Irradiation	kWh/m²	Assessment of Concentrated PV (CPV) and Concentrated Solar Power (CSP) technologies. It is also important for simulation of flat-plate PV and tracking technologies.	Long-term annual and monthly averages
DIF	Diffuse Horizontal Irradiation	kWh/m²	Complementary parameter to GHI and DNI	Long-term yearly and monthly average of daily totals
GTI	*Global Irradiation at optimum tilt*	*kWh/m²*	*Assessment of solar resource for PV technologies*	*Long-term yearly and monthly average of daily totals*
OPTA	*Optimum angle*	*°*	*Optimum tilt to maximize yearly PV production*	-
PVOUT	*Photovoltaic power potential*	*kWh/kWp*	*Assessment of power production potential for a PV power plant with free-standing fixed-c-Si modules, mounted at optimum tilt to maximize yearly PV production*	*Long-term yearly and monthly average of daily*

Note: in italics, we indicate data layer that have been accuracy enhanced indirectly from GHI and DNI

The regionally adapted maps provide accurate and reliable information on GHI and DNI solar resources for Pakistan. Lower uncertainty reduces financial risk and improves the engineering quality of the solar power plants. The regionally adapted Solargis model is capable of delivering more accurate solar resource data. For a project-specific site, the uncertainty can be further reduced by site adaptation.

Tables 1.1 to 1.4 describe the primary data layers GHI, DIF and DNI that have been processed by the regionally adapted solar model as part of the country data delivery. The other data layers, such as GTI and PVOUT were re-computed using the accuracy enhanced GHI, DIF and DNI.

Metadata

The complete metadata for the final outputs is delivered in separate files in accordance with the ISO 19139 standard. The metadata is located in the metadata folder, in two formats:

- PDF – human readable format
- XML – exchange format

Table 1.2: General information about GIS data layers supplied for Pakistan

Geographical extent	Land area (with buffer 15 km towards the ocean) between 38°N and 23°N, 60°E and 80°E, covering the Islamic Republic of Pakistan (approx. 2 980 000 km^2)
Map projection	Geographic (Latitude/Longitude), datum WGS84 (also known as *GCS_WGS84*; *EPSG: 4326*)
Data format	ESRI ASCII raster data format *(asc)* GeoTIFF raster data format (*tif*)

Table 1.3: Technical specification of the accuracy enhanced GIS data layers

Acronym	Full name	Data format	Spatial resolution	Time representation	No. of data layers
GHI	Global Horizontal Irradiation	Raster	9 arc-sec. (approx. 240 x 275 m)	1999 - 2016	12+1
DNI	Direct Normal Irradiation	Raster	9 arc-sec. (approx. 240 x 275 m)	1999 - 2016	12+1
DIF	Diffuse Horizontal Irradiation	Raster	9 arc-sec. (approx. 240 x 275 m)	1999 – 2016	12+1
GTI	Global Irradiation at optimum tilt	Raster	9 arc-sec. (approx. 240 x 275 m)	1999 – 2016	12+1
OPTA	Optimum angle	Raster	2 arc-min (approx. 3300x3700 m)	-	1
PVOUT	Photovoltaic power potential	Raster	9 arc-sec. (approx. 240 x 275 m)	1999 – 2016	12+1

Table 1.4: Characteristics of the final raster GIS data files

Characteristics	Range of values
West – East	60:00:00E – 80:00:00E
North – South	38:00:00S – 23:00:00S
Resolution (GHI, DNI, GTI, DIF, PVOUT)	00:00:09 (8000 columns x 6000 rows)
Resolution (OPTA)	00:02 (600 columns x 450 rows)
Data type	Float
No data value	-9999, NaN

* http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/ESRI_ASCII_raster_format/009t0000000z000000/

Explanation:

- MM: month of data – from 01 to 12
- Ext: file extension (asc or tif)

Data layers are provided as separate files in a tree structure, organized according to

- File format (ASCII or GEOTIF)
- Time summarization (yearly and monthly)

Complementary files:

- Project files (*.prj) complement ESRI ASCII grid files (*.asc)
- World files (*.tfw) complement GeoTIFF files (*.tif)

2 Solargis database

2.1 Solar resource data calculated by satellite-based solar model

Solar radiation is calculated by numerical models, which are parameterized by a set of inputs characterizing the cloud transmittance, state of the atmosphere and terrain conditions. A comprehensive **overview of the Solargis model** is made available in a recent book publication [1]. The methodology is also described in [2, 3]. The related uncertainty and requirements for bankability are discussed in [4, 5].

In the Solargis approach, the **clear-sky irradiance** is calculated by the simplified SOLIS model [6]. This model allows fast calculation of clear-sky irradiance from the set of input parameters. Sun position is a deterministic parameter, and is described by the algorithms with satisfactory accuracy. Stochastic variability of clear-sky atmospheric conditions is determined by changing concentrations of atmospheric constituents, namely aerosols, water vapour and ozone. Global atmospheric data, representing these constituents, are routinely calculated by world atmospheric data centres:

- In Solargis, the new generation **aerosol data set** representing Atmospheric Optical Depth (AOD) is used. The calculation accuracy is strongly determined by the quality of aerosols, especially for cloudless conditions. The aerosol data set is developed and regularly updated by MACC-II/CAMS project (ECMWF) [7, 8]), and delivered at a spatial resolution of about 75 km and 125 km. An important feature of this AOD data set is its ability to capture the daily variability of aerosols, which allows simulating events with an extreme atmospheric load of aerosol particles more precisely. Thus, it reduces the uncertainty of instantaneous estimates of GHI and especially DNI and allows for improved statistical distribution of irradiance values [9, 10]. Data for the period from 2003 to 2012 comes from this database; the data representing a period from year 2013 to the present time is derived from a near-real time operational model. It is to be noted that coverage of high frequency (daily) aerosol data by MAC-II/CAMS model is limited to the period from 2003 onwards. For the remaining years (1999 to 2002) daily data values from the MERRA-2 model (NASA) [11] are used. Before its use, MERRA-2 aerosol data is harmonized with MACC-II/CAMS data.
- **Water vapour** is also highly variable in space and time, but it has a lower impact on the values of solar radiation, compared to aerosols. The GFS and CFSR databases (NOAA NCEP) are used in Solargis, and the data represent the daily variability from 1999 to the present time [12, 13, 14].
- **Ozone** absorbs solar radiation at wavelengths shorter than 0.3 µm, thus having negligible influence on the broadband solar radiation.

The clouds are the most influencing factor, modulating clear-sky irradiance. Effect of clouds is calculated from the satellite data in the form of a **cloud index** (cloud transmittance). The cloud index is derived by relating the radiance recorded by the satellite in three spectral channels and surface albedo to the cloud optical properties. Spatial resolution of the Meteosat MFG IODC satellite data, used in Solargis, is about 2.7 to 3.2 km in Pakistan. The time step is 30 minutes for whole period from 1999 to 2016 [15]. In Solargis, the modified calculation scheme of Cano has been adopted to retrieve cloud optical properties from the satellite data [16]. A number of improvements have been introduced to better cope with specific situations such as snow, ice, or high albedo areas (arid zones and deserts), as well as complex terrain.

To calculate **all-sky irradiance** in each time step, the clear-sky global horizontal irradiance is coupled with the cloud index.

Direct Normal Irradiance (DNI) is calculated from Global Horizontal Irradiance (GHI) using a modified Dirindex model [17]. Diffuse irradiance for tilted surfaces is calculated by the Perez model [18]. The calculation procedure also includes terrain disaggregation, while the spatial resolution is enhanced with use of the digital terrain model to 250 meters [19].

Solargis model version 2.1 has been used. Table 2.1 summarizes technical parameters of the model inputs and of the primary data outputs. This model was enhanced by regional adaptation based on the ground solar measurements (Chapter 4)

Table 2.1: Input data used in the Solargis model and related GHI and DNI outputs for Pakistan

Inputs into the Solargis model	Source of input data	Time representation	Original time	Approx. grid resolution
Cloud index	Meteosat MFG IODC (EUMETSAT)	1999 to date	30 minutes	3.2 x 3.3 km
Atmospheric optical depth (aerosols)*	MACC/CAMS* (ECMWF) MERRA-2 (NASA)	2003 to date 1999 to 2002	3 hours 1 hour	75 km and 125 km 50 km
Water vapour	CFSR/GFS (NOAA)	1999 to date	1 hour	35 and 55 km
Elevation and horizon	SRTM-3 (SRTM)	-	-	1 km
Solargis primary data outputs (GHI and DNI)	-	**1999 to date**	**30 minutes**	**1 km**

** Aerosol data for 2003-2012 come from the reanalysis database; the data representing years 2013-present are derived from near- real time (NRT) operational model*

2.2 Combined use of satellite model and ground measurements

The fundamental difference between a satellite observation and a ground measurement is that signal received by the satellite radiometer integrates a large area, while a ground station represents a pinpoint measurement. This results in a mismatch when comparing instantaneous values from these two observation instruments, mainly during intermittent cloudy weather and changing aerosol load. Nearly half of the hourly Root Mean Square Deviation (RMSD) for GHI and DNI can be attributed to this mismatch (value at sub-pixel scale), which is also known as the "nugget effect" [20].

The satellite pixel is not capable of describing the inter-pixel variability in complex regions, where within one pixel, diverse natural conditions could vary (e.g. fog in narrow valleys or along the coast). In addition, the coarse spatial resolution of atmospheric databases such as aerosols or water vapour is not capable of describing local patterns of the state of the atmosphere. These features can be seen in the satellite DNI data by increased bias due to an imperfect description of aerosol load. Satellite data have inherent inaccuracies, which have a certain degree of geographical and time variability.

DNI is particularly sensitive to the variability of cloud information, aerosols, water vapour, and terrain shading. The relationship between the uncertainty of global and direct irradiance is nonlinear. Often, a negligible error in global irradiance may have a high impact on the direct irradiance component.

The solar energy projects require representative and accurate GHI and DNI time series. The satellite-derived databases are used to describe long-term solar resource for a specific site. However, their problem when compared to the high-quality ground measurements is a slightly higher bias and partial disagreement of frequency distribution functions, which may limit their potential to record the occurrence of extreme situations (e.g. very low atmospheric turbidity resulting in a high DNI and GHI). A solution is to correlate satellite-derived data with ground measurements to understand the source of the discrepancy, and subsequently, to improve the accuracy of the resulting time series.

The Solargis satellite-derived data are correlated with ground measurement data with two objectives:

- Improvement of the overall bias (removal of systematic deviations)
- Improvement of the fit of the frequency distribution of values.

The relationship between the uncertainty of global and direct irradiance is nonlinear. Often, a negligible error in global horizontal irradiance may trigger a larger error in the direct irradiance component. Limited spatial and temporal resolution of the input data, and the simplified nature of the models results in the occurrence of systematic and random deviations of the model outputs when compared to the ground observations. The deviations in the satellite-computed data, which have a *systematic nature*, can be reduced by site adaptation or regional adaptation methods.

2.2.1 Site adaptation vs. regional adaptation

The terminology related to the procedure of improving the accuracy of the satellite data is not harmonized, and various terms are used:

- Correlation of ground measurements and satellite-based data;
- Calibration of the satellite model (its inputs and parameters);
- Site adaptation or regional adaptation of satellite based data.

The term *site adaptation* or *regional adaptation* is more general and best explains the concept of adapting the satellite-based model (by correlation, calibration, fitting and recalculation) to the ground measured data.

- **Site adaptation** aims to adapt the characteristics of the satellite-based time series to the site-specific conditions described by local measurements.
- **Regional adaptation** aims to identify systematic patterns of deviation at the regional scale and correct them rather than focusing on a specific site.

In this study, we apply a regional adaptation of the Solargis model to improve its performance at the regional level. Its advantage is that the database in the given region has reduced uncertainty. To obtain the best accuracy for a specific location, it is preferred to apply the site adaptation of the model, as it focuses on matching the model outputs to the specific local climate conditions described by the ground measurements.

It is implicit that the more accurate the satellite-based model is, the lower the difference expected between the calculation results of the original model and its regionally adapted version.

2.2.2 Conditions to be met

Four conditions are important for successful adaptation of the satellite-based model:

1. High quality DNI and GHI ground measurements for at least 12 months must be available; optimally data for 2 or 3 years should be used;
2. For regional-adaptation, the sites should be distributed over the whole territory, to provide information for the major climatic regions;
3. High quality satellite data must be used, with consistent quality over the whole period of data;
4. There has to be a systematic difference identified between both data sources.

Systematic difference can be measured by two characteristics:

- Bias (offset)
- Systematic deviation in the distribution of hourly or daily values (in the histogram)

Systematic difference can be stable over the year or it can slightly change seasonally for certain meteorological conditions (e.g. typical cloud formation during a day, seasonal air pollution). The data analysis should distinguish systematic differences from those arising at occasional events, such as extreme sand storms or forest fires. The episodically occurring differences may mislead the results of adaptation, especially if short period of ground measurements is only available.

If one of the four above-mentioned conditions is not fulfilled, the model adaptation will not provide the expected results. In fact, such an attempt may provide even worse results.

For the quantitative assessment of the accuracy enhancement procedures, the following metrics are used:

- Metrics based on the comparison of all pairs of the hourly daytime data values: Mean Bias, Root Mean Square Deviation (RMSD) and histogram in an absolute and relative form (divided by the daytime mean DNI values);
- Metrics based on the difference of the cumulative distribution functions: KSI (Kolmogorov-Smirnov test Integral) [21]

The normalized KSI is defined as an integral of absolute differences of two cumulative distribution functions *D* normalized by the integral of critical value $a_{critical}$.

$$KSI\% = \frac{\int_{x\,min}^{x\,max} D_n dx}{a_{critical}} * 100$$

$$a_{critical} = V_c * (x_{max} - x_{min})$$

$$V_c = \frac{1.63}{\sqrt{N}}, \quad N \geq 35$$

where critical value depends on the number of the data pairs *N*. As the KSI value is dependent on the size of the sample, the KSI measure may be used only for the relative comparison of fit of cumulative distribution of irradiance values.

More about the Solargis site adaptation can be found in [22] and more general description is in [23].

3 Ground measurements in Pakistan

3.1 Inventory of available solar meteorological stations and data

Data from the measuring stations in Pakistan was collected and harmonized with the objective of acquiring reference solar radiation data for reducing the uncertainty of the model. The quality data from nine meteorological stations were available for this assessment (Tables 3.1 and 3.2, Figure 3.1). Positions and detailed information about measurement sites is also available on the Global Solar Atlas website: http://globalsolaratlas.info/?c=30.183122,71.279297,6&e=1. The information about the instruments is summarized in Tables 3.3 to 3.5.

Table 3.1: Summary of information for installed measurement stations in Pakistan

Project name	Renewable Energy Resource Mapping and Geospatial Planning - Pakistan
Project ID	P146140
Project framework	Energy Sector Management Assistance Program (ESMAP)
Project administrator	The World Bank Group
Data measurement points	2 stations TIER 1, 7 stations TIER 2, total of 9 measurement sites
Measurement service provider	CSP Services GmbH
Maintenance service provider	PITCO, Pakistan

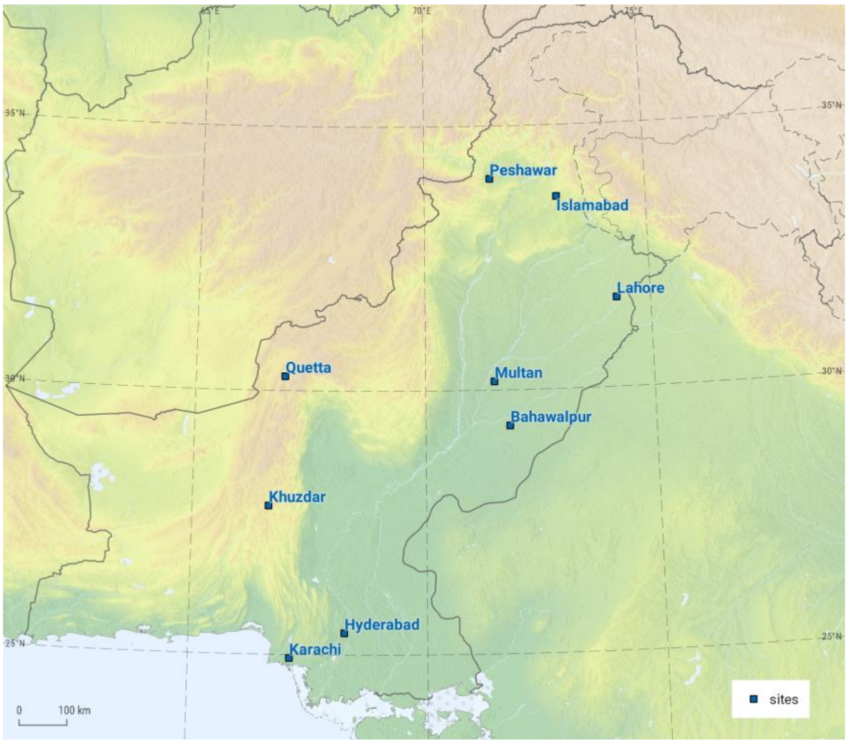


Figure 3.1: Position of the solar meteorological stations

Table 3.2: Overview information on measurement stations operated in Pakistan

No.	Site name	Site ID	Latitude [°]	Longitude [°]	Altitude [m a.s.l.]	Measurement station host
1	Peshawar	PES	34.00170	71.48540	370	Mechanical Engineering Department, University of Engineering and Technology, Peshawar
2	Islamabad	ISB	33.64191	72.98380	500	National University of Science and Technology campus, Islamabad
3	Lahore	LAH	31.69458	74.24410	220	Center of Energy Research at the University of Engineering and Technology campus, Kala Shah Kaku
4	Quetta	QUE	30.27080	66.93980	1590	Baluchistan University of Information Technology, Engineering and Management Sciences campus, Quetta
5	Multan	MUL	30.16540	71.49780	95	M. Nawaz Sharif University of Engineering and Technology, Multan
6	Bahawalpur	QASP	29.32542	71.81877	120	Quaid-e-Azam Solar Park, Bahawalpur, Punjab Province
7	Khuzdar	KHU	27.81780	66.62940	1260	Baluchistan University of Engineering and Technology campus, Khuzdar
8	Hyderabad	HYD	25.41340	68.25950	60	Mehran University of Engineering and Technology, Jamshoro
9	Karachi	KAR	24.93340	67.11160	40	IM Department building of NED University of Engineering and Technology, Karachi

Table 3.3: Instruments used for measuring solar radiation

No.	Site name	Site ID	Station	DNI	GHI	DIF
1	Peshawar	PES	TIER 2	Twin-RSI, CSP Services	CMP 10, Kipp & Zonen Twin-RSI, CSP	Twin-RSI, CSP Services
2	Islamabad	ISB	TIER 1	CHP 1, Kipp & Zonen	CMP 21, Kipp & Zonen	CMP 21, Kipp & Zonen
3	Lahore	LAH	TIER 2	Twin-RSI, CSP Services	CMP 10, Kipp & Zonen Twin-RSI, CSP	Twin-RSI, CSP Services
4	Quetta	QUE	TIER 2	Twin-RSI, CSP Services	CMP 10, Kipp & Zonen Twin-RSI, CSP Services	Twin-RSI, CSP Services
5	Multan	MUL	TIER 2	Twin-RSI, CSP Services	CMP 10, Kipp & Zonen Twin-RSI, CSP	Twin-RSI, CSP Services
6	Bahawalpur	QASP	TIER 1	CHP 1, Kipp & Zonen	CMP 21, Kipp & Zonen	CMP 21, Kipp & Zonen
7	Khuzdar	KHU	TIER 2	Twin-RSI, CSP Services	CMP 10, Kipp & Zonen Twin-RSI, CSP	Twin-RSI, CSP Services
8	Hyderabad	HYD	TIER 2	Twin-RSI, CSP Services	CMP 10, Kipp & Zonen Twin-RSI, CSP Services	Twin-RSI, CSP Services
9	Karachi	KAR	TIER 2	Twin-RSI, CSP Services	CMP 10, Kipp & Zonen Twin-RSI, CSP	Twin-RSI, CSP Services

Table 3.4: Instruments installed at the solar meteorological stations TIER 1

Parameter	Instrument	Type	Manufacturer	Uncertainty
GHI*	Secondary standard pyranometer	CMP 21	Kipp & Zonen	< ±2 % (daily)
DNI	First class thermopile pyrheliometer	CHP 1	Kipp & Zonen	< ±1 % (daily)
DIF*	Secondary standard pyranometer	CMP 21	Kipp & Zonen	-
TEMP	Temperature probe	CS215	Campbell Scientific	± 0.9 ºC
RH	Relative humidity probe	CS215	Campbell Scientific	± 4 % RH
WS	3 cup anemometer	RNRG 40C	NRG Systems	± 0.45 m/s at 10 m/s
WD	Wind vane	RNRG 200P	NRG Systems	< ±1 %
AP	Barometric pressure sensor	278	Setra	± 2 mb
-	Data logger	CR 1000	Campbell Scientific	± (0.06% of reading + offset), 0° to 40°C

** Instruments are assembled with a Kipp & Zonen ventilation unit CVF4*

Table 3.5: Instruments installed at the solar meteorological stations TIER 2

Parameter	Instrument	Type	Manufacturer	Uncertainty
GHI	Secondary standard pyranometer	CMP 10	Kipp & Zonen	< ±2 % (daily)
GHI_RSI	Rotation Shadowband Irradiometer	Twin-RSI	CSP Services	< ±3.5 % (instant.) *
DNI_RSI	Rotation Shadowband Irradiometer	Twin-RSI	CSP Services	< ±3.5 % (instant.) *
DIF_RSI	Rotation Shadowband Irradiometer	Twin-RSI	CSP Services	< ±3.5 % (instant.) *
TEMP	Temperature probe	CS215	Campbell Scientific	± 0.9 ºC
RH	Relative humidity probe	CS215	Campbell Scientific	± 4 % RH
WS	3 cup anemometer	RNRG 40C	NRG Systems	± 0.45 m/s at 10 m/s
WD	Wind vane	RNRG 200P	NRG Systems	< ±1 %
AP	Barometric pressure sensor	278	Setra	± 2 mb
-	Data logger	CR 1000	Campbell Scientific	± (0.06% of reading + offset), 0° to 40°C

** Specified by Customer, after corrections*

Solar parameters at TIER 1 stations are measured by high quality instruments (Table 3.4), i.e. first class pyrheliometers (for DNI) and secondary standard pyranometers (for GHI and DIF). TIER 2 stations use secondary standard pyranometer (for GHI, high quality and accuracy) and Twin-RSI instruments for GHI, DNI and DIF, which measure with less accuracy (Table 3.5). Overview of the data availability, time step and measured parameters is shown in Table 3.6 and Figure 3.2.

Table 3.6: Overview information on solar meteorological stations operating in the region

No.	Site name	Site ID	Parameters	Time step	Period of data used in this study
1	Peshawar	PES	GHI, GHI_RSI, DNI_RSI, DIF_RSI	10 min	10 Apr 2015 – 31 Dec 2016
2	Islamabad	ISB	GHI, DNI, DIF	10 min	25 Oct 2014 – 31 Jan 2017
3	Lahore	LAH	GHI, GHI_RSI, DNI_RSI, DIF_RSI	10 min	22 Oct 2014 – 31 Jan 2017
4	Quetta	QUE	GHI, GHI_RSI, DNI_RSI, DIF_RSI	10 min	17 Sep 2015 – 31 Jan 2017
5	Multan	MUL	GHI, GHI_RSI, DNI_RSI, DIF_RSI	10 min	20 Oct 2014 - 31 Jan 2017
6	Bahawalpur	QASP	GHI, DNI, DIF	10 min	17 Oct 2014 – 31 Jan 2017
7	Khuzdar	KHU	GHI, GHI_RSI, DNI_RSI, DIF_RSI	10 min	22 Sep 2015 – 31 Jan 2017
8	Hyderabad	HYD	GHI, GHI_RSI, DNI_RSI, DIF_RSI	10 min	21 Apr 2015 – 31 Jan 2017
9	Karachi	KAR	GHI, GHI_RSI, DNI_RSI, DIF_RSI	10 min	22 Apr 2015 – 31 Jan 2017

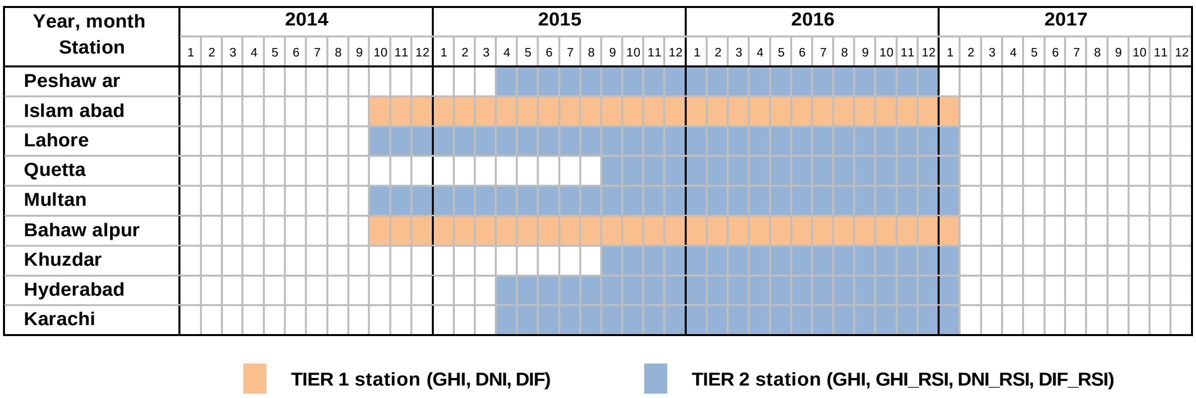


Figure 3.2: Solar resource data availability (GHI, DNI and DIF).

3.2 Summary of meteorological stations operation

The TIER 1 stations were installed by CSP Services staff in October 2014. The TIER 2 stations have been installed by PITCO staff in 2014 and 2015. During the measurement campaign, at least one maintenance visit for inspection and test for correct operation has been performed for each meteorological station by PITCO technical staff. Moreover, routine cleaning and maintenance has been performed by local personnel, as summarized in Table 3.7. Some issues, detected during operation, are summarized in Table 3.8.

Data measured by Twin-RSI and pyrheliometers from all measured stations has been corrected for soiling. GHI measured by pyranometers was not corrected. The TIER 1 stations have installed ventilation units CVF4, which suppress soiling. Pyranometer measurements on TIER 2 stations show non-uniform soiling patterns, so correcting the data could reduce its quality. Significantly degraded GHI readings by soiling from TIER 2 stations were excluded from further processing (Chapter 3.3).

Table 3.7: Summary of maintenance visits

No.	Site name	Site ID	Comment
1	Peshawar	PES	PITCO maintenance and test visits: 16 Dec 2015, 19 May 2016 Local maintenance frequency usually every 1 – 2 weeks, rarely up to 1 month
2	Islamabad	ISB	PITCO maintenance and test visits: 13 Mar 2015, 18 Jun 2015, 15 Dec 2015, 20 May 2016 Local maintenance frequency usually every 1 – 5 days, rarely up to 9 days
3	Lahore	LAH	PITCO maintenance and test visits: 12 Aug 2015, 11 Feb 2016 Local maintenance frequency variable from 1 – 2 weeks up to 1 month
4	Quetta	QUE	PITCO maintenance and test visit: 28 Mar 2016 No local maintenance during first 3 months, then usually 1 week, later degrading to 2 – 3 weeks
5	Multan	MUL	PITCO maintenance and test visits: 18 Nov 2015, 27 Apr 2016 Local maintenance frequency usually every 1 – 2 weeks, rarely up to 3 weeks
6	Bahawalpur	QASP	PITCO maintenance and test visits: 30 Dec 2014 (sun tracker defect), 17 Mar 2015, 17 Nov 2015, 28 Apr 2016 (power supply repair) Local maintenance frequency usually every 1 – 2 days, rarely up to 1 week, gaps during tracker and power supply failure
7	Khuzdar	KHU	PITCO maintenance and test visit: 16 Mar 2016 Variable local maintenance from 1 – 3 week period up to 3 months
8	Hyderabad	HYD	PITCO maintenance and test visits: 7 Oct 2015, 2 Jun 2016 Local maintenance frequency usually every 3 – 5 days, later degrading up to 3 - 4 weeks
9	Karachi	KAR	PITCO maintenance and test visits: 8 Oct 2015, 3 Jun 2016 Local maintenance frequency regularly between 1 – 2 weeks, one 3 weeks gap in beginning

Table 3.8: Summary of operation issues

No.	Site name	Site ID	Issues list
1	Peshawar	PES	No significant issues, generally well running station
2	Islamabad	ISB	Tracker loaned to QASP site in period from March to June 2015 Ventilation units CVF4 ventilator problems (corrosion) from August to December 2015 Main power failure on 17 April 2016 (1 day) Tracker misalignment since December 2016
3	Lahore	LAH	RSI sensor temperature measurement difficulties from October 2014 to August 2015 (in this period, the temperature was calculated from air temperature, wind speed and irradiance) Significant battery degradation since November 2016, thus some days of measurements are missing in December 2016 and January 2017
4	Quetta	QUE	No significant issues, generally well running station
5	Multan	MUL	No significant issues, generally well running station
6	Bahawalpur	QASP	Shading influence from newly built mosque in the morning after station installation Tracker error since 12 December 2014, repaired at 17 March 2015 Main power supply failure in periods 26 February 2016 – 28 April 2016, 24 - 26 May 2016 and 19 - 20 June 2016 Tracker misalignment since December 2016
7	Khuzdar	KHU	No significant issues, generally well running station
8	Hyderabad	HYD	No significant issues, generally well running station
9	Karachi	KAR	No significant issues, generally well running station

3.3 Quality control and harmonization of solar measurements

3.3.1 Method

Prior to the comparison with satellite-based solar resource data, the ground-measured irradiance was quality-controlled by Solargis. Quality Control (QC) is based on methods defined in SERI QC procedures, Younes et al. and Long et al. [24, 25, 26, 27] and implemented in-house by the company Solargis. The tests are applied in two runs: (i) the automatic tests are run to identify the obvious issues; next (ii) by visual inspection we identify and flag inconsistencies, which are of a more complex nature. Visual inspection is an iterative and time-consuming process.

The **automatic QC tests** include:

- Identification of missing values
- Correction of time shifts
- Evaluation of measurements against sun position (Sun below and above horizon)
- Comparing the data with possible minimum and maximum irradiance limits
- Evaluation of consistency of GHI, DIF and DNI by comparing the redundant measurements (if available).

The **visual quality control** aims to identify and flag the following erroneous patterns:

- Shading from nearby objects (near shading) or mountains (far shading)
- Regular data error patterns
- Irregular anomalies
- Comparison of measurements from different instruments.

Data readings not passing one or more QC tests were flagged and excluded from further analysis. Overview of the quality control results for all nine meteorological stations can be found in Chapter 3.3.2.

The QC for Tier 1 and Tier 2 stations is slightly different due to instrument setup. A very sensitive consistency test (consistency between GHI, DNI and DIF) can be used only for data from Tier 1 stations, as these measurements are conducted by three independent instruments. For measurements from Tier 2 stations the consistency test is irrelevant, as the DNI is calculated from the GHI and DIF using Twin RSI instrument. In addition to measurements from Twin RSI, the Tier 2 stations measure GHI by an independent low uncertainty secondary standard thermopile pyranometer CMP10 that allows inter-comparison of GHI with measurements from Twin RSI.

3.3.2 Results of quality control

The detailed results of quality control for all GHI and DNI instruments are presented in the Annex. The overview of results in Table 3.9 shows only a small amount of data readings not passing the QC. The data from Tier 1 stations have slightly more excluded values due to more complex and more sensitive quality control tests as well as the presence of local shading (in Islamabad).

Despite the lower nominal uncertainty compared to Twin RSI, the thermopile based pyranometers and pyrheliometers are more sensitive to proper cleaning, thus soiling is usually more visible in these data (e.g. QC results for Quetta station). This results in slightly higher number of identified problems in data from the CMP10 pyranometer from Tier 2 stations. Moreover, the smaller soiling problems in Twin RSI data were corrected by the service provider (CSP Services).

Table 3.9: Result of quality control. Percentage of valid data records

	Valid DNI records [%]	Valid GHI records [%]
Peshawar* (Tier 2)	100.0	99.2 / 99.0
Islamabad (Tier 1)	98.0	95.9
Lahore* (Tier 2)	100.0	99.3 / 99.4
Quetta* (Tier 2)	100.0	93.3 / 100.0
Multan* (Tier 2)	98.3	97.8 / 98.1
Bahawalpur (Tier 1)	99.7	98.3
Khuzdar* (Tier 2)	100.0	99.8 / 100.0
Hyderabad* (Tier 2)	100.0	96.7 / 99.8
Karachi* (Tier 2)	100.0	99.4 / 99.8

** Tier 2 stations - GHI measurements from thermophile CMP10 pyranometer / Twin RSI*

3.3.3 Findings

The QC results show several types of issues related to the measured radiation. This especially applies to the degradation of some data due to instrument soiling and shading.

The main findings:

- In general, the thermopile pyranometers are more susceptible to proper cleaning, thus soiling is usually more visible in these data. The visual data inspection shows that cleaning was not always sufficient at some stations. Several short periods with measurements affected by soiling were identified in GHI measured by CMP10 at Tier 2 stations (Figure 3.2). Such data were excluded from further processing.
- The GHI and DNI measurements from the Islamabad station were influenced by shading (Figure 3.3). This issue was identified only in the morning hours of summer 2016, while in previous years the instruments were not shaded. The shaded data readings were flagged and excluded from further processing.
- Several very short periods with unrealistically low GHI and DIF were identified in measurements from Islamabad and Bahawalpur (Figure 3.4). These data readings were excluded from further processing.
- A very small proportion of data from Tier 1 stations did not pass the consistency test between GHI, DNI and DIF measurements (Figure 3.5). These issues occurred mostly during low sun elevation angle conditions and were excluded from further processing. Even after exclusion of these readings a residual inconsistency between the measured DNI and DNI calculated from GHI and DIF remains in the data. For the Islamabad station, the measured DNI is lower by -2.8% compared to calculated DNI. For Bahawalpur, this difference is -3.5 %. These residual inconsistencies are usually the result of small effects of insufficient cleaning and other operational issues, and/or the accuracy of the instruments. The differences are within acceptable limits for in-field measurements, as well as quality control tests.
- The inter-comparison of measurements from two GHI instruments indicates a systematic disagreement at all Tier 2 stations (Figure 3.6). The measurements from CMP10 pyranometers are always higher than GHI from Twin RSI (after exclusion of soiling issues). The disagreement is negligible for low irradiance values, and increases for high irradiance values. For the period of available measurements, the average difference for all stations ranges from -0.7 - -3.9. Maximum noontime difference has a high day-by-day variability and may exceed -6 to -8% (Figure 3.7). There is no clear explanation for these features, but it can be related to different spectral sensitivity of instruments (silicon sensor used in RSI has narrower spectral range), operation of the instruments in specific climate or possible issues in calibration. The

data and metadata inspection did not provide a clear indication that would help to identify the main source of this feature of GHI measurements. As the thermopile pyranometers have lower nominal uncertainty, a higher priority was given to these data in the regional-adaptation of Solargis data (Chapter 4)

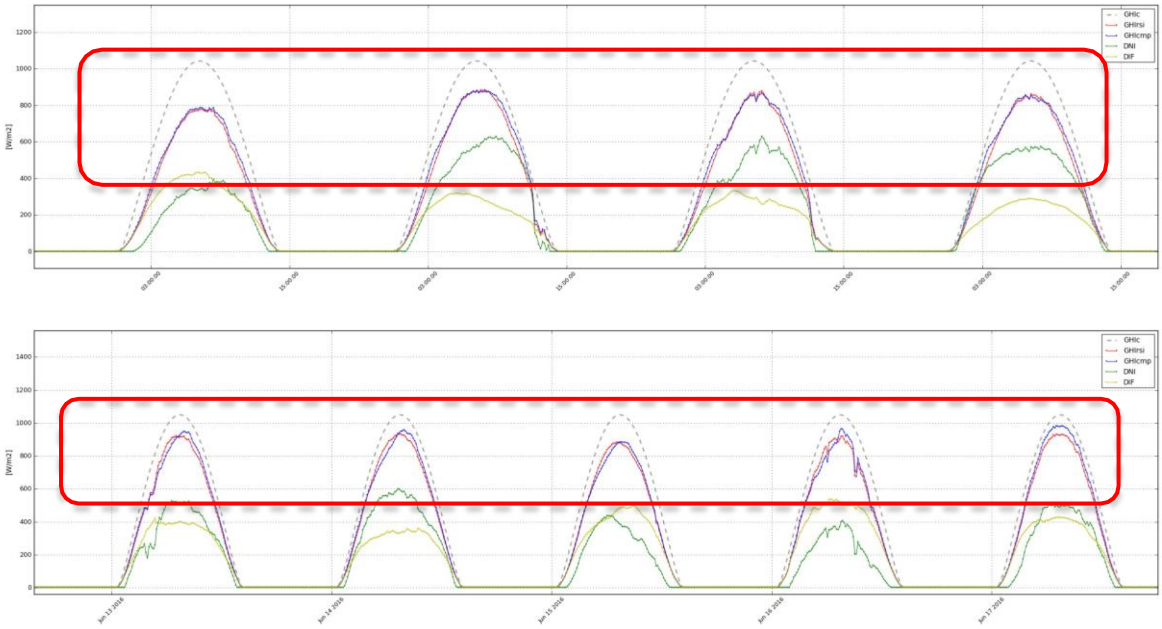


Figure 3.2: Effect of soiling on GHI measurements

Top: Mulan; bottom: Hyderabad
blue: GHI (CMP10), red: GHI (Twin RSI); green: DNI; yellow: DIF

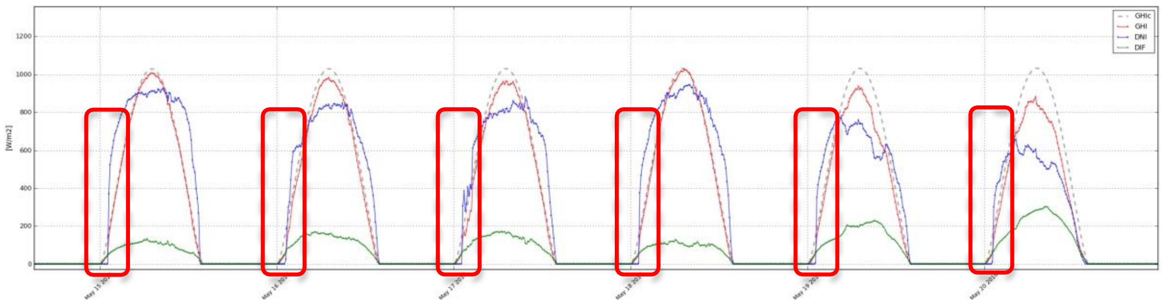


Figure 3.3: Shaded DNI and GHI measurements at Islamabad station

blue: GHI (CMP10); red: GHI (Twin RSI); green: DNI; yellow: DIF

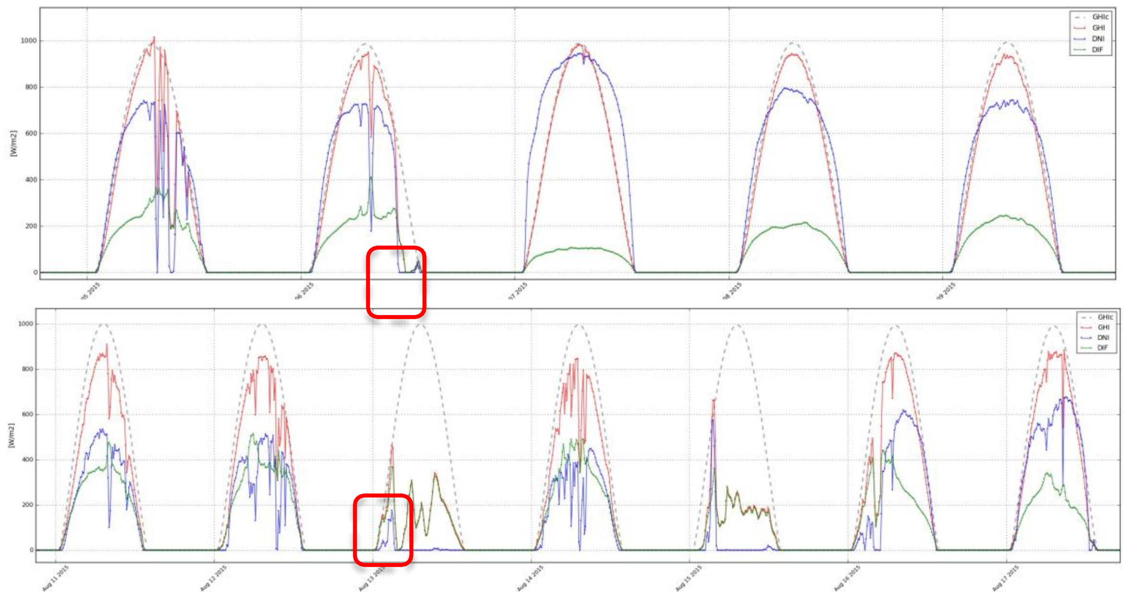


Figure 3.4: Short periods with unrealistically low GHI and DIF measurements

Top: Bahawalpur; bottom: Islamabad
blue: GHI (CMP10), red: GHI (Twin RSI); green: DNI; yellow: DIF

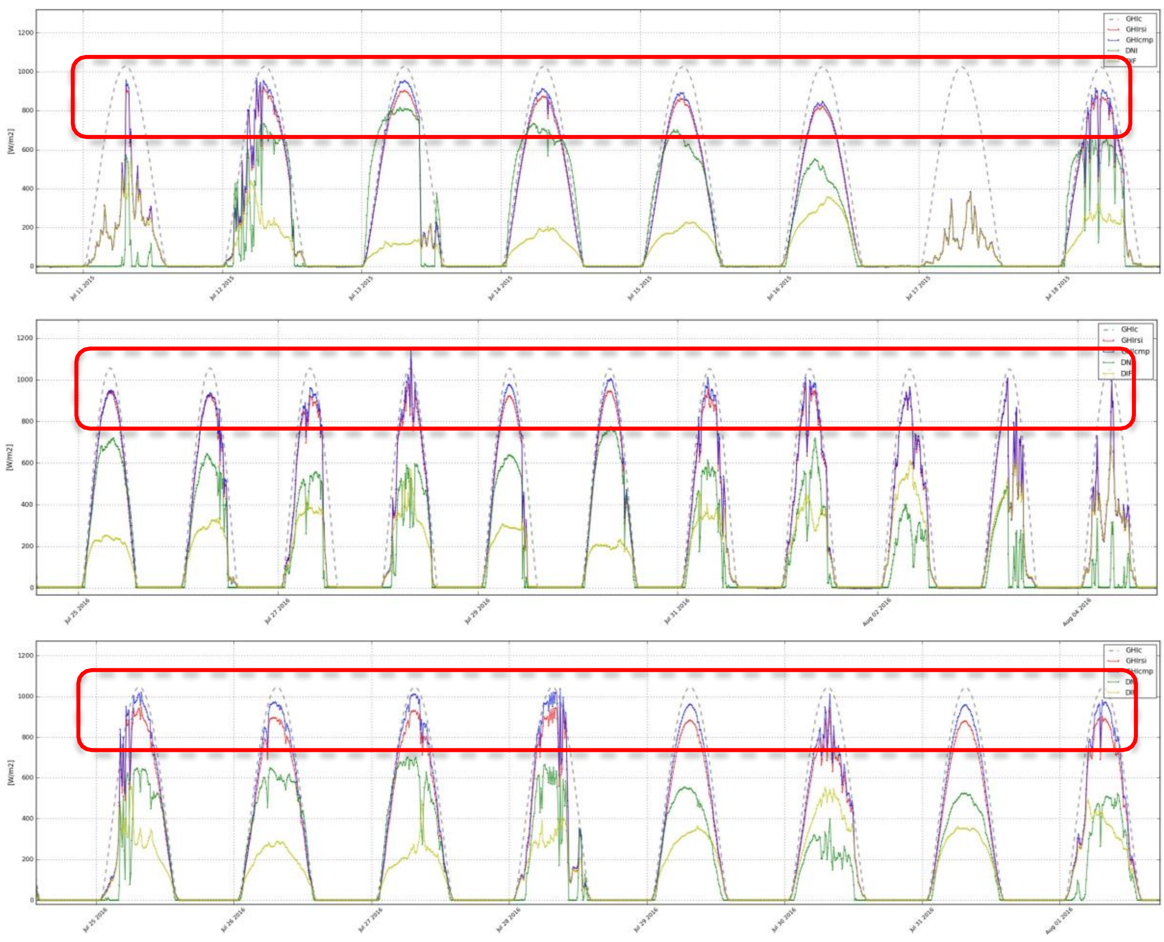


Figure 3.5: GHI difference between thermopile pyranometer CMP10 and Twin RSI

Top: Peshawar; middle: Khuzda; bottom: Hyderabad
blue: GHI (CMP10), red: GHI (Twin RSI); green: DNI; yellow: DIF

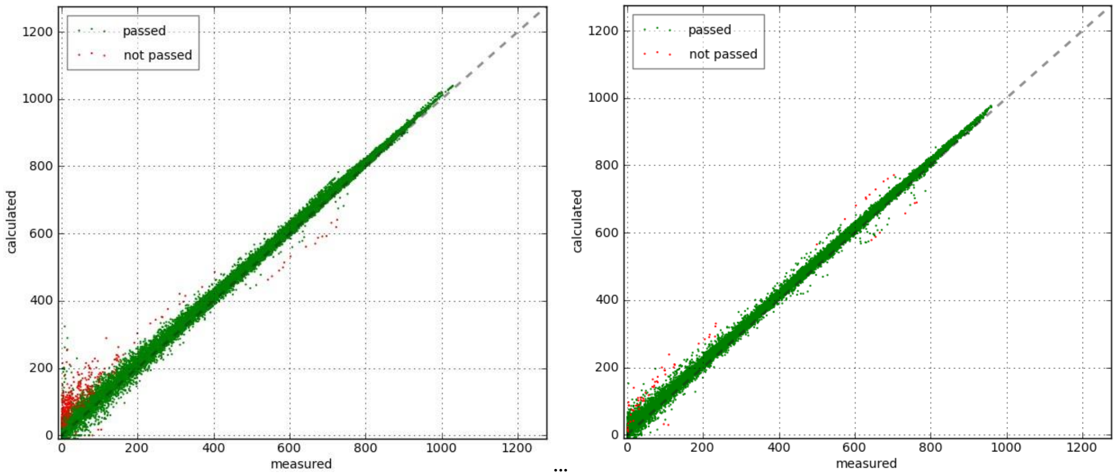
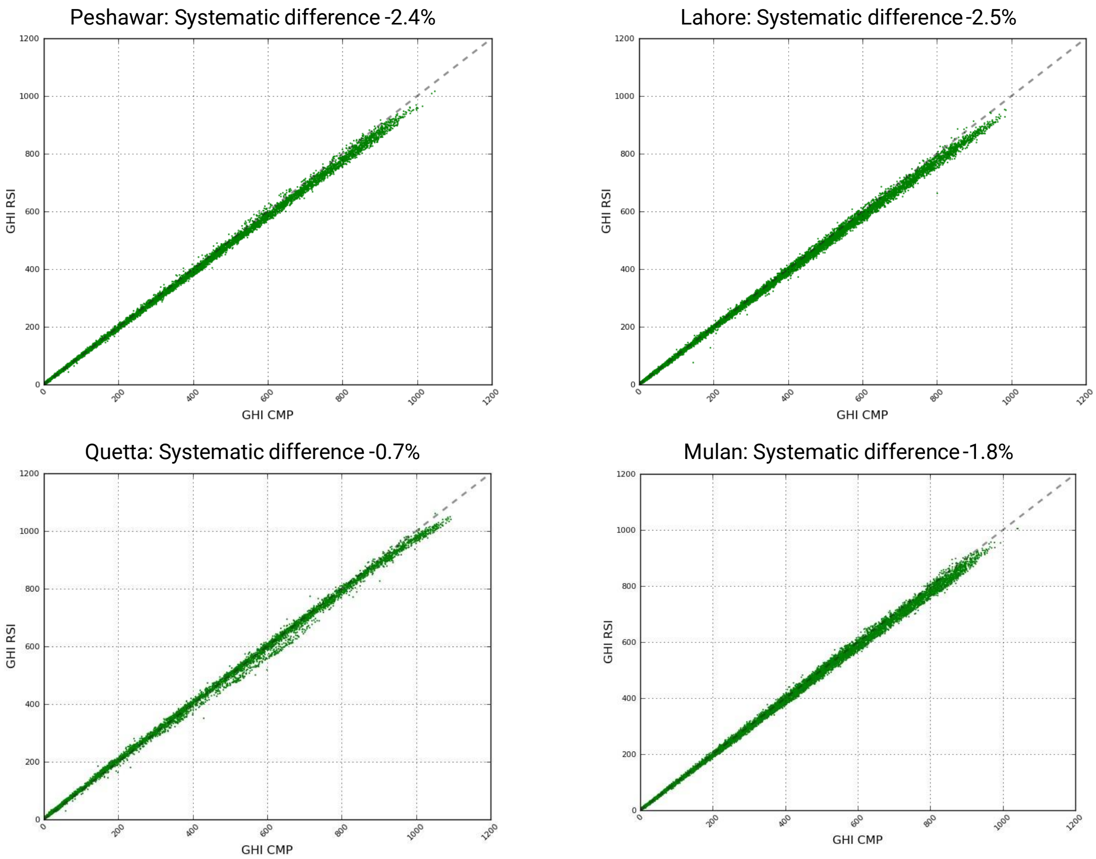


Figure 3.6: Consistency test for Tier 1 stations

Left: Islamabad; right: Bahawalpur
X-axis: measured DNI, Y-axis: DNI calculated from GHI and DIF



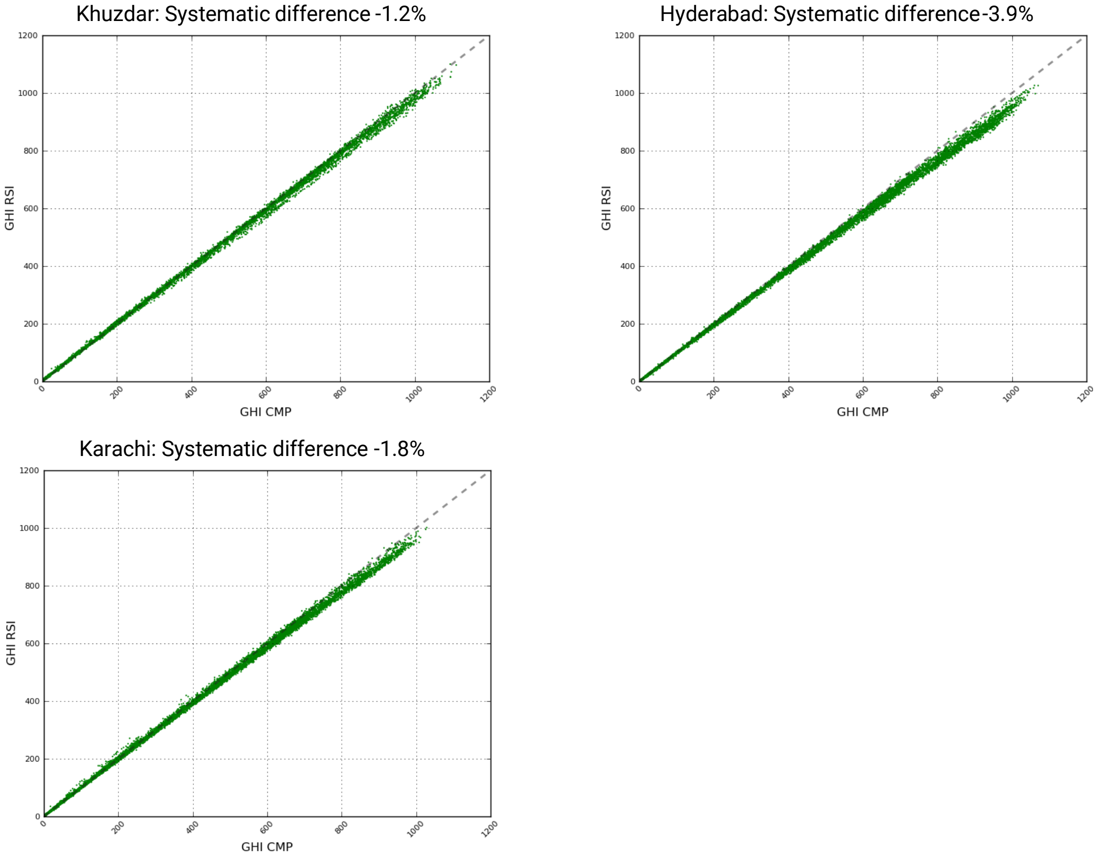


Figure 3.7: Difference between GHI measured by pyranometer CMP10 and by Twin RSI

X-axis: pyranometer CMP10, Y-axis: Twin RSI

3.3.4 Conclusions

Based on the quality control results (Table 3.10) we conclude that the solar radiation measurements come from high quality instruments with good operation and maintenance. Occasional operational issues (such as incorrect tracking, battery failure) were resolved by the service provider in a short time. Issues with insufficient cleaning reported for some stations were recorded and partly resolved by the service provider.

Several smaller issues were identified by Solargis quality control and the data records were excluded from further processing. The residual uncertainty of measurements after quality control is indicated by:

- Discrepancy between data from two types of GHI instruments for Tier 2 stations
- Slight inconsistence between GHI, DNI and DIF measurements for Tier 1 stations.

In both cases the residual issues are within the expectations given by the combined nominal uncertainties of sensors. These issues have direct implication on the achievable uncertainty of ground measurements, and subsequently, on the results of the regional-adaptation of the Solargis data (Chapter 4).

In evaluation of the uncertainty of measurements several factors are considered:

1. Thermopile pyranometer CMP10 has lower nominal uncertainty than the Twin RSI instrument.
2. The thermopile pyranometers are more susceptible to soiling, which is very relevant for the geographical conditions of Pakistan.
3. Instruments are used in challenging environmental conditions (high temperature, humidity, etc.).

Table 3.10: Quality control summary

Description	Peshawar	Islamabad	Lahore	Quetta	Multan	Bahawalpur	Khuzdar	Hyderabad	Karachi
Station description, metadata									
Instrument accuracy									
Instrument calibration									
Data structure									
Cleaning and maintenance information									
Time reference									
Quality control complexity									
Qualit y control									
Time period									
Other issues		Ventilation, tracking, shading	Sensor temp.			Tracker, shading, power			

Legend: Quality flag

Very good	Good	Sufficient	Problematic	Insufficient	Not specified

4 Regional adaptation of Solargis model

Ground measurements from nine solar meteorological stations in Pakistan are used for the regional adaptation of the Solargis model (Figure 4.1). Additionally, the data from four AERONET stations, operating in the region, is used for the improvement of Aerosol Optical Depth data used in the Solargis model (Chapter 2.1). Ultimately, the model has been run to produce verified and accuracy enhanced GHI and DNI data layers and maps of Pakistan.

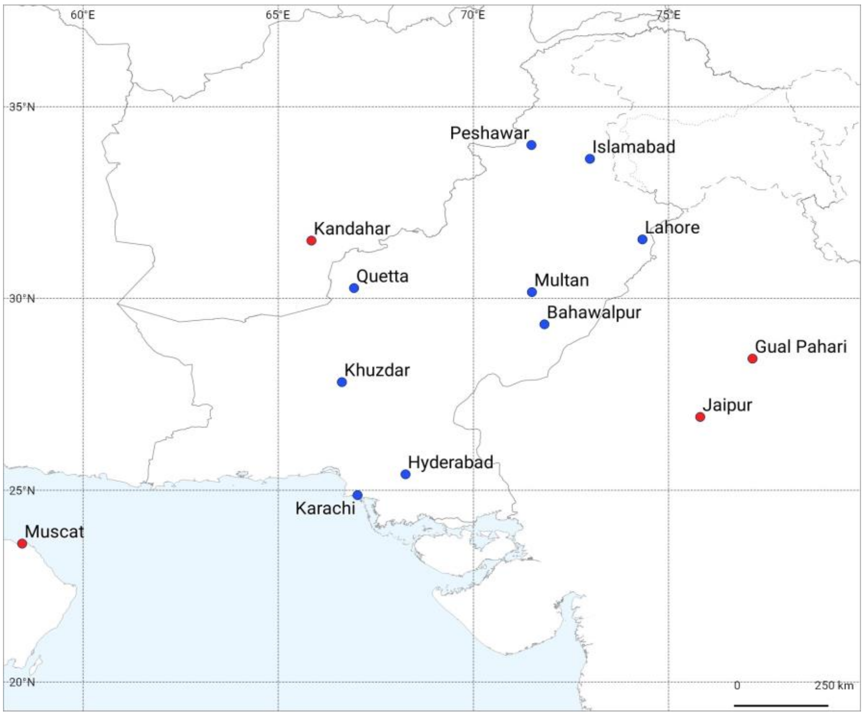


Figure 4.1: Position of stations used for regional adaptation

Aerosol measurements (red) – AERONET; Solar radiation measurements (blue) – ESMAP

4.1 Solargis method

Solargis regional model adaptation aims at reducing the mean bias (systematic deviation), RMSD (random deviation) and KSI (difference between frequency distribution of the measured and satellite-based data).

The comparison of data produced by the original Solargis model to the ground measurements shows a relatively good fit in various aspects: under cloudy as well as cloudless conditions, solar radiation intermittency, and also for day-by-day variability. A deficiency in some regions is due to a systematic underestimation or overestimation of daily profiles for cloudless conditions (Figure 4.2). This indicates possible issues in the calculation of clear-sky model (cloudless conditions), which is mainly controlled by aerosol data input. Other sources of deviation, such as the accuracy of cloud identification, have a much lower impact on the model results. Therefore, we have decided to focus the regional adaptation on accuracy improvement of the model inputs, namely Aerosol Optical Depth (AOD).

More about the Solargis site adaptation in [22].

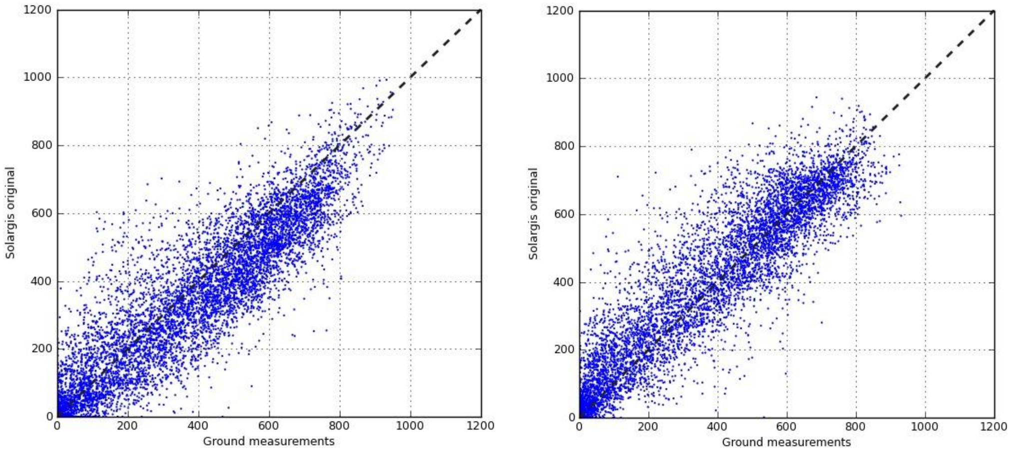


Figure 4.2: Comparison of original satellite-model DNI with ground measurements.
Left: underestimation - Bahawalpur; right: slight overestimation - Karachi

The method was conducted in two steps:

1. Coefficients for adaptation of Aerosol Optical Depth were derived for individual ESMAP meteorological stations and AERONET stations (Chapter 4.1.1)
2. Model adaptation coefficients were interpolated between the solar meteorological and AERONET sites to extend over the territory of Pakistan (Chapter 4.1.2).

4.1.1 Reduction of systematic deviation between satellite and measured data

Deviations between raw satellite-based data and ground-measurements were analysed individually for each site. We focussed on the understanding of differences between these two datasets:

1. Deviation for the whole period of measurements
2. Seasonal patterns of deviation
3. Spread of deviations for various weather situations
4. Differences in the cumulative distribution of values.

Based on these results, **correction factors for AOD** (Atmospheric Optical Depth) values were proposed. They were developed by comparing the cloudless situations with theoretical clear-sky profiles. The input aerosols were corrected separately for low and medium to high aerosol load concentrations, and for each group a separate set of monthly correction factors was identified.

In the next step, these **factors were harmonized** to avoid abrupt month-by-month changes that may be the result of insufficient representation of some situations by ground measurements, rather than a systematic problem in the data. In this phase, correction factors from the neighbouring sites were also compared to avoid issues in a spatial context, especially if sites are located in similar geographic conditions. In a case of nearby sites with contradicting deviations (e.g. Karachi with positive and Hyderabad with negative bias), the corrections for individual sites have to be balanced to avoid high spatial mismatch. The bias is removed only partially to maintain the spatial consistency of corrections. This approach helps to avoid steep changes of correction coefficients in the area.

In the last step, the **satellite-based model was recalculated** using proposed aerosol correction factors and the results were again evaluated. Larger residual deviations were removed in the second iteration of the same procedure.

For each site, the adaptation procedure results in a set of monthly correction values for low and medium AOD load. The correction was developed to respect the seasonal and spatial context of the data.

During the data evaluation, the availability and quality of ground measurements were considered, as the ground measurements may not provide a comprehensive representation of specific situations or may contain residual errors. The satellite data is available in 30-minute time step; ground-measured data is available in 10-minute time steps. To reduce the conceptual difference of point and satellite pixel measurements, all the measures are calculated using aggregated data in *hourly timestep*.

After adaptation of AOD, the model is fully recalculated for the period of 18 years. Thus, the consistency of GHI, DNI and DIF components is maintained.

The aerosol adaptation method removes the major source of discrepancies between satellite-model data and the ground measurements. Because of fundamental difference between the modelling using satellite data and ground measurements, even if bias was minimized, some mismatch between the data may still be present – mainly in the *frequency distribution* of values. The main focus of the regional adaptation is to determinate correction coefficients for individual sites that are used to remove seasonal and annual deviations in the regional context. The residual discrepancies present in the regionally adapted data can be removed only in the local context as they are sourced by locally specific features such as pollution in big cities. Such residual discrepancies cannot be extrapolated and they are not addressed in the regional model adaptation.

4.1.2 Spatial interpolation of aerosol correction coefficients

The main objective of this step is to extend the correction coefficients identified at the sites to the territory of Pakistan. To achieve this goal, a complex interpolation technique, incorporating orographic barriers, was used. The selection of interpolator is based on the assumption that the spatial distribution of aerosols is controlled by air mass movement influenced by orographic features.

Applying the spatial interpolation, with use of the digital elevation model, we extended the correction factors information from the solar (and aerosol) measuring stations to the whole territory of Pakistan and neighbouring areas. The interpolation was applied separately for each month and two aerosol load conditions. The output of the interpolation is a set of 24 aerosol correction layers.

Finally, aerosol correction layers were used for regional re-calibration of the Solargis model inputs and full recalculation of the 18-year database for the territory of Pakistan.

4.2 Results and validation

4.2.1 Accuracy estimate of DNI and GHI at the solar measuring stations

The original Solargis data show a regional pattern of over- and underestimation, compared to the ground measurements. The model adaptation allowed the removal of a large part of mismatch between satellite-based data and ground measurements, especially for DNI.

In semi-arid and desert conditions, the clouds have lower importance and it is mainly AOD that determines the mismatch between ground-measured and satellite data. However, AOD has a strong non-linear effect, especially on the DNI output. Therefore, more accurate results (reduction of RMSD and KSI) were achieved by adaptive adjustment of the AOD values.

Tables 4.1 to 4.4 summarize validation of the regional adaptation results for all solar measuring stations. The original Solargis data represent output of the model, which is based on a standard calculation scheme without consideration of any corrections derived from the measurements available for Pakistan. The regionally adapted model includes corrections based on ESMAP project measurements in Pakistan. The GHI validation statistics (Table 4.2 and Table 4.4) show a comparison to measurements from thermopile pyranometers. This type of instrument has lower nominal uncertainty and is more stable for various spectral irradiance conditions (Chapter 3.3). Terms are explained in Glossary. Absolute values of bias are calculated for daytime hours only.

Table 4.1: Direct Normal Irradiance: bias and KSI before and after regional model adaptation

Meteo station	Original DNI model data			DNI after regional adaptation		
	Bias [kWh/m^2]	Bias [%]	KSI [-]	Bias [kWh/m^2]	Bias [%]	KSI [-]
Quetta	17	3.0	123	-4	-0.7	124
Peshawar	-34	-10.1	154	-6	-1.8	51
Khuzdar	-25	-4.3	165	1	0.1	75
Multan	-1	-0.3	58	8	2.6	43
Lahore	8	3.0	62	0	-0.1	46
Hyderabad	-33	-7.7	192	-1	-0.2	113
Karachi	18	5.2	93	5	1.3	65
Islamabad	-17	-4.8	93	-6	-1.7	49
Bahawalpur	-37	-9.8	167	-13	-3.5	66
Mean	-12	-2.9	123	-1.8	-0.4	70
Standard deviation	22	5.8		6.3	1.8	

Table 4.2: Global Horizontal Irradiance: bias and KSI before and after regional model adaptation

Meteo station	Original GHI model data			GHI after regional adaptation		
	Bias [kWh/m^2]	Bias [%]	KSI [-]	Bias [kWh/m^2]	Bias [%]	KSI [-]
Quetta	1	0.1	20	-3	-0.6	26
Peshawar	8	1.9	35	18	4.3	76
Khuzdar	-6	-1.2	40	-4	-0.7	34
Multan	7	1.6	40	11	2.5	56
Lahore	14	3.7	70	11	2.7	56
Hyderabad	-24	-4.6	111	-17	-3.2	81
Karachi	3	0.7	27	-1	-0.3	22
Islamabad	6	1.5	33	12	2.9	57
Bahawalpur	-8	-1.7	38	-1	-0.2	17
Mean	0.1	0.2	46	2.9	0.8	47
Standard deviation	11.4	2.4		10.9	2.4	

Table 4.3: Direct Normal Irradiance: RMSD before and after regional model adaptation

Meteo station	RMSD of original DNI data			RMSD of DNI after regional adaptation		
	Hourly [%]	Daily [%]	Monthly [%]	Hourly [%]	Daily [%]	Monthly [%]
Quetta	25.1	18.1	8.6	22.0	14.1	3.0
Peshawar	35.9	29.4	19.1	30.7	23.1	5.4
Khuzdar	23.5	17.3	7.4	21.6	14.7	3.2
Multan	31.1	24.1	11.0	28.2	20.7	6.2
Lahore	38.6	30.2	12.4	33.9	25.1	6.2
Hyderabad	27.9	21.7	9.7	24.3	17.0	2.1
Karachi	28.4	20.4	8.8	25.8	17.1	3.2
Islamabad	33.4	25.7	10.1	29.3	20.2	3.3
Bahawalpur	29.4	23.7	13.7	26.0	19.6	7.1
Mean	30.4	23.4	11.2	26.9	19.1	4.4

Table 4.4: Global Horizontal Irradiance: RMSD before and after regional model adaptation

Meteo station	RMSD of original GHI data			RMSD of GHI after regional adaptation		
	Hourly [%]	Daily [%]	Monthly [%]	Hourly [%]	Daily [%]	Monthly [%]
Quetta	11.7	5.7	2.1	11.6	5.7	1.5
Peshawar	14.3	9.2	4.6	14.3	9.1	4.9
Khuzdar	11.5	5.9	1.7	11.5	6.0	1.8
Multan	13.5	9.5	4.7	13.0	8.7	4.0
Lahore	17.2	12.2	5.7	16.4	11.4	4.9
Hyderabad	10.2	7.3	5.6	9.5	6.3	4.5
Karachi	10.8	6.6	4.4	10.5	6.4	4.3
Islamabad	15.2	8.6	2.9	15.2	8.4	3.3
Bahawalpur	14.1	10.6	6.3	13.1	9.1	4.6
Mean	13.2	8.4	4.2	12.8	7.9	3.8

As a result, at the level of individual sites in Pakistan, the mean bias of the adapted values stays close to zero, which means that the model is well balanced to simulate all type of weather and geographical conditions. The standard deviation of bias values is 1.8% and 2.4% for DNI and GHI, respectively, which is low compared to the inherent uncertainty of ground sensors.

The most significant improvement was achieved for Direct Normal Irradiance (DNI) (Tables 4.1 and 4.3). The average bias of DNI for all stations dropped from -2.9% to -1.4% and the standard deviation was considerably reduced from 5.8% to 1.8% by regional adaptation. This confirms removal of specific regional problems where the DNI bias of original data was close to -10%. The spread of DNI values and fit of cumulative distributions expressed by RMSD and KSI were also improved. Slightly higher bias after adaptation can still be seen in two nearby stations - Multan and Bahawalpur (in distance of ca. 90 km) - where there was a stronger contradiction in the original bias.

To preserve the spatial consistency of the adapted data, the original bias was reduced in these sites only partially, as the regional adaptation aims for the reduction of regional features, not local ones.

The average Global Horizontal Irradiance (GHI) bias for all stations after adaptation is 0.8% with standard deviation of 2.4% (Table 4.2). There is a small increase of the bias compared to original data, but the change is below the uncertainty of ground measurements used for adaptation. Slight improvement can be seen in RMSE where the spread of values was reduced (Table 4.4). In most cases, we can see an improvement of bias at the level of individual sites, with exception of a few stations (e.g. Peshawar and Islamabad) that had a slight bias increase. This increase is the result of the regional adaptation method where the consistency of DNI and GHI data is preserved. For these stations, the original regional bias of DNI was much higher, and its reduction resulted in a slight increase of the GHI bias.

The regionally-adapted model values better represent the geographical variability of DNI and GHI solar resource and they also improve the distribution and match of hourly values. Some residual discrepancies are still identified in the output data, but their removal is beyond the possibilities of the regional adaptation. The residuals can only be removed for the locations of the meteorological stations in the context of the site-adaptation. Moreover, the residual discrepancies should be evaluated within the context of the quality of ground measurements (Chapter 3.3).

The coefficients of the regional adaptation were derived for the period with overlapping ground data with model data. These coefficients were implemented to the model to recalculate the full-time series of solar radiation. The effect of regional adaptations on the long-term average annual sum of DNI and GHI is summarized in Table 4.5.

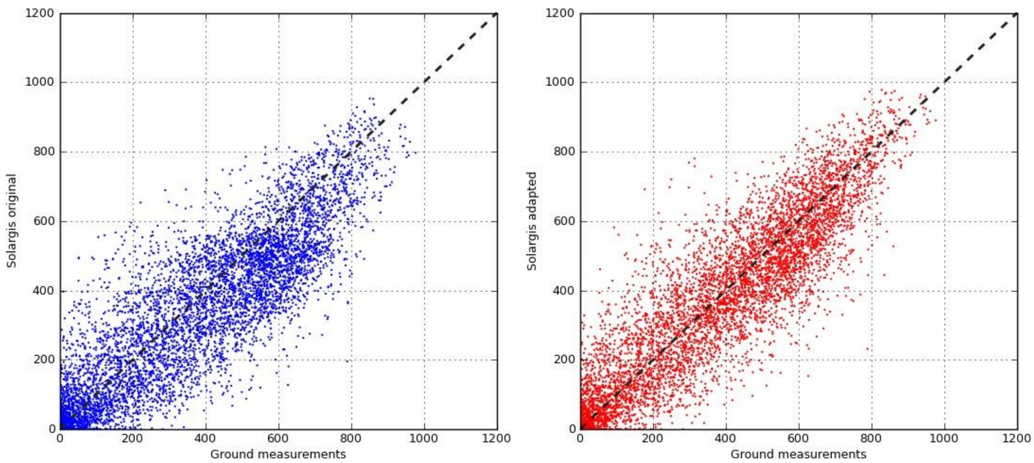


Figure 4.3: Adaptation of DNI hourly values for Peshawar

Left: original Solargis data, right: regionally adapted Solargis data.
The X-axis represents the measured DNI and the Y-axis represents the satellite-derived DNI.

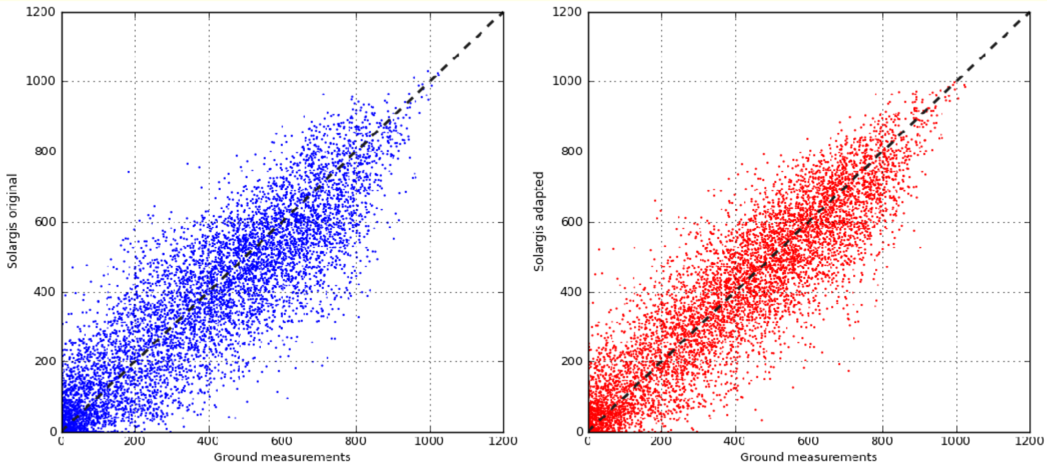


Figure 4.4: Adaptation of DNI hourly values for Islamabad

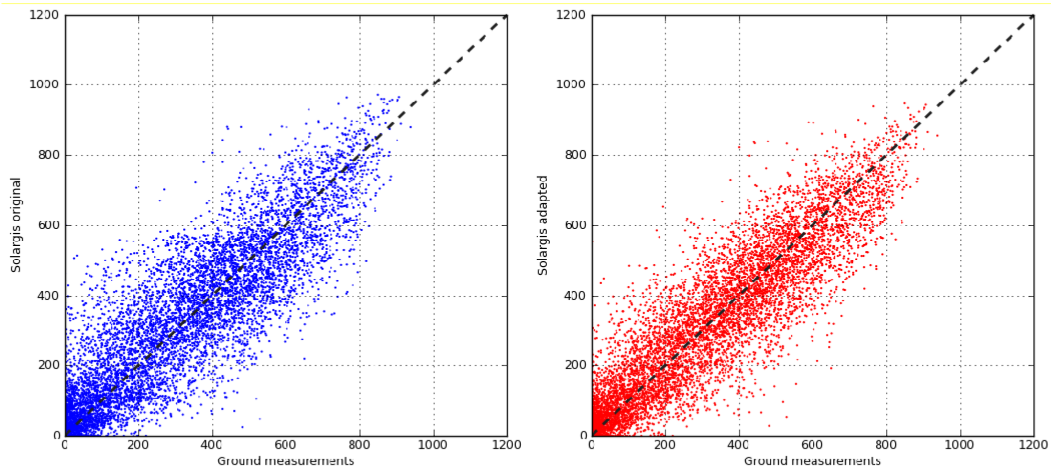


Figure 4.5: Adaptation of DNI hourly values for Lahore

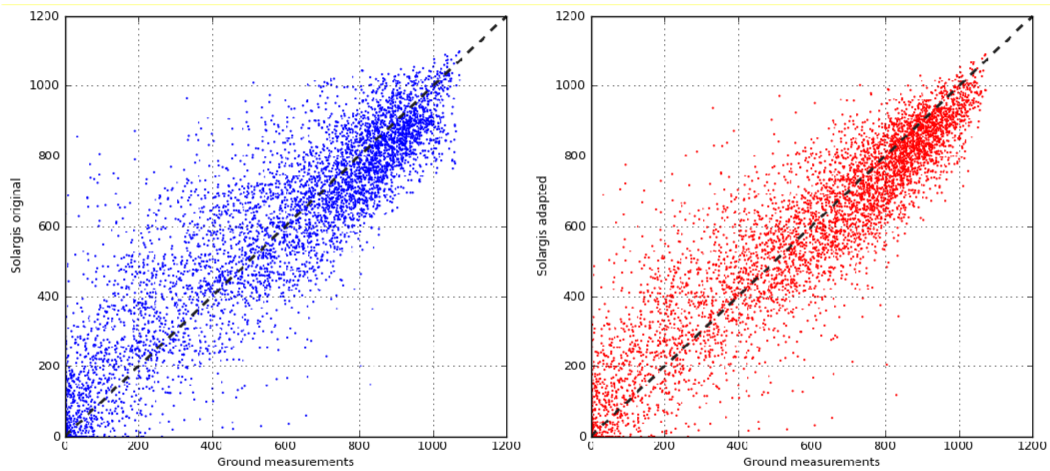


Figure 4.6: Adaptation of DNI hourly values for Quetta

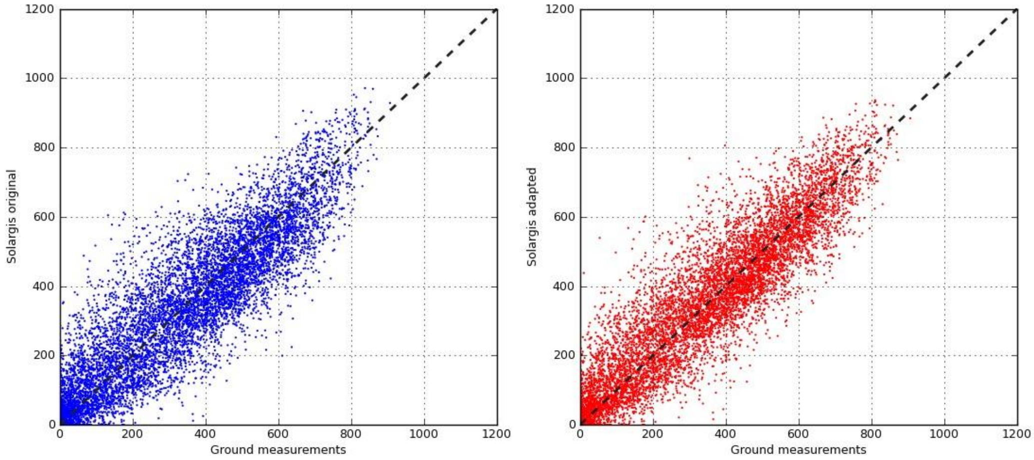


Figure 4.7: Adaptation of DNI hourly values for Multan

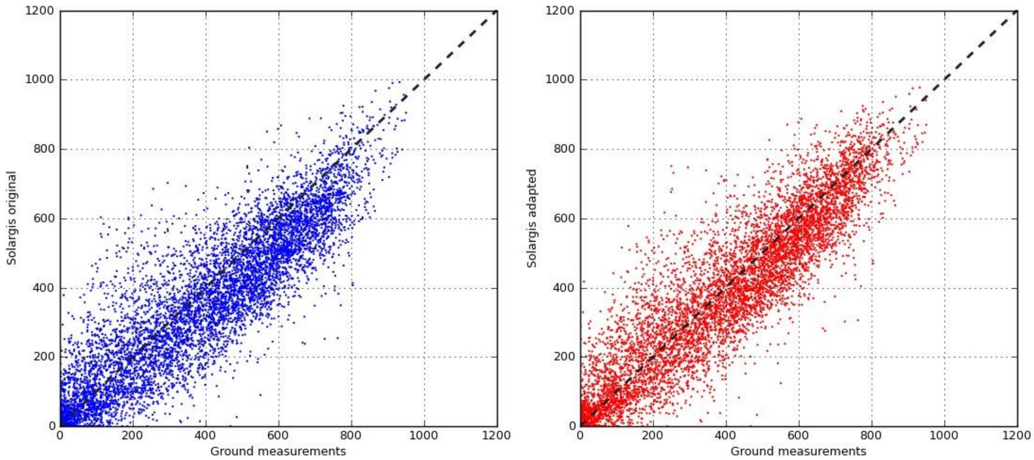


Figure 4.8: Adaptation of DNI hourly values for Bahawalpur

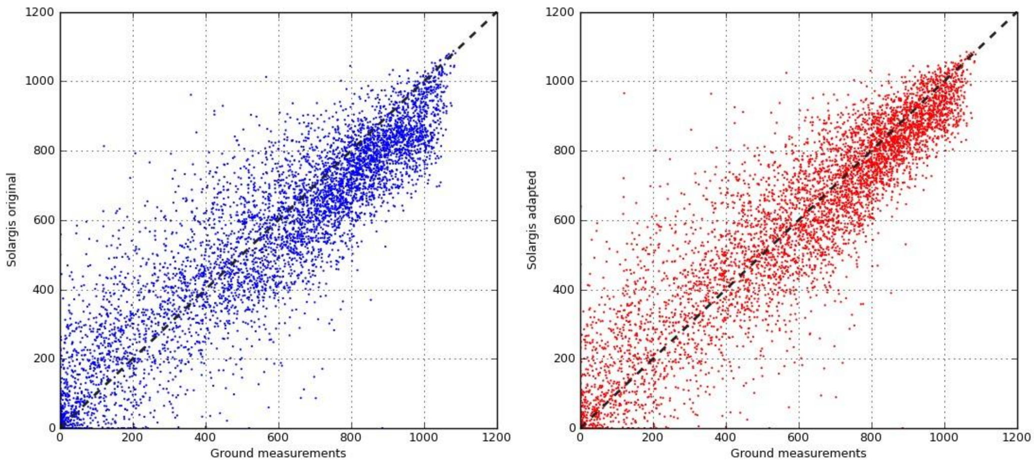


Figure 4.9: Adaptation of DNI hourly values for Khuzdar

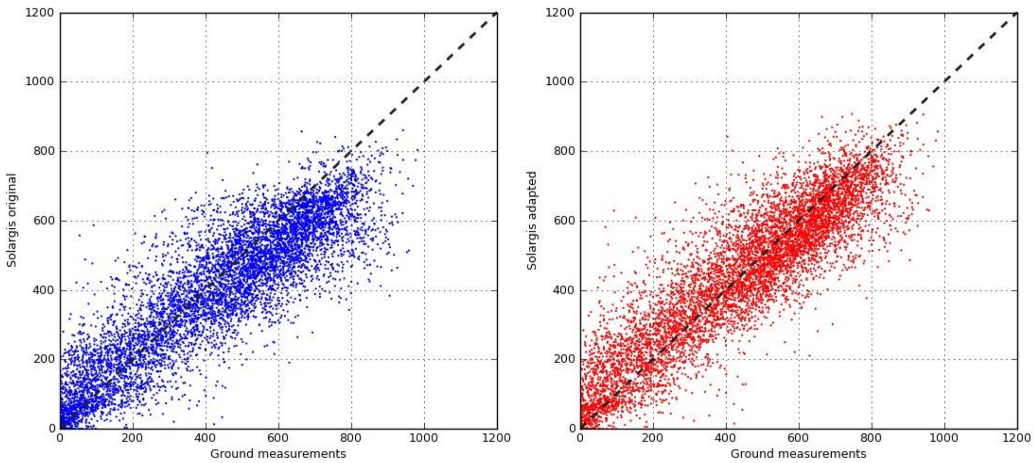


Figure 4.10: Adaptation of DNI hourly values for Hyderabad.

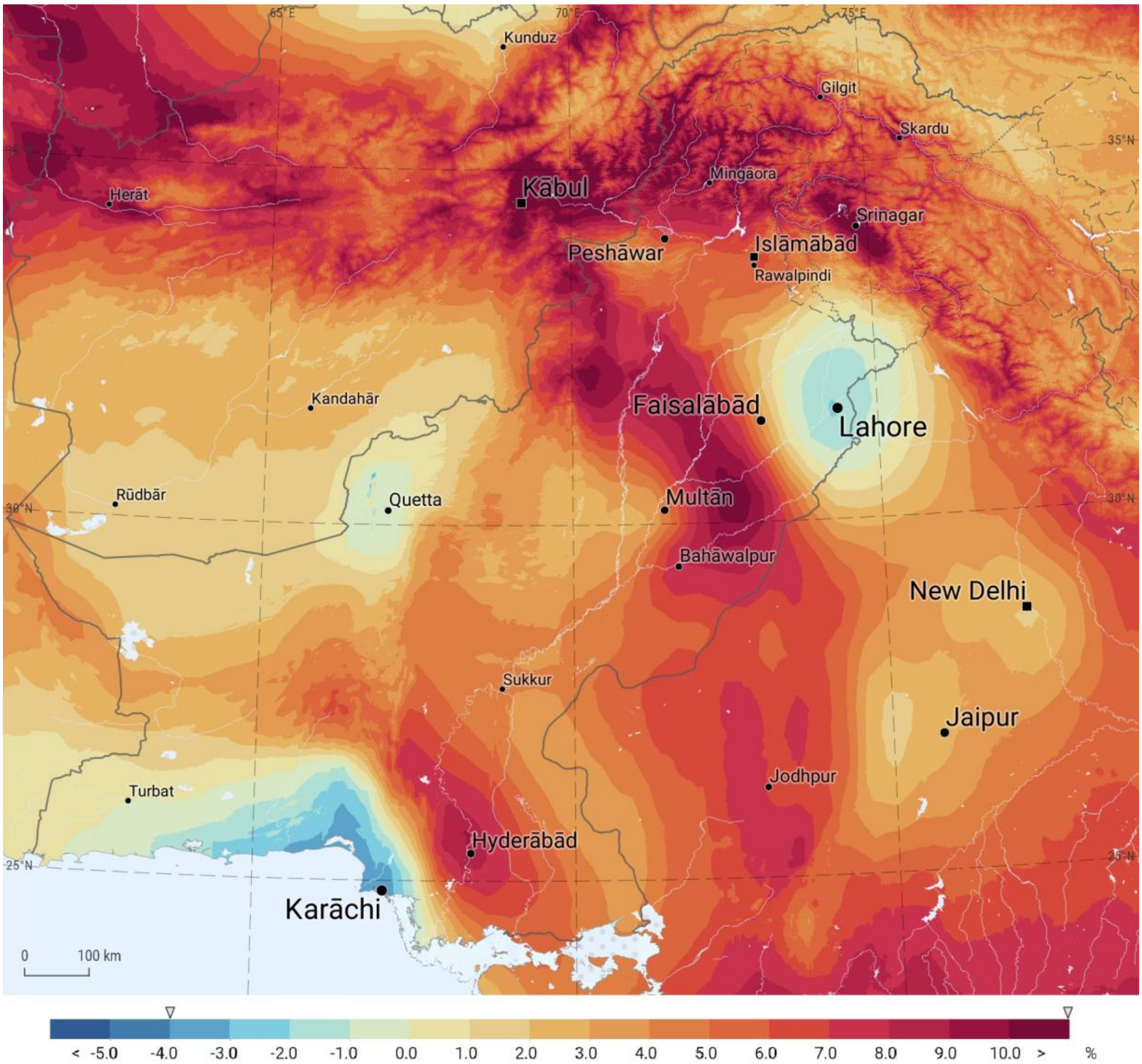
Table 4.5: Comparison of long-term average of yearly summaries of original and regionally-adapted values

Meteo station	DNI annual values*			GHI annual values*		
	Original [kWh/m^2]	Adapted [kWh/m^2]	Difference [%]	Original [kWh/m^2]	Adapted [kWh/m^2]	Difference [%]
Quetta	2301	2277	-1.0	2183	2191	0.4
Peshawar	1244	1320	6.1	1650	1695	2.7
Khuzdar	2133	2239	5.0	2183	2213	1.4
Multan	1309	1369	4.6	1822	1858	2.0
Lahore	1205	1181	-2.0	1698	1700	0.1
Hyderabad	1618	1758	8.7	2069	2108	1.9
Karachi	1576	1522	-3.4	1962	1949	-0.7
Islamabad	1389	1468	5.7	1722	1758	2.1
Bahawalpur	1455	1569	7.8	1909	1959	2.6

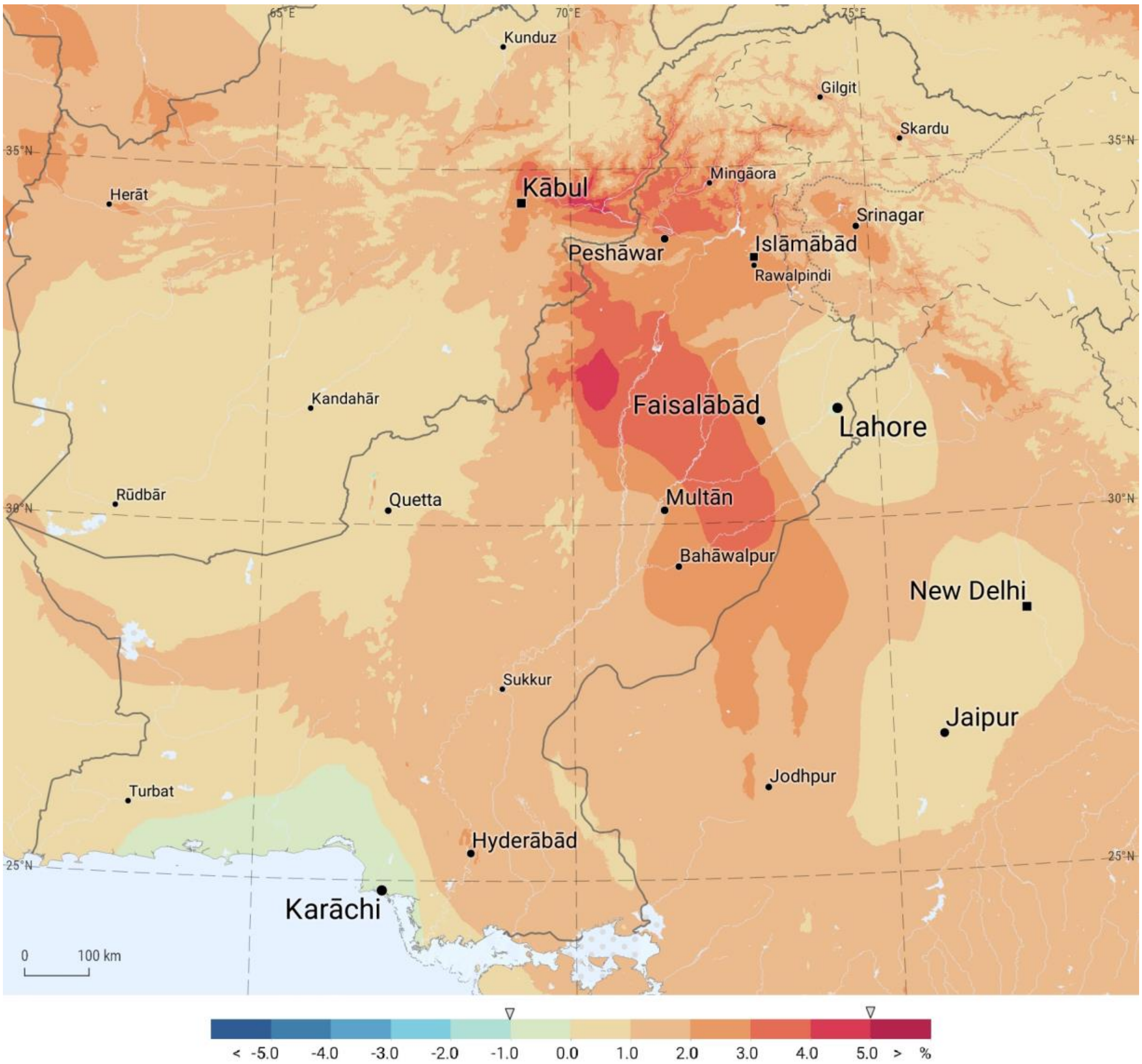
** Values represent GIS data layers, which may slightly deviate from the site-specific data values due to spatial resolution and disaggregation.*

4.2.2 Accuracy-enhanced DNI and GHI maps

Regionally adapted DNI and GHI long-term averages are higher in the majority of the area, compared to the original data (Fig 4.4 and 4.5). The absolute change of DNI is higher, compared to GHI, as DNI is more sensitive to changes of aerosol load introduced in the first step of the regional adaptation.



Map 4.4: Map of yearly DNI differences between the original and adapted model outputs



Map 4.5: Map of yearly GHI differences between the original and adapted model outputs

5 Solar resource maps of Pakistan

5.1 Accuracy enhanced maps of DNI and GHI long-term averages

After the regional adaptation of the Solargis satellite model, full time series representing a period of 18 years (1999 to 2016) are aggregated into long-term yearly averages of DNI and GHI (Figure 5.1 and 5.2). Important outcomes of this exercise are two maps with reduced uncertainty (Chapter 5.2).

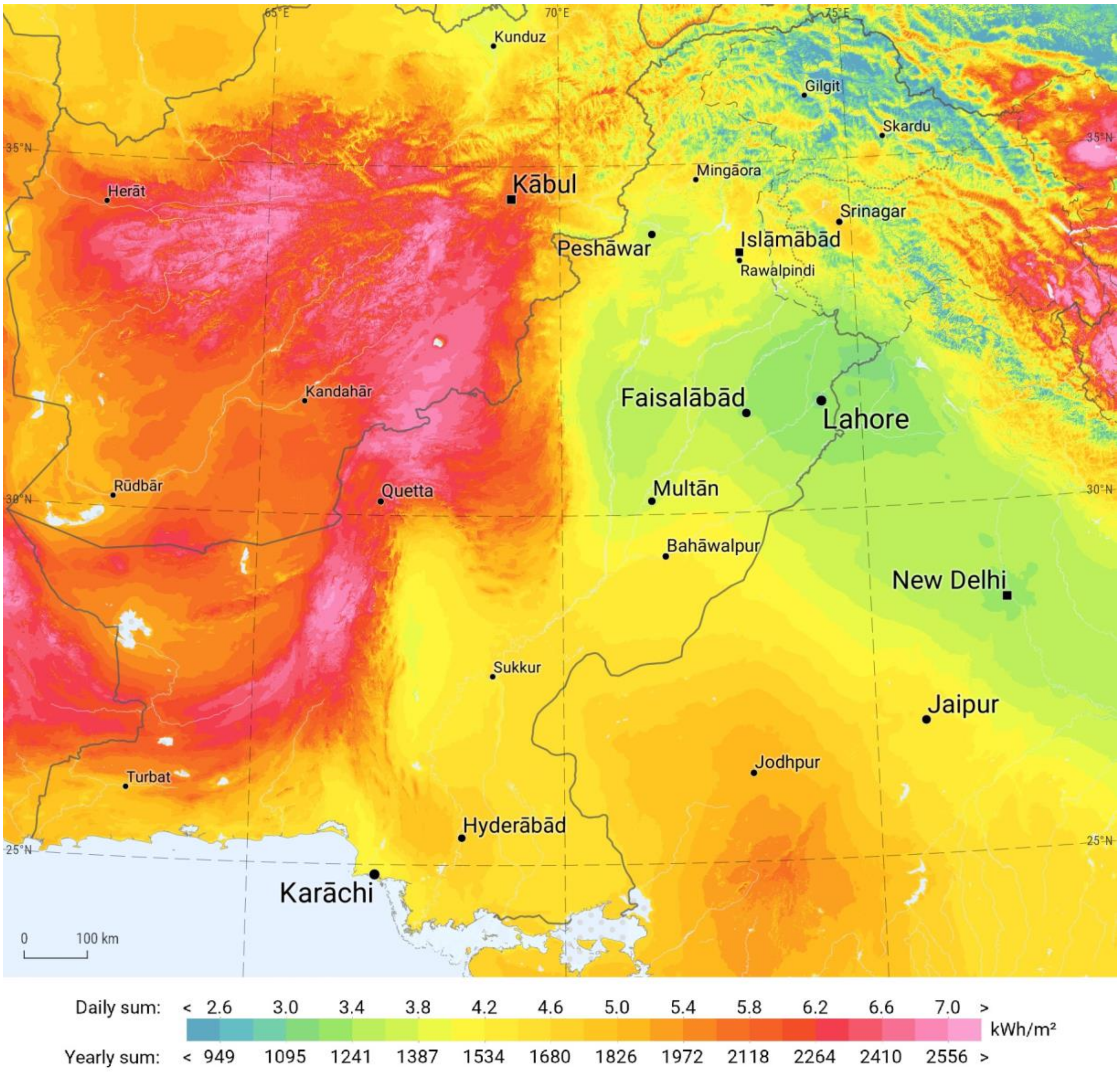


Figure 5.1: Accuracy-enhanced map of DNI yearly long-term average

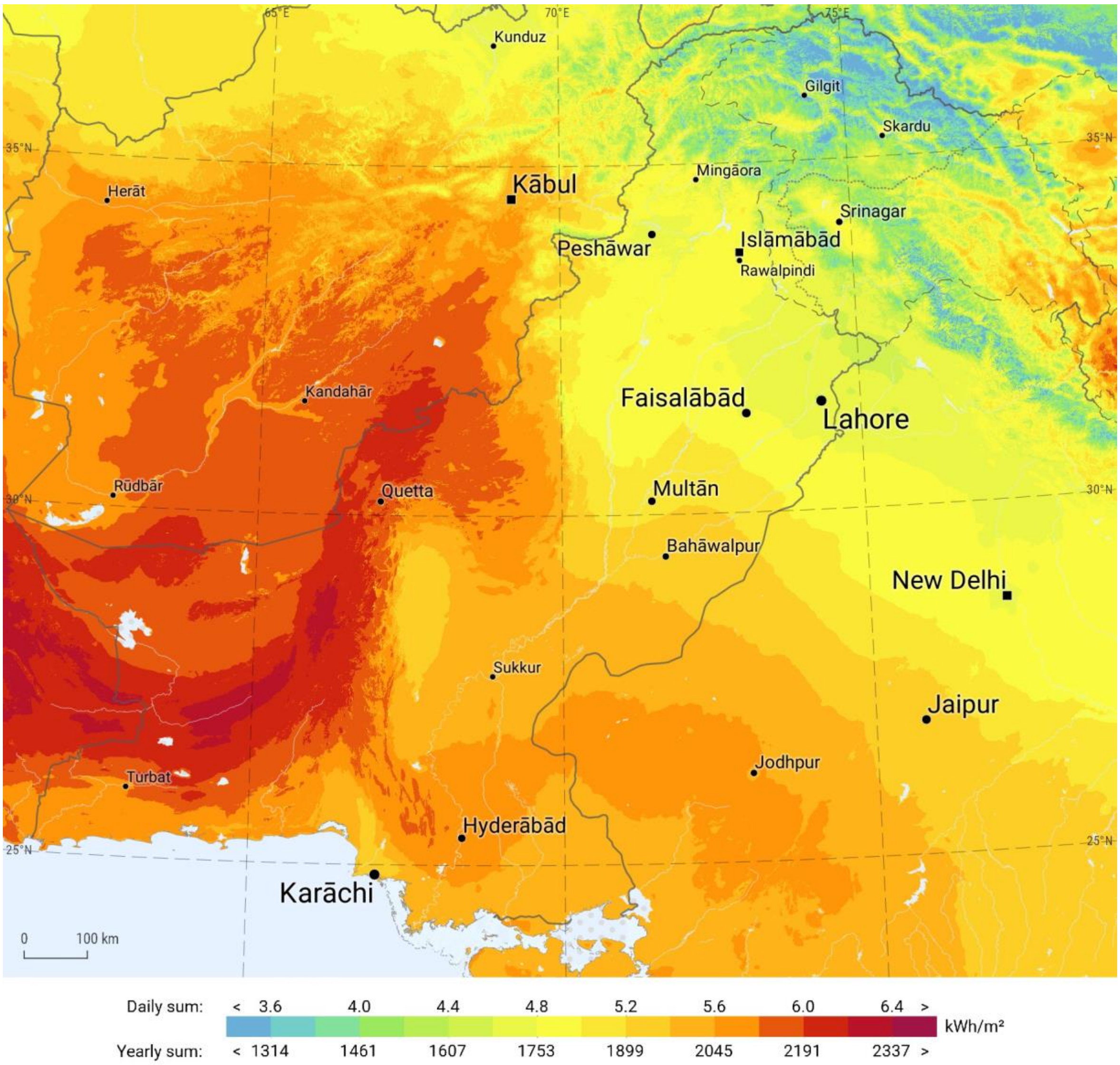


Figure 5.2: Accuracy-enhanced map of GHI yearly long-term average

5.2 Uncertainty of solar resource maps

Solargis model is based on the use of the best available algorithms and input data, and it has been calibrated and validated for all geographies. Therefore, the model has robust and uniform behaviour in all conditions. Validation sites in Pakistan and surrounding regions show consistent bias and RMSD within the expected range (Tables 4.1 to 4.4). Bias (systematic deviation) of yearly GHI and DNI is found within acceptable limits of accuracy, in the majority of cases close to the uncertainty that is typical for long-term estimates calculated from medium quality measuring instruments.

For practical use, the statistical measures of accuracy have to be converted into uncertainty, which better characterizes probabilistic nature of a possible error of the model estimate. Uncertainty is based on the assumption of normal distribution of solar radiation model errors, which is a simplification given by availability of validation data and limited geographical knowledge.

Typically, a long-term yearly average estimate is required, often denoted as P50 value (in case of normal distribution equivalent to median). Besides P50, project developers, technical consultants and the finance industry inquire about uncertainty of long-term yearly GHI or DNI. P90 values are calculated by subtracting uncertainty from P50 value.

The uncertainty in this report is calculated for 80% probability of occurrence, thus P90 value indicates an estimate at 90% probability of exceedance.

The **uncertainty of regionally adapted satellite-based DNI and GHI** is determined by uncertainty of the model, ground measurements, and the model adaptation method. More specifically it depends on:

1. Parameterization and adaptation of **numerical models integrated in Solargis** for the given data inputs and their ability to generate accurate results for various geographical and time-variable conditions:
    - Data inputs into Solargis model: accuracy of Meteosat satellite data, MACC-II/CAMS and MERRA2 aerosols and GFS/CFSR water vapour
    - Solis clear-sky model and its capability to properly characterize various states of the atmosphere
    - Simulation accuracy of the Solargis cloud transmittance algorithms, being able to properly distinguish different states of various surface types, albedo, clouds and fog
    - Diffuse and direct decomposition by Perez model
    - Transposition from global horizontal to in-plane irradiance (for GTI) by Perez model
    - Terrain shading and disaggregation by Ruiz-Arias model
2. Uncertainty of the **ground-measurements**, which is determined by:
    - Accuracy of the instruments
    - Maintenance practices, including sensor cleaning, service and calibration
    - Data post-processing and quality control procedures.
3. Uncertainty of the **model adaptation** at regional scale and residual uncertainty after adaptation

The uncertainty of the estimate $Uncert_{est}$ in this study is estimated from the model uncertainty of the Solargis model $Uncert_{model}$, the uncertainty of the measurements $Uncert_{meas}$, and the uncertainty of the model adaptation method $Uncert_{adapt}$:

$$Uncert_{est} = \sqrt{Uncert_{model}^2 + Uncert_{meas}^2 + Uncert_{adapt}^2}$$

Combined uncertainty of the yearly estimate $Uncert_{est}$ is estimated empirically, based on the experience and accuracy evaluation of the model and measurements (Chapter 3 and Chapter 4.2.1). We consider it to have probabilistic nature and is estimated based on the statistical measures calculated for seven solar measuring stations. The expert estimate of the combined user uncertainty in this report assumes 80% probability of occurrence of values, i.e. 90% probability of exceedance. Tables 5.1 and 5.2 summarize the estimated uncertainty.

The uncertainty from the interannual variability of solar resource is not considered in this study.

Based on today's knowledge and experience, we assume that the lowest achievable uncertainty (assuming uncertainty of the model and measurements at P90) of satellite-based long-term estimates is indicatively ±2.5% for GHI and ±3.5% for DNI. The uncertainty at best possible limits can only be achieved if the following conditions are met:

- Best available models and approaches are applied
- Input data (satellite, atmospheric, etc.) are quality controlled and homogenized
- Satellite model is adapted for local geography by high quality ground measurements, available for a period of at least 3 to 4 years
- Ground measurements are available for GHI, DNI and DIF, measured by high-standard meteorological instruments and equipment, applying best operation and maintenance practices.

The lowest uncertainty levels can only be achieved by site-adaptation for a very local region around meteorological stations with site-specific microclimatic conditions recorded in ground measurements. In the case of the regional adaptation used in this study, the uncertainty is usually higher because it describes uncertainty of any selected location in the broader region. Moreover, a residual discrepancy between ground measurements, and the model data can be found after regional adaptation (Tables 4.1 and 4.2). This adaptation approach is designed to correct only regional discrepancy patterns, not to resolve site-specific issues.

The uncertainty levels of regionally adapted data (Table 5.1) are higher than the best achievable results by site-specific adaptation. It is estimated that the majority of Pakistan territory has uncertainty of the model-adapted data at the level of ±4.0 to 4.5% for GHI and ±6.0% for DNI. We expect that higher uncertainty, in some regions, occurs, which is partly a result of uncertainty of ground measurements and non-optimal distribution of measurement stations throughout Pakistan (e.g. there is no data in Western and Northern Pakistan). These uncertainties can be reduced by the use of longer periods of high quality ground measurements and measurements from more meteorological stations located in regions where no ground data is available.

The expected uncertainty regions of Pakistan and associated uncertainty are presented in Table 5.2 and Figure 5.3. This map was derived by expert evaluation of the distribution and the quality of ground measurements within the context of regional geography and regional adaptation method capabilities.

Table 5.1: Uncertainty of the model estimate for original and regionally-adapted annual GHI and DNI and how does it compare to the best-achievable case.

	Direct Normal Irradiation			Global Horizontal Irradiation		
	Low	Medium	Exceptional	Low	Medium	Exceptional
Original data	< ±10%	< ±16%	±18% +	< ±6%	< ±8%	±8% +
After adaptation	±6% to ±8%	< ±14%	±18% +	±4% to ±4.5%	< ±6%	±8% +
Best-achievable	±4.0%	-	-	±2.5%	-	-

Table 5.2: Geographic distribution of the model uncertainty

Model uncertainty yearly estimates	**Region**	**GHI**	**GTI**	**DNI**
Low	Lowlands and monotonous terrain	±4% to ±4.5%	< ±5.5%	±6% to ±8%
Medium	Mountains and West Pakistan, large urban areas	< ±6%	< ±7 %	< 14%
High (exceptional)	High mountains	±8% +	±9% +	±18% +

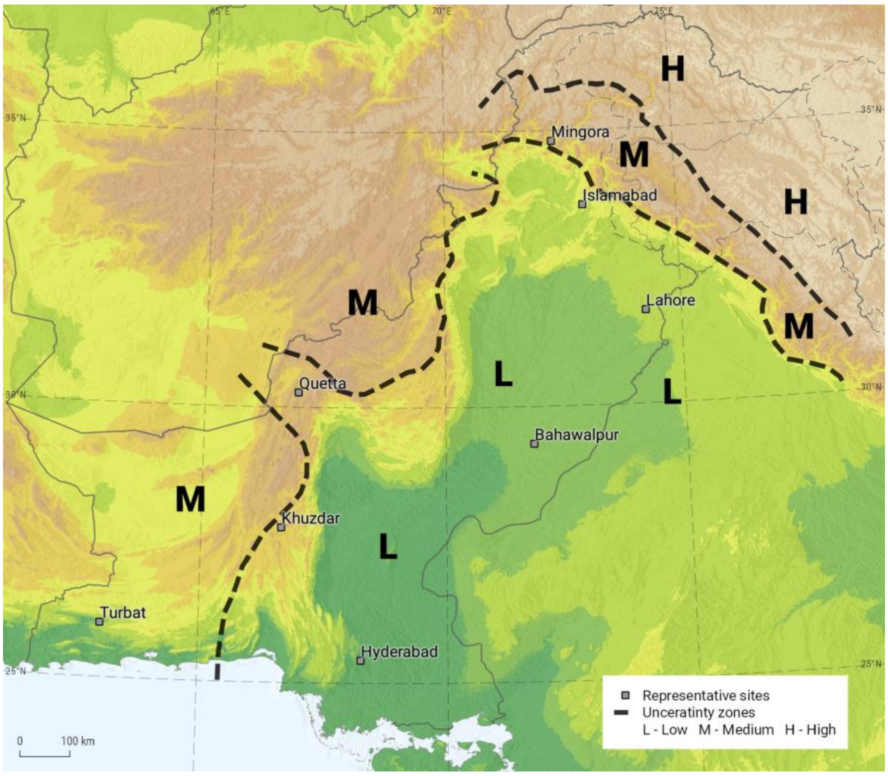


Figure 5.3: Geographic distribution of the model uncertainty
L: low; M: medium; H: high

6 Conclusions

This project reduced the uncertainty of the DNI and GHI solar resource database and the resulting maps representing the territory of Pakistan.

It is the result of systematic work on (i) setting up a network of solar meteorological stations with high-standard equipment and on (ii) implementation of rigorous practices in operation and maintenance of solar equipment. Well-linked to this infrastructure is satellite-based solar radiation model Solargis, which has proven quality and reliability of time series, and derived site-specific data products and map-based outputs.

Reduced uncertainty

The typical uncertainty of the Solargis model estimate has been reduced from the original range of ±10.0% to ±15.5% for **DNI** yearly values to the range of ±6.0% to ±14.0% for accuracy enhanced values. For yearly **GHI** the uncertainty reduction is seen from the original range of ±5.0% to ±6.5% to the range of ±4.0% to ±6.0% for the accuracy enhanced values. There is no reduction of expected uncertainty in the high mountains of the North, because no ground measurements were available.

Besides reducing systematic deviation (bias), the regional model adaptation also results in the improvement of other data quality indicators such as reducing random deviation (measurable by Root Mean Square Deviation) and by improving the probability distribution of hourly values (measurable by Kolmogorov-Smirnoff Index). Higher-quality DNI and GHI data have substantial benefits in energy simulation, which can in turn be used for more reliable financial predictions.

Role of solar measuring stations in maintaining sustainable solar data infrastructure

Receiving data from a number of high-quality measuring stations enables an improved understanding of the geographical and temporal variability of solar resource in regions of Pakistan, as a country with a diverse climate.

Even though regional adaptation reduced uncertainty, it is important to maintain the operation of the solar meteorological stations, with special focus on the following cases:

- For new sites, relevant to any larger solar power project, it is important to operate a solar meteorological station to reduce uncertainty to an achievable minimum (see Table 5.1) of the site-specific **long-term estimates.**
- For existing sites, the meteorological stations together with satellite data make it possible to maintain high quality and bankability of solar resource and meteorological data for sustainable **performance assessment** of solar power plants in the region.
- Keeping solar measuring stations is of strategic importance to maintain quality of satellite models and of solar power **forecasts**.

7 List of figures

8 List of tables

9 References

[1] Perez R., Cebecauer T., Šúri M., 2013. Semi-Empirical Satellite Models. In Kleissl J. (ed.) Solar Energy Forecasting and Resource Assessment. Academic press.

[2] Cebecauer T., Šúri M., Perez R., High performance MSG satellite model for operational solar energy applications. ASES National Solar Conference, Phoenix, USA, 2010.

[3] Šúri M., Cebecauer T., Perez P., Quality procedures of Solargis for provision site-specific solar resource information. Conference SolarPACES 2010, September 2010, Perpignan, France.

[4] Cebecauer T., Suri M., Gueymard C., Uncertainty sources in satellite-derived Direct Normal Irradiance: How can prediction accuracy be improved globally? Proceedings of the SolarPACES Conference, Granada, Spain, 20-23 Sept 2011.

[5] Šúri M., Cebecauer T., 2014. Satellite-based solar resource data: Model validation statistics versus user's uncertainty. ASES SOLAR 2014 Conference, San Francisco, 7-9 July 2014.

[6] Ineichen P., A broadband simplified version of the Solis clear sky model, 2008. Solar Energy, 82, 8, 758-762.

[7] Morcrette J., Boucher O., Jones L., Salmond D., Bechtold P., Beljaars A., Benedetti A., Bonet A., Kaiser J.W., Razinger M., Schulz M., Serrar S., Simmons A.J., Sofiev M., Suttie M., Tompkins A., Uncht A., GEMS-AER team, 2009. Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part I: Forward modelling. Journal of Geophysical Research, 114.

[8] Benedictow A. et al. 2012. Validation report of the MACC reanalysis of global atmospheric composition: Period 2003-2010, MACC-II Deliverable D83.1.

[9] Cebecauer T., Šúri M., Accuracy improvements of satellite-derived solar resource based on GEMS re-analysis aerosols. Conference SolarPACES 2010, September 2010, Perpignan, France.

[10] Cebecauer T., Perez R., Šúri M., Comparing performance of Solargis and SUNY satellite models using monthly and daily aerosol data. Proceedings of the ISES Solar World Congress 2011, September 2011, Kassel, Germany.

[11] Molod, A., Takacs, L., Suarez, M., and Bacmeister, J., 2015: Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2, Geosci. Model Dev., 8, 1339-1356, doi:10.5194/gmd-8-1339-2015

[12] GFS model. http://www.nco.ncep.noaa.gov/pmb/products/gfs/

[13] CFSR model. https://climatedataguide.ucar.edu/climate-data/climate-forecast-system-reanalysis-cfsr/

[14] CFSv2 model http://www.cpc.ncep.noaa.gov/products/CFSv2/CFSv2seasonal.shtml

[15] Meteosat satellites MFG and MSG, EUMETSAT. http://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Meteosat/index.html

[16] Cano D., Monget J.M., Albuisson M., Guillard H., Regas N., Wald L., 1986. A method for the determination of the global solar radiation from meteorological satellite data, Solar Energy, 37, 31-39.

[17] Perez R., Ineichen P., Maxwell E., Seals R. and Zelenka A., 1992. Dynamic global-to-direct irradiance conversion models. ASHRAE Transactions-Research Series, pp. 354-369.

[18] Perez, R., Seals R., Ineichen P., Stewart R., Menicucci D.,1987. A new simplified version of the Perez diffuse irradiance model for tilted surfaces. Solar Energy, 39, 221-232.

[19] Ruiz-Arias J.A., Cebecauer T., Tovar-Pescador J., Šúri M., 2010. Spatial disaggregation of satellite-derived irradiance using a high-resolution digital elevation model. Solar Energy, 84,1644-57.

[20] Zelenka A., Perez R., Seals R., Renne D., 1997. Effective Accuracy of Satellite-Derived Hourly Irradiances. Theoretical Appl. Climatology, 62, 199-207.

[21] Espinar B., Ramírez L., Drews A., Beyer H. G., Zarzalejo L. F., Polo J., Martín L., 2009. Analysis of different comparison parameters applied to solar radiation data from satellite and German radiometric stations. Solar Energy, 83, 1, 118-125.

[22] Cebecauer T., Suri M., 2016. Site-adaptation of satellite-based DNI and GHI time series: overview and Solargis approach. SolarPACES conference, September 2016, Cape Town, South Africa.

[23] Polo J., Wilbert S., Ruiz-Arias J.A., Meyer R., Gueymard C., Šúri M., Martín L., Mieslinger T., Blanc P., Grant I., Bolan J., Ineichen P., Remund J., Escobar R., Troccoli A., Sengupta M., Nielsen K.P., Renné D., Geuder N., Cebecauer T., 2016. Preliminary survey on site-adaptation techniques for satellite-derived and reanalysis solar radiation datasets. Solar Energy, 132, 25–37.

[24] NREL, 1993. User's Manual for SERI_QC Software-Assessing the Quality of Solar Radiation Data. NREL/TP-463-5608. Golden, CO: National Renewable Energy Laboratory.

[25] Younes S., Claywell R. and Munner T, 2005. Quality control of solar radiation data: Present status and proposed new approaches. Solar Energy 30, 1533-1549.

[26] Long C.N., Dutton E.G., 2002. BSRN Global Network recommended QC tests, V2.0. Rapport technique BSRN. http://epic.awi.de/30083/1/BSRN_recommended_QC_tests_V2.pdf

[27] Long C.N., Shi Y., 2008. An automated quality assessment and control algorithm for surface radiation measurements. Open Atmos. Sci. J., 2, 23-37.

Annex: Quality control at solar meteorological stations

The colours indicate the following flags:

- Green: data passed all tests
- Grey: sun below horizon
- Blue: data excluded by visual inspection - mainly soiling, shading, incorrect tracking
- White and brown: missing data
- Red: GHI, DNI and DIF consistency problem
- Violet: data out of physical limits

Peshawar

Table I: Occurrence of data readings for Peshawar meteorological station

Data availability	GHI CMP10		GHI, DNI Twin_RSI	
Sun below horizon	44 216	48.6%	44 216	48.6%
Sun above horizon	46 693	51.4%	46 693	51.4%
Total data samples	**90 909**	**100.0%**	**90 909**	**100.0%**

Table II: Excluded ground measurements after quality control (Sun above horizon) - Peshawar

Type of test	Occurrence of data samples (Sun above horizon)					
	GHI CMP10		DNI Twin_RSI		GHI Twin_RSI	
Physical limits test	356	0.8%	0	0.0%	466	1.0%
Consistency test (GHI – DNI – DIF)	-	-	-	-	-	-
Visual test (incorrect data)	0	0.0%	0	0.0%	0	0.0%
Other (non valid data)	0	0.0%	0	0.0%	0	0.0%
Total excluded data samples	**356**	**0.8%**	**0**	**0.0%**	**466**	**1.0%**
Total samples	**46 693**	**100.0%**	**46 693**	**100.0%**	**46 693**	**100.0%**

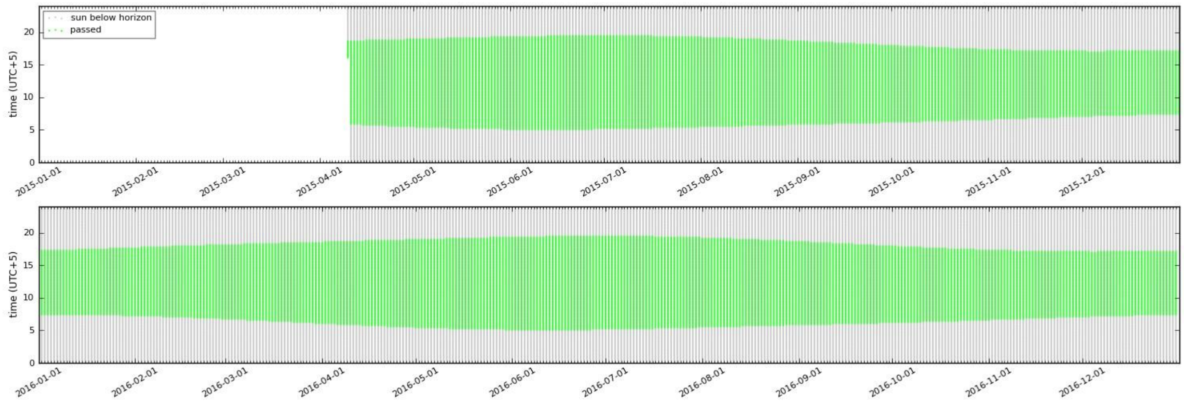


Figure I: Quality control for DNI (Twin RSI)

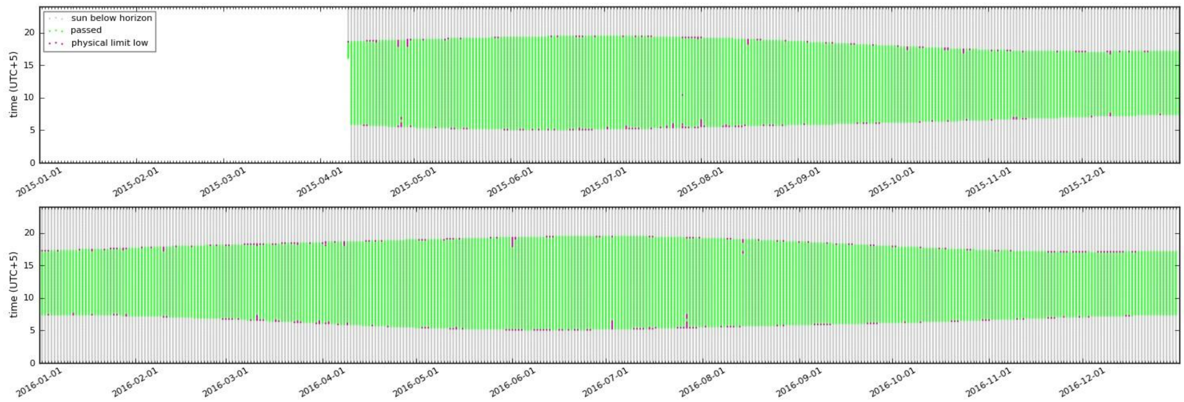


Figure II: Quality control for GHI (Twin RSI)

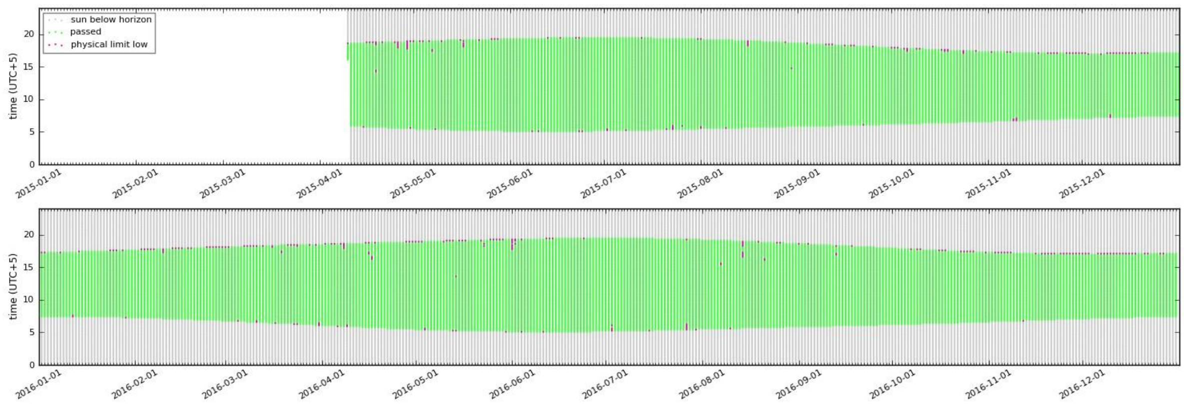


Figure III: Quality control for GHI (CMP10)

Islamabad

Table III: Occurrence of data readings for Islamabad meteorologicalstation

Data availability	GHI CMP21		DNI CHP1	
Sun below horizon	59 839	50.3%	59 839	50.3%
Sun above horizon	59 104	49.7%	59 104	49.7%
Total data samples	118 943	100.0%	118 943	100.0%

Table IV: Excluded ground measurements after quality control (Sun above horizon) - Islamabad

Type of test	Occurrence of data samples (Sun above horizon)			
	GHI CMP21		DNI CHP1	
Physical limits test	1 286	2.2%	0	0.0%
Consistency test (GHI – DNI – DIF)	312	0.5%	312	0.5%
Visual test (incorrect data)	766	1.3%	867	1.5%
Other (non valid data)	51	0.1%	7 743	-
Total excluded data samples	**2 415**	**4.1%**	**8 922**	**2.0%**
Total samples	**59 104**	**100.0%**	**59 104**	**100.0%**

Figure IV: Quality control for DNI (CHP1)

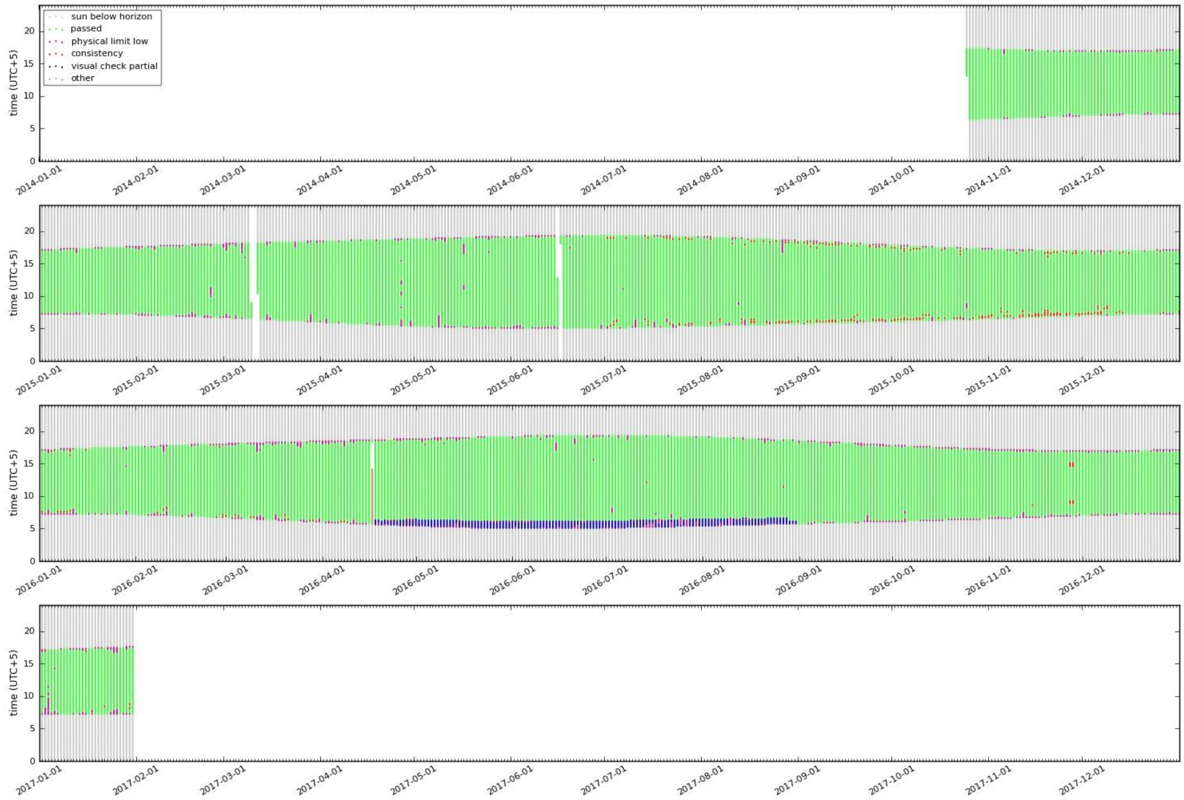


Figure V: Quality control for GHI (CMP21)

Lahore

Table V: Occurrence of data readings for Lahore meteorological station

Data availability	GHI CMP10		GHI, DNI Twin_RSI	
Sun below horizon	58 166	50.0%	58 166	50.0%
Sun above horizon	58 109	50.0%	58 109	50.0%
Total data samples	**116 275**	**100.0%**	**116 275**	**100.0%**

Table VI: Excluded ground measurements after quality control (Sun above horizon) - Lahore

Type of test	Occurrence of data samples (Sun above horizon)					
	GHI CMP10		DNI Twin_RSI		GHI Twin_RSI	
Physical limits test	305	0.5%	0	0.0%	344	0.6%
Consistency test (GHI – DNI – DIF)	-	-	-	-	-	-
Visual test (incorrect data)	123	0.2%	0	0.0%	0	0.0%
Other (non valid data)	0	0.0%	0	0.0%	0	0.0%
Total excluded data samples	**428**	**0.7%**	**0**	**0.0%**	**344**	**0.6%**
Total samples	**58 109**	**100.0%**	**58 109**	**100.0%**	**58 109**	**100.0%**

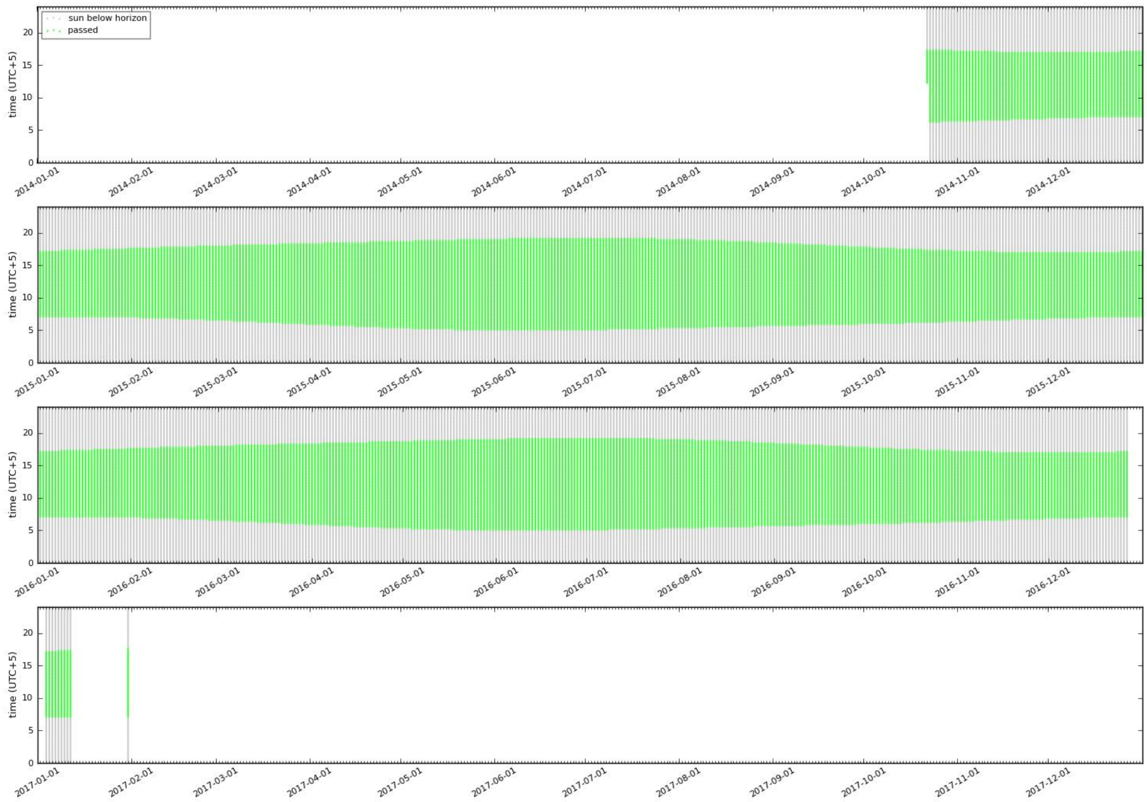


Figure VI: Quality control for DNI (Twin RSI)

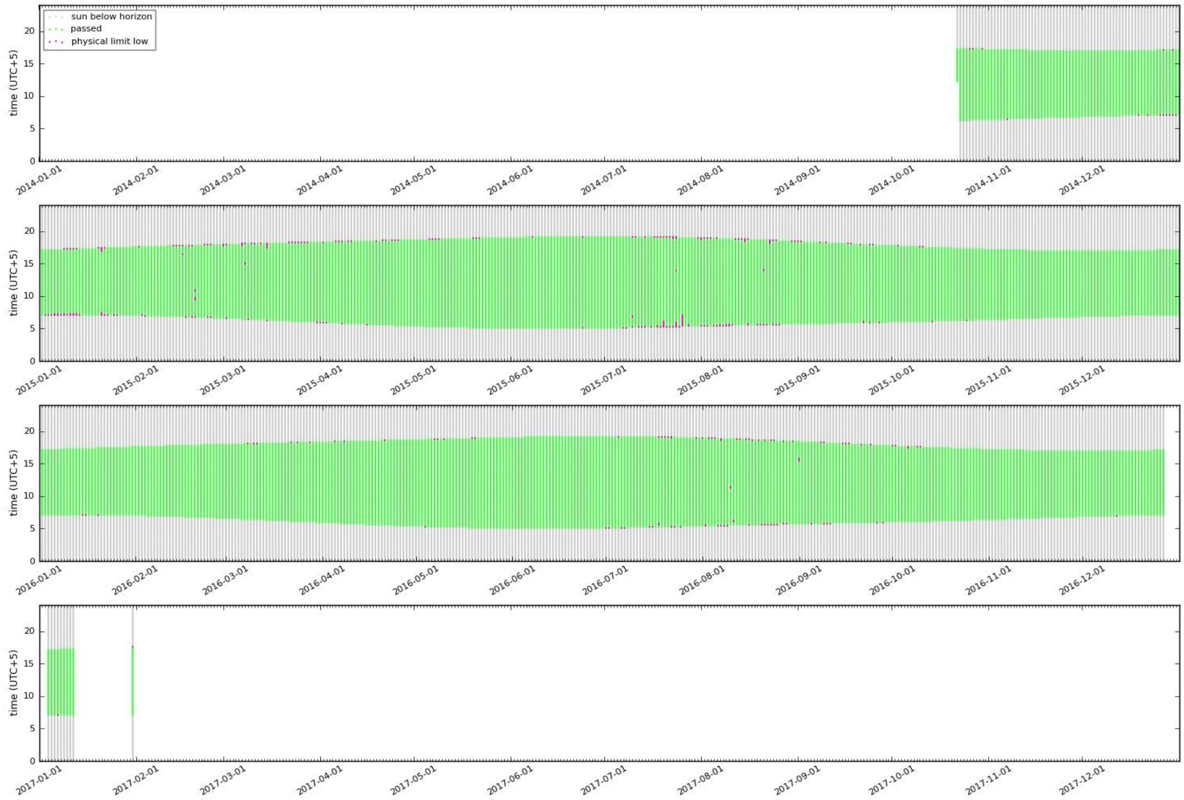


Figure VII: Quality control for GHI (Twin RSI)

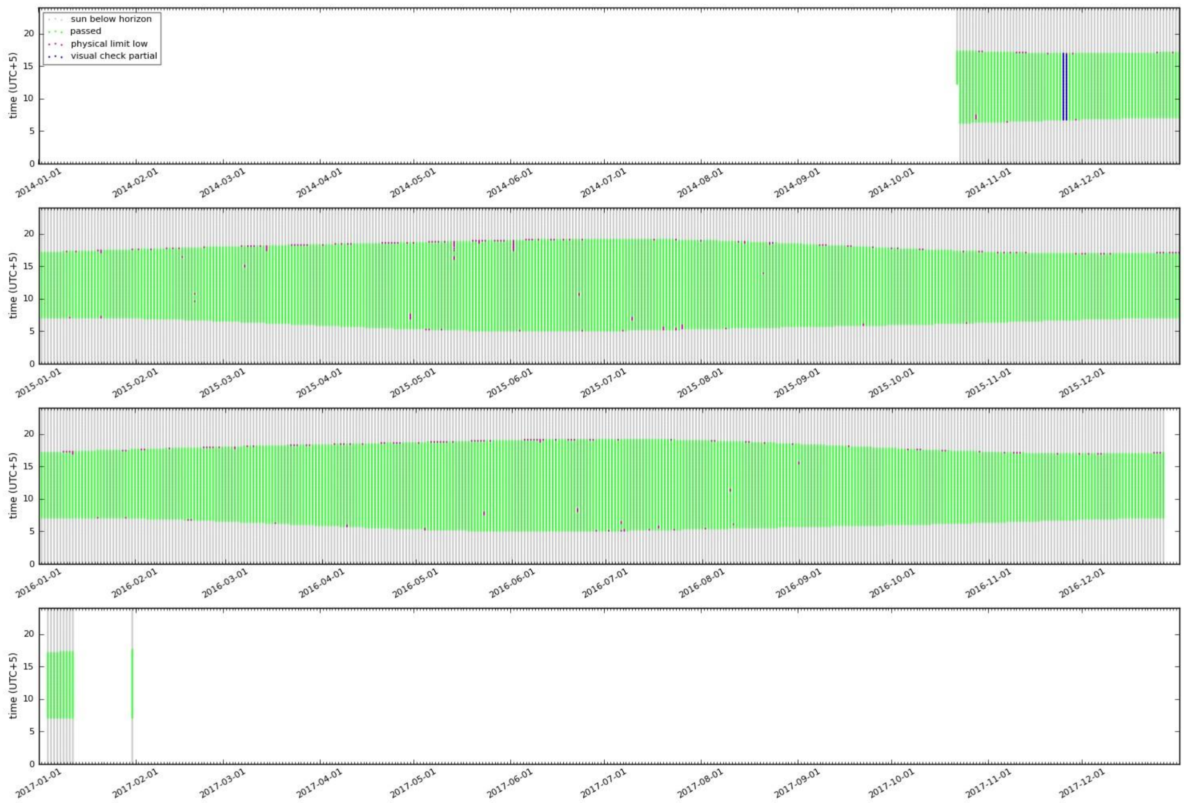


Figure VIII: Quality control for GHI (CMP10)

Quetta

Table VII: Occurrence of data readings for Quetta meteorological station

Data availability	GHI CMP10		GHI, DNI Twin_RSI	
Sun below horizon	36 815	50.9%	36 815	50.9%
Sun above horizon	35 518	49.1%	35 518	49.1%
Total data samples	**72 333**	**100.0%**	**72 333**	**100.0%**

Table VIII: Excluded ground measurements after quality control (Sun above horizon) - Quetta

Type of test	Occurrence of data samples (Sun above horizon)					
	GHI CMP10		DNI Twin_RSI		GHI Twin_RSI	
Physical limits test	143	0.4%	0	0.0%	8	0.0%
Consistency test (GHI – DNI – DIF)	-	-	-	-	-	-
Visual test (incorrect data)	2 235	6.3%	0	0.0%	0	0.0%
Other (non valid data)	0	0.0%	0	0.0%	0	0.0%
Total excluded data samples	**2 378**	**6.7%**	**0**	**0.0%**	**8**	**0.0%**
Total samples	**35 518**	**100.0%**	**35 518**	**100.0%**	**35 518**	**100.0%**

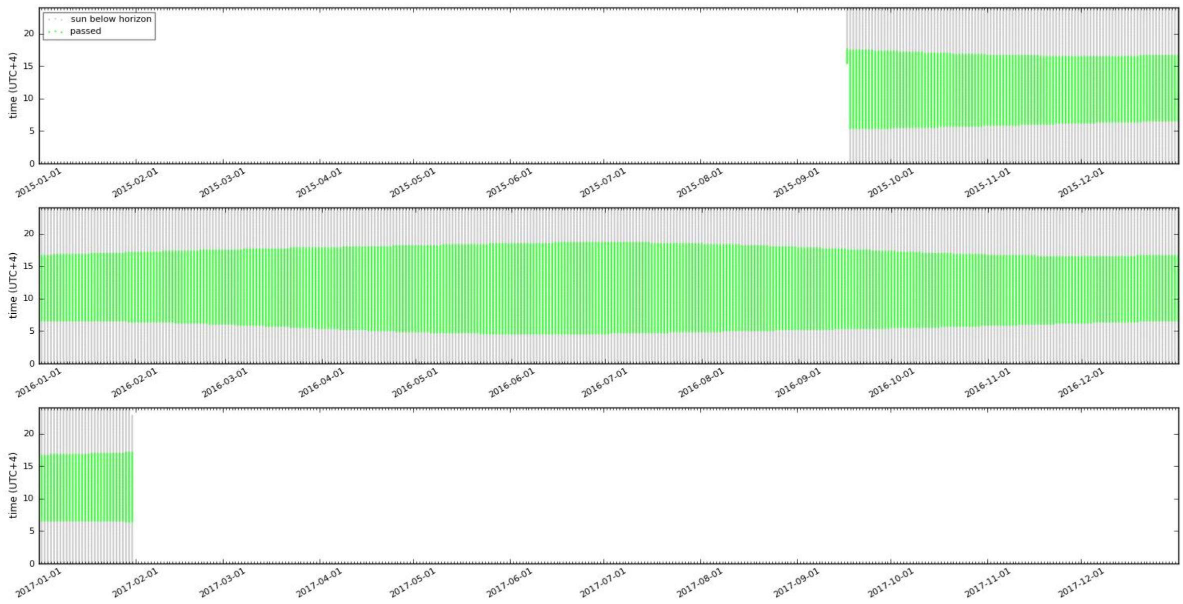


Figure IX: Quality control for DNI (Twin RSI)

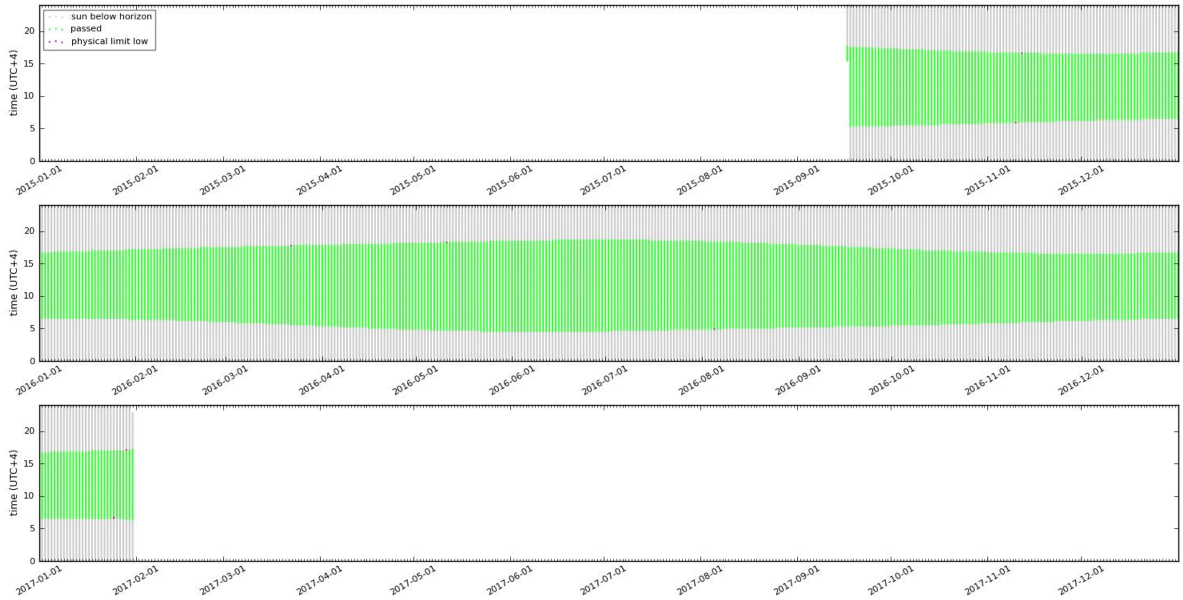


Figure X: Quality control for GHI (Twin RSI)

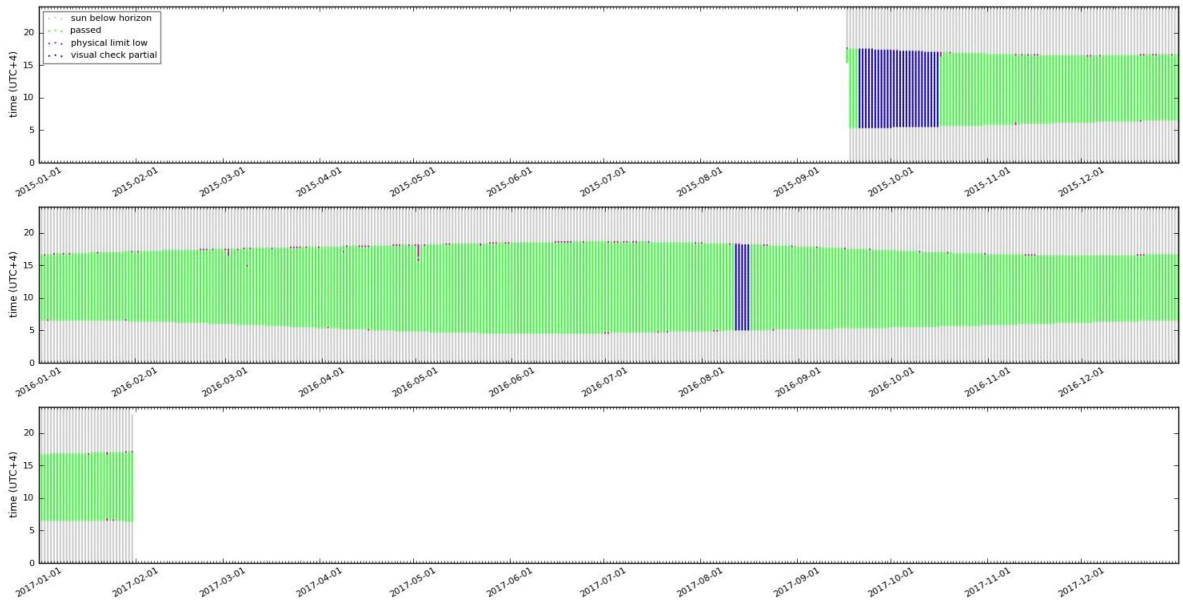


Figure XI: Quality control for GHI (CMP10)

Multan

Table IX: Occurrence of data readings for Multan meteorological station

Data availability	GHI CMP10		GHI, DNI Twin_RSI	
Sun below horizon	60 318	50.2%	60 318	50.2%
Sun above horizon	59 843	49.8%	59 843	49.8%
Total data samples	**120 161**	**100.0%**	**120 161**	**100.0%**

Table X: Excluded ground measurements after quality control (Sun above horizon) - Multan

Type of test	Occurrence of data samples (Sun above horizon)					
	GHI CMP10		DNI Twin_RSI		GHI Twin_RSI	
Physical limits test	292	0.5%	0	0.0%	123	0.2%
Consistency test (GHI – DNI – DIF)	-	-	-	-	-	-
Visual test (incorrect data)	1 003	1.7%	1 008	1.7%	1 008	1.7%
Other (non valid data)	0	0.0%	0	0.0%	0	0.0%
Total excluded data samples	**1 295**	**2.2%**	**1 008**	**1.7%**	**1 131**	**1.9%**
Total samples	**59 843**	**100.0%**	**59 843**	**100.0%**	**59 843**	**100.0%**

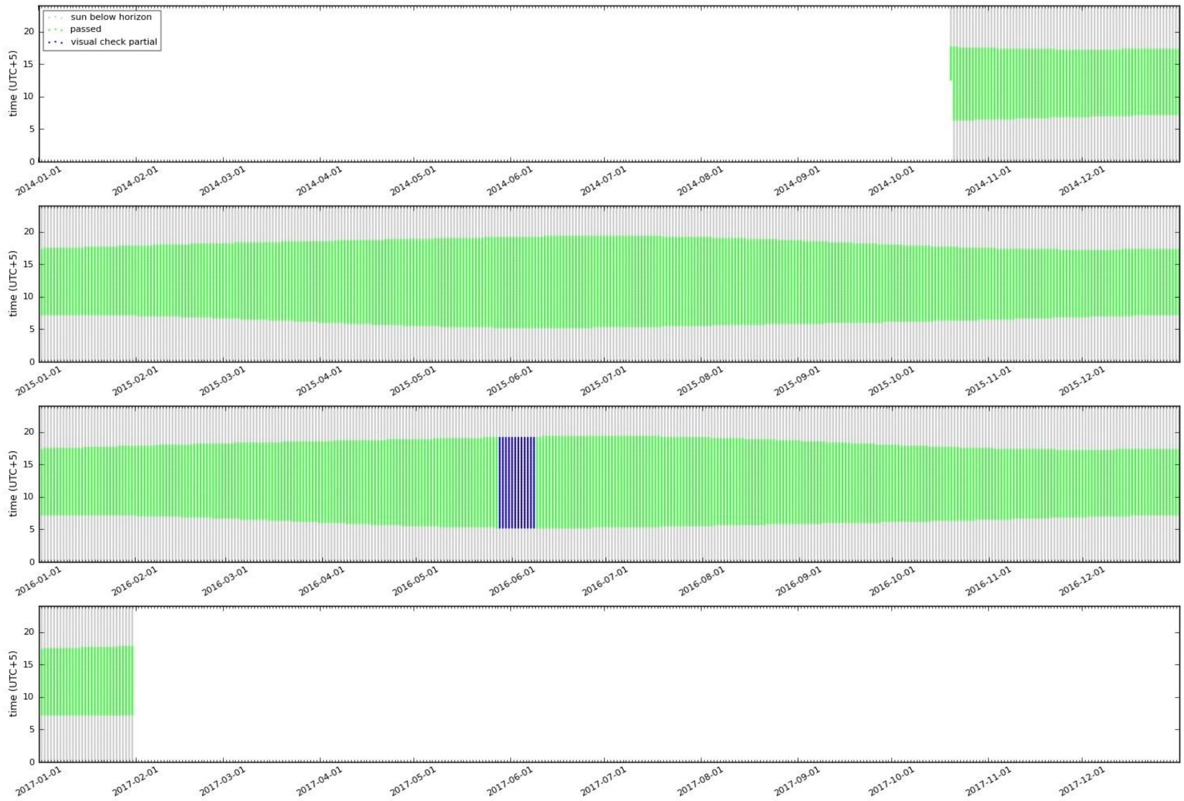


Figure XII: Quality control for DNI (Twin RSI)

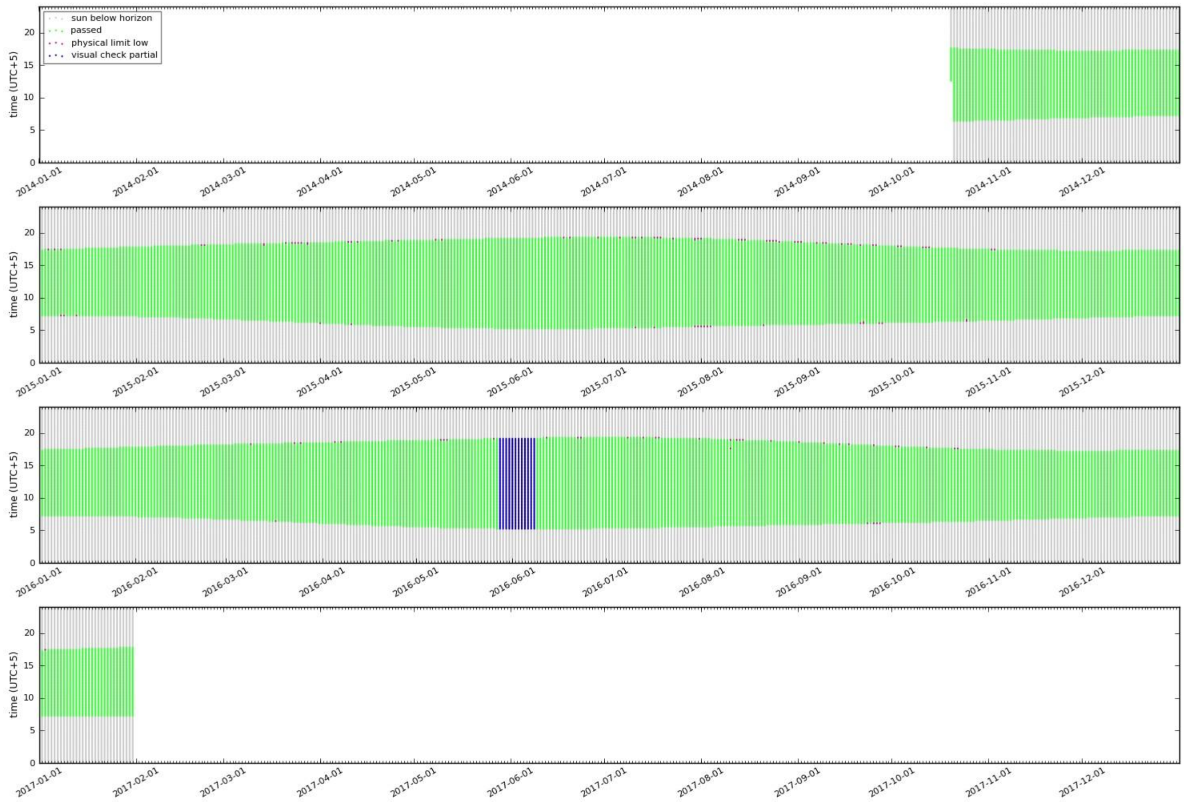


Figure XIII: Quality control for GHI (Twin RSI)

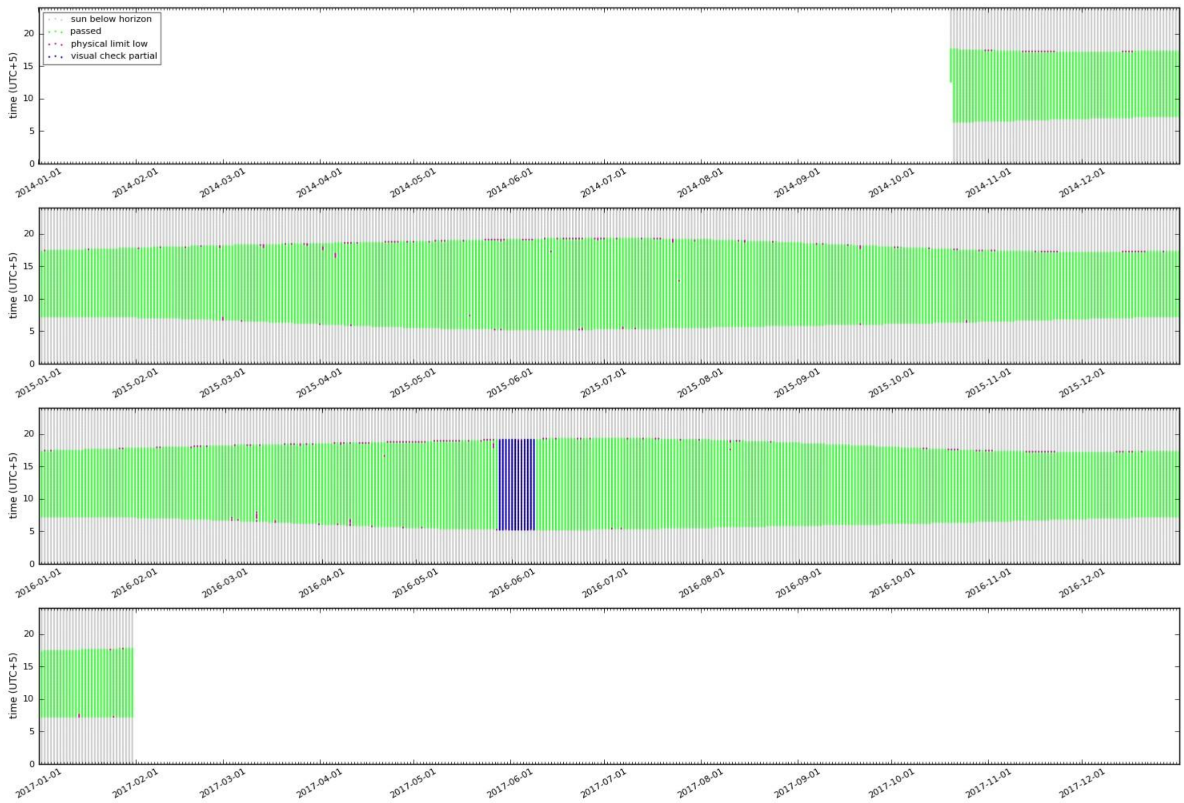


Figure XIV: Quality control for GHI (CMP10)

Bahawalpur

Table XI: Occurrence of data readings for Bahawalpur meteorologicalstation

Data availability	GHI CMP21		DNI CHP1	
Sun below horizon	56 063	50.4%	56 063	50.4%
Sun above horizon	55 182	49.6%	55 182	49.6%
Total data samples	**111 245**	**100.0%**	**111 245**	**100.0%**

Table XII: Excluded ground measurements after quality control (Sun above horizon) - Bahawalpur

Type of test	Occurrence of data samples (Sun above horizon)			
	GHI CMP21		DNI CHP1	
Physical limits test	633	1.1%	0	0.0%
Consistency test (GHI – DNI – DIF)	73	0.1%	73	0.1%
Visual test (incorrect data)	113	0.2%	113	0.2%
Other (non valid data)	131	0.2%	6 438	-
Total excluded data samples	**950**	**1.7%**	**6 624**	**0.3%**
Total samples	**55 182**	**100.0%**	**55 182**	**100.0%**

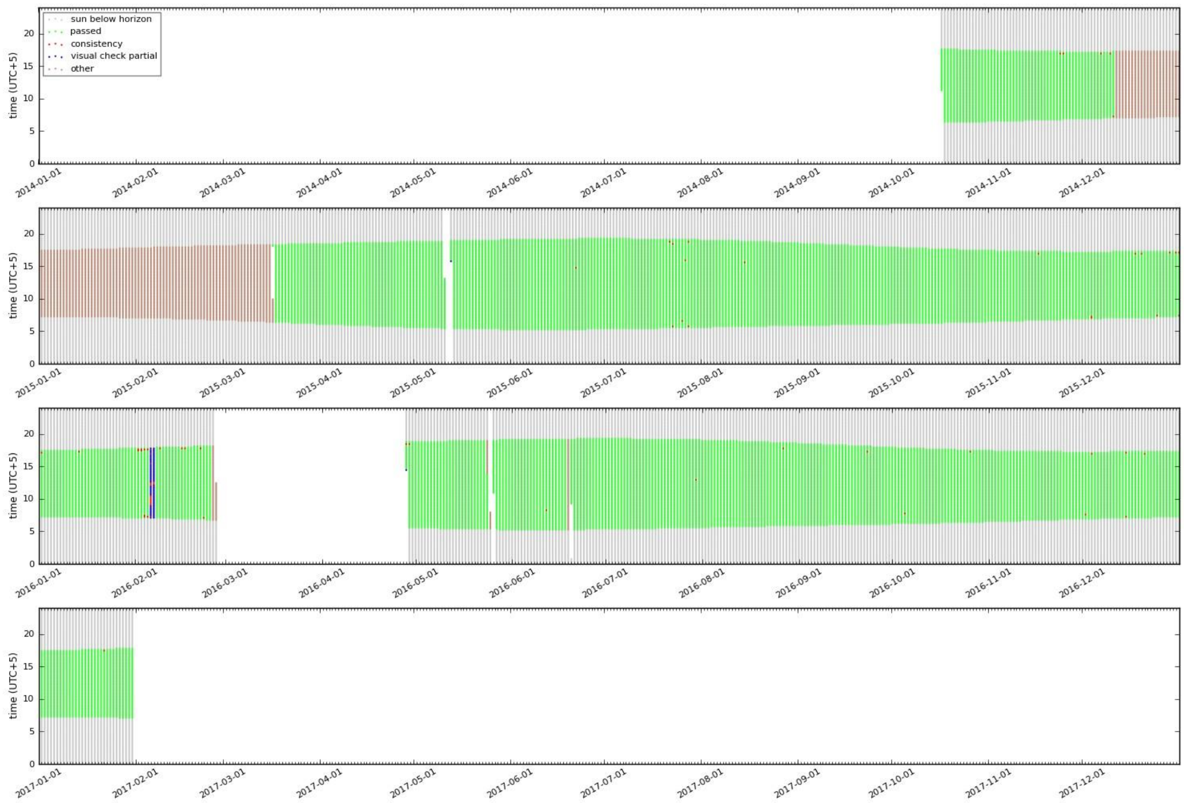


Figure XV: Quality control for DNI (CHP1)

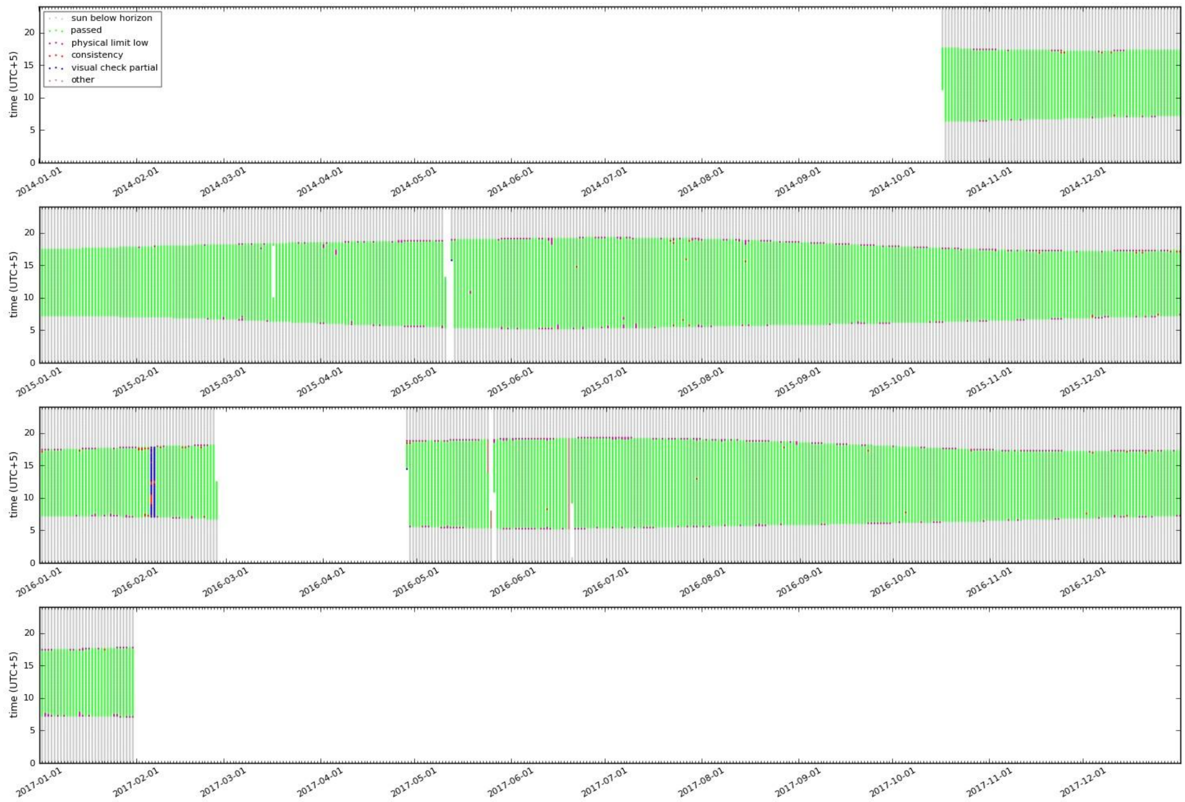


Figure XVI: Quality control for GHI (CMP21)

Khuzdar

Table XIII Occurrence of data readings for Khuzdar meteorological station

Data availability	GHI CMP10		GHI, DNI Twin_RSI	
Sun below horizon	36 374	50.8%	36 374	50.8%
Sun above horizon	35 239	49.2%	35 239	49.2%
Total data samples	**71 613**	**100.0%**	**71 613**	**100.0%**

Table XIV: Excluded ground measurements after quality control (Sun above horizon) - Khuzdar

Type of test	Occurrence of data samples (Sun above horizon)					
	GHI CMP10		DNI Twin_RSI		GHI Twin_RSI	
Physical limits test	87	0.2%	0	0.0%	8	0.0%
Consistency test (GHI – DNI – DIF)	-	-	-	-	-	-
Visual test (incorrect data)	0	0.0%	0	0.0%	0	0.0%
Other (non valid data)	0	0.0%	0	0.0%	0	0.0%
Total excluded data samples	**87**	**0.2%**	**0**	**0.0%**	**8**	**0.0%**
Total samples	**35 239**	**100.0%**	**35 239**	**100.0%**	**35 239**	**100.0%**

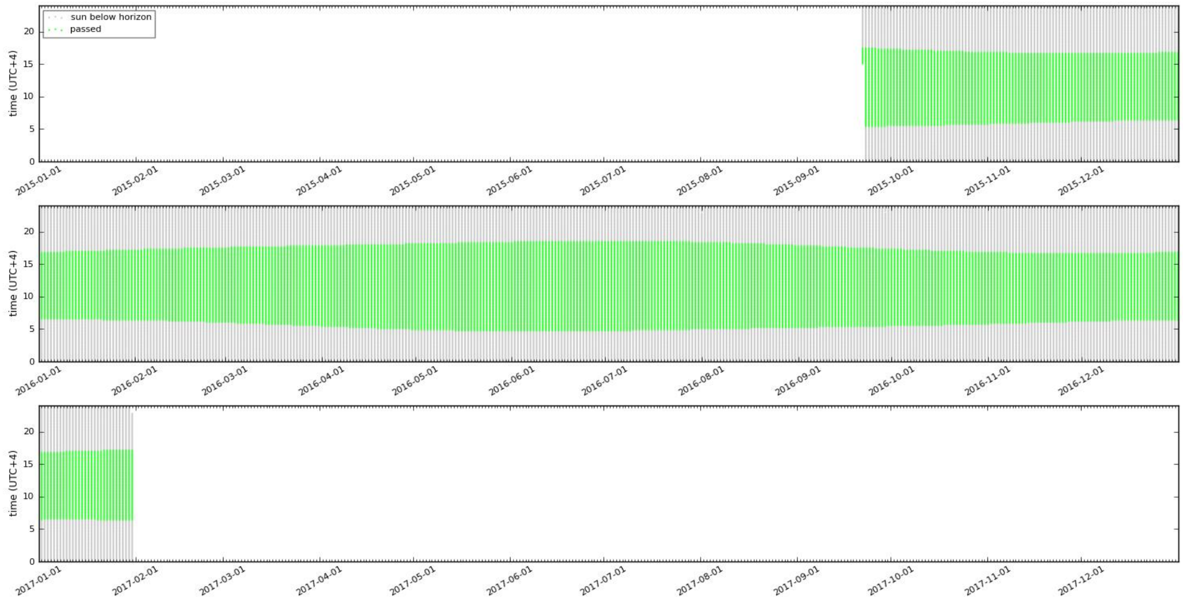


Figure XVII: Quality control for DNI (Twin RSI)

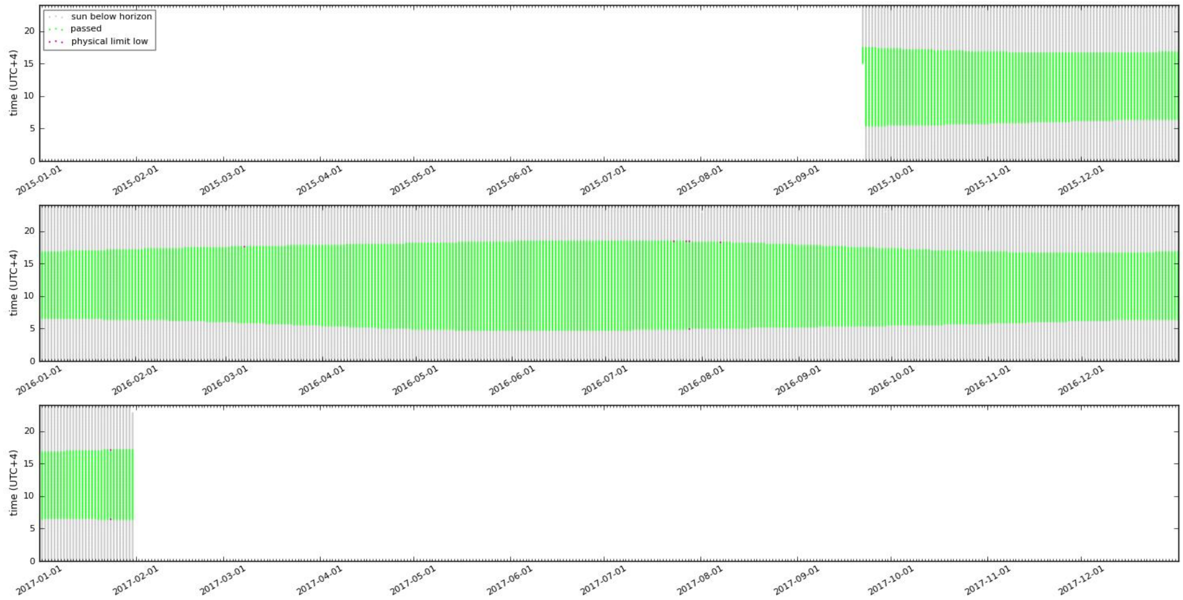


Figure XVIII: Quality control for GHI (Twin RSI)

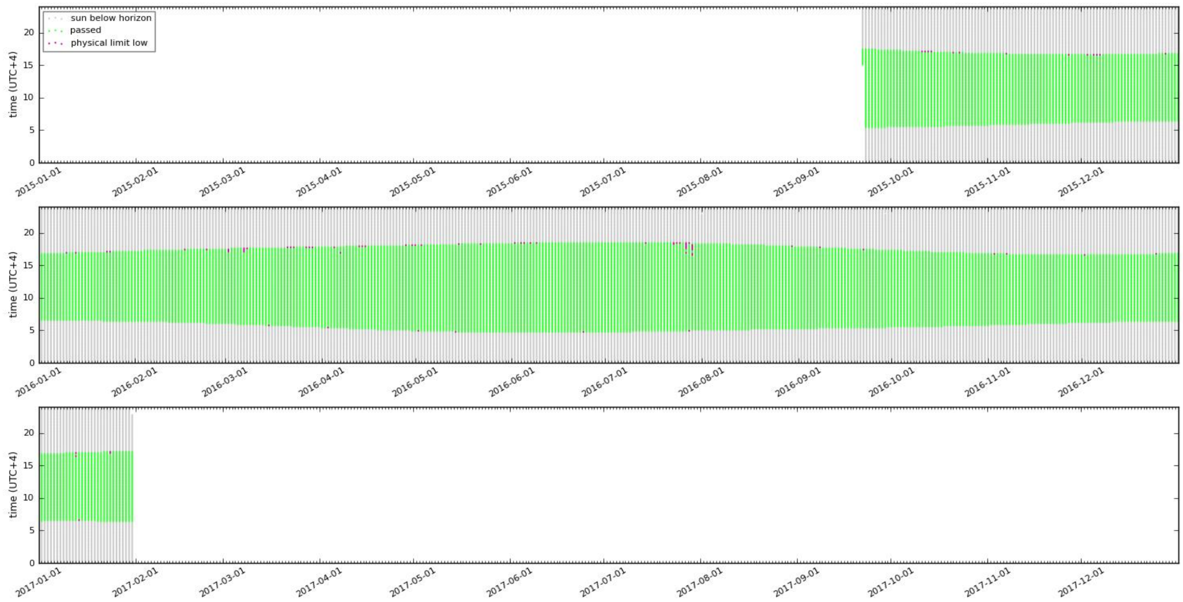


Figure XIX: Quality control for GHI (CMP10)

Hyderabad

Table XV: Occurrence of data readings for Hyderabad meteorologicalstation

Data availability	GHI CMP10		GHI, DNI Twin_RSI	
Sun below horizon	46 211	49.3%	46 211	49.3%
Sun above horizon	47 564	50.7%	47 564	50.7%
Total data samples	**93 775**	**100.0%**	**93 775**	**100.0%**

Table XVI: Excluded ground measurements after quality control (Sun above horizon) - Hyderabad

Type of test	Occurrence of data samples (Sun above horizon)					
	GHI CMP10		DNI Twin_RSI		GHI Twin_RSI	
Physical limits test	71	0.1%	0	0.0%	92	0.2%
Consistency test (GHI – DNI – DIF)	-	-	-	-	-	-
Visual test (incorrect data)	1 496	3.1%	0	0.0%	0	0.0%
Other (non valid data)	0	0.0%	0	0.0%	0	0.0%
Total excluded data samples	**1 567**	**3.3%**	**0**	**0.0%**	**92**	**0.2%**
Total samples	**47 564**	**100.0%**	**47 564**	**100.0%**	**47 564**	**100.0%**

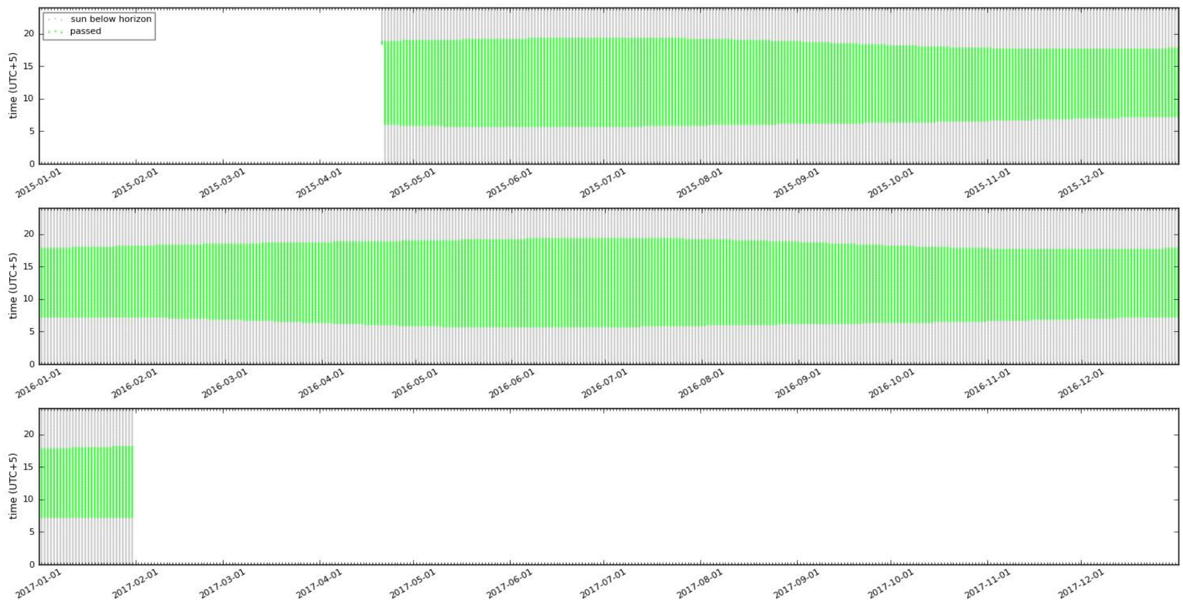


Figure XX: Quality control for DNI (Twin RSI)

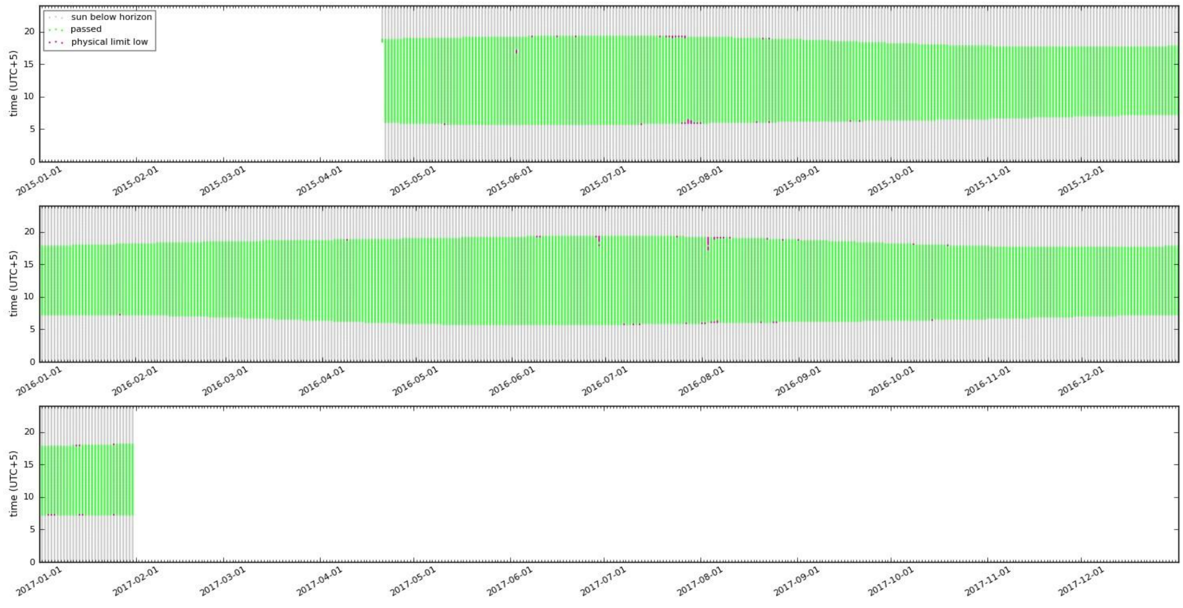


Figure XXI: Quality control for GHI (Twin RSI)

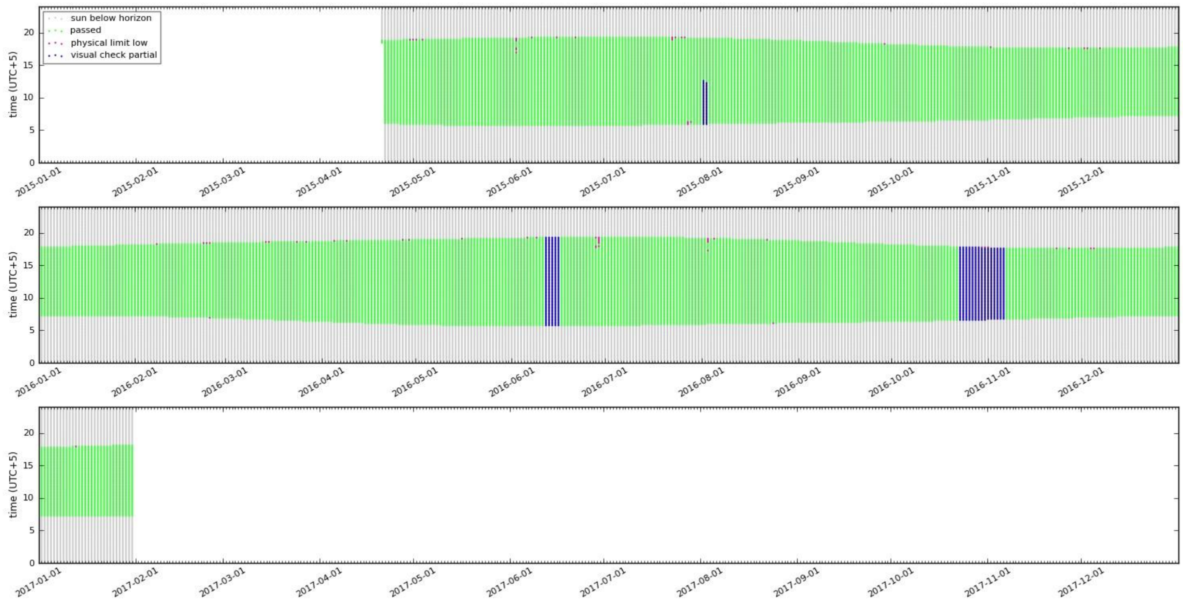


Figure XXII: Quality control for GHI (CMP10)

Karachi

Table XVII: Occurrence of data readings for Karachi meteorological station

Data availability	GHI CMP10		GHI, DNI Twin_RSI	
Sun below horizon	46 206	49.3%	46 206	49.3%
Sun above horizon	47 448	50.7%	47 448	50.7%
Total data samples	**93 654**	**100.0%**	**93 654**	**100.0%**

Table XVIII: Excluded ground measurements after quality control (Sun above horizon) - Karachi

Type of test	Occurrence of data samples (Sun above horizon)					
	GHI CMP10		DNI Twin_RSI		GHI Twin_RSI	
Physical limits test	52	0.1%	0	0.0%	85	0.2%
Consistency test (GHI – DNI – DIF)	-	-	-	-	-	-
Visual test (incorrect data)	243	0.5%	0	0.0%	0	0.0%
Other (non valid data)	0	0.0%	0	0.0%	0	0.0%
Total excluded data samples	**295**	**0.6%**	**0**	**0.0%**	**85**	**0.2%**
Total samples	**47 448**	**100.0%**	**47 448**	**100.0%**	**47 448**	**100.0%**

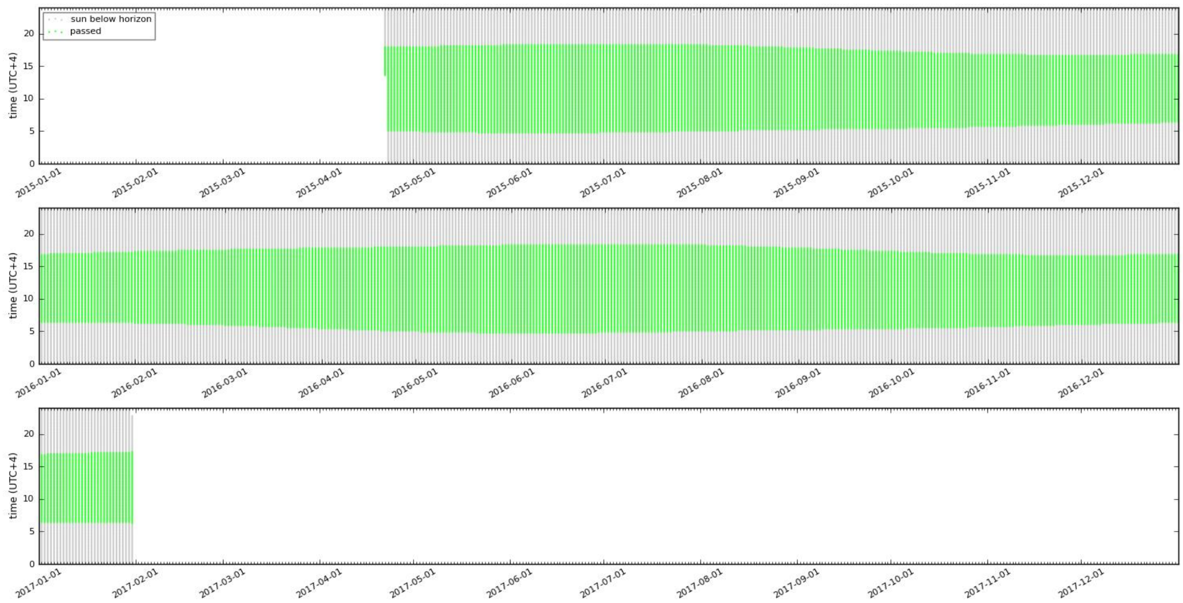


Figure XXIII: Quality control for DNI (Twin RSI)

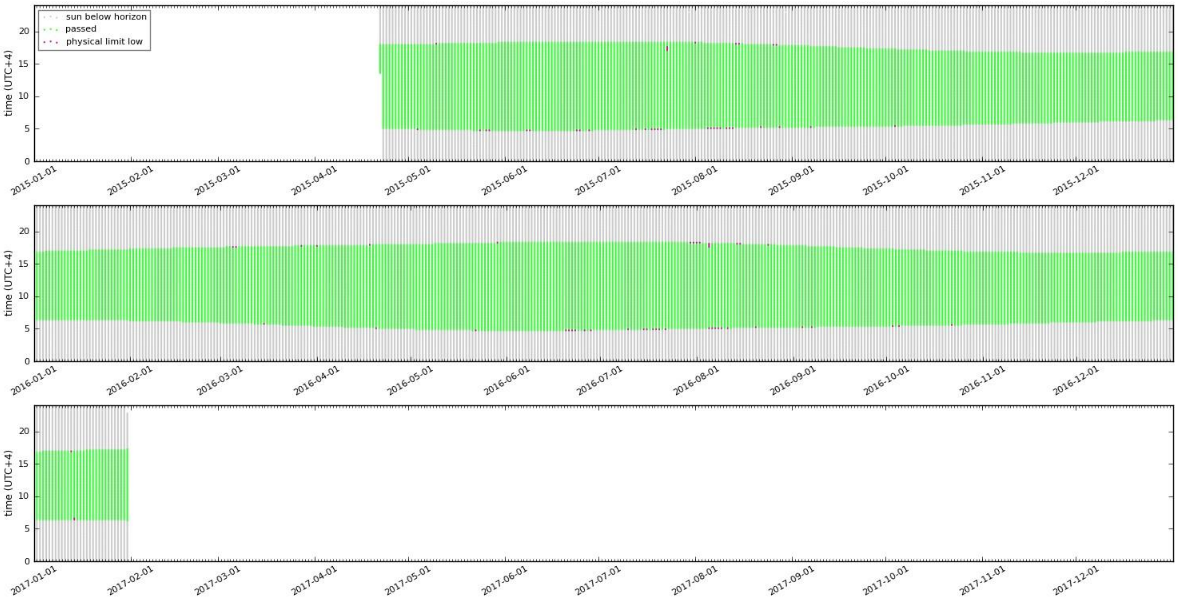


Figure XXIV: Quality control for GHI (Twin RSI)

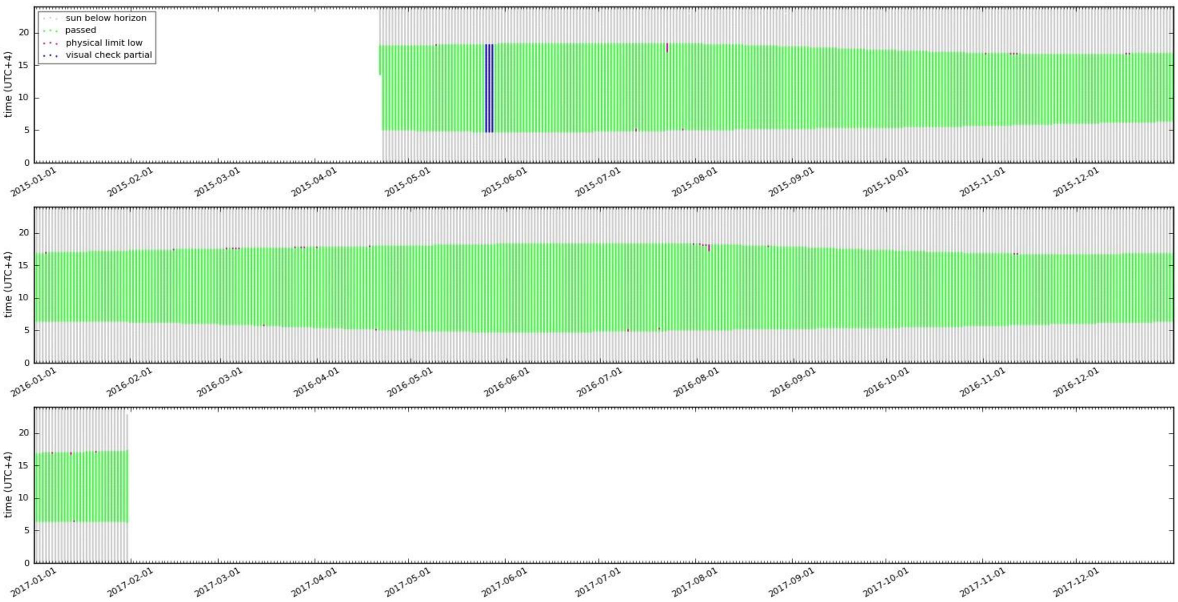


Figure XXV: Quality control for GHI (CMP10)

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

ESMAP
Energy Sector Management Assistance Program

SOLARGIS