# CAMS Solar radiation products Regular Validation Report, Issue #34, March-May 2021

Mireille Lefèvre

ECMWF COPERNICUS REPORT

Copernicus Atmosphere Monitoring Service

# D1.3.1
# Regular Validation Report
# Issue #34
# M-A-M 2021

## CAMS2-73
## Solar radiation products

Issued by: Armines / M. Lefèvre

Date: 14/01/2022

Ref: CAMS2_73_2021SC1_D1.3.1-2021Q4_RAD_validation_report_MAM2021_v1

IMPLEMENTED BY ECMWF

# Contributors

**ARMINES**
Mireille Lefèvre

# Table of Contents

# Executive summary

The CAMS Radiation Service (CRS) is continuously monitored against ground-based irradiance observations. This includes stations in various geographical locations and climates (see maps in Fig. 2.1 and description in section 4.2). Several stations being affected by large satellite viewing angles are included by intention to provide quantitative insight where the CRS may become unreliable for some applications (Tab. 2.1). From the #31 issue onwards, a new table of the ground datasets performance is presented with the amount of data rejected / retained for each station by the quality check (Tab. 2.2).

From #30 issue onwards, the protocol of validation (see Annex A.5) has been changed regarding the threshold applied for the usable measurements. For the consistency, the same new protocol has been applied to every quarter in the multi-annual plots. From #33 issue onwards, the validation of McClear (the cloud-free module of the CRS) is performed along with the CAMS-RAD module for cloudy skies. McClear can be computed for all sites at the global scale, thus stations outside the Meteosat field of view can be added. For this issue five BSRN stations (Tateno in Japan, Granite Island and Langley in United States, Syowa in Antarctica and Ny-Ålesund in Spitzbergen) are selected. More stations will be further added in order to increase the number of available datasets for the validation of McClear. Indeed, the procedure for selecting cloud-free instants in the period (see Annex A.8) requires stations to have at least 2 products *G*, *D* and/or *B* and an integration time of 1 min. These constraints eliminate many stations used for the CAMS-RAD validation. Moreover, for a station eligible, the filter on cloud-free instants may eliminate many values especially in winter time. Stations with a too small number of samples for the computation of statistical indicators have been eliminated. The table 2.1 lists the stations retained for the period March to May (MAM) 2021 with information on the availability of a validation for the modules McClear or CAMS-RAD.

For MAM 2021, the correlation is very good for global irradiances (*G*, Fig. 3.4) except for Sonnblick located at a mountain top, and excellent for the cloud-free irradiances $G_c$ for all stations. The correlation is good for diffuse irradiances (*D*, Fig. 3.5) and beam *(B*, Fig. 3.6) except for Sonnblick and Izaña both located on mountain tops. Cloud-free skies correlations $B_c$ and $D_c$ are generally lower than those for *B* and *D*.
Regarding the clearness indices, the all-skies beam $KT_B$ are at the same level as for *B* for all stations. Apart from this case and whatever the period, the clearness indices correlations are lower than those for the irradiances, and lower for the cloud-free skies compared to all-skies.

**Global irradiance**
The bias for the global irradiance is null or low for 13 stations and noticeable or large for 18 stations out of 31 (Tab. 3.1). It is a better performance than in the previous period MAM 2020 (with respectively 11 and 18 stations out of 29), considering that the worse performances in 2021

occured at stations not present in 2020, i.e., Izaña (-9.4 %), Sonnblick (-19.7 %) and Reunion Island (-8.4 %), 2 of them being at a high elevation. The highest biases are negative (Fig. 3.1) as observed in DJF 2021 (issue #33). During MAM 2020 (issue #30 with CRS v3.2) a clear trend from highest positive biases to lowest negative biases from North to South were observed. The trend is not observed anymore since DJF 2021 and this can be attributed to the upgrade to CRS v4.0. This improvement is also well observed on the frequency distributions. With the previous CRS v3.2 stations with frequent cloudy skies (low $KT_G$) overestimated *G* (positive bias), while stations with frequent clear skies (high $KT_G$) underestimated *G* (negative bias). With the CRS v4.0 such a trend is not visible anymore: the distribution of *G* is more homogeneous between stations and the discrepancy with ground measurement distributions is reduced (Fig.3.7). As a whole, the frequency distributions of all-skies *G* and cloud-free $G_c$ are good, but those for $KT_G$ present some discrepancies. The distribution of clearness index $KT_G$ in a large range [0.1, 0.7] is slightly over-represented at the expense of lowest $KT_G < 0.1$ and mainly highest $KT_G > 0.7$ (Fig. 3.8) leading to an underestimation for most stations. Regarding McClear this distribution is good except as expected for stations Izaña and Sonnblick at high elevations, and Ny-Ålesund at a high latitude.
The multi-annual biases (Fig. 3.16b) highlight a better overall performance for the 2017-2020 MAM quarters than in previous issue #30 with CRS v3.2. No large discrepancy is observed for the MAM quarters over 5 years, except for enerMENA stations which exhibit recurrent problem of maintenance (see section 4.4).
It should be noted that mountain top stations (Sonnblick and Izaña) observe a better McClear performance (1.7 % and -0.9 % respectively) with the IFS cycle CY47r1 although all-skies performance remains bad (-20 % and -9 % respectively). McClear has good overall performance except at extreme latitudes (Syowa in Antarctica and Ny-Ålesund in Spitzbergen) whatever the year, and at enerMENA stations.
The monthly means of the estimated *G* (Fig. 3.13) are identical or close to those of the measurements at all stations. The monthly standard deviations that highlight the dynamics of the dataset are also close to those of measurements for both McClear and CAMS-RAD.

**Diffuse irradiance**
The bias for the diffuse irradiance (Fig. 3.2) is null or low for 11 stations out of 18, and noticeable or large for 7 stations. This is a similar result as previous period MAM 2020 (CRS v3.2) with respectively 9 and 6 out of 15. The all-skies multi-annual biases for CRS v4.0 (Fig. 3.17b) show an overall slight shift towards lower or negative values compared to CRS v3.2 (Fig 3.17b in issue #30), thus an improvement of the cloud detection in the new CRS. Highest biases are negative. Only Sonnblick at a high elevation presents a very large and negative bias.
The trend from negative bias for stations with frequent cloudy skies, to positive bias for stations with frequent clear skies, observed with CRSv3.2 is not observed anymore. Actually, the frequency distributions of *D* (fig. 3.9) shows a good adequation between measurements and model outputs.
The monthly means of the estimated *D* (Fig. 3.14) display some slight discrepancies with a slightly weaker dynamic compared to measurements. The distribution of $KT_D$ (Fig. 3.10) follows a different

pattern than *D*: medium $KT_D$ are always over-represented, at the expense of higher $KT_D$ and lower $KT_D$. That illustrates a lack of dynamic for the $KT_D$. contrary to *D*.

Regarding McClear (Fig. 3.2, Fig. 3.14) all biases are positive (overestimation with a higher dynamic compared to measurements), except top mountain stations Izaña and Sonnblick (underestimation with a lower dynamic compared to measurements), and high latitude stations Syowa and Ny-Ålesund (underestimation with a good representativity of the dynamic). Apart these "extreme" areas, McClear overestimates the diffuse component. It should be noted that the larger biases concern enerMENA stations. The frequency distribution (Fig. 3.10) of $KT_{Dc}$ is bad for enerMENA stations, while the distribution is correct for most of the BSRN stations. It is thus difficult to attribute the bias to McClear or to enerMENA stations. More stations, and validation of other seasons are needed to conclude on this point.
Contrary to *D*, results are not satisfactory for $D_c$ whatever the year (Fig. 3.17b). As the diffuse component is a minor contribution to the global clear-skies irradiance $G_c$, that does not highly impact $G_c$.

**Direct normal irradiance**
The trends observed for *B* and $B_c$ are opposite to those observed for *D* and $D_c$. The bias for *B* (Fig. 3.3) is mostly positive (vs negative for *D*), also illustrated with the frequency distributions (Fig. 3.11). An over-representation of medium range [100-900] W $m^{-2}$ is obtained at the expense of lower values or higher values with an overall overestimation and loss of dynamic. The *B* multi-annual biases (Fig. 3.18b) for CRS v4.0 compared to CRS v3.2 (Fig.3.18b in issue #30) show an overall shift towards positive values leading to reduced absolute values.
Contrary to all-skies, clear-skies $B_c$ are underestimated (vs overestimated for $D_c$) except for top mountain stations of Sonnblick and Izaña. No great discrepancy over the five years for McClear can be observed for the MAM quarter. The monthly means of the estimated *B* and $B_c$ (Fig. 3.15) present slight discrepancies at most stations, with a dynamic close to that of measurements for both CAMS-RAD and McClear.

Overall, CAMS-RAD exhibits for this MAM quarter better performance with the CRS v4.0 (2021) compared to CRS v3.2 for the 3 components. Results are satisfactory except for stations at high elevations, i.e., Sonnblick (3109 m), Izaña (2373 m), and for stations at high latitudes, i.e., Syowa (69° S) and Ny-Ålesund (78.9° N). Besides these "extreme" area, bad results persist at the enerMENA stations which exhibit recurrent problems of maintenance but also offer insight into aerosol-affected regions with sparse coverage from other networks. Another main result concerns the bias trends from North to South, or from stations with frequent cloudy skies to stations with frequent cloud-free skies. These trends were not observed in DJF 2021 and MAM 2021 confirms this result, due to an improvement of the cloud detection in the new CRS v4.0. This is in accordance with a better representation of the frequency distributions with CRS v4.0 for *G* and *B* irradiances for all-skies and for cloud-free skies, although this is not the case for *D* and $D_c$.

Regarding McClear, performance for $B_c$ are much better than for diffuse $D_c$. The opposite trends in diffuse and direct components combine each other to finally exhibit good results for the global irradiance.

# 1 Introduction

The CAMS Radiation Service, abbreviated as CRS, delivers estimates of the solar radiation arriving at ground level on a horizontal surface. Outcomes of the CRS have to be validated on a periodic basis. Following practices in CAMS, this validation is performed every quarter. Following current practices, the CRS irradiances are tested against qualified ground measurements measured at several ground-based stations serving as reference. These ground measurements are coincident in time with the CRS estimates.

On 28 June 2021, version 4.0 of the CRS was introduced with a new version APOLLO Next Generation (APOLLO_NG) provided by the DLR that improves cloud detection and cloud physical parameters; includes an improved handling of circumsolar irradiances; and includes the most recent CAMS IFS cycle update for cloud-free parameters aerosols, water vapour and ozone. For consistency, this new version 4.0 is used from now on back to 2016 for the multi-annual assessments.

This report is issue #34 of a regular report performed with version 4.0 of the CRS. It deals with hourly means of global, diffuse, and direct irradiances for the period March to May 2021, abbreviated as MAM 2021.

In #24 of this report a multi-annual assessment was introduced for each station as a 3-years times-series in Annex C. This assessment is provided once a year for the yyyyQ3 report (DJF datasets) when in-situ measurements are completed for the year. This has made the quarterly validation reports very large. Therefore, from issue #25 onwards detailed results are partly provided in a zip archive supplement as individual files in a well-organized folder structure.

From #27 (2019Q3 report) onwards, multi-annual statistical indicators of the respective season are given in section 3.

From #31 (2021Q1 report) onwards, a summary of the quality check operated on each station is given in section 2.

From #33 (2021Q3 report), the validation is performed on both cloud-free and all-sky components of the CRS, namely McClear and CAMS-RAD modules (see https://atmosphere.copernicus.eu/solar-radiation and the User Guide published there). Contrary

to CAMS-RAD which is available at sites inside the Meteosat field of view only, McClear radiation can be computed at the global scale.

Figure 1.1 summarizes the evolution of the ground observation availability since 2014. All validation reports are online at https://atmosphere.copernicus.eu/supplementary-services.

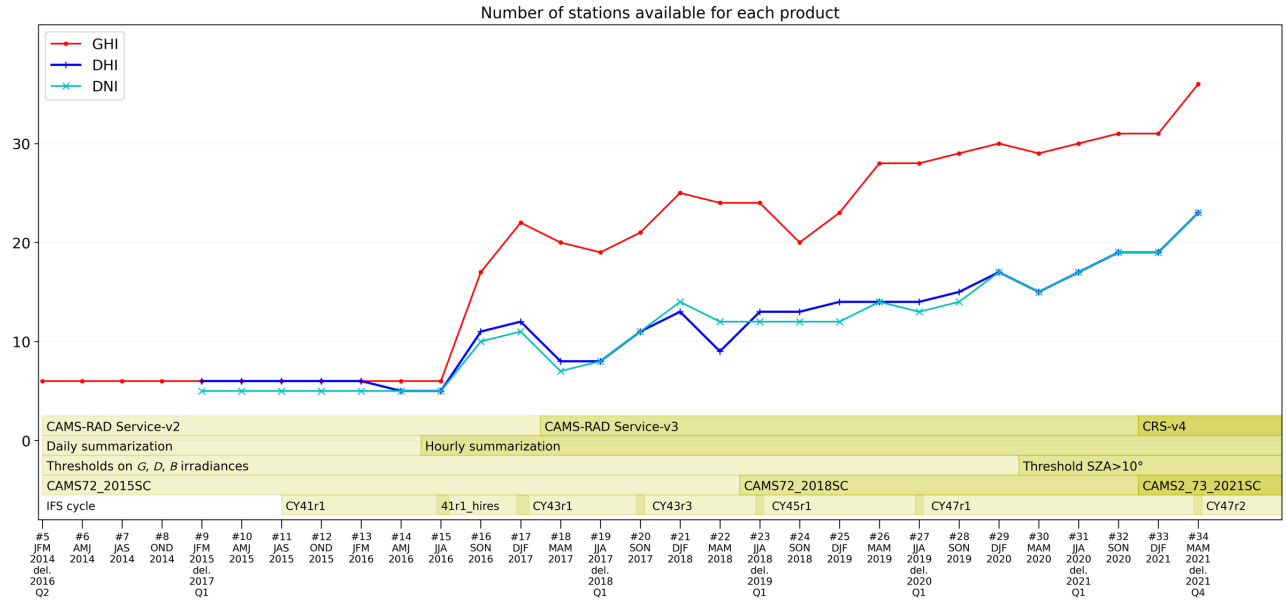



Figure 1.1. Evolution of the number of stations used for each product from 2014.
On the x-axis are the validation reports identified by their issue number,
the CAMS Radiation Service version as used in the previously published validation reports,
along with the months and years corresponding to the measurements.

# 2 Stations

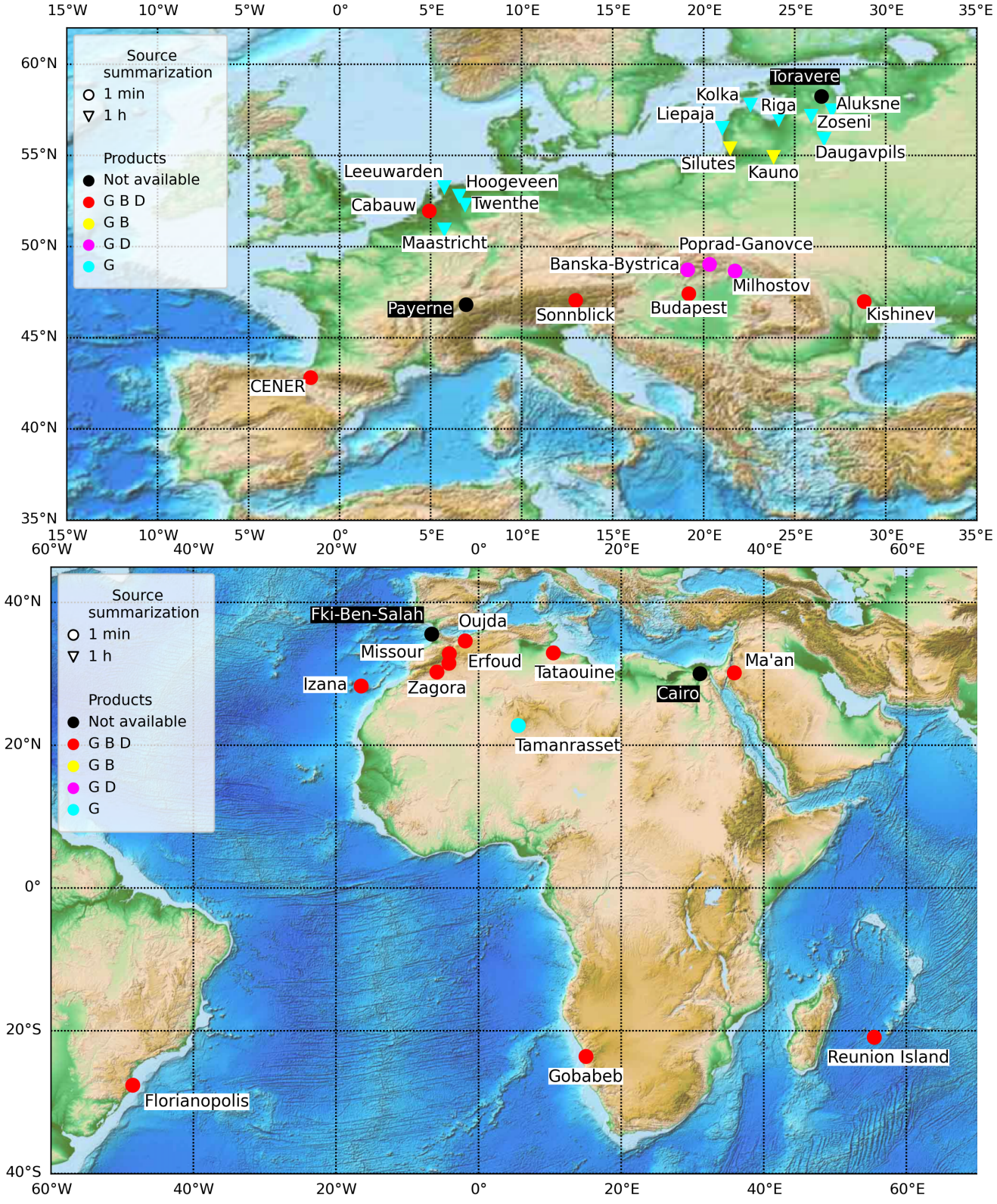
The source of ground-based measurements is given in section 4, along with a description of each station, its provider, and the products provided: *G*, *B*, *D* are global, beam and diffuse irradiances respectively, and $KT_G$, $KT_B$, $KT_D$ are global, beam and diffuse clearness indices, respectively. These parameters are defined in Annex A.7.

## 2.1 Maps

Figure 2.1 shows the location of the stations as reported in Table 4.2. Symbols code the initial integration time of the data as reported in Table 4.1: circle for 1 min, and downward triangle for 1 h. Colours code the type of data at each site: red for (*G, B, D*), yellow for (*G, B*), magenta for (*G, D*) and cyan for *G*.

In the #27 report, some stations have been removed from the set of usual stations, and some others have been added (see Fig. 2.3 in #27). Stations too close to the sea like Hoorn (see 4.1.1 in #27 for the report of the problem) and Vlissingen in the Netherlands have been removed and replaced by Leeuwarden and Maastricht (Fig. 2.3). Stations have been closed by the providers like Carpentras and Dobele in December 2018, Rucava in September 2018, or are not updated regularly like PSA (year 2017 was provisioned in December 2018, years 2018 and 2019 was provisioned in April 2020). New stations offset these removals like CENER in Spain, or Aluksne and Kolka, two stations provided monthly by the Latvian network instead of Dobele and Rucava. Since November 2019, the enerMENA network provides a new station Fki Ben Salah in Morocco. In the #31 report, the KNMI station of Cabauw has been replaced with the same station provided by the BSRN network that offers the 3 components *G*, *D* and *B* at 1 min of resolution. The BSRN station of Izaña has been added to illustrate restricts for validation at a mountain top with restricted representativity for its surroundings. In the #32 report, the new BSRN stations of Budapest and Reunion Island have been added.

In this #34 issue, the BSRN stations Granite Island and Langley in United States, Syowa in Antarctica, and Ny-Ålesund in Spitzbergen have been added for the validation of McClear outside the satellite field of view. They come along with station Tateno added in the #33 issue. Such stations will be further integrated.


Source summarization
1 min
1 h
Products
Not available
G B D
G B
G D
G
Toravere
Kolka
Riga
Aluksne
Liepaja
Zoseni
Daugavpils
Silutes
Kauno
Leeuwarden
Hoogeveen
Cabauw
Twenthe
Maastricht
Poprad-Ganovce
Banska-Bystrica
Milhostov
Payerne
Sonnblick
Budapest
Kishinev
CENER
Fki-Ben-Salah
Oujda
Missour
Erfoud
Tataouine
Ma'an
Izana
Zagora
Cairo
Tamanrasset
Gobabeb
Reunion Island
Florianopolis

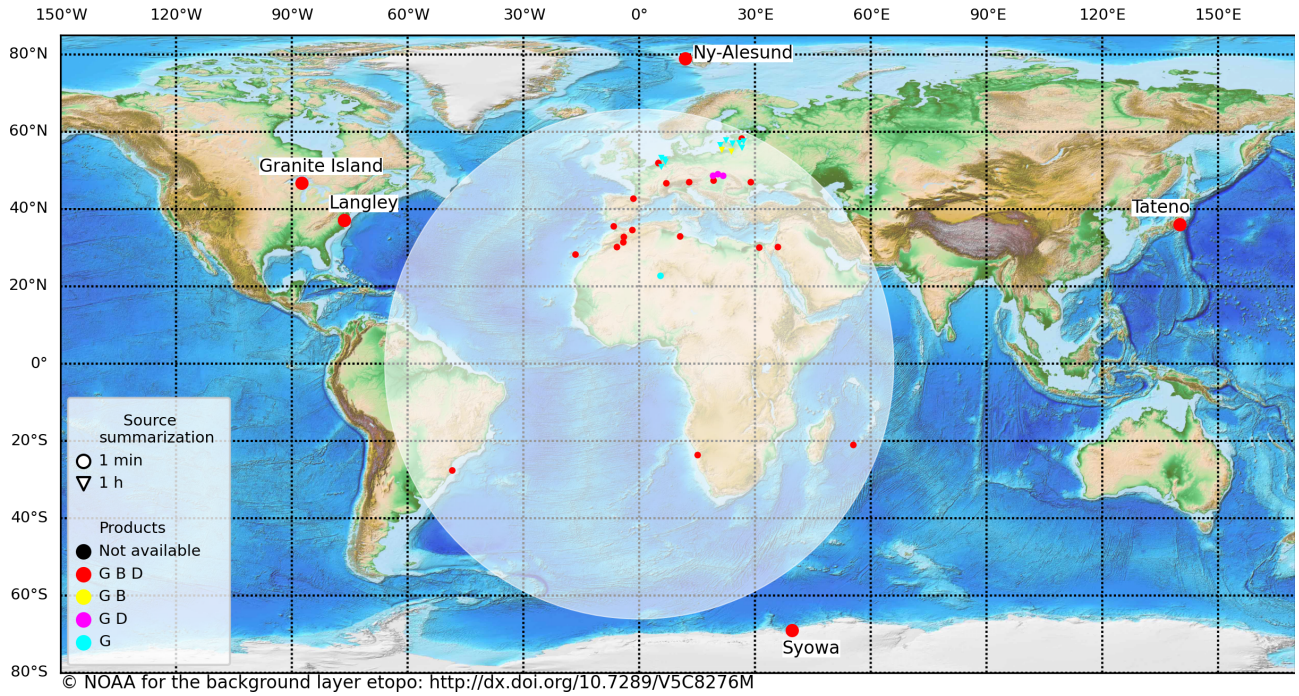



Figure 2.1. Maps showing the 35 stations inside the satellite field of view: 22 stations in Europe (top), 13 stations outside Europe (middle) and 5 stations outside the satellite field of view (bottom). Symbols code the initial integration time: circle for 1 min, and downward triangle for 1 h. Colours code the type of data at each site: red for (*G, B, D*), yellow for (*G, B*), magenta for (*G, D*) and cyan for *G*.

## 2.2 List of the stations retained for this quarter

Depending on the provision of fresh data, possible problems affecting measuring instruments, possible rejection of some data by the quality control, and other causes, it is not always possible to use the same set of stations to perform the quarterly validation. Table 2.1 lists the 36 stations that have been retained for this quarter, along with the 4 stations having no available data for this quarter. More details on the stations and their providers are presented in Table 4.1 in the section 4 “Sources of data”.

Due to low satellite viewing angles, lower CRS quality is expected for stations in Lettland and Lithuania (Aluksne, Daugavpils, Kauno, Kolka, Liepaja, Riga, Silutes, Toravere) and Brazil (Florianopolis). Nevertheless, stations in Lettland and Lithuania provide some insight on the area where the CRS may become unreliable and therefore, this information is kept in this user-oriented report. The same reason motivates the addition of Izaña and Sonnblick at mountain tops and not representative for their surroundings.

A summary of the quality check operated on each station is given in Table 2.2 (section 2.3).

Table 2.1. List of stations retained or excluded for this regular validation report, along with the number of hourly samples for global (*G*), diffuse (*D*) and beam (*B*) components remaining after the quality check process, the solar zenith angle threshold, and the selection of cloud-free instants for McClear.

| Colour status: | Dataset eliminated with Nb of obs. <=30 | Data source missing: --- | Comparison *unavailable*:<br>CAMS-RAD: station out of MSG field of view<br>McClear: D or B missing or source integration time is > 1 min |
|---|---|---|---|

| **Station** | **Provider (Table 4.1)** | **Nb of 1 h samples** | | **Issue for the period** |
|---|---|---|---|---|
| | | **all sky** | **cloud free** | |
| Aluksne | LEGMC | G:629 D: --- B: --- | *Unavailable* | Reduced service accuracy expected due to low satellite viewing angle. |
| Banska-Bystrica | SHMI | G:1046 D:1046 B: --- | G+D: 55 | |
| Budapest | BSRN | G:1044 D:1044 B:1045 | G+D+B: 83 | |
| Cabauw | BSRN | G:1005 D:1005 B:1006 | G+D+B: 31 | |
| *Cairo* | *enerMENA* | *G: --- D: --- B: ---* | *G+D+B: ---* | tracker broken --> Only GHI data accurate. **Station doubtfully maintained, eliminated by the quality check (see § 4.4.1).** |
| CENER | BSRN | G:1022 D:1022 B:1026 | G+D+B: 81 | |
| Daugavpils | LEGMC | G:632 D: --- B: --- | *Unavailable* | Reduced service accuracy expected due to low satellite viewing angle. |
| Erfoud | enerMENA | G:1036 D:1036 B:1036 | G+D+B: 182 | Good quality |
| *Fki Ben Salah* | *enerMENA* | *G: --- D: --- B: ---* | *G+D+B: ---* | *G* data not available. *D* data good quality. DNI pyrheliometer sometimes misaligned in the morning. Pb sensor blocked. Maintenance visit and installation of new wind vane in Feb. **Station eliminated by the quality check in the absence of *G* (see § 4.4.5).** |
| Florianopolis | BSRN | G:333 D:333 B:333 | *G+D+B: 2* | Reduced service accuracy expected due to low satellite viewing angle. **No data available since April 2021.** |
| Gobabeb | BSRN | G:922 D:922 B:922 | G+D+B: 356 | |
| Granite Island | BSRN | *Unavailable* | G+D+B: 117 | **New station outside satellite field of view.** |
| Hoogeveen | KNMI | G:1042 D: --- B: --- | *Unavailable* | |
| Izaña | BSRN | G:1031 D:1031 B:1033 | G+D+B: 223 | Reduced service accuracy expected due to altitude 2372 m and its location at a mountain top. |

| Station | Provider (Table 4.1) | Nb of 1 h samples | | Issue for the period |
|---|---|---|---|---|
| | | all sky | cloud free | |
| Kauno | LHMS | G:1048 D: --- B:1048 | *Unavailable* | Reduced service accuracy expected due to low satellite viewing angle. |
| Kishinev | WRDC | G:1028 D:1028 B:1028 | *Unavailable* | |
| Kolka | LEGMC | G:629 D: --- B: --- | *Unavailable* | Reduced service accuracy expected due to low satellite viewing angle. |
| Langley | BSRN | *Unavailable* | G+D+B: 144 | **New station outside satellite field of view.** |
| Leeuwarden | KNMI | G:1041 D: --- B: --- | *Unavailable* | |
| Liepaja | LEGMC | G:632 D: --- B: --- | *Unavailable* | Reduced service accuracy expected due to low satellite viewing angle. |
| Ma'an | enerMENA | G:988 D:988 B:995 | G+D+B: 326 | Quality ok. Tracker unaligned in the morning. |
| Maastricht | KNMI | G:1046 D: --- B: --- | *Unavailable* | |
| Milhostov | SHMI | G:1027 D:1027 B: --- | G+D: 67 | |
| Missour | enerMENA | G:983 D:983 B:989 | G+D+B: 148 | Good quality. Week-daily cleaning. Tracker off in April until 14th. |
| Ny-Ålesund | BSRN | *Unavailable* | G+D+B: 93 | **New station outside satellite field of view.** |
| Oujda | enerMENA | G:446 D:446 B:456 | G+D+B: 60 | Good quality in March. Tracker off in April and May: only GHI available. Cleaning every ~3 days. |
| *Payerne* | *BSRN* | *G: --- D: --- B: ---* | *: ---* | No data available from February 2020 onwards. |
| Poprad-Ganovce | SHMI | G:1047 D:1047 B:1048 | G+D+B: 51 | |
| Reunion Island | BSRN | G:934 D:934 B:935 | G+D+B: 52 | |
| Riga | LEGMC | G:629 D: --- B: --- | *Unavailable* | Reduced service accuracy expected due to low satellite viewing angle. |
| Silutes | LHMS | G:1041 D: --- B:1041 | *Unavailable* | Reduced service accuracy expected due to low satellite viewing angle. |
| Sonnblick | BSRN | G:1019 D:1019 B:1024 | G+D+B: 91 | Reduced service accuracy expected due to altitude 3109 m and its location at a mountain top. |
| Syowa | BSRN | *Unavailable* | G+D+B: 40 | **New station outside satellite field of view.** |
| Tamanrasset | BSRN | G:995 D: --- B: --- | *Unavailable* | Only *G* is available. *D* and *B* unavailable from March 2018 onwards. |
| Tataouine | enerMENA | G:1033 D:1033 B:1033 | G+D+B: 128 | Good quality, cleaning every ~3days. |
| Tateno | BSRN | *Unavailable* | G+D+B: 57 | |

| Station | Provider (Table 4.1) | Nb of 1 h samples | | Issue for the period |
|---|---|---|---|---|
| | | all sky | cloud free | |
| *Toravere* | *BSRN* | *G: --- D: --- B: ---* | *: ---* | **No data available from June 2020 onwards.** Reduced service accuracy expected due to low satellite viewing angle. |
| Twenthe | KNMI | G:1044 D: --- B: --- | *Unavailable* | |
| Zagora | enerMENA | G:333 D:333 B:333 | G+D+B: 71 | Good quality. |
| Zoseni | LEGMC | G:629 D: --- B: --- | *Unavailable* | Reduced service accuracy expected due to low satellite viewing angle. |

## 2.3 Quality check

Table 2.2 summarizes for each station:

- The number of usable values in the ground dataset expressed as a percentage of the number of daylight instants in the MAM period, i.e., the number of instants corresponding to the source resolution (1 h or 1 min) for a solar zenith angle SZA <= 89°.
- The number of values rejected by the quality check (QC) expressed as a percentage of the usable values in the ground dataset.
- The number of retained values in the ground dataset after the quality check, expressed as a percentage of the number of daylight instants in the MAM 2021 period. That represents the dimension of the ground dataset before the integration time to hourly resolution and the application of the SZA threshold.

Table 2.2. Quantities of ground measurements rejected or retained for each station in the MAM 2021 period.

| Station | Variable | Daylight instants during MAM 2021 | Observation >= 0 W $m^{-2}$ during daylight before QC | | QC Test | Data rejected by the QC during daylight | | Observation >= 0 W $m^{-2}$ during daylight after QC and before SZA threshold | |
|---|---|---|---|---|---|---|---|---|---|
| | | N | N | % | | N | % | N | % |
| Aluksne | G | 1261 | 766 | 60.7 | BSRN-1C | 0 | 0.000 | 766 | 60.7 |
| Banska-Bystrica | G | 73139 | 73130 | 100.0 | BSRN-2C | 2 | 0.003 | 73119 | 100.0 |

| Station | Variable | Daylight instants during MAM 2021 | Observation >= 0 W m$^{-2}$ during daylight before QC | | QC Test | Data rejected by the QC during daylight | | Observation >= 0 W m$^{-2}$ during daylight after QC and before SZA threshold | |
|---|---|---|---|---|---|---|---|---|---|
| | | N | N | % | | N | % | N | % |
| | D | | 73127 | 100.0 | | 6 | 0.008 | 73119 | 100.0 |
| Budapest | G | 72796 | 72511 | 99.6 | BSRN-3C | 633 | 0.873 | 70131 | 96.3 |
| | B | | 70517 | 96.9 | | 171 | 0.242 | 70131 | 96.3 |
| | D | | 72545 | 99.7 | | 633 | 0.873 | 70131 | 96.3 |
| CENER | G | 71711 | 71678 | 100.0 | BSRN-3C | 88 | 0.123 | 68686 | 95.8 |
| | B | | 68783 | 95.9 | | 88 | 0.128 | 68686 | 95.8 |
| | D | | 71680 | 100.0 | | 88 | 0.123 | 68686 | 95.8 |
| Cabauw | G | 74100 | 73941 | 99.8 | BSRN-3C | 63 | 0.085 | 70136 | 94.7 |
| | B | | 71006 | 95.8 | | 51 | 0.072 | 70136 | 94.7 |
| | D | | 71132 | 96.0 | | 71 | 0.100 | 70136 | 94.7 |
| Cairo | | | | | | | | | |
| Daugavpils | G | 1256 | 762 | 60.7 | BSRN-1C | 0 | 0.000 | 762 | 60.7 |
| Erfoud | G | 69625 | 69624 | 100.0 | BSRN-3C | 0 | 0.000 | 69598 | 100.0 |
| | B | | 69624 | 100.0 | | 0 | 0.000 | 69598 | 100.0 |
| | D | | 69624 | 100.0 | | 26 | 0.037 | 69598 | 100.0 |
| Fki Ben Salah | | | | | | | | | |
| Florianopolis | G | 61699 | 22010 | 35.7 | BSRN-3C | 0 | 0.000 | 21684 | 35.1 |
| | B | | 21760 | 35.3 | | 0 | 0.000 | 21684 | 35.1 |
| | D | | 21963 | 35.6 | | 29 | 0.132 | 21684 | 35.1 |
| Gobabeb | G | 62368 | 61951 | 99.3 | BSRN-3C | 16 | 0.026 | 61831 | 99.1 |
| | B | | 62269 | 99.8 | | 12 | 0.019 | 61831 | 99.1 |
| | D | | 62339 | 100.0 | | 16 | 0.026 | 61831 | 99.1 |
| Granite Island | G | 72607 | 68611 | 94.5 | BSRN-3C | 129 | 0.188 | 68122 | 93.8 |
| | B | | 68586 | 94.5 | | 120 | 0.175 | 68122 | 93.8 |
| | D | | 71942 | 99.1 | | 121 | 0.168 | 68122 | 93.8 |
| Hoogeveen | G | 1231 | 1231 | 100.0 | BSRN-1C | 0 | 0.000 | 1231 | 100.0 |
| Izana | G | 69137 | 67409 | 97.5 | BSRN-3C | 27 | 0.040 | 67070 | 97.0 |
| | B | | 67715 | 97.9 | | 26 | 0.038 | 67070 | 97.0 |
| | D | | 67423 | 97.5 | | 27 | 0.040 | 67070 | 97.0 |
| Kauno | G | 1252 | 1252 | 100.0 | BSRN-1C | 1 | 0.080 | 1251 | 99.9 |
| | B | | 1252 | 100.0 | | 16 | 1.278 | 1236 | 98.7 |
| Kishinev | G | 1209 | 1208 | 99.9 | BSRN-3C | 17 | 1.407 | 1188 | 98.3 |
| | B | | 1208 | 99.9 | | 36 | 2.980 | 1172 | 96.9 |

| Station | Variable | Daylight instants during MAM 2021 | Observation >= 0 W m$^{-2}$ during daylight before QC | | QC Test | Data rejected by the QC during daylight | | Observation >= 0 W m$^{-2}$ during daylight after QC and before SZA threshold | |
|---|---|---|---|---|---|---|---|---|---|
| | | N | N | % | | N | % | N | % |
| | D | | 1208 | 99.9 | | 19 | 1.573 | 1188 | 98.3 |
| Kolka | G | 1264 | 767 | 60.7 | BSRN-1C | 0 | 0.000 | 767 | 60.7 |
| Langley | G | 70614 | 70094 | 99.3 | BSRN-3C | 150 | 0.214 | 66298 | 93.9 |
| | B | | 66748 | 94.5 | | 148 | 0.222 | 66298 | 93.9 |
| | D | | 70286 | 99.5 | | 153 | 0.218 | 66298 | 93.9 |
| Leeuwarden | G | 1237 | 1237 | 100.0 | BSRN-1C | 0 | 0.000 | 1237 | 100.0 |
| Liepaja | G | 1258 | 764 | 60.7 | BSRN-1C | 0 | 0.000 | 764 | 60.7 |
| Ma'an | G | 69405 | 69269 | 99.8 | BSRN-3C | 414 | 0.598 | 66856 | 96.3 |
| | B | | 69269 | 99.8 | | 356 | 0.514 | 66856 | 96.3 |
| | D | | 68559 | 98.8 | | 1703 | 2.484 | 66856 | 96.3 |
| Maastricht | G | 1226 | 1226 | 100.0 | BSRN-1C | 1 | 0.082 | 1225 | 99.9 |
| Milhostov | G | 73114 | 71979 | 98.4 | BSRN-2C | 0 | 0.000 | 71722 | 98.1 |
| | D | | 71729 | 98.1 | | 7 | 0.010 | 71722 | 98.1 |
| Missour | G | 69849 | 69810 | 99.9 | BSRN-3C | 899 | 1.288 | 63752 | 91.3 |
| | B | | 69810 | 99.9 | | 894 | 1.281 | 63752 | 91.3 |
| | D | | 67418 | 96.5 | | 3664 | 5.435 | 63752 | 91.3 |
| Ny-Ålesund | G | 99088 | 97469 | 98.4 | BSRN-3C | 432 | 0.443 | 89908 | 90.7 |
| | B | | 90780 | 91.6 | | 273 | 0.301 | 89908 | 90.7 |
| | D | | 97853 | 98.8 | | 426 | 0.435 | 89908 | 90.7 |
| Oujda | G | 70146 | 30185 | 43.0 | BSRN-3C | 0 | 0.000 | 29556 | 42.1 |
| | B | | 30596 | 43.6 | | 0 | 0.000 | 29556 | 42.1 |
| | D | | 29592 | 42.2 | | 0 | 0.000 | 29556 | 42.1 |
| Payerne | | | | | | | | | |
| Poprad-Ganovce | G | 73222 | 73205 | 100.0 | BSRN-3C | 100 | 0.137 | 71611 | 97.8 |
| | B | | 71742 | 98.0 | | 87 | 0.121 | 71611 | 97.8 |
| | D | | 73178 | 99.9 | | 119 | 0.163 | 71611 | 97.8 |
| Reunion Island | G | 62772 | 62549 | 99.6 | BSRN-3C | 35 | 0.056 | 61868 | 98.6 |
| | B | | 62117 | 99.0 | | 3 | 0.005 | 61868 | 98.6 |
| | D | | 62306 | 99.3 | | 35 | 0.056 | 61868 | 98.6 |
| Riga | G | 1262 | 765 | 60.6 | BSRN-1C | 0 | 0.000 | 765 | 60.6 |
| Silutes | G | 1255 | 1255 | 100.0 | BSRN-1C | 0 | 0.000 | 1255 | 100.0 |
| | B | | 1255 | 100.0 | | 21 | 1.673 | 1234 | 98.3 |
| Sonnblick | G | 72704 | 67168 | 92.4 | BSRN-3C | 351 | 0.523 | 62953 | 86.6 |

| Station | Variable | Daylight instants during MAM 2021 | Observation >= 0 W m$^{-2}$ during daylight before QC | | QC Test | Data rejected by the QC during daylight | | Observation >= 0 W m$^{-2}$ during daylight after QC and before SZA threshold | |
|---|---|---|---|---|---|---|---|---|---|
| | | N | N | % | | N | % | N | % |
| | B | | 69564 | 95.7 | | 16 | 0.023 | 62953 | 86.6 |
| | D | | 66926 | 92.1 | | 3455 | 5.162 | 62953 | 86.6 |
| Syowa | G | 40826 | 36746 | 90.0 | BSRN-3C | 24 | 0.065 | 33770 | 82.7 |
| | B | | 34863 | 85.4 | | 0 | 0.000 | 33770 | 82.7 |
| | D | | 35489 | 86.9 | | 225 | 0.634 | 33770 | 82.7 |
| Tamanrasset | G | 68339 | 68244 | 99.9 | BSRN-1C | 0 | 0.000 | 68244 | 99.9 |
| Tataouine | G | 69855 | 69847 | 100.0 | BSRN-3C | 55 | 0.079 | 69761 | 99.9 |
| | B | | 69847 | 100.0 | | 54 | 0.077 | 69761 | 99.9 |
| | D | | 69826 | 100.0 | | 65 | 0.093 | 69761 | 99.9 |
| Tateno | G | 70423 | 48679 | 69.1 | BSRN-3C | 0 | 0.000 | 48094 | 68.3 |
| | B | | 48094 | 68.3 | | 0 | 0.000 | 48094 | 68.3 |
| | D | | 48679 | 69.1 | | 0 | 0.000 | 48094 | 68.3 |
| Toravere | | | | | | | | | |
| Twenthe | G | 1229 | 1229 | 100.0 | BSRN-1C | 0 | 0.000 | 1229 | 100.0 |
| Zagora | G | 69431 | 21794 | 31.4 | BSRN-3C | 0 | 0.000 | 21792 | 31.4 |
| | B | | 21794 | 31.4 | | 0 | 0.000 | 21792 | 31.4 |
| | D | | 21794 | 31.4 | | 2 | 0.009 | 21792 | 31.4 |
| Zoseni | G | 1262 | 766 | 60.7 | BSRN-1C | 0 | 0.000 | 766 | 60.7 |
| | D | | 35489 | 86.9 | | 225 | 0.634 | 33770 | 82.7 |
| Tamanrasset | G | 68339 | 68244 | 99.9 | BSRN-1C | 0 | 0.000 | 68244 | 99.9 |

# 3 Overview of the results

Following the ISO standard (1995), the deviations are computed by subtracting observations for each instant from the product estimations (CRS - measurements), and are summarized by usual statistical quantities such as the bias or the root mean square error. The validation procedure is described in Annex A. The selection of the cloud-free instants for the validation of McClear is described in Annex A.8. Detailed results for the report's season are given for each station in Annexes B.

## 3.1 Bias and standard deviation of errors

Detailed tables for the MAM 2021 quarter summarize the performance of the CRS for hourly means of global irradiance (in W $m^{-2}$, Tab. 3.1), corresponding clearness index (Tab. 3.2), and the performances relative to the mean of measurements (in percent, Tab. 3.3). Similar detailed tables are given for the *D* component (Tab. 3.4 to 3.6) and the *B* component (Tab. 3.7 to 3.9). They are available as an excel file in a folder of the supplemental zip archive:

CAMS2_73_2021SC1_D1.3.1-2021Q4_section3_statistical_metrics/
CAMS2_73_2021SC1_D1.3.1-2021Q4_section3_statistical_metrics_MAM2021.xlsx

Table 3.1 summarizes the performance of CRS for the hourly mean of irradiances. The following empirical rules adopted for the bias are derived from the uncertainty (20 W $m^{-2}$) of the measurements of hourly irradiation of good quality from the recommendations of the WMO (see Annex A.6).

Table 3.1. Number of stations in each category of bias (in W $m^{-2}$)

| Rules for the bias (in W m ²) | | CAMS-RAD | | McClear | |
|---|---|---|---|---|---|
| | | *G* | *D* | *G* | *D* |
| Null bias | Absolute value of the bias ≤ 5 | 4 | 7 | 4 | 0 |
| Low bias | 5 < absolute value of the bias ≤ 10 | 9 | 4 | 5 | 1 |
| Noticeable bias | 10 < absolute value of the bias ≤ 20 | 11 | 3 | 9 | 4 |
| Large bias | 20 < absolute value of the bias ≤ 60 | 5 | 3 | 3 | 11 |
| Very large bias | absolute value of the bias> 60 | 2 | 1 | 0 | 5 |
| Total stations with nb of obs.>30 | | 30 | 31 | 18 | 21 |

Relative biases and RMSE are given in Figures 3.1, 3.2 and 3.3 for respectively *G*, *D* and *B*.

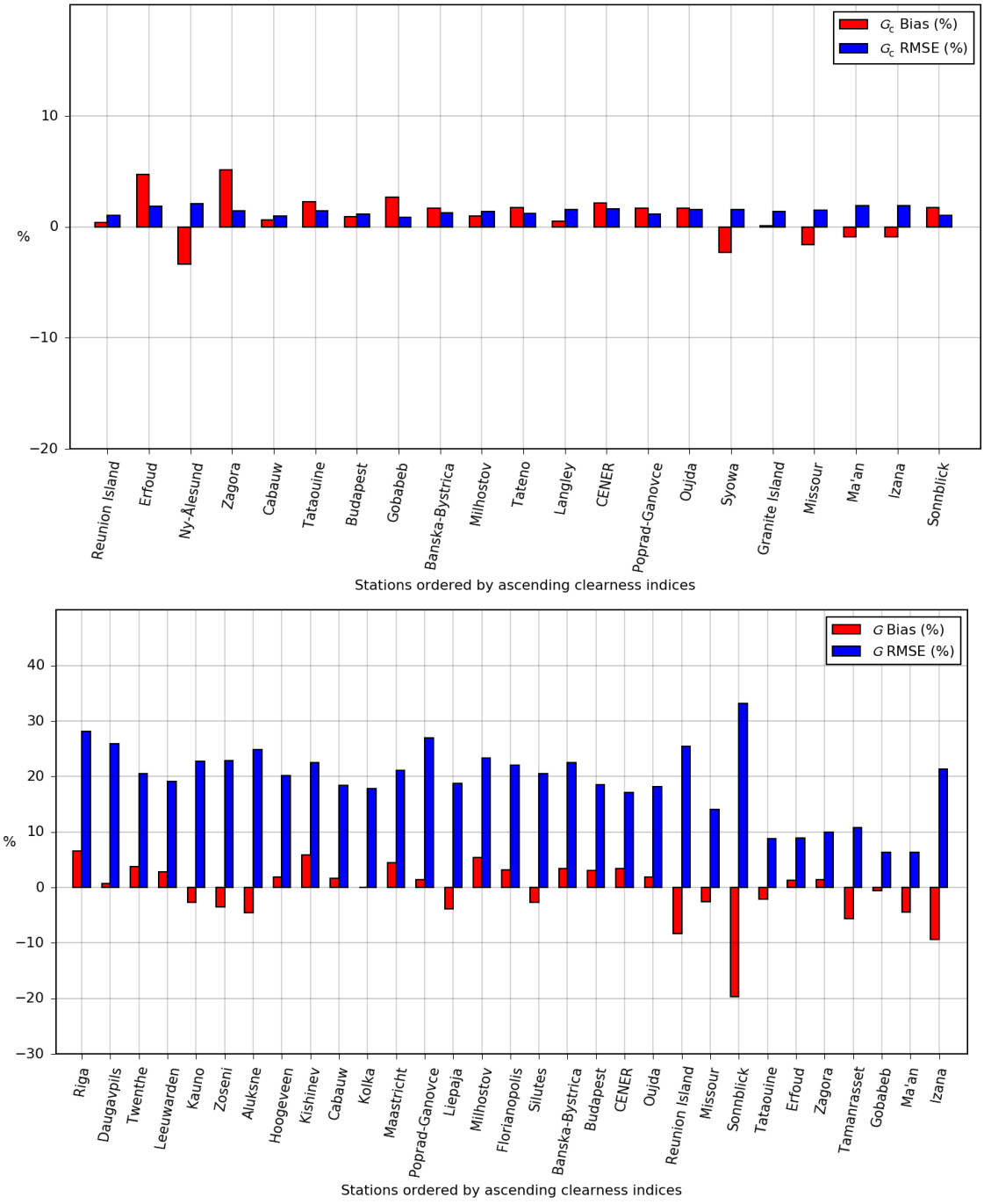



Figure 3.1. Relative bias and RMSE for global irradiance. Top: McClear. Bottom: CAMS-RAD.

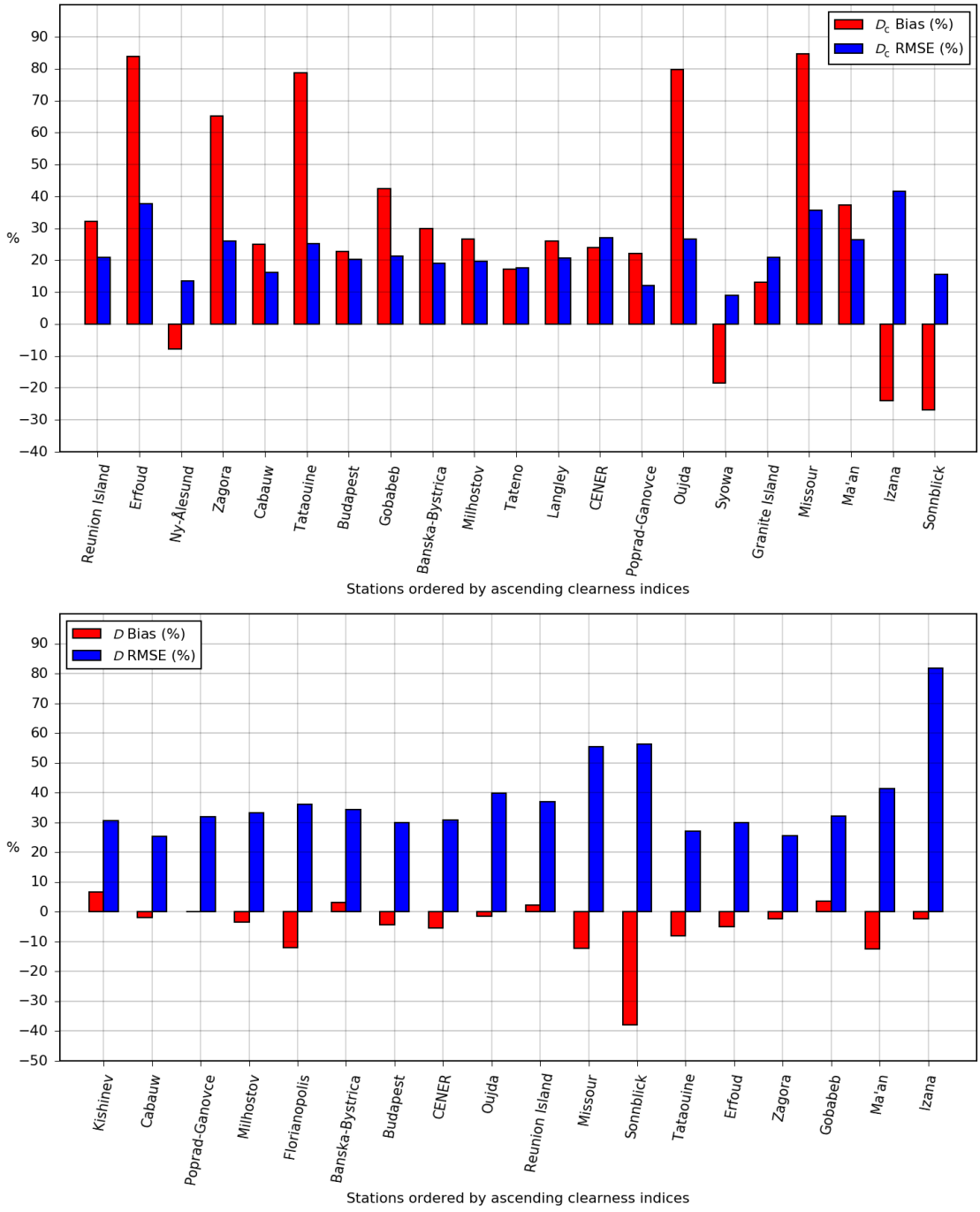



Figure 3.2. Relative bias and RMSE for diffuse irradiance. Top: McClear. Bottom: CAMS-RAD.

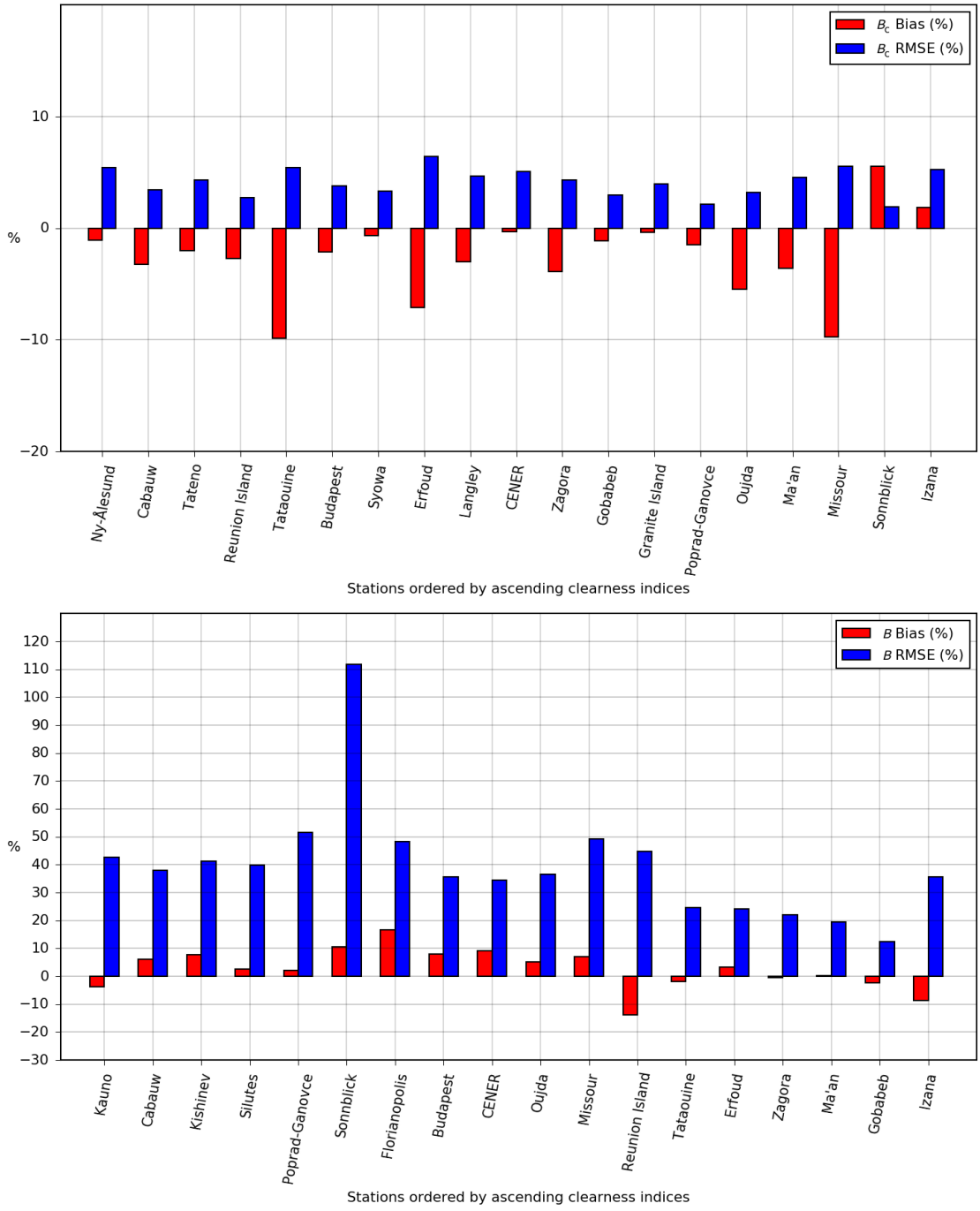



Figure 3.3. Relative bias and RMSE for direct irradiance. Top: McClear. Bottom: CAMS-RAD.

## 3.2 Ability to reproduce the intra-day variability

The correlation coefficients of the MAM 2021 quarter are displayed for irradiances and clearness indices: *G* and $KT_G$ (Fig. 3.4), *D* and $KT_D$ (Fig. 3.5), *B* and $KT_B$ (Fig. 3.6).

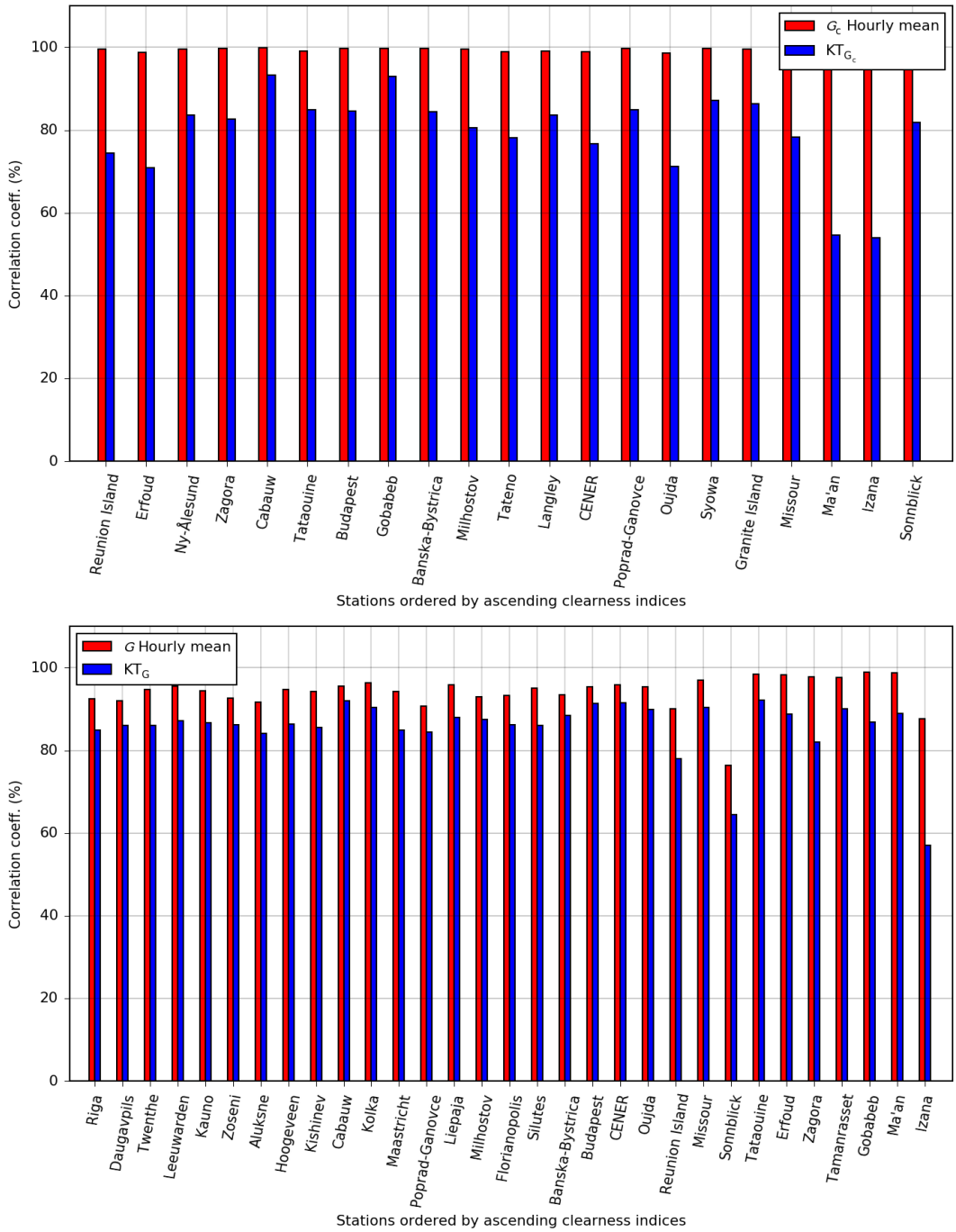



Figure 3.4. Correlation coefficients for global irradiance and clearness index.

Top: McClear. Bottom: CAMS-RAD.

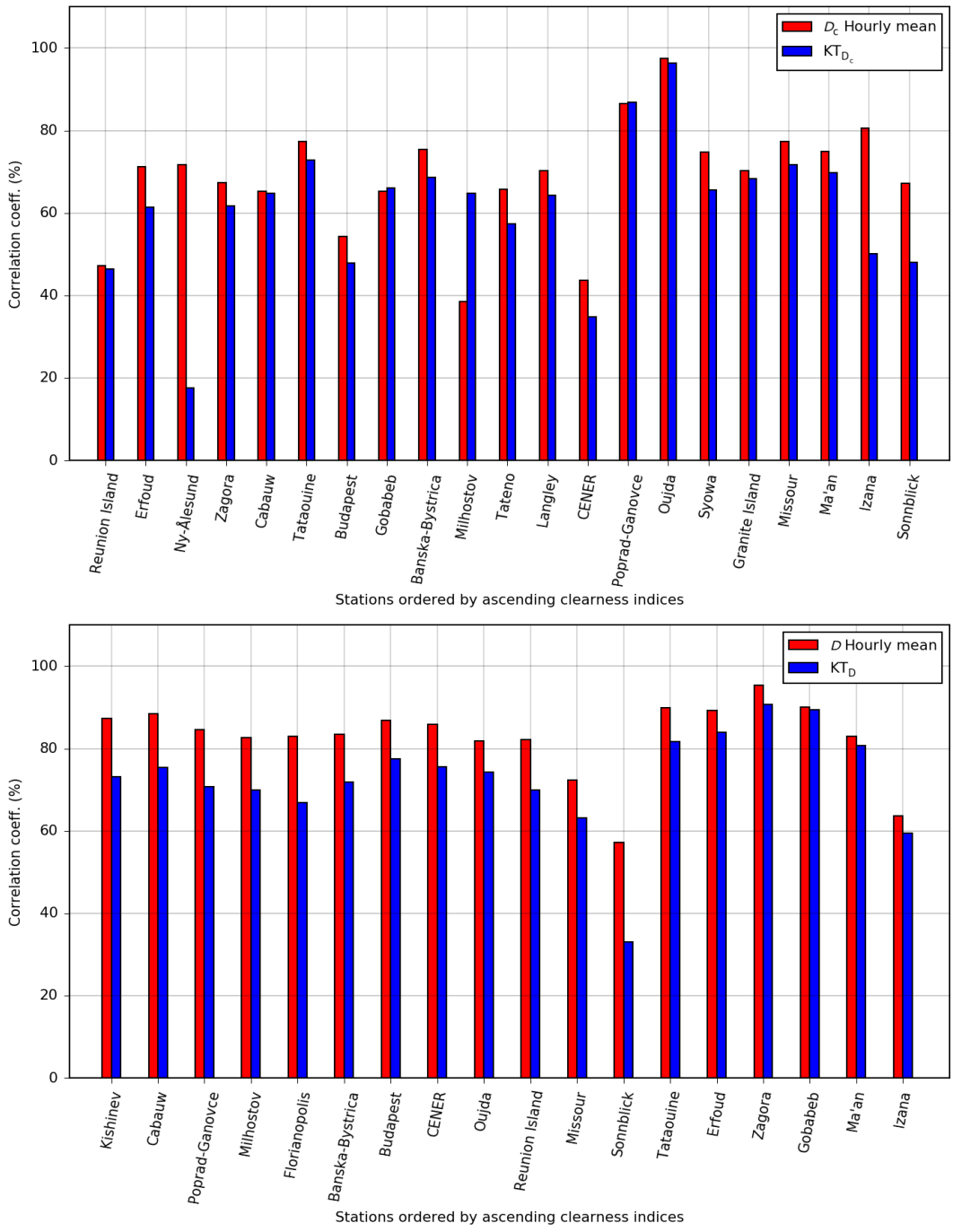



Figure 3.5. Correlation coefficients for diffuse irradiance and clearness index. Top: McClear. Bottom: CAMS-RAD.

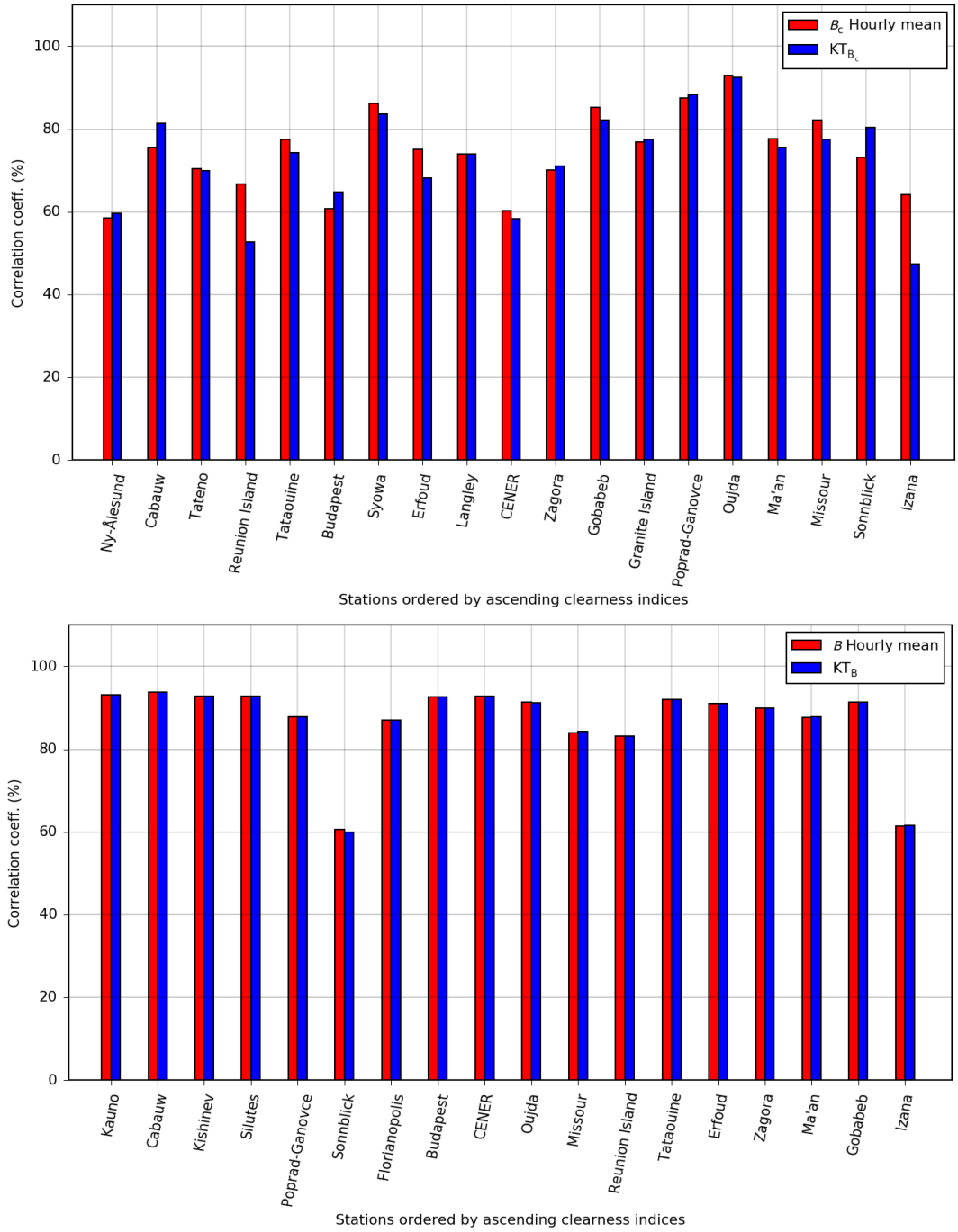



Figure 3.6. Correlation coefficients for direct normal irradiance and clearness index. Top: McClear. Bottom: CAMS-RAD.

## 3.3 Ability to reproduce the frequency distributions of measurements

The frequency distributions of G and $KT_G$ (Fig. 3.7 and 3.8), D and $KT_D$ (Fig. 3.9 and 3.10), B and $KT_B$ (Fig. 3.11 and 3.12) are given for several classes of irradiances or *KT*. The objective of these figures is to display in which bins of values an over- or underestimation of the frequencies may be observed.

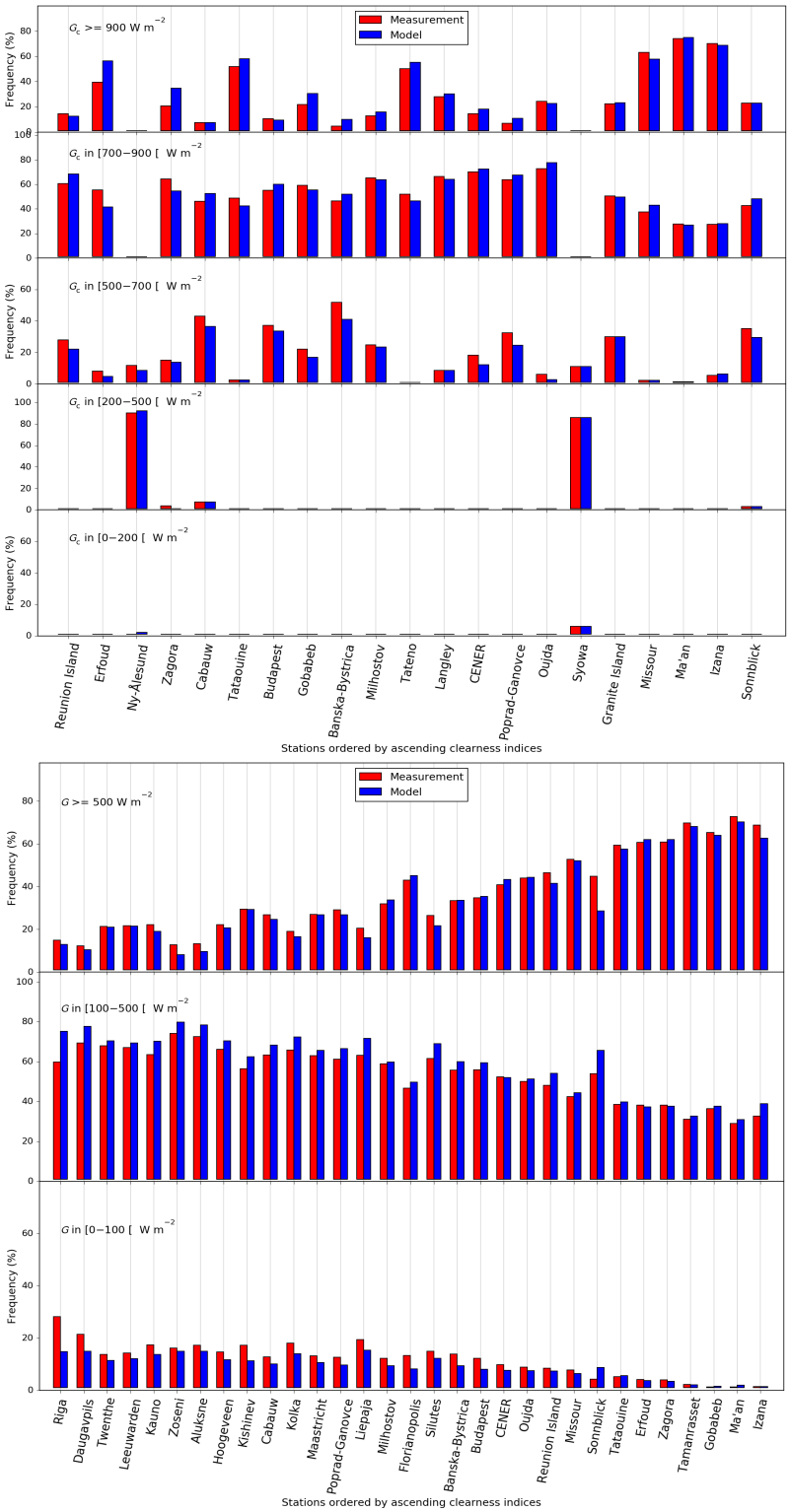



Figure 3.7. Frequency distribution of measurements for global irradiance. Top: McClear. Bottom: CAMS-RAD.

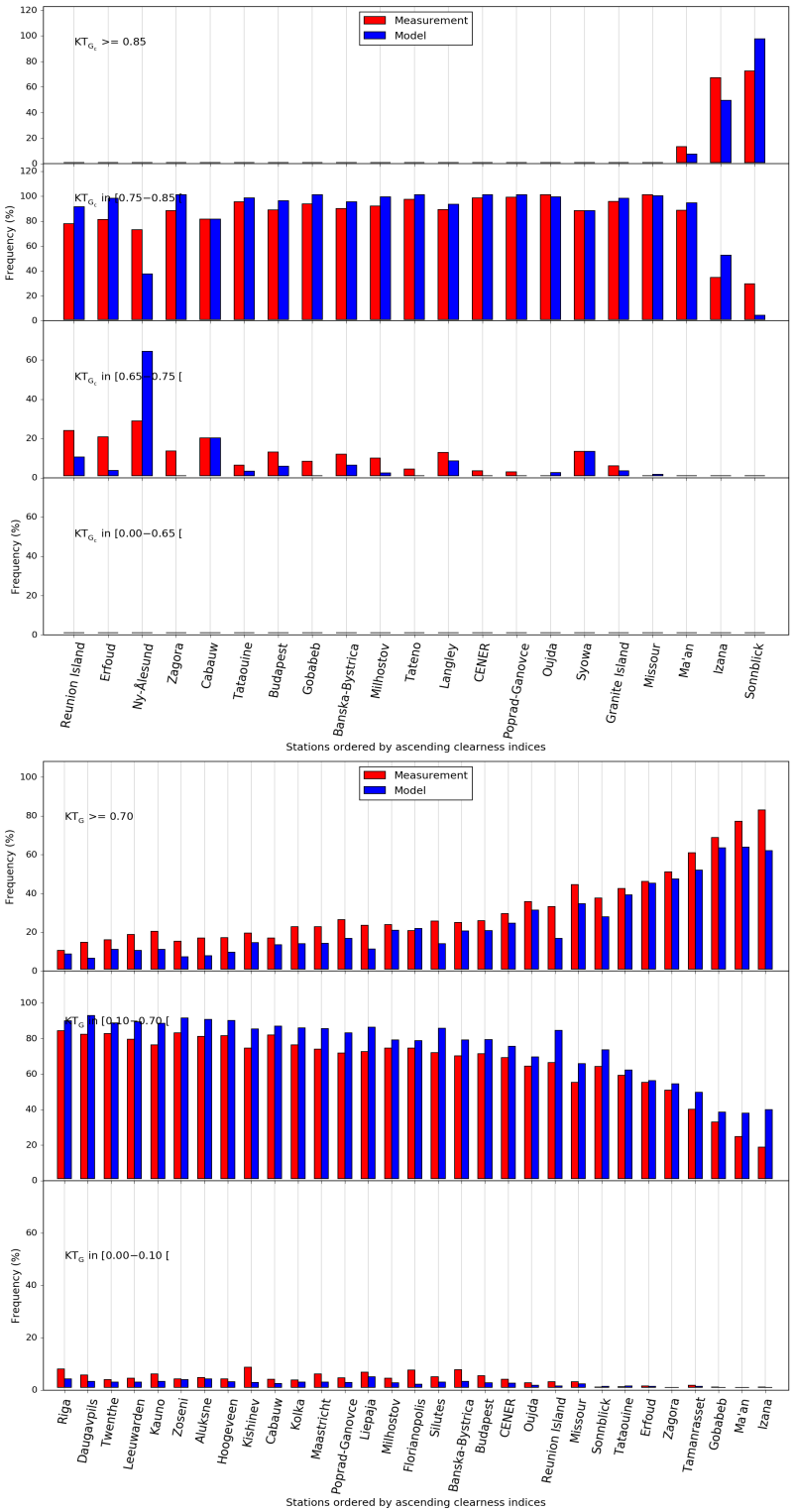



Figure 3.8. Frequency distribution of measurements for global clearness index. Top: McClear. Bottom: CAMS-RAD.

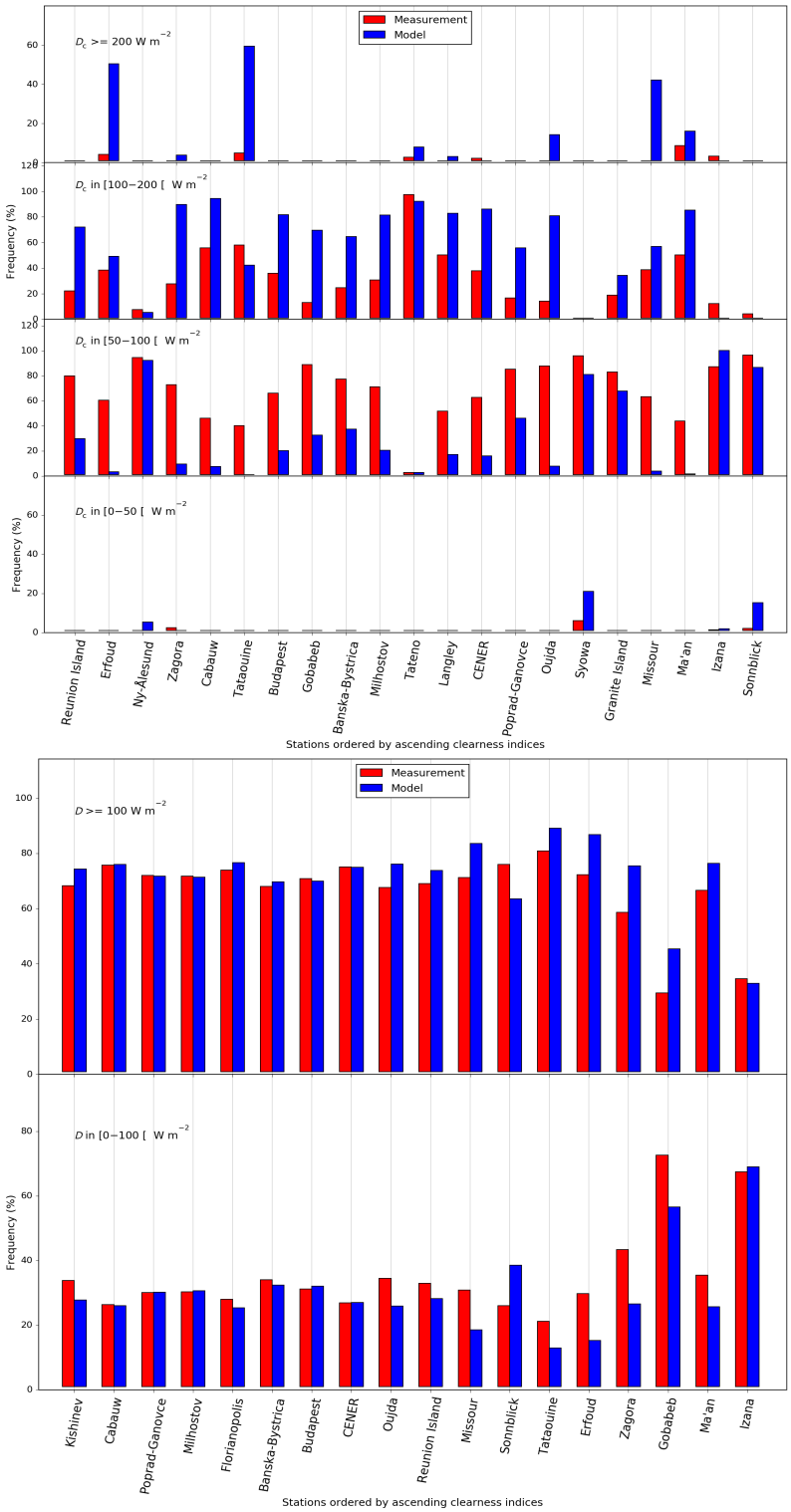



Figure 3.9. Frequency distribution of measurements for diffuse irradiances. Top: McClear. Bottom: CAMS-RAD.

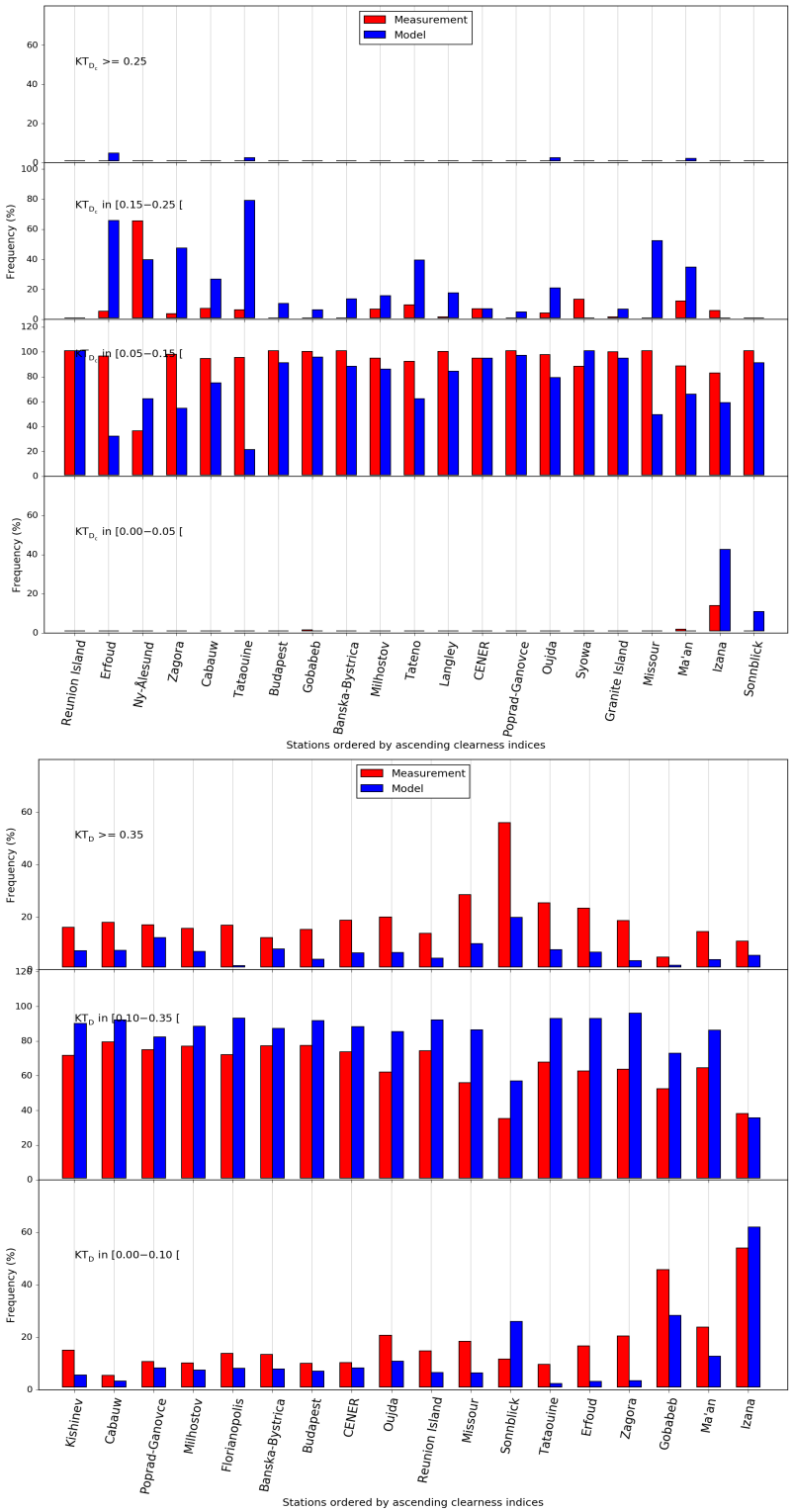



Figure 3.10. Frequency distribution of measurements for diffuse clearness index. Top: McClear. Bottom: CAMS-RAD.

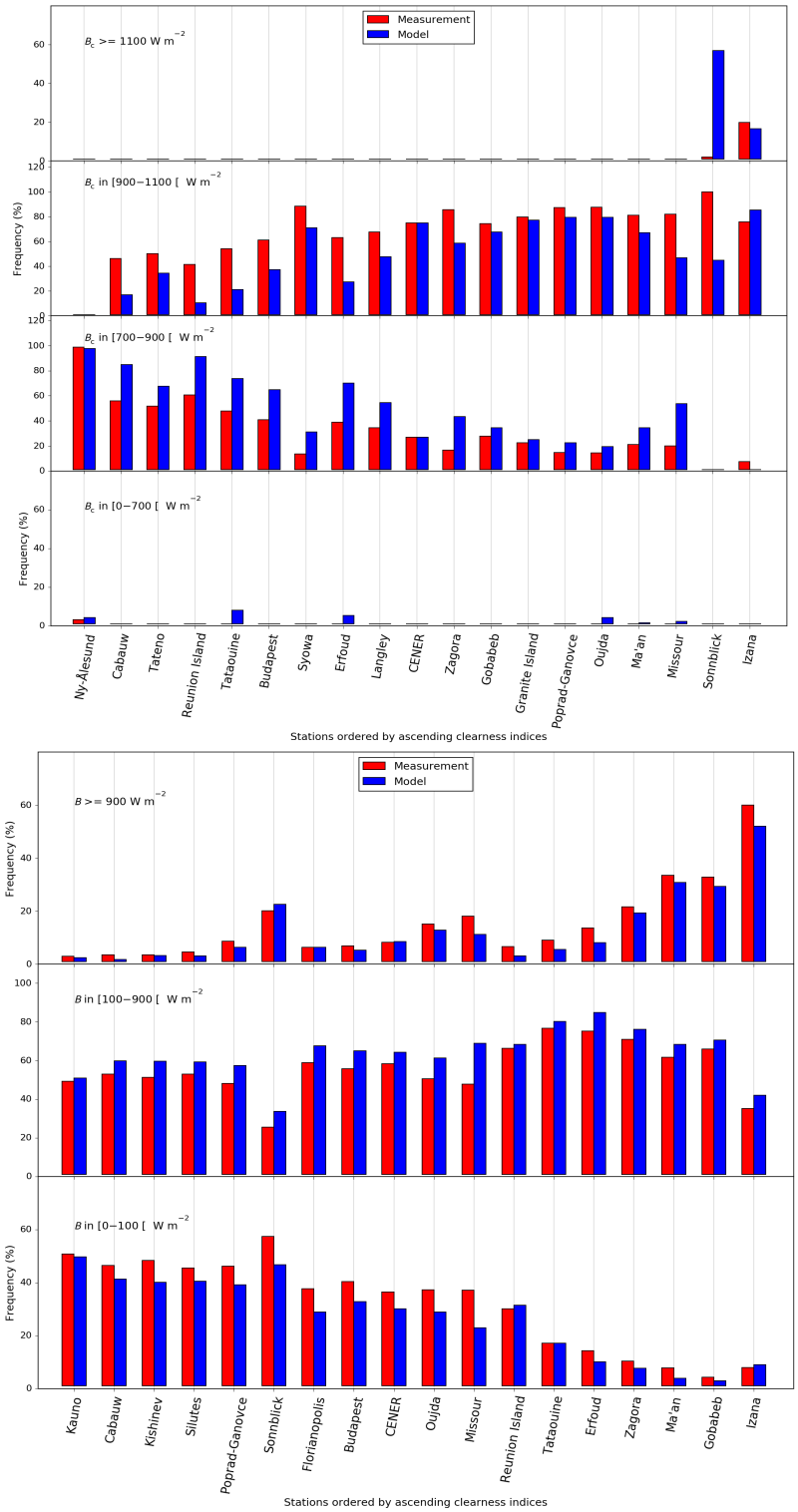



Figure 3.11. Frequency distribution of measurements for direct irradiances. Top: McClear. Bottom: CAMS-RAD.

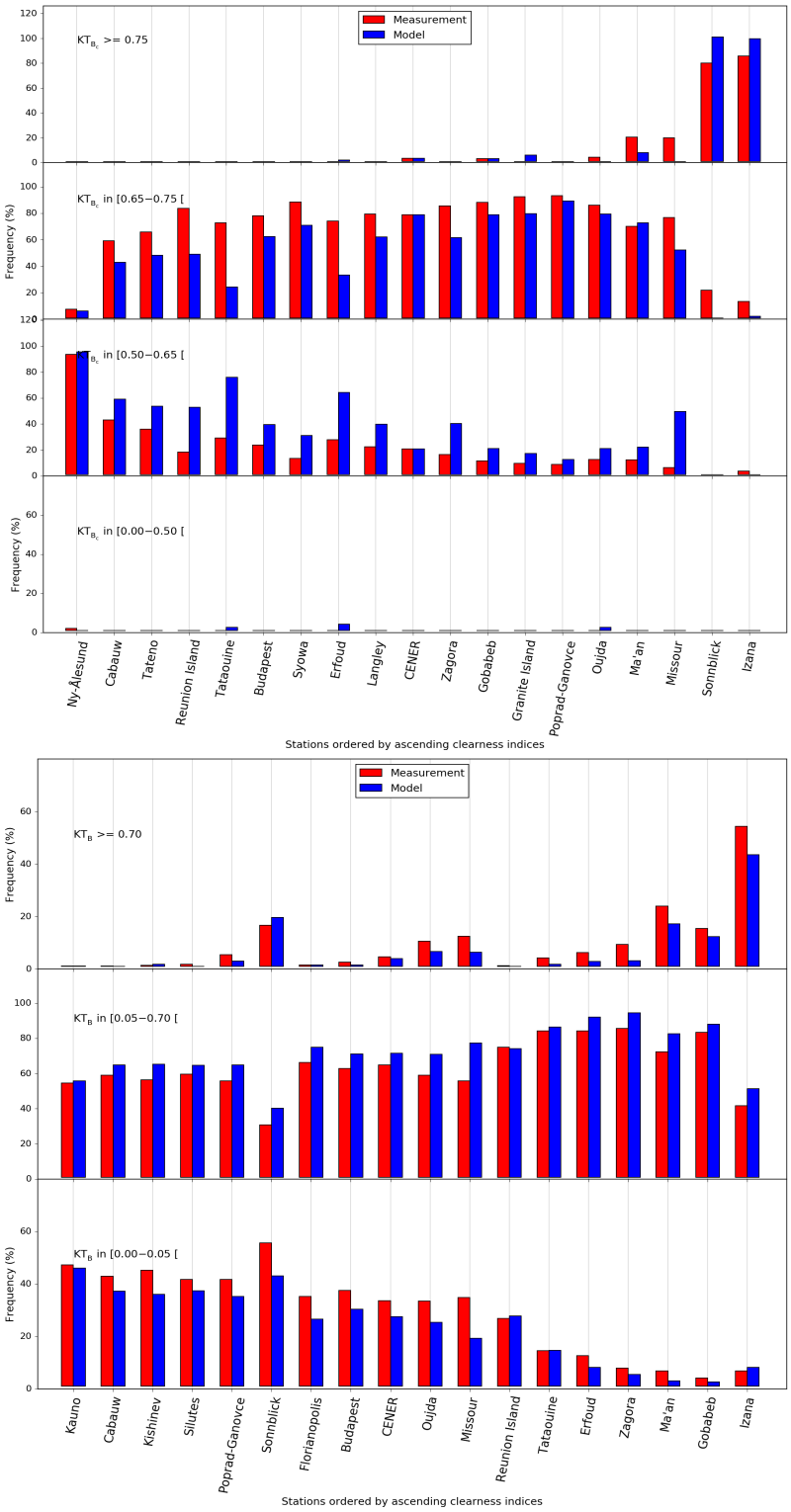



Figure 3.12. Frequency distribution of measurements for direct clearness index. Top: McClear. Bottom: CAMS-RAD.

## 3.4 Ability to reproduce the monthly means and standard deviation for the period

The final batch of analyses deals with the capability of CRS to reproduce the monthly means of the irradiance for each month of the period and its variability within a month, expressed as the standard-deviation of the hourly values (estimates and observations) within this month. Figures 3.13 to 3.15 are respectively for *G*, *D* and *B*.

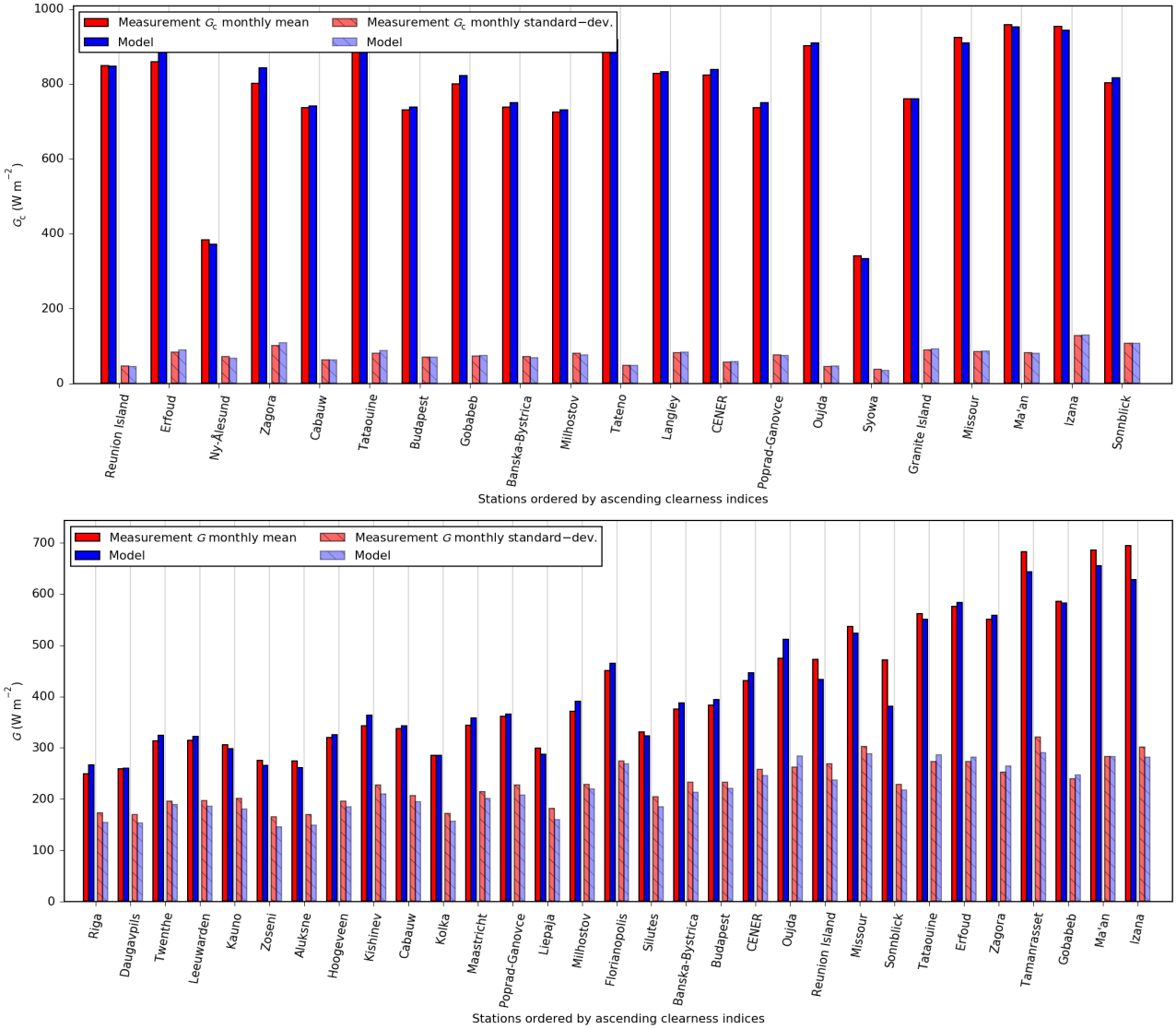



Figure 3.13. Monthly means and standard-deviations for global irradiance.
Top: McClear. Bottom: CAMS-RAD.

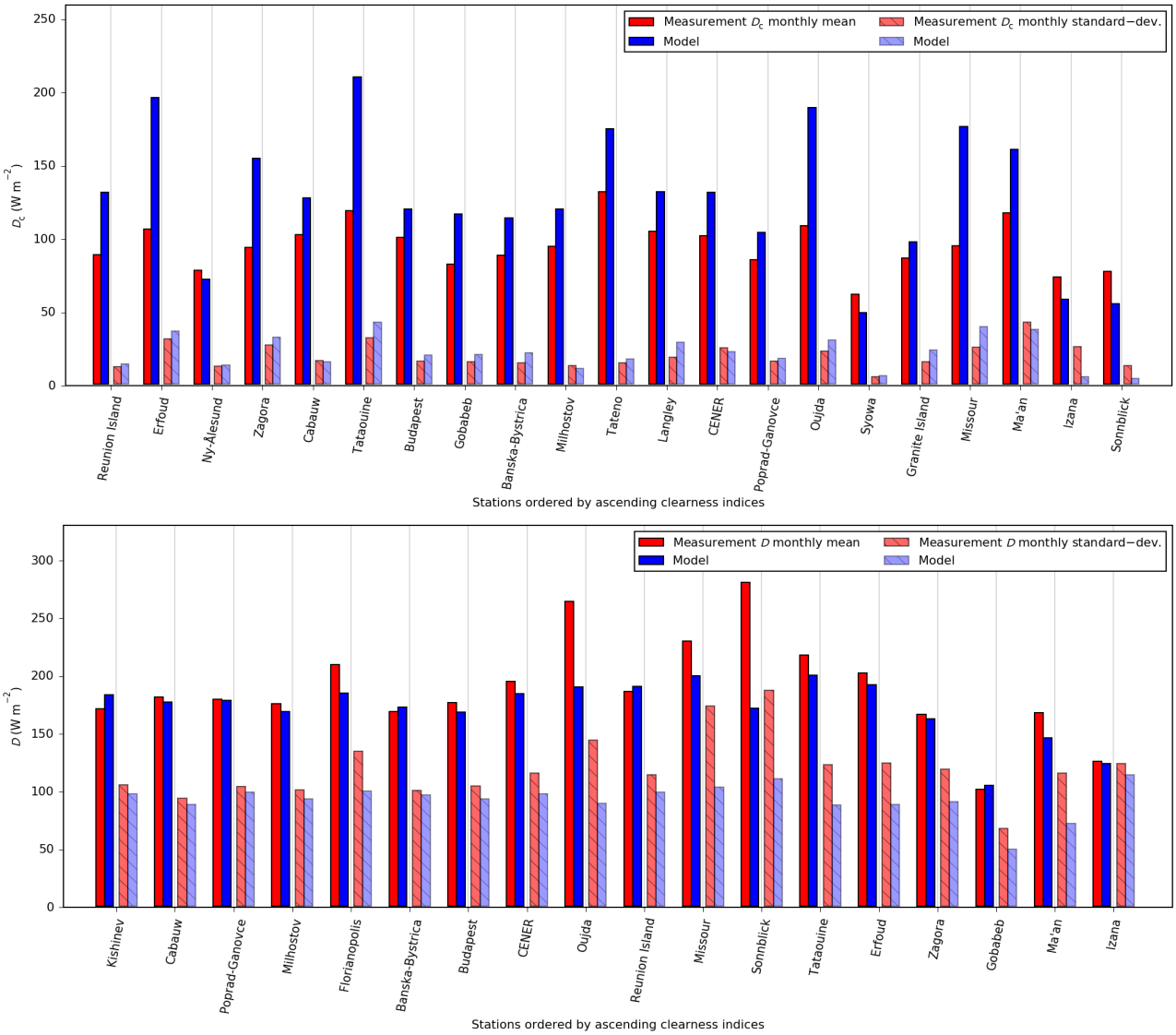



Figure 3.14. Monthly means and standard-deviations for diffuse irradiance.
Top: McClear. Bottom: CAMS-RAD.

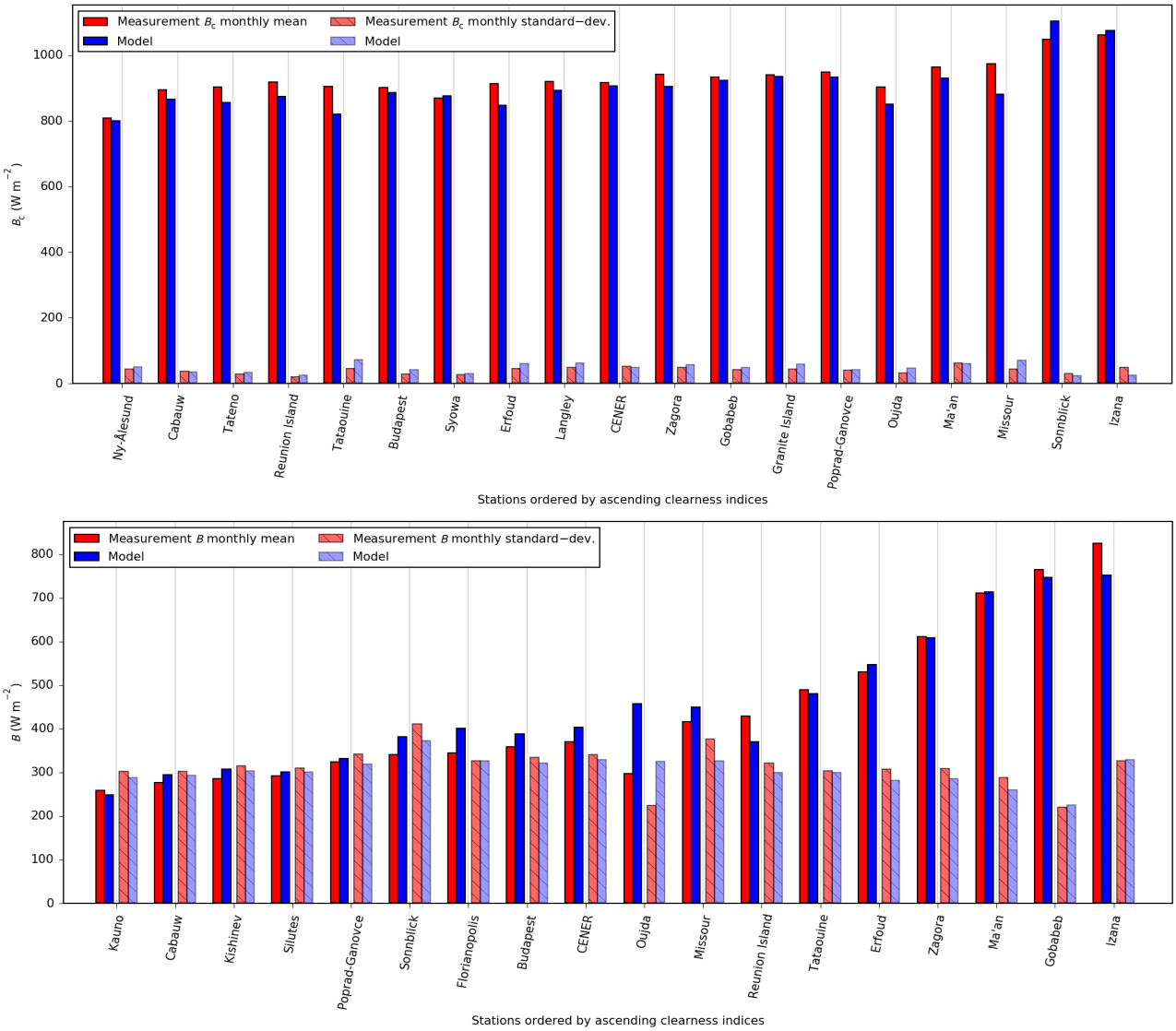
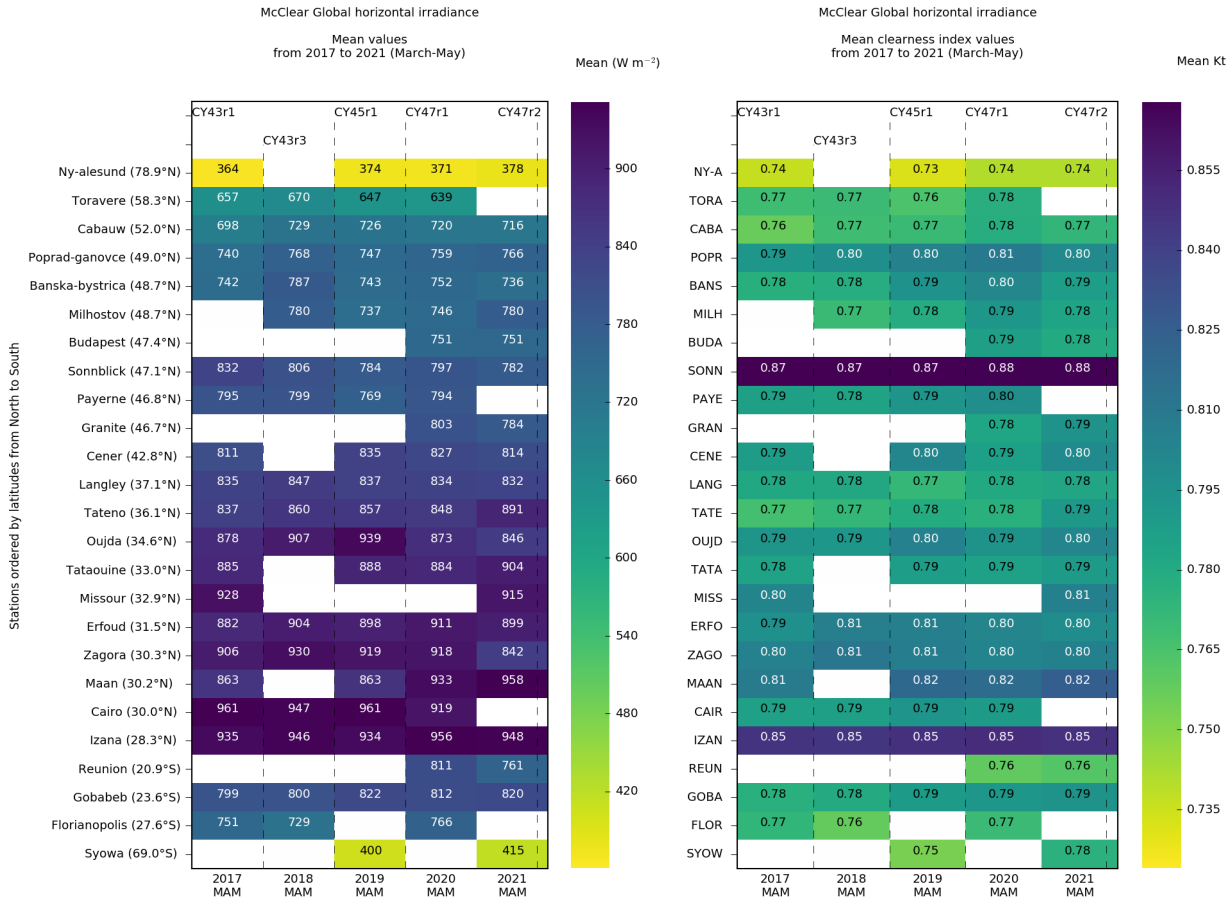



Figure 3.15. Monthly means and standard-deviation for direct normal irradiance. Top: McClear. Bottom: CAMS-RAD.

## 3.5 Multi-annual statistical indicators

Statistical indicators (mean irradiance, mean KT, relative bias and relative RMSE) have been computed for the MAM period and for the years 2017 to 2021. Figures 3.16 to 3.18 are respectively for *G*, *D* and *B*.

## 3.5.1 Global irradiance



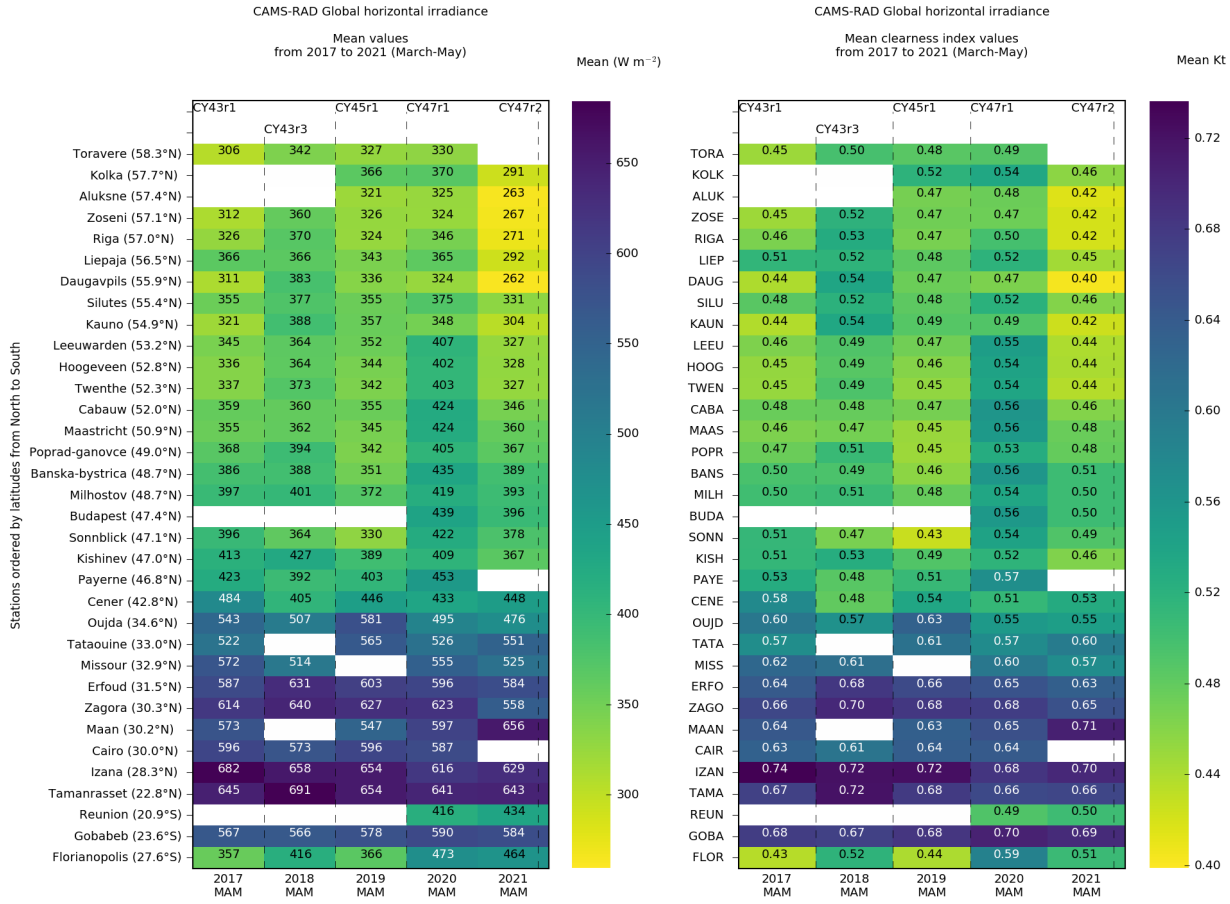
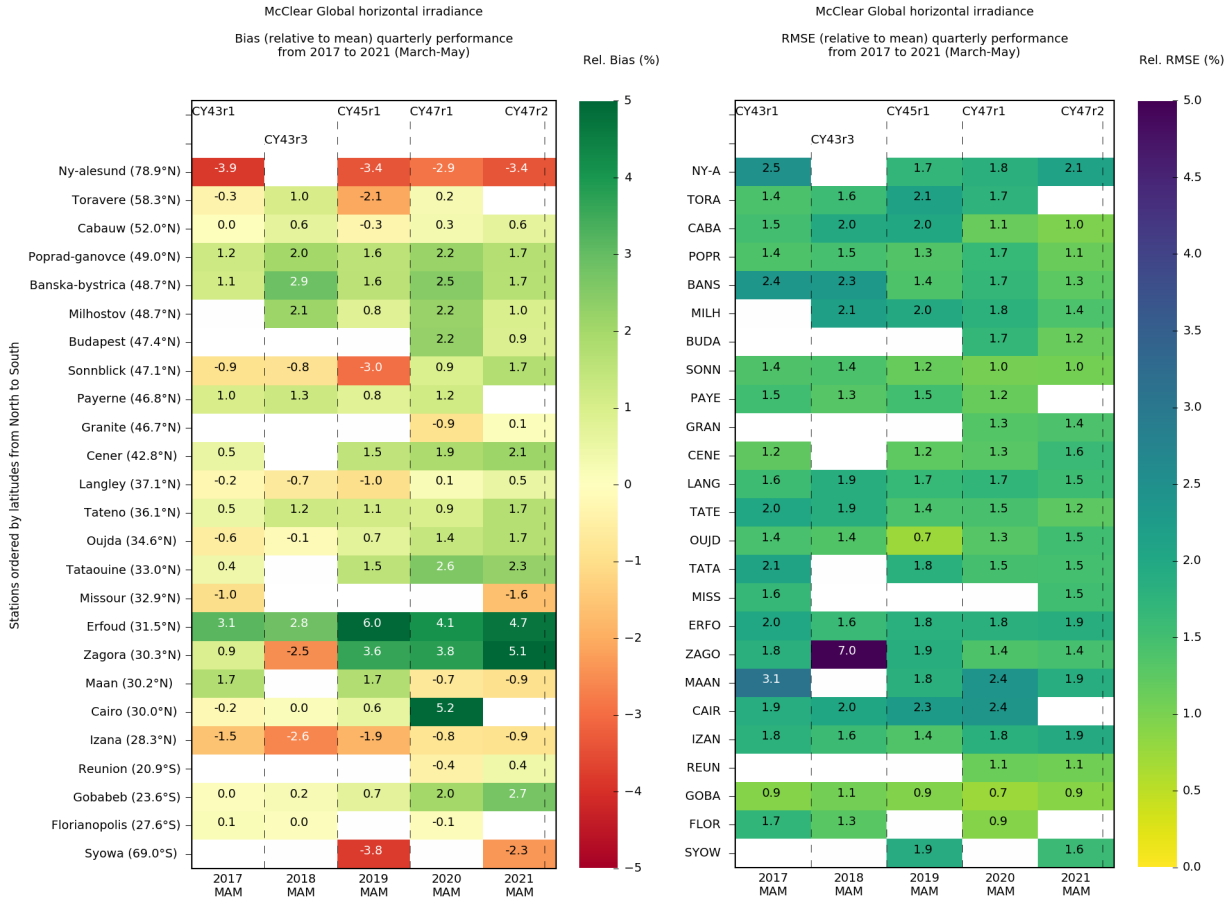



Figure 3.16a. Multi-annual means for global horizontal irradiance (left) and clearness index (right). Top: McClear. Bottom: CAMS-RAD.


McClear Global horizontal irradiance
Bias (relative to mean) quarterly performance from 2017 to 2021 (March-May)
Rel. Bias (%)
McClear Global horizontal irradiance
RMSE (relative to mean) quarterly performance from 2017 to 2021 (March-May)
Rel. RMSE (%)
Stations ordered by latitudes from North to South
CY43r1
CY43r3
CY45r1
CY47r1
CY47r2
2017 MAM
2018 MAM
2019 MAM
2020 MAM
2021 MAM

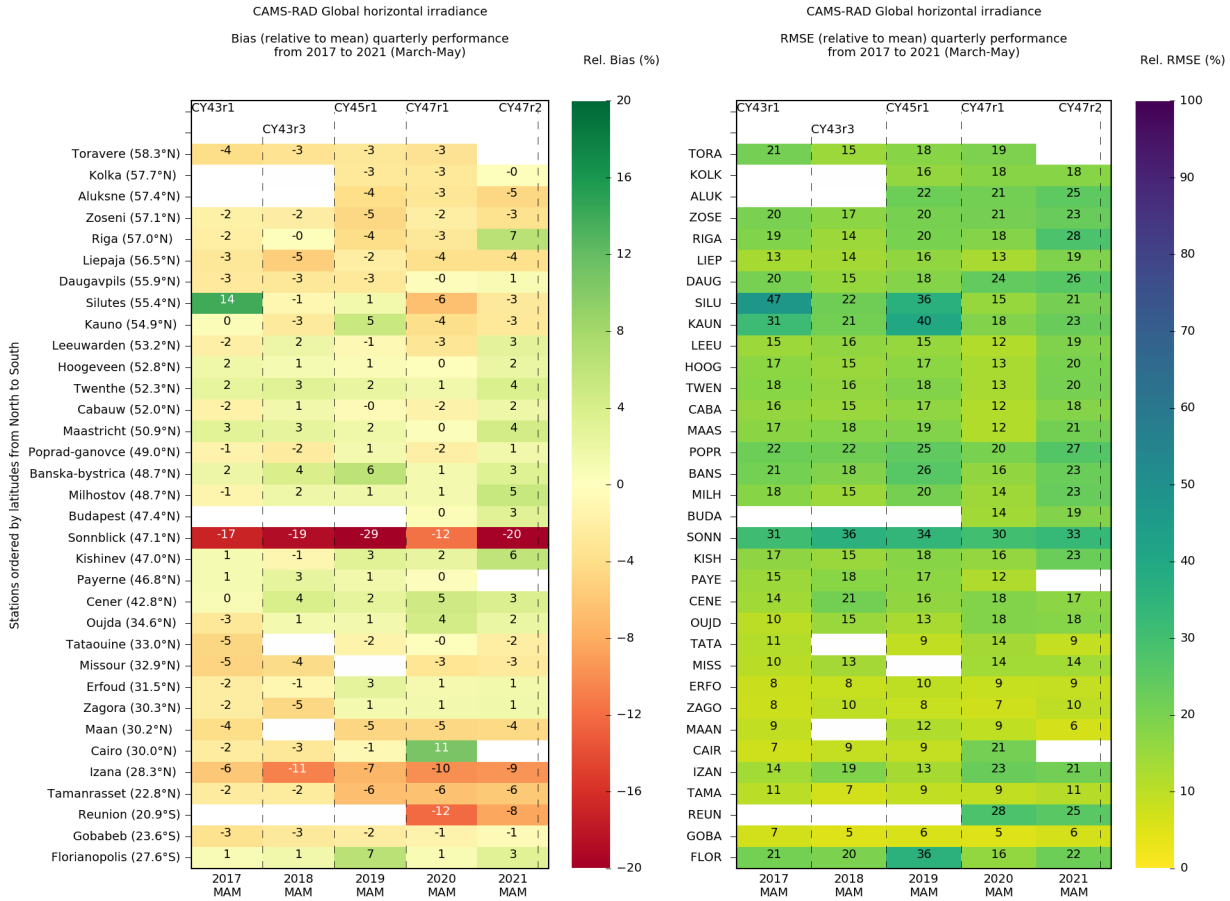
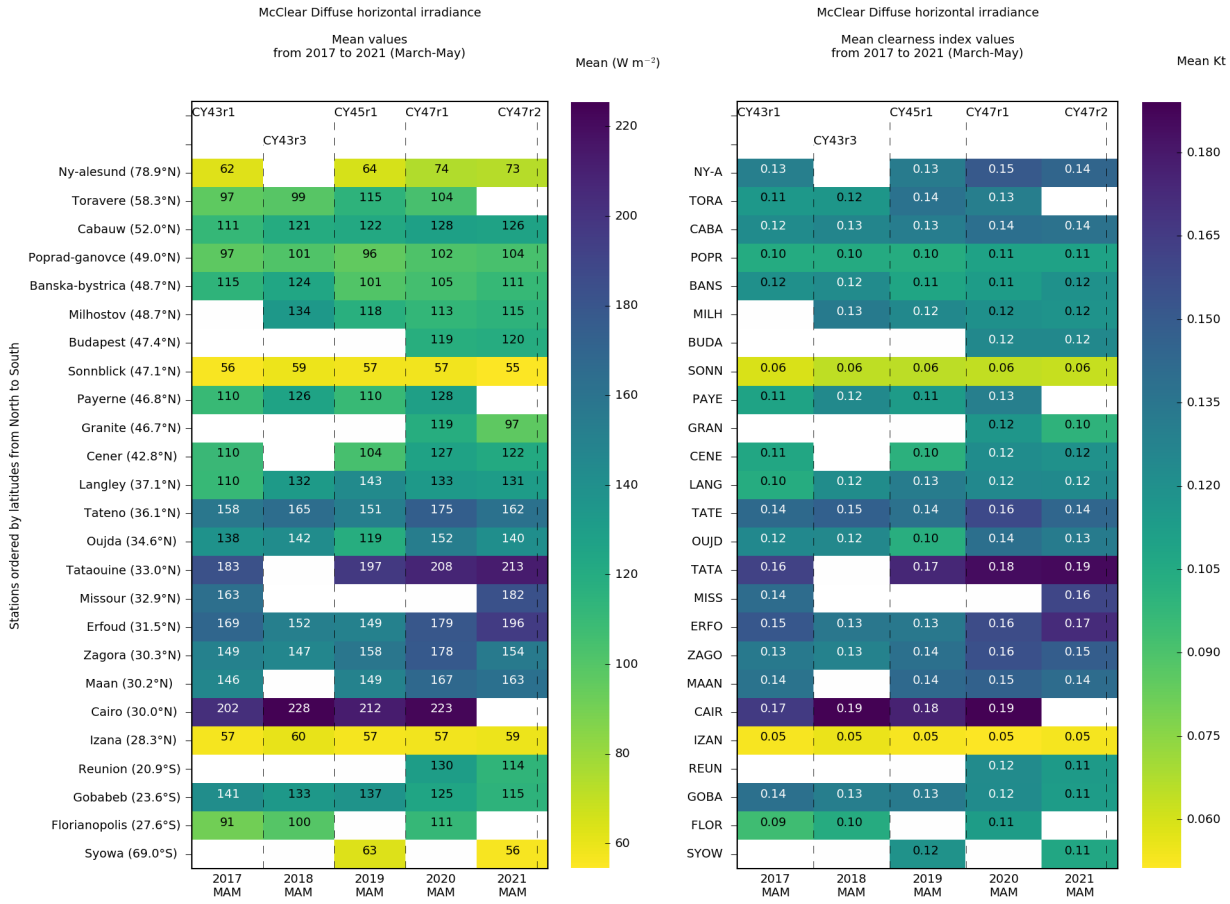



Figure 3.16b. Multi-annual global horizontal irradiance relative bias (left) and RMSE (right). Top: McClear. Bottom: CAMS-RAD.

## 3.5.2 Diffuse irradiance



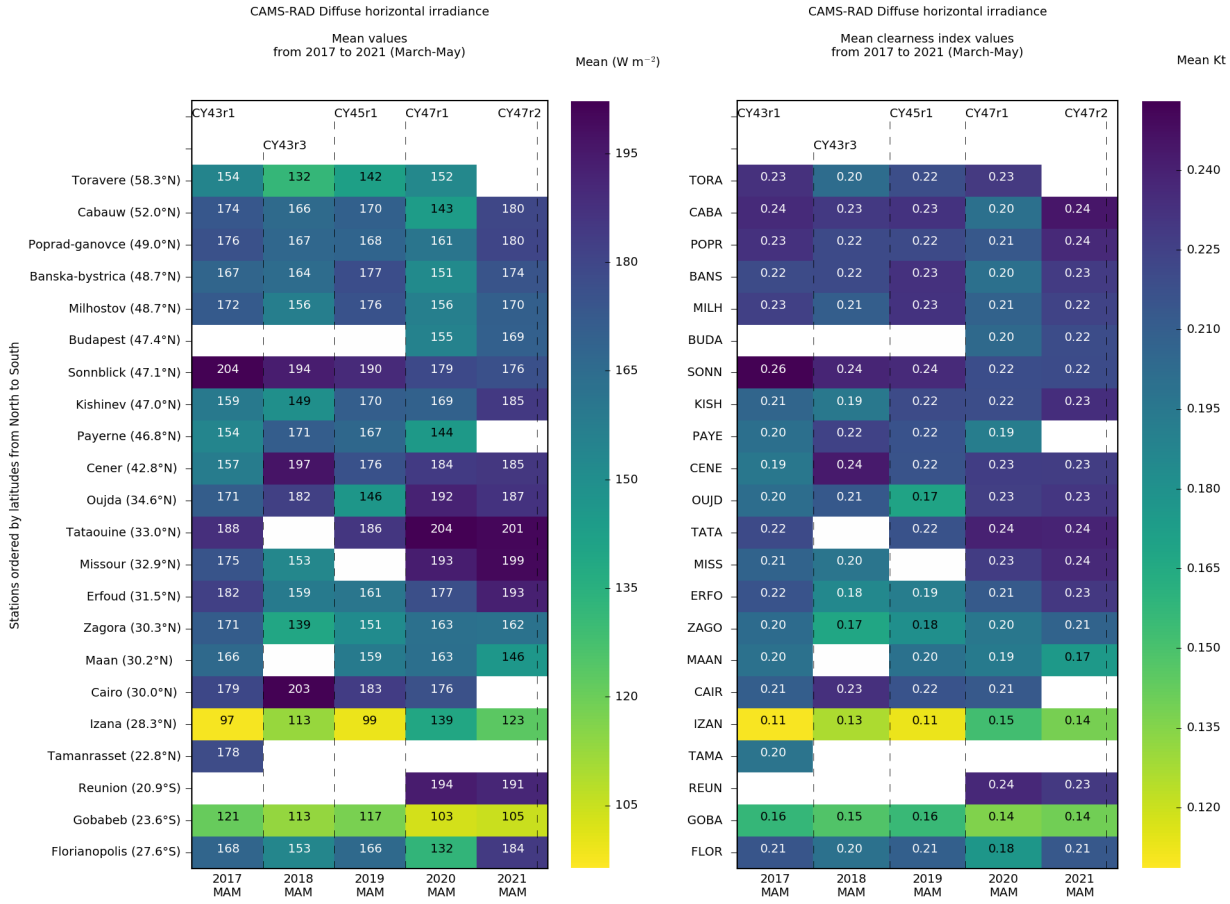



Figure 3.17a. Multi-annual means for diffuse horizontal irradiance (left) and clearness index (right). Top: McClear. Bottom: CAMS-RAD.

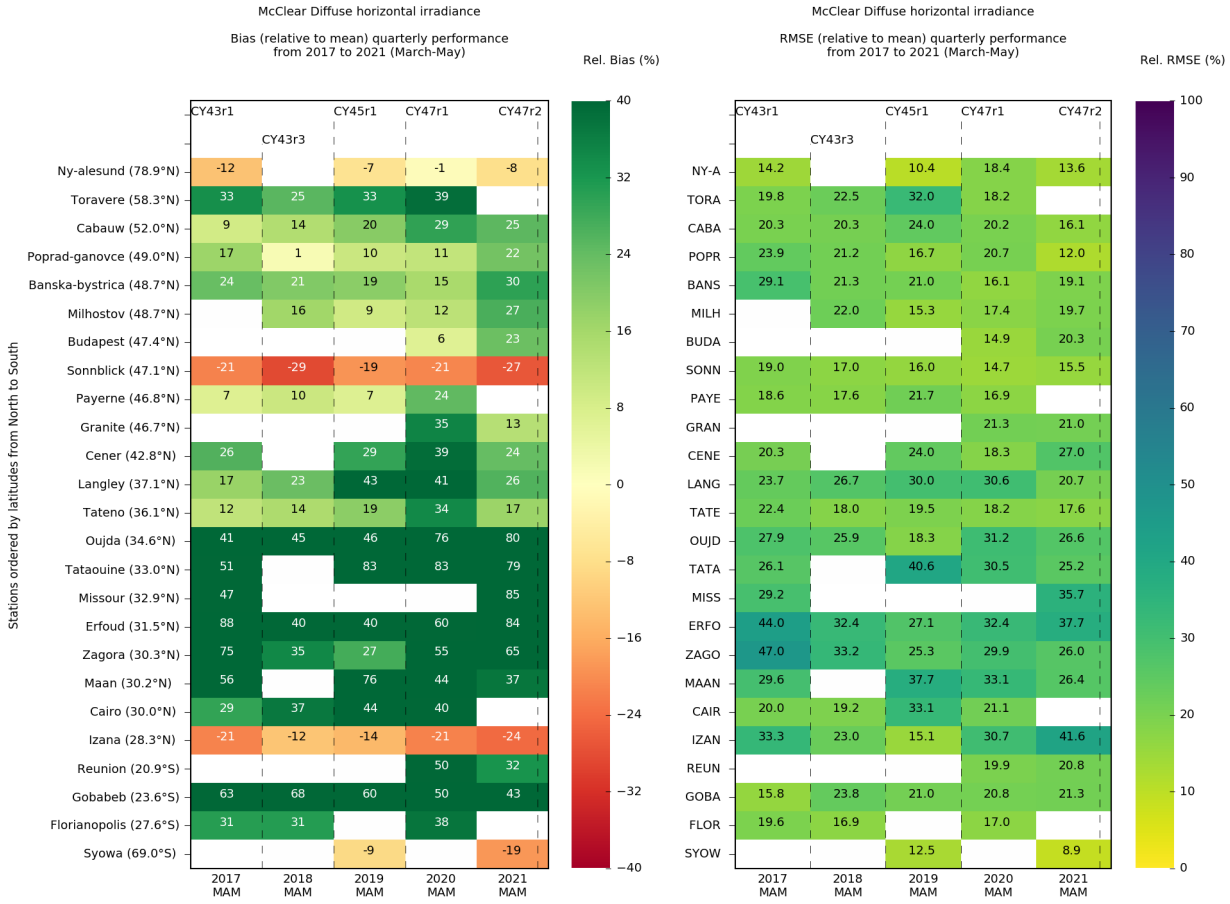
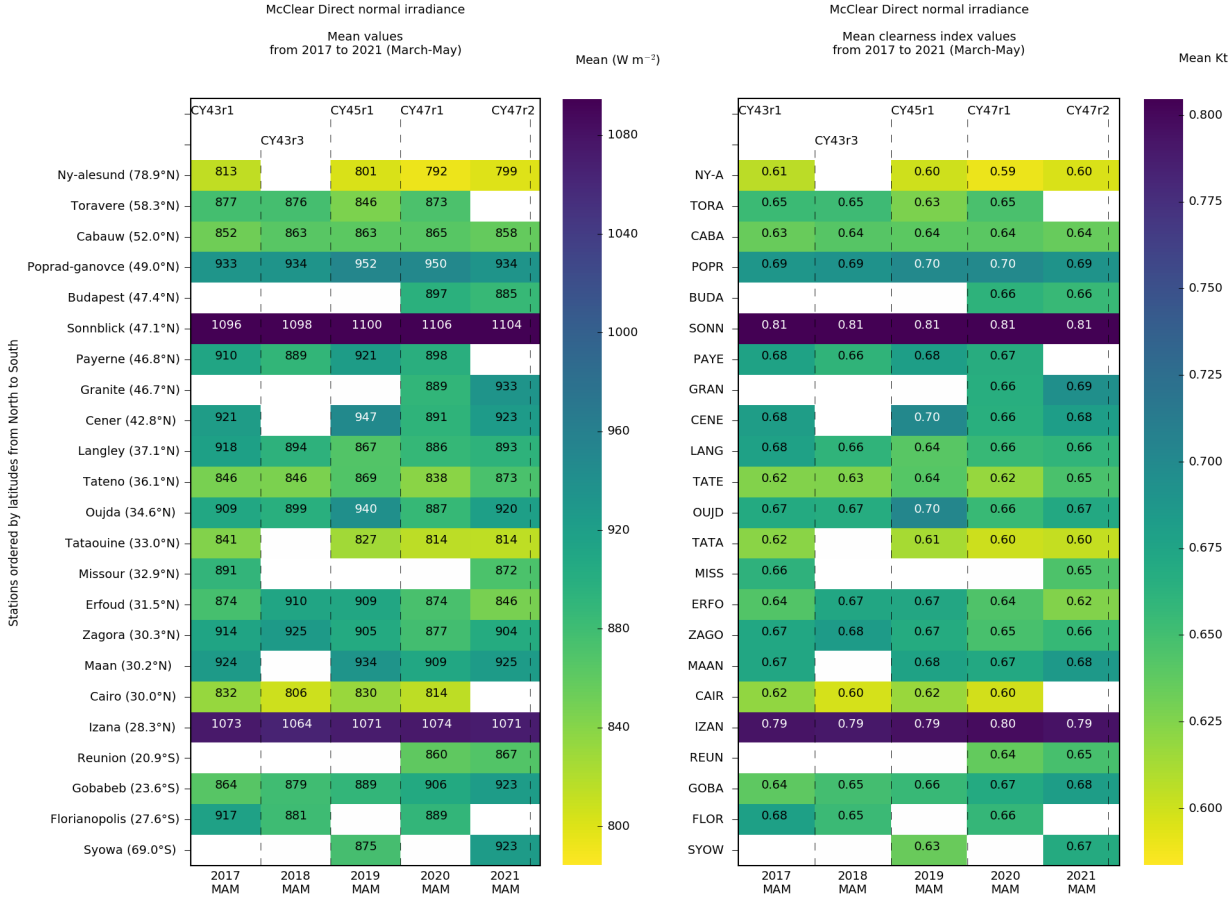

McClear Diffuse horizontal irradiance
Bias (relative to mean) quarterly performance from 2017 to 2021 (March-May)
Rel. Bias (%)
McClear Diffuse horizontal irradiance
RMSE (relative to mean) quarterly performance from 2017 to 2021 (March-May)
Rel. RMSE (%)
Stations ordered by latitudes from North to South
CY43r1
CY43r3
CY45r1
CY47r1
CY47r2
2017 MAM
2018 MAM
2019 MAM
2020 MAM
2021 MAM
Ny-alesund (78.9°N) -12 -7 -1 -8
Toravere (58.3°N) 33 25 33 39
Cabauw (52.0°N) 9 14 20 29 25
Poprad-ganovce (49.0°N) 17 1 10 11 22
Banska-bystrica (48.7°N) 24 21 19 15 30
Milhostov (48.7°N) 16 9 12 27
Budapest (47.4°N) 6 23
Sonnblick (47.1°N) -21 -29 -19 -21 -27
Payerne (46.8°N) 7 10 7 24
Granite (46.7°N) 35 13
Cener (42.8°N) 26 29 39 24
Langley (37.1°N) 17 23 43 41 26
Tateno (36.1°N) 12 14 19 34 17
Oujda (34.6°N) 41 45 46 76 80
Tataouine (33.0°N) 51 83 83 79
Missour (32.9°N) 47 85
Erfoud (31.5°N) 88 40 40 60 84
Zagora (30.3°N) 75 35 27 55 65
Maan (30.2°N) 56 76 44 37
Cairo (30.0°N) 29 37 44 40
Izana (28.3°N) -21 -12 -14 -21 -24
Reunion (20.9°S) 50 32
Gobabeb (23.6°S) 63 68 60 50 43
Florianopolis (27.6°S) 31 31 38
Syowa (69.0°S) -9 -19
NY-A 14.2 10.4 18.4 13.6
TORA 19.8 22.5 32.0 18.2
CABA 20.3 20.3 24.0 20.2 16.1
POPR 23.9 21.2 16.7 20.7 12.0
BANS 29.1 21.3 21.0 16.1 19.1
MILH 22.0 15.3 17.4 19.7
BUDA 14.9 20.3
SONN 19.0 17.0 16.0 14.7 15.5
PAYE 18.6 17.6 21.7 16.9
GRAN 21.3 21.0
CENE 20.3 24.0 18.3 27.0
LANG 23.7 26.7 30.0 30.6 20.7
TATE 22.4 18.0 19.5 18.2 17.6
OUJD 27.9 25.9 18.3 31.2 26.6
TATA 26.1 40.6 30.5 25.2
MISS 29.2 35.7
ERFO 44.0 32.4 27.1 32.4 37.7
ZAGO 47.0 33.2 25.3 29.9 26.0
MAAN 29.6 37.7 33.1 26.4
CAIR 20.0 19.2 33.1 21.1
IZAN 33.3 23.0 15.1 30.7 41.6
REUN 19.9 20.8
GOBA 15.8 23.8 21.0 20.8 21.3
FLOR 19.6 16.9 17.0
SYOW 12.5 8.9


Figure 3.17b. Multi-annual diffuse horizontal relative bias (left) and RMSE (right). Top: McClear. Bottom: CAMS-RAD.

## 3.5.3 Direct normal irradiance



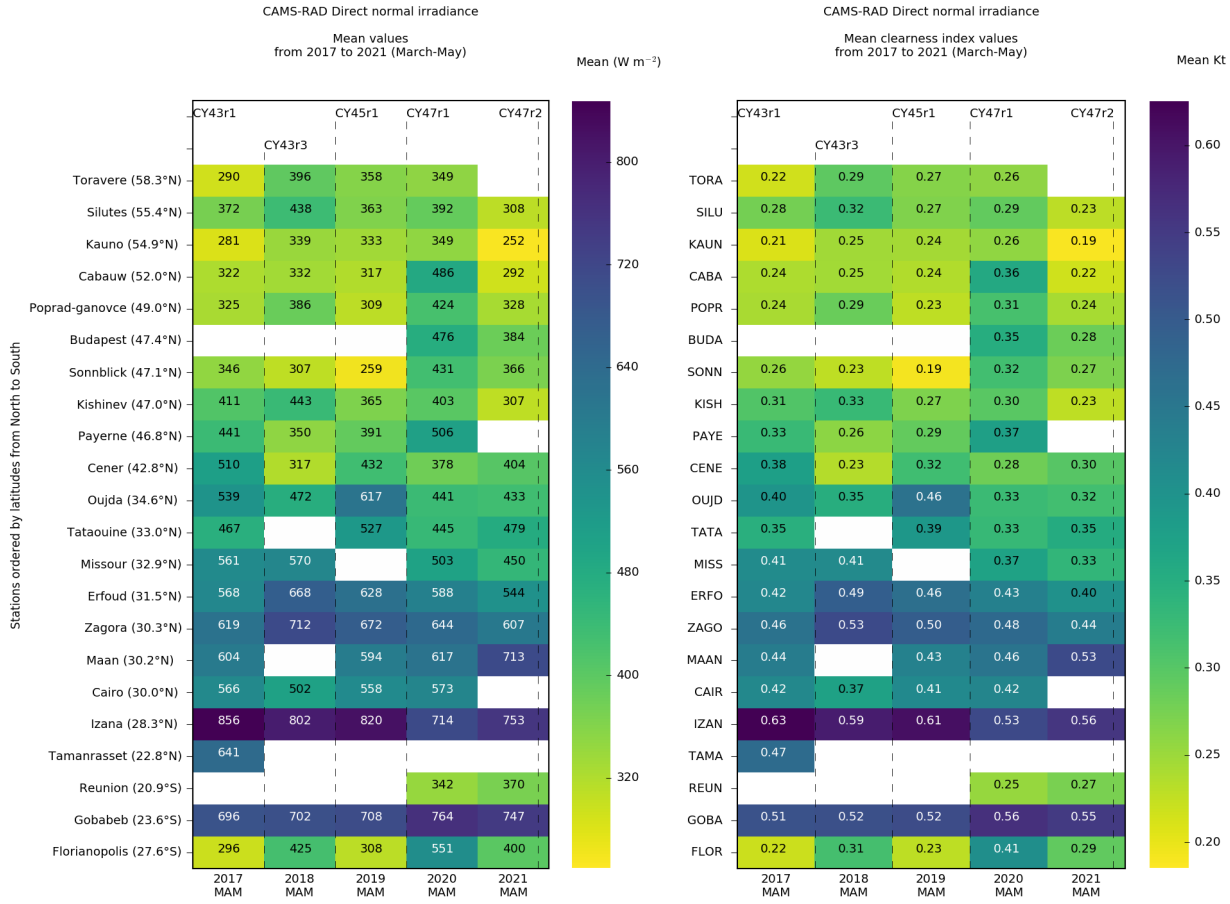



Figure 3.18a. Multi-annual means for direct normal irradiance (left) and clearness index (right). Top: McClear. Bottom: CAMS-RAD.

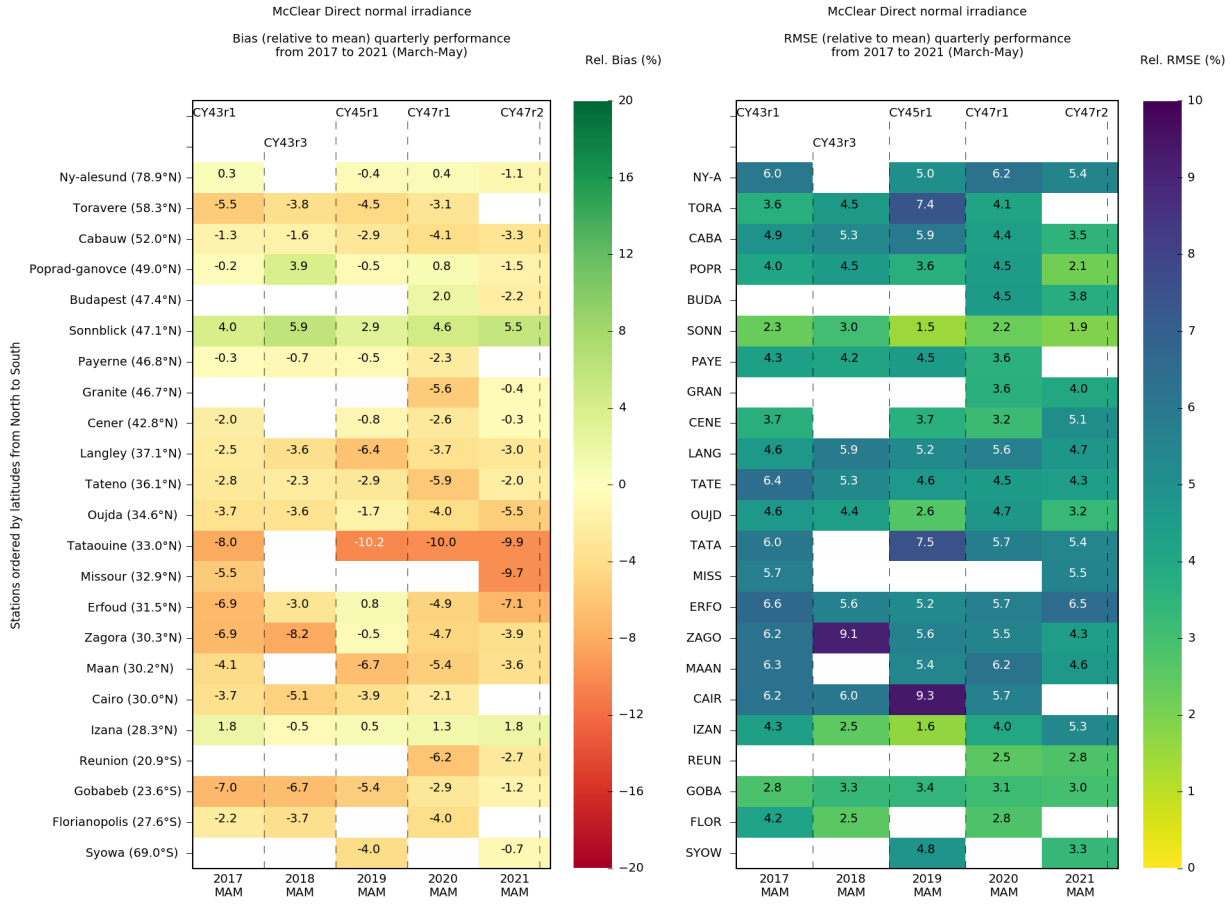

McClear Direct normal irradiance
Bias (relative to mean) quarterly performance from 2017 to 2021 (March-May)
Rel. Bias (%)
Stations ordered by latitudes from North to South
CY43r1
CY43r3
CY45r1
CY47r1
CY47r2
Ny-alesund (78.9°N) 0.3 -0.4 0.4 -1.1
Toravere (58.3°N) -5.5 -3.8 -4.5 -3.1
Cabauw (52.0°N) -1.3 -1.6 -2.9 -4.1 -3.3
Poprad-ganovce (49.0°N) -0.2 3.9 -0.5 0.8 -1.5
Budapest (47.4°N) 2.0 -2.2
Sonnblick (47.1°N) 4.0 5.9 2.9 4.6 5.5
Payerne (46.8°N) -0.3 -0.7 -0.5 -2.3
Granite (46.7°N) -5.6 -0.4
Cener (42.8°N) -2.0 -0.8 -2.6 -0.3
Langley (37.1°N) -2.5 -3.6 -6.4 -3.7 -3.0
Tateno (36.1°N) -2.8 -2.3 -2.9 -5.9 -2.0
Oujda (34.6°N) -3.7 -3.6 -1.7 -4.0 -5.5
Tataouine (33.0°N) -8.0 -10.2 -10.0 -9.9
Missour (32.9°N) -5.5 -9.7
Erfoud (31.5°N) -6.9 -3.0 0.8 -4.9 -7.1
Zagora (30.3°N) -6.9 -8.2 -0.5 -4.7 -3.9
Maan (30.2°N) -4.1 -6.7 -5.4 -3.6
Cairo (30.0°N) -3.7 -5.1 -3.9 -2.1
Izana (28.3°N) 1.8 -0.5 0.5 1.3 1.8
Reunion (20.9°S) -6.2 -2.7
Gobabeb (23.6°S) -7.0 -6.7 -5.4 -2.9 -1.2
Florianopolis (27.6°S) -2.2 -3.7 -4.0
Syowa (69.0°S) -4.0 -0.7
2017 MAM
2018 MAM
2019 MAM
2020 MAM
2021 MAM
20
16
12
8
4
0
−4
−8
−12
−16
−20
McClear Direct normal irradiance
RMSE (relative to mean) quarterly performance from 2017 to 2021 (March-May)
Rel. RMSE (%)
CY43r1
CY43r3
CY45r1
CY47r1
CY47r2
NY-A 6.0 5.0 6.2 5.4
TORA 3.6 4.5 7.4 4.1
CABA 4.9 5.3 5.9 4.4 3.5
POPR 4.0 4.5 3.6 4.5 2.1
BUDA 4.5 3.8
SONN 2.3 3.0 1.5 2.2 1.9
PAYE 4.3 4.2 4.5 3.6
GRAN 3.6 4.0
CENE 3.7 3.7 3.2 5.1
LANG 4.6 5.9 5.2 5.6 4.7
TATE 6.4 5.3 4.6 4.5 4.3
OUJD 4.6 4.4 2.6 4.7 3.2
TATA 6.0 7.5 5.7 5.4
MISS 5.7 5.5
ERFO 6.6 5.6 5.2 5.7 6.5
ZAGO 6.2 9.1 5.6 5.5 4.3
MAAN 6.3 5.4 6.2 4.6
CAIR 6.2 6.0 9.3 5.7
IZAN 4.3 2.5 1.6 4.0 5.3
REUN 2.5 2.8
GOBA 2.8 3.3 3.4 3.1 3.0
FLOR 4.2 2.5 2.8
SYOW 4.8 3.3
2017 MAM
2018 MAM
2019 MAM
2020 MAM
2021 MAM
10
9
8
7
6
5
4
3
2
1
0


Figure 3.18b. Multi-annual direct normal irradiance relative bias (left) and RMSE (right). Top: McClear. Bottom: CAMS-RAD.

## 3.6 Recommendations

As already reported in previous reports, there is room to improve the CRS for large solar zenith angles. There are several cases with large *B* values observed in the ground observations which are underestimated in CRS. The cloud analysis from Meteosat images indicates a fully cloudy pixel with cloud coverage of 100 % and a low cloud optical depth. In the previous used scheme for CRS v3.2 the cloud optical depth was set to a clipping value of 0.5 even if the retrieval provides a smaller value. This was a meaningful setting for the original application field of the used cloud retrieval package, but is not appropriate for surface radiation calculations. The exact value of the cloud optical depth plays a greater role when the sun is low above horizon which happens very often in winter at great latitude, and at the beginning and end of the day in any case. This clipping value has been removed in the current version of the cloud retrieval package and therefore in CRS v4.0.

The CRS is composed of two models (see https://atmosphere.copernicus.eu/solar-radiation and the User Guide published there). The sky model of the CRS - called McClear - exploits the CAMS service information on aerosol properties and total column content in water vapour and ozone. It provides irradiances that would be observed in cloud-free conditions. McClear generally performs well as shown by several publications (Eissa et al., 2015; Lefèvre et al., 2013, Lefèvre, Wald, 2016; Marchand et al., 2017). However, detailed analyses of the deviations for CRS reveal discrepancies that may be large also for cloud-free cases. These discrepancies may be traced back to the over- or underestimation of the occurrences of cloud-free cases or to any gross errors in aerosol conditions modelled as input to McClear. Note should be taken that there is no means in this study to discriminate the cases of underestimation of the occurrences of overcast cases and those of underestimation of the optical depth. Both cases appear as an underestimation by CRS of the frequency of low clearness indices. Similarly, there is no means to discriminate the cases of overestimation of the occurrences of medium skies cases and those of underestimation of the optical depth of the optically thick clouds or overestimation of the optical depth of the optically thin clouds. These cases appear as an overestimation by CRS of the frequency of medium clearness indices. Finally, there is no means to discriminate the cases of underestimation of the occurrences of cloud-free cases and those of overestimation of the optical depth of the optically thin clouds. These cases appear as an underestimation by CRS of the frequency of large clearness indices.

# 4 Sources of data

Measurements are taken from various sources and measuring stations that are discussed in this section. Note that a more detailed presentation is done in the ground measurement catalogue updated every year which is accessible via the CAMS help desk:

CAMS2_73_2021SC1-D1.4.1-2021Q4_ground_catalogue_202112_v2.pdf for December 2021

## 4.1 Sources of data

Efforts are made to build the quarterly validation reports with the same set of stations to better follow and monitor the quality of the irradiance products delivered by the CRS, though this is difficult as discussed later.

Measurements originate from different networks as reported in Table 4.1. They have been acquired in different time systems (UT: Universal Time, TST: True Solar Time, GMT: Greenwich Meridian Time). No change in time system is performed during this validation. The handling of the different time systems is described in the annex A describing the procedure for validation.

Table 4.1. Source of data for each station, time system and type of data (*G, B, D* stands respectively for global, direct at normal incidence and diffuse).

| Station | Source of data | Time system | Initial time integration interval | Type of data acquired |
|---|---|---|---|---|
| Aluksne | Latvian Environment, Geology and Meteorology Centre (LEGMC) | TST | 1 h | G - - |
| Kolka | | | | G - - |
| Dobele | | | | G B - |
| Riga | | | | G - - |
| Zoseni | | | | G - - |
| Liepaja | | | | G - - |
| Rucava | | | | G - - |
| Daugavpils | | | | G - - |
| Silutes | Lithuanian Hydrometeorological Service (LHMS) | UT | 1 h | G B - |
| Kauno | | | | |
| Leeuwarden | KNMI | UT | 1 h | G - - |
| Hoogeveen | | | | |
| Twenthe | | | | |
| Maastricht | | | | |
| Poprad-Ganovce | Slovak Hydrometeorological Institute (SHMI) | UT | 1 min | G B D |
| Banska-Bystrica | | | | G - D |
| Milhostov | | | | G - D |
| Kishinev | WRDC | GMT+2 | 1 h | G B D |
| Oujda | EnerMENA | UT | 1 min | G B D |
| Tataouine | | UT+1 | | |
| Fki Ben Salah | | UT | | |
| Missour | | UT | | |
| Erfoud | | UT | | |
| Cairo | | UT+2 | | |
| Zagora | | UT | | |
| Ma'an | | UT+2 | | |
| Ny-Ålesund | BSRN | UT | 1 min | G B D |
| Toravere | | | | |
| Cabauw | | | | |
| Budapest | | | | |
| Sonnblick | | | | |
| Payerne | | | | |
| Granite Island | | | | |
| CENER | | | | |
| Langley | | | | |
| Tateno | | | | |
| Izaña | | | | |
| Tamanrasset | | | | |
| Reunion island | | | | |
| Gobabeb | | | | |
| Florianopolis | | | | |
| Syowa | | | | |

## 4.2 Short description of the stations selected for the validation

The geographical coordinates of the selected stations (see maps in Figure 2.1) are given in Table 4.2.

Table 4.2. List of stations used to realize the validation report in general, and their coordinates, ordered from North to South.

| Country | Station | Latitude | Longitude | Elevation a.s.l. (m) |
|---|---|---|---|---|
| Norway (Spitzbergen) | Ny-Ålesund | 78.925 | 11.930 | 11 |
| Estonia | Toravere | 58.254 | 26.462 | 70 |
| Latvia | Aluksne | 57.440 | 27.035 | 197 |
| Latvia | Kolka | 57.743 | 22.584 | 30 |
| Latvia | Zoseni | 57.135 | 25.906 | 188 |
| Latvia | Riga | 56.951 | 24.116 | 6 |
| Latvia | Dobele | 56.620 | 23.320 | 42 |
| Latvia | Liepaja | 56.475 | 21.021 | 4 |
| Latvia | Rucava | 56.162 | 21.173 | 19 |
| Latvia | Daugavpils | 55.870 | 26.617 | 98 |
| Lithuania | Silutes | 55.352 | 21.447 | 5 |
| Lithuania | Kauno | 54.884 | 23.836 | 77 |
| The Netherlands | Hoorn | 53.393 | 5.346 | 0 |
| The Netherlands | Leeuwarden | 53.225 | 5.755 | 1 |
| The Netherlands | Hoogeveen | 52.750 | 6.575 | 16 |
| The Netherlands | Twenthe | 52.273 | 6.897 | 34 |
| The Netherlands | Cabauw | 51.972 | 4.927 | -1 |
| The Netherlands | Vlissingen | 51.442 | 3.596 | 8 |
| The Netherlands | Maastricht | 50.906 | 5.762 | 114 |
| United Kingdom | Camborne | 50.217 | -5.317 | 88 |
| Slovakia | Poprad-Ganovce | 49.035 | 20.324 | 709 |
| Slovakia | Banska-Bystrica | 48.734 | 19.117 | 427 |
| Slovakia | Milhostov | 48.663 | 21.722 | 105 |
| Hungary | Budapest | 47.429 | 19.182 | 139 |
| Austria | Sonnblick | 47.054 | 12.958 | 3109 |
| Moldova | Kishinev | 47.000 | 28.817 | 205 |
| Switzerland | Payerne | 46.815 | 9.944 | 491 |
| United States | Granite Island | 46.721 | -87.411 | 208 |
| France | Carpentras | 44.083 | 5.059 | 100 |
| Spain | CENER | 42.816 | -1.601 | 471 |
| United States | Langley | 37.104 | -76.387 | 3 |

| Country | Station | Latitude | Longitude | Elevation a.s.l. (m) |
|---|---|---|---|---|
| Spain | PSA | 37.091 | -2.358 | 500 |
| Japan | Tateno | 36.058 | 140.126 | 25 |
| Morocco | Oujda | 34.650 | -1.900 | 617 |
| Morocco | Missour | 32.860 | -4.107 | 1107 |
| Morocco | Fki Ben Salah | 32.578 | -6.622 | 485 |
| Morocco | Erfoud | 31.491 | -4.218 | 859 |
| Morocco | Zagora | 30.272 | -5.852 | 783 |
| Egypt | Cairo | 31.036 | 31.009 | 104 |
| Tunisia | Tataouine | 32.974 | 10.485 | 210 |
| Jordan | Ma'an | 30.172 | 35.818 | 1012 |
| Spain | Izaña | 28.309 | -16.499 | 2373 |
| Algeria | Adrar | 27.878 | -0.270 | 262 |
| Algeria | Tamanrasset | 22.790 | 5.529 | 1385 |
| France | Reunion island | -20.9014 | 55.4836 | 116 |
| Namibia | Gobabeb | -23.561 | 15.042 | 407 |
| Brazil | Florianopolis | -27.605 | -48.523 | 11 |
| Antarctica (Cosmonaut Sea) | Syowa | -69.005 | 39.589 | 18 |

The selected stations are located in several different climates as reported in Table 4.3. The description of climates is taken from the updated world map of the Köppen-Geiger climate classification by Peel et al. (2007).

Table 4.3. List of climates and corresponding stations

| Climate | Stations |
|---|---|
| **As**: Tropical savanna climate, dry summer | Reunion island |
| **BWh**: Arid and hot climate of desert type | Tataouine, Erfoud, Cairo, Zagora, Ma'an, Tamanrasset, Fki Ben Salah |
| **BWk**: Arid and cold climate of desert type | Gobabeb |
| **Cfa**: Temperate climate without dry season and hot summer | Florianopolis, Tateno |
| **Cfb**: Temperate climate without dry season and warm summer | Hoorn, Hoogeveen, Twenthe, Cabauw, Leeuwarden, Maastricht, Vlissingen, Camborne, Payerne, CENER |
| **Csa**: Temperate climate with dry and hot summer | Oujda, Missour |
| **Csb**: Temperate climate with dry and warm summer (Mediterranean climate) | Carpentras, PSA, Izaña, Langley |
| **Dfa**: Continental wet. Cold climate without dry season and hot summer | Poprad-Ganovce, Banska-Bystrica, Milhostov |
| **Dfb**: Continental wet. Cold climate without dry season and warm summer | Toravere, Zoseni, Riga, Dobele, Daugavpils, Liepaja, Aluksne, Kolka, Rucava, Silutes, Kauno, Kishinev, Budapest, Granite Island |
| **Dfc:** Subarctic climate | Sonnblick |
| **ET**: Polar tundra climate with no true summer | Ny-Ålesund |
| **EF**: Ice cap climate | Syowa |

Among the set of stations, are several stations, such as Toravere, which are at the edge of the field of view of the Meteosat Second Generation satellites and most likely at the edge of physical assumptions used when retrieving cloud properties. This validation is meant to include extreme cases into the station list.

One may note that though the validation aims at validating the variables *G*, *B*, and *D* delivered by the CRS, several stations are included that measure only the global irradiance *G*. They have been selected in order to check the spatial consistency of the quality of the CRS products within the same network and same climate. Figure 2.1 shows several groups of stations that are close to each other within the same climate: The Eastern Baltic area, The Netherlands, and Slovakia. One expects similar performances of CRS within a group.

It should be underlined that the validation process cannot be automated in its present form. Hence, the number of selected stations and the choice of these stations is a trade-off between the desire to cover as much as possible various climatic conditions and the amount of human resources available.

## 4.3 Evolution of the list of stations

The selection of the stations considers the results published in scientific journals or conferences. Several authors have compared McClear or CRS estimates to ground-based measurements. These publications are analysed to see if there are differences in quality when compared to our own validations.

The conclusions of these analyses may have an impact on the choice and the number of stations selected for the quarterly validation. For example, if one or more authors report performances that are dissimilar to ours in geographical areas for which we do not have stations, we will try our best to obtain measurements from stations if they follow our constraints in quality and timely access. In another example, authors may have made in-depth studies of the performances of CRS in geographical areas where we have stations, such as the Netherlands where we are using several KNMI stations. Given this background, exploiting a single KNMI station may be sufficient now to monitor the performances of CRS in this area.

## 4.4 Stations under evaluation

Some stations providing unusual observations are commented below.
From the #28 issue, some issues are not commented anymore. It concerns (see last #27 issue for details):

- The KNMI station of Hoorn was removed from the set used for the quarterly validation. This station will be reported in the next ground data catalogue for information and as a warning to users at the coastal limit.
- The enerMENA station of Missour which gave abnormal overestimated measurements of diffuse irradiances during April and June 2018. This station was under maintenance since then and resume on October 2019.

From the #31 issue, some other stations are not commented anymore (see last #30 issue for details):

- The Lithuanian stations of Silutes and Kauno. Thanks to the renovation operated in January 2020, problems have been solved, included many times shifts.

Still under evaluation are stations Fki Ben Salah, Zagora, Erfoud and Oujda in Morocco, Cairo in Egypt. Some discrepancies between estimated CRS and measurements have been reported in previous reports.
In the #32 issue, the BSRN station of Florianopolis was illustrated regarding its bad results for the SON 2020 quarter.

From #33 issue, all the figures have been updated with the version 4.0 of the CRS.

### 4.4.1 Cairo (Egypt)

During the past 2019 and 2020 quarters some bad values for the *D* and *B* components appeared at Cairo which were not eliminated with the quality check process. Contrary to the global irradiance, the diffuse and direct irradiances present very large variations in Feb.-June 2020 (Fig. 4.1). A tracker failure along with a lack of cleaning during failure period explain these anomalies which disappear by mid-June before a new tracking failure in October. According to the provider *D* and *B* are not recommended for the validation.

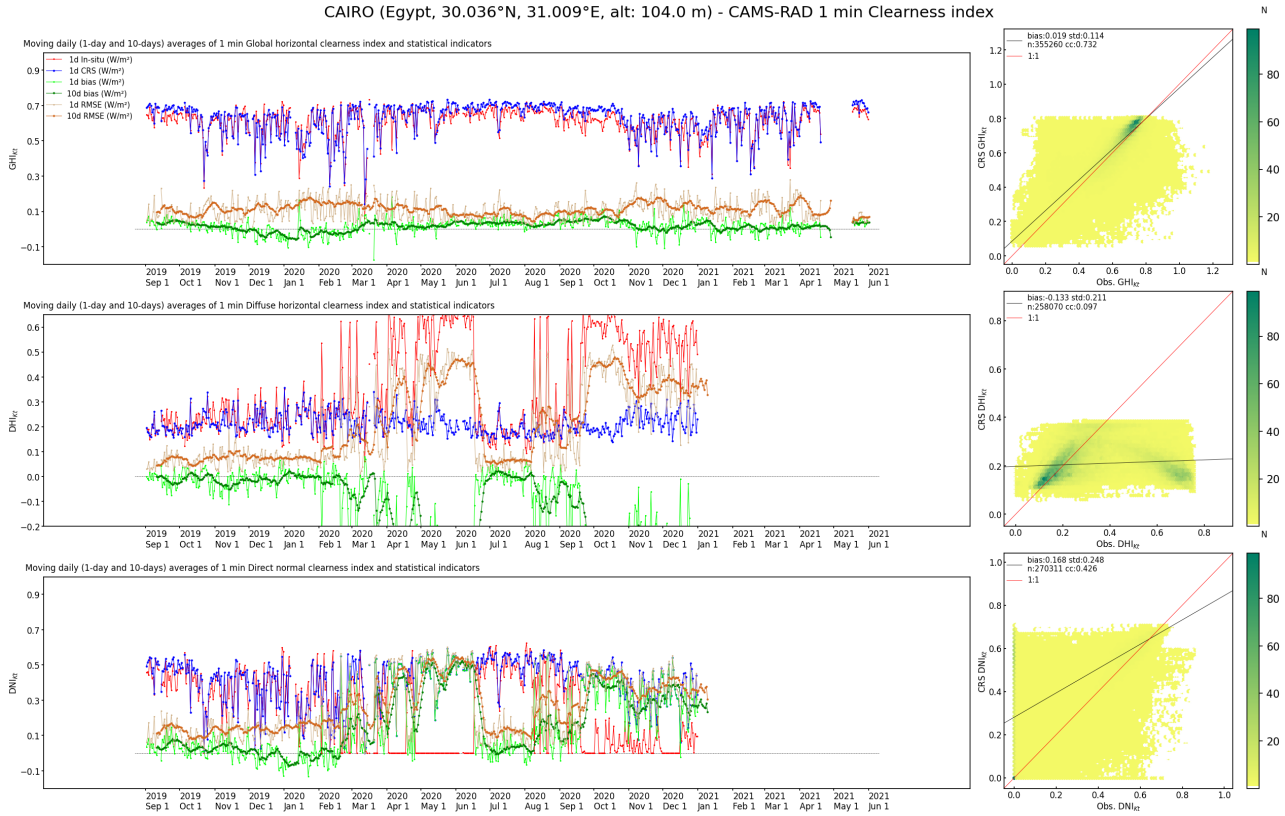



Figure 4.1. Station Cairo, daily and 10 days running average of clearness index from Sept. 2019 to May 2021.
Top: Global. Middle: Diffuse. Bottom: Direct normal clearness index.

### 4.4.2 Zagora (Morocco)

The RSI sensor has been exchanged in January 2020. A new calibration has been applied that corrects the *G* component discrepancy between measured and estimated that was observed before. Nevertheless, *B* which is calculated from *G* and *D*, presents since then and until mid-June 2020 a discrepancy with estimated CRS that was not observed before. This discrepancy re-appears since October (Fig. 4.2).

Although some discrepancies appear during March 2021 (data eliminated in April and May) for the *D* and *B* components, no problem was reported regarding the maintenance of in-situ sensors. Thus, the data have not been eliminated from the dataset.

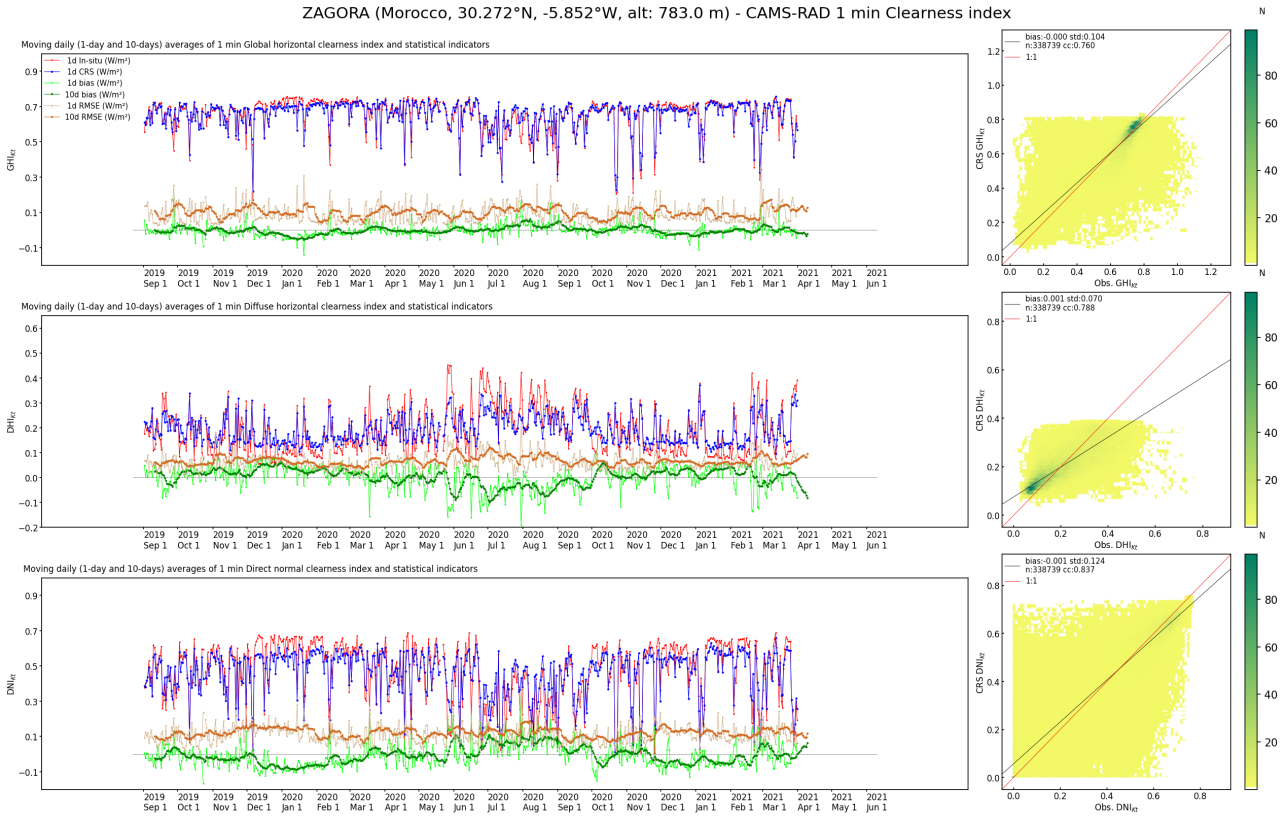



Figure 4.2. Station Zagora, daily and 10 days running average of clearness index from Sept. 2019 to May 2021.
Top: Global. Middle: Diffuse. Bottom: Direct normal clearness index.

### 4.4.3 Erfoud (Morocco)

The station Erfoud presents a decrease of *D* measurements from October 2019 to March 2020 and a sudden increase since then, with a concomitant variation of *B* in opposite way (Fig. 4.3). In this case of Erfoud, a remote site without regular cleaning, the continuously increasing deviations have been caused by dirt on the RSI sensor. Nevertheless, like Zagora, the discrepancies on *D* and *B* re-appear in October 2020 and remain until Feb. 2021 period, but disappear since then.

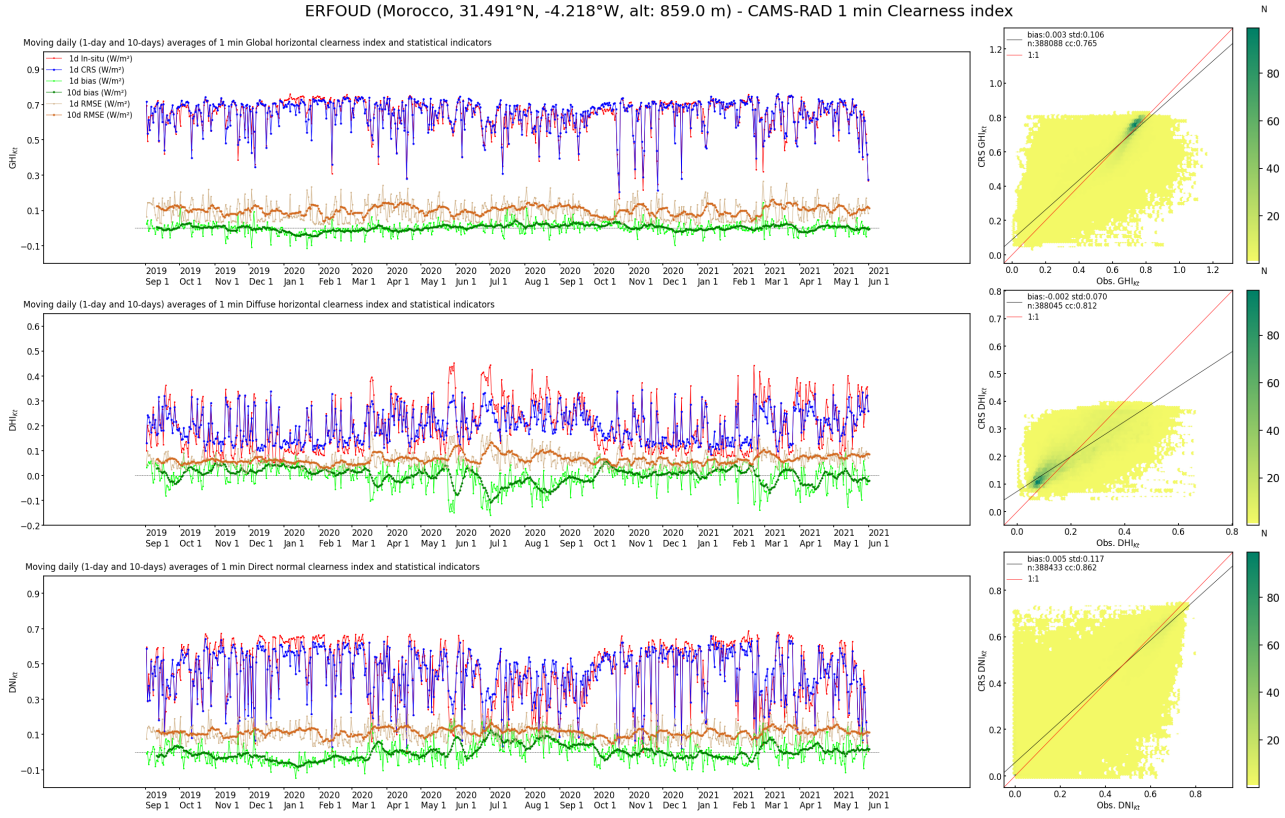



Figure 4.3. Station Erfoud, daily and 10 days running average of clearness index from Sept. 2019 to May 2021.
Top: Global. Middle: Diffuse. Bottom: Direct normal clearness index.

### 4.4.4 Oujda (Morocco)

Sensors were not cleaned during COVID-19 from March to May 2020, and a tracker failure occurred between 24 April and 3 May, and between 8-16 June which probably causes the discrepancy observed on the *D* component. No problem of maintenance was reported for March 2021, but tracker outages in April and May 2021.

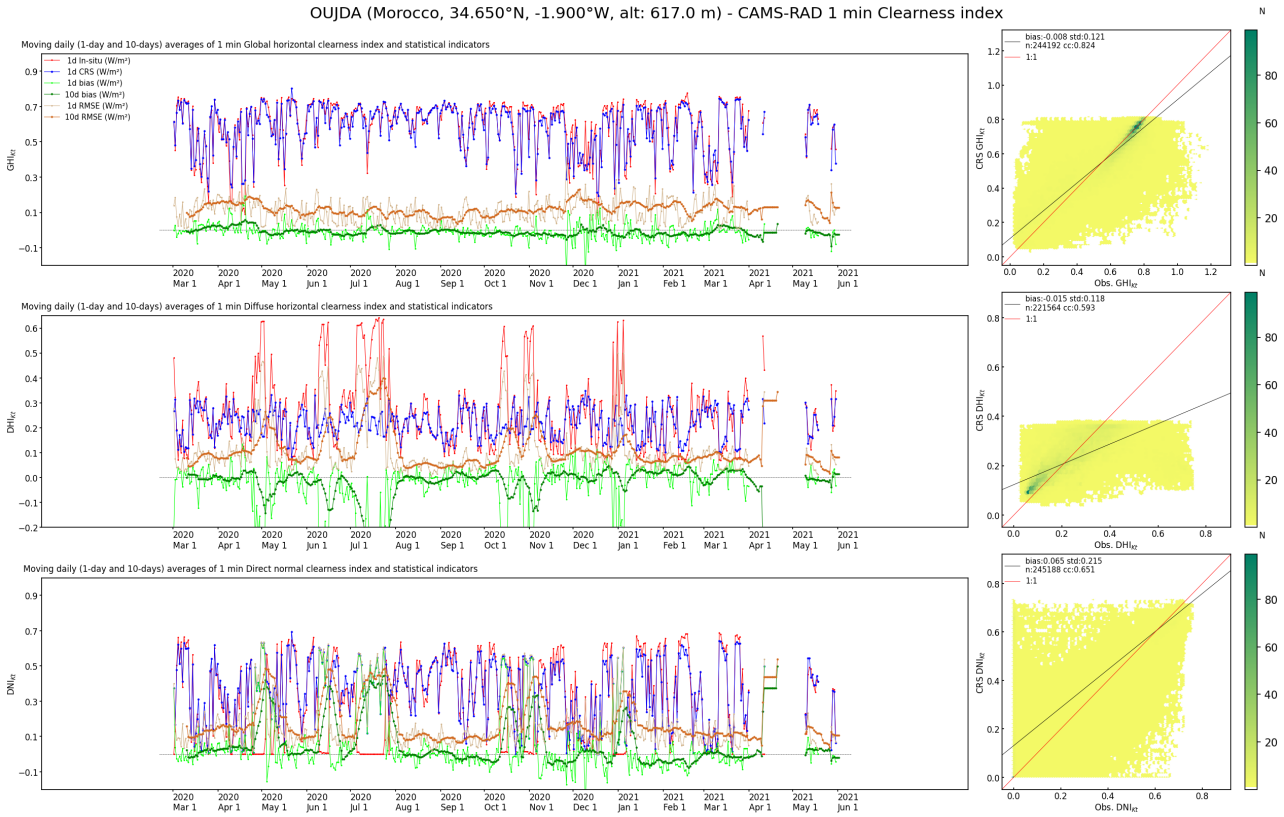



Figure 4.4. Station Oujda, daily and 10 days running average of clearness index from March 2020 to May 2021.
Top: Global. Middle: Diffuse. Bottom: Direct normal clearness index.

### 4.4.5 Fki Ben Salah (Morocco)

During 2020 pyrheliometer is sometimes misaligned in the afternoon due to erroneous mounting. This may explain the anomalies observed on fig. 4.5. Although misalignment still remains in 2021, the discrepancies tend to disappear.

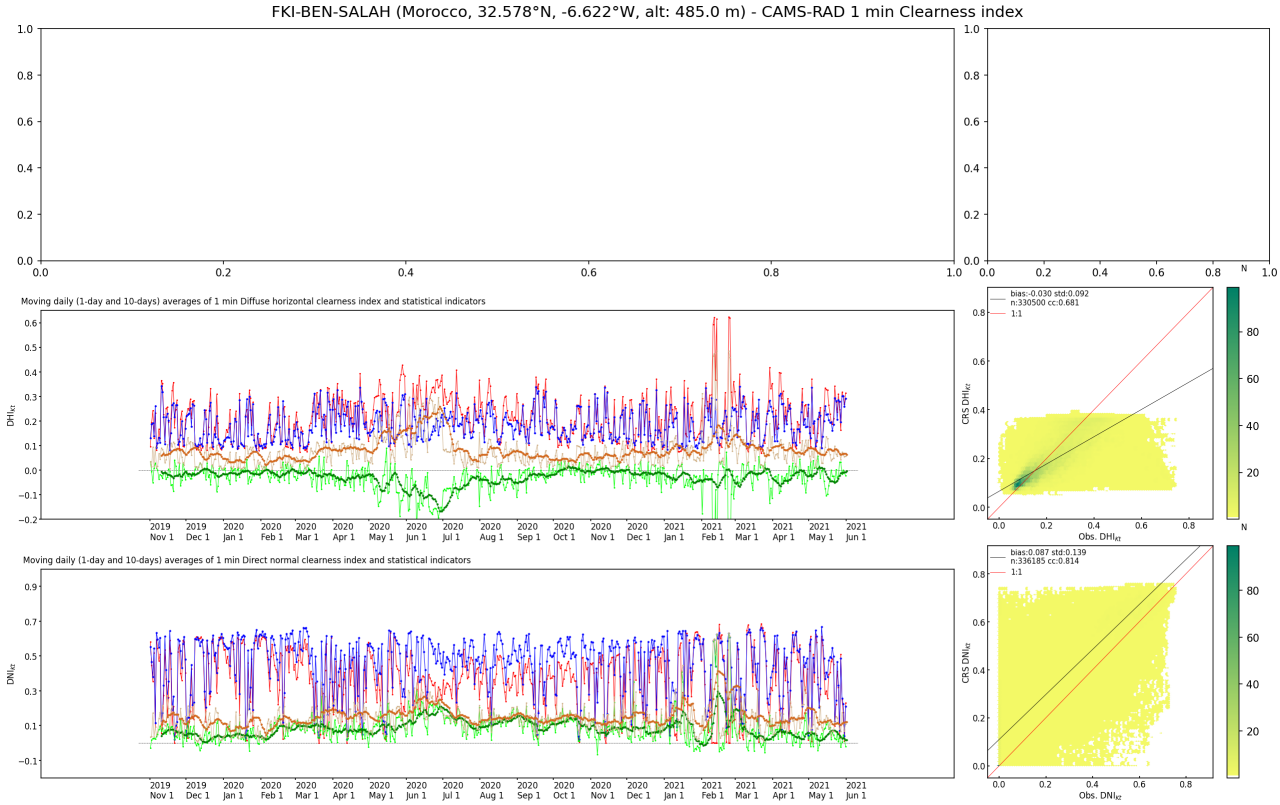



Figure 4.5. Station Fki Ben Salah, daily and 10 days running average of clearness index
From Nov. 2019 to May 2021.
Top: Global. Middle: Diffuse. Bottom: Direct normal clearness index.

### 4.4.6 Florianopolis (Brazil)

The BSRN station of Florianopolis exhibits unusual and large discrepancies during Sept. 2020 for the *D* and *B* components (Fig. 4.6). This is not observed since Oct. 2020.

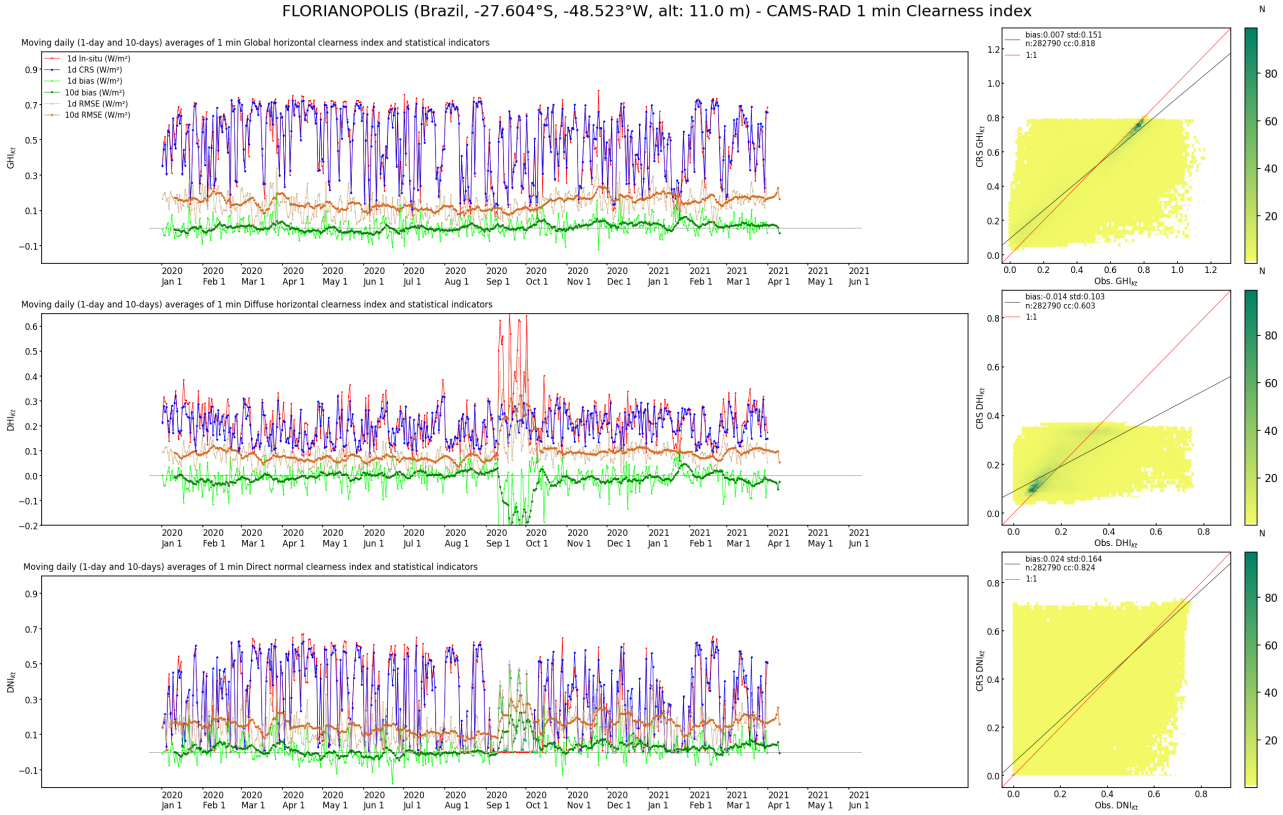



Figure 4.6. Station Florianopolis, daily and 10 days running average of clearness index From Jan. 2020 to May 2021.
Top: Global. Middle: Diffuse. Bottom: Direct normal clearness index.

# 5 Acknowledgements

The authors recognize the key role of the operators of ground stations in offering measurements of solar radiation for this validation. The authors thank all ground station operators of the Baseline Surface Radiation Network (BSRN) for their valuable measurements and the Alfred Wegener Institute for hosting the BSRN website. They also thank the Global Atmosphere Watch (GAW) program of the World Meteorological Organization (WMO) and the World Radiation Data Center (WRDC) a laboratory of the Voeikov Main Geophysical Observatory in Saint-Petersburg, Russia, for hosting the GAW website. The authors thank the University of Jordan, CRTEn and IRESEN for operating the stations of respectively Ma'an, Tataouine and Missour that belong to the EnerMENA Network as well as the German aerospace center DLR for graciously making the measurements

available, and Natalie Hanrieder for her precious information on station performance. The EnerMENA has been set up with an initial support of the German Foreign Office. The Latvian Environment, Geology and Meteorology Centre (LEGMC) and the Slovak Hydrometeorological Institute (SHMI) have kindly supplied measurements for respectively Latvia and Slovakia. The authors thank Alexandr A. Aculinin and his Atmospheric Research Group at the Institute of Applied Physics of the Academy of Sciences of Moldova for generously providing the measurements at Kishinev. Measurements for The Netherlands have been downloaded from the web site of the KNMI.

# 6 Reference documents

Blanc, P., Wald, L.: The SG2 algorithm for a fast and accurate computation of the position of the Sun. Solar Energy, 86, 3072-3083, doi: 10.1016/j.solener.2012.07.018, 2012.

Blanc, P., Espinar, B., Geuder, N., Gueymard, C., Meyer, R., Pitz-Paal, R., Reinhardt, B., Renne, D., Sengupta, M., Wald, L., Wilbert, S.: Direct normal irradiance related definitions and applications: the circumsolar issue. Solar Energy, 110, 561-577, doi: 10.1016/j.solener.2014.10.001, 2014.

Eissa, Y., Munawwar, S., Oumbe, A., Blanc, P., Ghedira, H., Wald, L., Bru, H., Goffe, D.: Validating surface downwelling solar irradiances estimated by the McClear model under cloud-free skies in the United Arab Emirates. Solar Energy, 114, 17-31, doi: 10.1016/j.solener.2015.01.017, 2015.

ISO Guide to the Expression of Uncertainty in Measurement: first edition, International Organization for Standardization, Geneva, Switzerland, 1995.

Greuell W., J. F. Meirink, and P. Wang (2013), Retrieval and validation of global, direct, and diffuse irradiance derived from SEVIRI satellite observations, J. Geophys. Res. Atmos., 118, 2340–2361, doi:10.1002/jgrd.50194.

Gschwind, B., Wald, L., Blanc, P., Lefèvre, M., Schroedter-Homscheidt, M. et al. Improving the McClear model estimating the downwelling solar radiation at ground level in cloud-free conditions–McClear-v3. Meteorologische Zeitschrift, Berlin: A. Asher & Co., 2019, 28 (2), pp.147-163. ⟨10.1127/metz/2019/0946⟩.

Korany M., M. Boraiy, Y. Eissa, Y. Aoun, M. M. Abdel Wahab, S. C. Alfaro, P. Blanc, M. El-Metwally, H. Ghedira, K. Hungershoefer, Wald, L.: A database of multi-year (2004-2010) quality-assured surface solar hourly irradiation measurements for the Egyptian territory. Earth System Science Data, 8, 105-113, doi: 10.5194/essd-8-105-2016, 2016.

Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Gschwind, B., Qu, Z., Wald, L., Schroedter-Homscheidt, M., Hoyer-Klick, C., Arola, A., Benedetti, A., Kaiser, J. W., Morcrette, J.-J.: McClear: a new model estimating downwelling solar radiation at ground level in clear-sky condition. Atmospheric Measurement Techniques, 6, 2403-2418, doi: 10.5194/amt-6-2403-2013, 2013.

Lefèvre, M., Wald, L.: Validation of the McClear clear-sky model in desert conditions with three stations in Israel. Advances in Science and Research, 13, 21-26, doi: 10.5194/asr-13-21-2016, 2016.

Marchand, M., Al-Azri, N., Ombe-Ndeffotsing, A., Wey, E., Wald, L.: Evaluating meso-scale change in performance of several databases of hourly surface irradiation in South-eastern Arabic Peninsula. Advances in Science and Research, 14, 7-15, doi:10.5194/asr-14-7-2017, 2017.

Peel, M. C., Finlayson, B. L., Mcmahon, T. A.: Updated world map of the Köppen-Geiger climate classification. Hydrol. Earth Syst. Sci. Discussions, 11(5), 1633-1644, 2007, hal-00305098.

Qu, Z., Oumbe, A., Blanc, P., Espinar, B., Gesell, G., Gschwind, B., Klüser, L., Lefèvre, M., Saboret, L., Schroedter-Homscheidt, M., and Wald L.., 2017. Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method, Meteorologische Zeitschrift, 26, 33-57, doi: 10.1127/metz/2016/0781.

Roesch, A., Wild, M., Ohmura, A., Dutton, E.G., Long, C.N., and Zhang, T.: Assessment of BSRN radiation records for the computation of monthly means, Atmos. Meas. Tech., 4, 339–354, doi:10.5194/amt-4-339-2011, 2011.

WMO: Technical Note No. 172, WMO-No. 554, World Meteorological Organization, Geneva, Switzerland, 121-123, 1981.

WMO: Guide to meteorological instruments and methods of observation, WMO-No 8, 2008 edition updated in 2010, World Meteorological Organization, Geneva, Switzerland, 2012.

# Annex A. Procedure for validation

The validation of a product is made by comparing high quality ground measurements acquired at a measuring station. These measurements are also called observations. There are several operations to perform *a)* to ensure that the observations are of sufficient quality so that they can be considered as a reference and *b)* to adapt the different time systems and samplings.

## A.1 Controlling the quality of the observations and taking care of the time system

Time series of observation at stations have been screened for their quality according to the WMO procedure (1981) with details given in Korany et al. (2016). The automated procedure checks whether the observations exceed physically possible and extremely rare limits as well as tests of consistency between the various components of the radiation whenever possible and flags them as suspicious. Then an additional visual check is performed to further remove suspicious outliers.

Observations have been acquired in different time systems (UT: Universal Time, or TST: True Solar Time). No resampling of observations is performed in the procedure for validation. In the case of observations acquired in TST system, the procedure for collecting corresponding CRS data is as follows. Given the time stamp in TST, the times in UT for the beginning and the end of the observation are computed using the SG2 library (Blanc, Wald, 2012). In parallel, the CRS data are requested with a time step of 1 min in the UT system. The corresponding CRS irradiance is computed by summing up the 1 min data for the instants comprised between the two time limits.

## A.2 Taking care of missing observations within an hour or one day

The validation may be performed at the sampling rate of the observations, e.g., every 1 min, 2 min, 10 min etc. It may be desirable to perform the validation at a time scale that is greater than the sampling rate, e.g. 1 h or 1 day. This necessitates summing up e.g. 60 observations at 1 min to yield the hourly irradiation.

Some of these observations will be flagged out by the quality check procedure. It comes out that some data is missing in a given hour and that the hourly irradiation cannot be computed with e.g. 60 observations made every 1 min within this hour but with less than 60. Hence, the sum of the valid observations is not the actual hourly irradiation; it will be equal or less.

One solution could be to reconstruct an hourly irradiation using e.g. the hourly profile of the extraterrestrial irradiation or of the irradiation in cloud-free case. This has been examined by the Task 36 “Solar Resource Knowledge Management” of the Solar Heating and Cooling Agreement of the International Energy Agency (2005-2010), which has recommended no to reconstruct hourly or daily irradiation from measurements with gaps.

The Task 36 has recommended instead constructing pseudo-hourly irradiation or irradiance by summing up the valid observations. A similar summation for the extraterrestrial irradiation is performed for exactly the same instants. This yields a pseudo-hourly extraterrestrial irradiation. The pseudo-hourly irradiation is valid only if the pseudo-hourly extraterrestrial irradiation is equal to or greater than 0.9 times the actual hourly extraterrestrial irradiation. This constraint is set to avoid extreme cases at sunrise and sunset. Invalid pseudo-hourly observations are rejected from the analysis. The same procedure applies to the daily irradiation if needed.

Pseudo-hourly irradiations from estimates are constructed in the same way.

### A.3 Pairing observations and estimates

At that stage, two data sets are available. The first one contains original observations, or pseudo-hourly or pseudo-daily irradiations, depending on the case. Only valid observations have been retained. The second one is made of the original estimates, or pseudo-hourly or pseudo-daily irradiations, depending on the case. Only valid observations and estimates have been retained in these data sets.

For the sake of the simplicity, observations and estimates, respectively, will denote either the original observations and estimates, or the pseudo-hourly or pseudo-daily irradiations.

For each instant of valid observation, an observation is paired to the estimate from the product made at the location of the station and this instant. Only pairs are kept for the validation.

### A.4 Overview of the procedure for validation

The procedure for validation comprises two parts. In the first one, differences between estimates and observations are computed and then summarized by classical statistical quantities. In the second part, statistical properties of estimates and observations are compared.

The procedure for validation applies to irradiation or irradiance, and clearness index. The changes in solar radiation at the top of the atmosphere due to changes in geometry, namely the daily course of the sun and seasonal effects, are usually well reproduced by models and lead to a de facto correlation between observations and estimates of irradiation. The clearness index is a stricter indicator of the performances of a model regarding its ability to estimate the optical state of the atmosphere. Though the clearness index is not completely independent of the position of the sun, the dependency is much less pronounced than for radiation.

### A.5 Computation of deviations and statistical quantities

This part of the present protocol of validation puts one more constraint on observations.

**Before 2020Q4 validation report (until DJF 2020):**
Since the lowest values can be noise and are therefore insignificant in a validation process, any observation should be greater than a minimum significant value. If there are not, the observations, and the corresponding estimates, are removed from the data sets and are not kept for the computation of the deviations.

The threshold is selected in such a way such that there is a 99.7 % chance that the actual irradiance is significantly different from 0 and that it can be used for the comparison. It is set to 1.5 times the uncertainty of measurements of good quality as reported by the WMO (2012).

The threshold is 30 W $m^{-2}$ (1.5 times 20 W $m^{-2}$) for the hourly (or intra-hourly) mean of global or diffuse irradiance and 7.5 W $m^{-2}$ (1.5 times 5 W $m^{-2}$) for the daily mean of global or diffuse irradiance. As for the direct irradiance at normal incidence, the threshold was set to 22 W $m^{-2}$ (1.5 times 15 W $m^{-2}$) for the hourly or intra-hourly mean and 7.5 W $m^{-2}$ (1.5 times 5 W $m^{-2}$) for the daily mean.

**December 2020 - change in the validation protocol on usage of lower thresholds:**
From #30 report (MAM 2020 dataset) onwards, no threshold on irradiance value is applied to *B* anymore. A new threshold on solar elevation (>10°) is applied on all components in order to avoid horizon obstructions and directional response of pyranometers. This new protocol is based on a recent study comparing the effect of different thresholds on the validation. The results for the *B* component are presented in the report CAMS72_2018SC2_D72.1.6.1-2020, the Service update report 2020.
From #31 report 2021Q1 (JJA 2020 dataset) onwards, no threshold on irradiance value is applied to *G* and *D* anymore, in order to be consistent with *B*. The quality check procedure used to remove all *G* under 50 W $m^{-2}$, for the reason that under this value the tests cannot be performed. That also automatically removes *D* under 15 W $m^{-2}$ approximately. Actually, it is unclear whether the data outside of the tested domain should be re-integrated into the dataset. A collaboration with the IEA PVPS task16 team for that (and other) unclear issues is foreseen. For the consistency of all components, the data outside the tested domain are re-integrated in the dataset.

Following the ISO standard (1995), the deviations are computed by subtracting observations for each instant from the estimates: deviation = estimate - observation. The set of deviations is summarized by a few quantities such as the bias or the root mean square error listed in next table. 2-D histograms between observations and estimates are drawn as well as histograms of the deviations.

Quantities summarizing the deviations

| | |
|---|---|
| Mean of measurements at station kept for validation | The mean of the measurements made at the station and kept for validation for this period. |
| Number of data pairs kept for validation | The number of couples of coincident data (CRS, ground measurements) used for validation. |
| Percentage of data pairs kept relative to the number of original measurements | The number of couples of coincident data (CRS, ground measurements) kept divided by the number of measurements available and greater than 0 from the station. |
| Bias (positive means overestimation) | The mean error for the period, i.e. the mean of the deviations. It is also equal to the differences between the mean of the CRS product and the mean of the ground measurements. The bias denotes a systematic error. Ideal value is 0. |
| Bias relative to the mean of measurements | The bias divided by the mean of measurements kept for validation, expressed in per cent. |
| RMSE | The root mean square error. Deviations are squared then averaged, and the RMSE is the root of this average. Ideal value is 0. |
| RMSE relative to the mean of measurements | The RMSE divided by the mean of measurements kept for validation, expressed in per cent. |
| Standard deviation | The bias is subtracted from each deviation. The result is squared and averaged. The standard deviation is the root of this average. It denotes the scattering of the deviations around the bias. Ideally, the standard deviation of deviations must be close to 0, and more exactly within the standard deviation of the errors of the measurements. |
| Relative standard deviation | The standard deviation divided by the mean of measurements kept for validation, expressed in per cent. |
| Correlation coefficient | The correlation coefficient between the CRS data and the ground measurements. It denotes how well the CRS product reproduces the change in measurements with time. The closer to 1 the correlation coefficient, the better the reproduction of the variability. |

Formula to compute the above-mentioned quantities

Formula

At instant $k$, observation is $x_k$ and estimate (model) is $y_k$

| | |
|---|---|
| Number of samples | $N$ pairs of coincident values $(x_k, y_k)$ |
| Mean observed value | $m_x = \frac{1}{N}\sum_{k=1}^{N} x_k$ |
| Mean of the estimates $y$ | $m_y = \frac{1}{N}\sum_{k=1}^{N} y_k$ |
| Deviation at $k$ | $\delta_k = (y_k - x_k)$ |
| Bias (mean deviation, systematic error) | $b = \frac{1}{N}\sum_{k=1}^{N} \delta_k$ |
| Relative bias | $rb = {}^{b}/_{m_x}$ |
| Root mean square error | $RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N} \delta_k^2}$ |
| Relative RMSE | $rRMSE = {}^{RMSE}/_{m_x}$ |
| Standard deviation of $\delta$ | $\sigma = \sqrt{\frac{1}{N}\sum_{k=1}^{N} (\delta_k - b)^2}$ |
| Relative standard deviation | $r\sigma = {}^{\sigma}/_{m_x}$ |
| Relation between $b$, $RMSE$ and $\sigma$ | $RMSE^2 = b^2 + \sigma^2$ |
| Standard deviation of $x$ | $\sigma_x = \sqrt{\frac{1}{N}\sum_{k=1}^{N} (x_k - m_x)^2}$ |
| Standard deviation of $y$ | $\sigma_y = \sqrt{\frac{1}{N}\sum_{k=1}^{N} (y_k - m_y)^2}$ |
| Covariance of $x$ and $y$ | $\sigma_{xy} = \frac{1}{N}\sum_{k=1}^{N} (x_k - m_x)(y_k - m_y)$ |
| Correlation coefficient | $CC = {}^{\sigma_{xy}}/_{\sigma_x \sigma_y}$ |

Assuming that the observations achieve the “moderate quality” pyranometer measurements defined by WMO (2008, rev. 2012) for hourly global radiation, one may ask if the CRS estimates are compliant with “moderate quality”. Defined as the 95 % probability (P95), the relative uncertainty for “moderate quality” should not exceed 20 %. The total uncertainty takes into account the uncertainty of observations and the uncertainty of the estimates. It can be expressed in a first approximation as the quadratic sum of both uncertainties. As a consequence, the total relative uncertainty should not exceed 28 % (P95), or 14 % (P66) if the estimates were of “moderate” quality.

## A.6 Typical uncertainty of measurements

The World Meteorological Organization (WMO, 2012) sets recommendations for achieving a given accuracy in measuring solar radiation. This document clearly states that “good quality measurements are difficult to achieve in practice, and for routine operations, they can be achieved only with modern equipment and redundant measurements.” The following Tables report the

typical uncertainty (95 % probability) that can be read in the WMO document. Uncertainties are expressed in J $m^{-2}$ in the original document. The following Tables report them in W $m^{-2}$ also.

Table A.1. Typical uncertainty (95 % probability) of measurements made by pyranometers (source: WMO 2012)

| | **Good quality** | **Moderate quality** |
|---|---|---|
| Hourly irradiation | 8 % if irradiation is greater than 0.8 MJ $m^{-2}$. Otherwise uncertainty is 0.06 MJ $m^{-2}$, i.e. 6 J $cm^{-2}$, or for irradiance approx. 20 W $m^{-2}$ | 20 % if irradiation is greater than 0.8 MJ $m^{-2}$. Otherwise uncertainty is 0.16 MJ $m^{-2}$, i.e. 16 J $cm^{-2}$, or for irradiance approx. 50 W $m^{-2}$ |
| Daily irradiation | 5 % if irradiation is greater than 8 MJ $m^{-2}$. Otherwise, uncertainty is set to 0.4 MJ $m^{-2}$, i.e. 40 J $cm^{-2}$, or for irradiance approx. 5 W $m^{-2}$ | 10 % if irradiation is greater than 8 MJ $m^{-2}$. Otherwise, uncertainty is set to 0.8 MJ $m^{-2}$, i.e. 80 J $cm^{-2}$, or for irradiance approx. 9 W $m^{-2}$ |

Table A.2. Typical uncertainty (95 % probability) of measurements made by pyrheliometers (source: WMO 2012)

| | **High quality** | **Good quality** |
|---|---|---|
| 1 min irradiation | 0.9 % 0.56 kJ $m^{-2}$, or approx. 9 W $m^{-2}$ | 1.8 % 1 kJ $m^{-2}$, or approx. 17 W $m^{-2}$ |
| Hourly irradiation | 0.7 % 21 kJ $m^{-2}$, or approx. 6 W $m^{-2}$ | 1.5 % 54 kJ $m^{-2}$, or approx. 15 W $m^{-2}$ |
| Daily irradiation | 0.5 % 200 kJ $m^{-2}$, or approx. 2 W $m^{-2}$ | 1.0 % 400 kJ $m^{-2}$, or approx. 5 W $m^{-2}$ |

## A.7 Definitions of a few quantities in solar radiation

The hourly global irradiation $G_{energy}h$ is the amount of energy received during 1 h on a horizontal plane at ground level. It is also known as hourly global horizontal irradiation, or hourly surface solar irradiation. The hourly diffuse irradiation $D_{energy}h$ is the amount of energy received from all directions of the sky vault, except that of the sun, during 1 h on a horizontal plane at ground level, and the hourly direct (or beam) irradiation $B_{energy}h$ is the amount of energy received from the direction of the sun during 1 h on this horizontal plane.

The hourly global irradiation $G_{energy}h$ is the sum of $B_{energy}h$ and $D_{energy}h$:
$G_{energy}h = B_{energy}h + D_{energy}h$

The hourly mean of global irradiance *G*h, respectively direct irradiance $B_{horizontal}$h and diffuse irradiance *D*h, is equal to $G_{energy}$h, respectively $B_{energy}$h and $D_{energy}$h, divided by 3600 s. If the irradiation is expressed in Wh $m^{-2}$, then the dividing duration is not 3600 s, but 1 h, yielding irradiance in W $m^{-2}$.

The hourly mean of direct irradiance at normal incidence *B*h is the irradiance received from the direction of the sun during one hour on a plane always normal to the direction of the sun. See Blanc et al. (2014) for more details on the definition of the direct irradiance at normal incidence and the incidence of the circumsolar radiation.

For the sake of simplicity, the notation *h* is abandoned in this text from now on. The hourly means of global and diffuse irradiances are noted *G* and *D*, and the hourly mean of the direct irradiance at normal incidence is noted *B*.

The hourly clearness index *KT* is defined as the ratio of *G* to the hourly extra-terrestrial irradiance *G0*: *KT* = *G* / *G0*. The extra-terrestrial irradiance is computed here by the means of the SG2 algorithm (Blanc, Wald, 2012). The direct clearness index and the diffuse clearness index are defined in a similar way. Because the ratio of the direct horizontal to the direct normal is equal to the cosine of the solar zenith angle at both ground level and top of atmosphere, it comes that the direct clearness index is the same than the direct normal clearness index.

## A.8 Selection of cloud-free observations

Following the recommendations of Roesch et al. (2011), only ground measurements are kept which obey the following constraints:

if $\Theta_S \leq 75°$, $1.08 \geq (D+B)/G \geq 0.92$
if $\Theta_S > 75°$, $1.15 \geq (D+B)/G \geq 0.85$
with $\Theta_S$ the solar zenith angle, D and B the diffuse and direct components of the global irradiance.

Two filters have been applied on the remaining ground measurements in order to retain reliable cloud-free instants (Lefèvre et al., 2013). The first one was a constraint on the amount of diffuse irradiance with respect to the global irradiance since the direct irradiance is prominent in case of clear-sky. Only those minutes for which D/G < 0.3, i.e. when the diffuse component is much less than the direct one, have been retained. The second filter dealt with the temporal variability of the irradiance. If there is no cloud, the sky should be clear for a long period. Checking this would avoid cases of broken clouds or noticeable spatial heterogeneity around the site if ergodicity is assumed. The first step of this filter was to retain only periods with enough measurements that have passed the first filter. A given instant t, expressed in min, was kept only if at least 30 % of the 1 min observations in both intervals [t-90, t] and [t, t+90] have been retained after the first filter. As a consequence, only are kept instants between sunrise +90 min and sunset –90 min.

## Annex B. Station validation reports for this quarter

The validation reports provide histograms of irradiances and clearness indices computed for both observations and CRS estimates, as well as monthly means and standard deviations of hourly means of irradiance for each month of the period.

This document only includes an example of the validation report for this quarter and for a single station (CABAUW). Similar reports for all stations listed in Table 2.1 are available in pdf format in the supplement zip archive:

CAMS2_73_2021SC1_D1.3.1-2021Q4_annexB_stationwise/
CAMS2_73_2021SC1_D1.3.1-2021Q4_annex_B_MAM2021_STATIONNAME.pdf

## Annex B. Station CABAUW



SOLAR RADIATION VALIDATION REPORT

**CAMS Radiation Service (CRSv4.0) - Hourly Mean of Irradiance**
**Cabauw - The Netherlands**
**Latitude: 51.972°N; Longitude: 4.927°E; Elevation a.s.l.: -0.7 m**
**From 2021-Mar to 2021-May**

This document reports on the performance of the product CAMS Radiation Service (CRSv4.0) when compared to high quality measurements of solar radiation made at the station of Cabauw from 2021-Mar to 2021-May using a standard validation protocol. Thresholds for the statistic indicators are GHI, DHI and DNI >= 0 W/m² with solar zenith angle <= 80°.

*Report generated on 2021-11-26*

---

# I. Summary of performance

*Summary of the performances of the CRSv4.0 product for Hourly Mean of Irradiance at Cabauw*

| | McClear | | | CAMS-RAD | | | |
|---|---|---|---|---|---|---|---|
| | **Global** | **Diffuse** | **Direct Normal** | **Global** | **Diffuse** | **Direct Normal** | **Unit** |
| Mean of measurements at station kept for validation | 712 | 101 | 887 | 340 | 183 | 276 | W/m² |
| Number of data pairs kept for validation | 31 | 31 | 31 | 1005 | 1005 | 1006 | |
| Percentage of data pairs kept relative to the number of data >0 in the period | 100 | 100 | 100 | 100 | 100 | 100 | |
| Percentage of data pairs kept relative to the number of valid instants in the period | 3 | 3 | 3 | 96 | 96 | 96 | % |
| Bias (positive means overestimation; ideal value is 0) | 4 | 25 | -29 | 6 | -4 | 17 | W/m² |
| Bias relative to the mean of measurements | 1 | 25 | -3 | 2 | -2 | 6 | % |
| RMSE (ideal value is 0) | 8 | 30 | 42 | 63 | 47 | 106 | W/m² |
| RMSE relative to the mean of measurements | 1 | 16 | 3 | 18 | 25 | 38 | % |
| Standard deviation (ideal value is 0) | 7 | 16 | 31 | 62 | 47 | 105 | W/m² |
| Relative standard deviation | 1 | 30 | 5 | 18 | 26 | 39 | % |
| Correlation coefficient (ideal value is 1) | 0.998 | 0.652 | 0.756 | 0.956 | 0.884 | 0.938 | |

## II. 2-D histograms (scatter density plots) - Histogram of deviations

The 2-D histogram, also known as scatter density plot, indicates how well the estimates given by CRSv4.0 match the coincident measurements on a one-to-one basis. Colors depict the number of occurrences of a given pair (measurement, estimate). In the following, yellow is used for the least frequent pairs, with blue for intermediate frequencies and blue for the highest-frequency pairs. Ideally, the dots should lie along the red line. Dots above the red line mean an overestimation. Dots below the red line denote an underestimation. The mean of the measurements, the bias, the standard-deviation and the correlation coefficient are reported. The blue line is the affine function obtained by the first axis of inertia minimizing the bias and the standard-deviation. Ideally, this line should overlay the red line. The blue line shows the trend in error when values are far off the mean of the measurements.

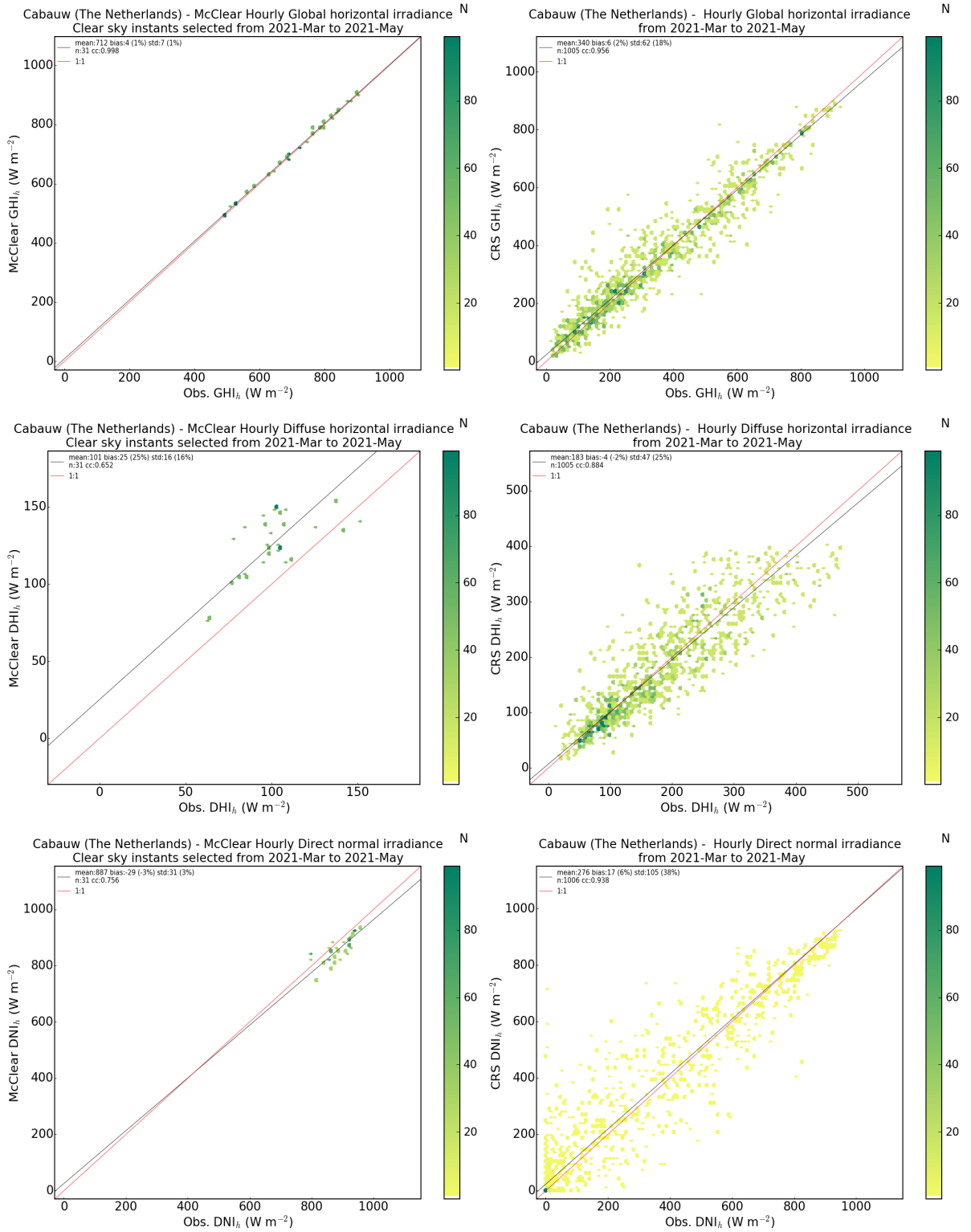
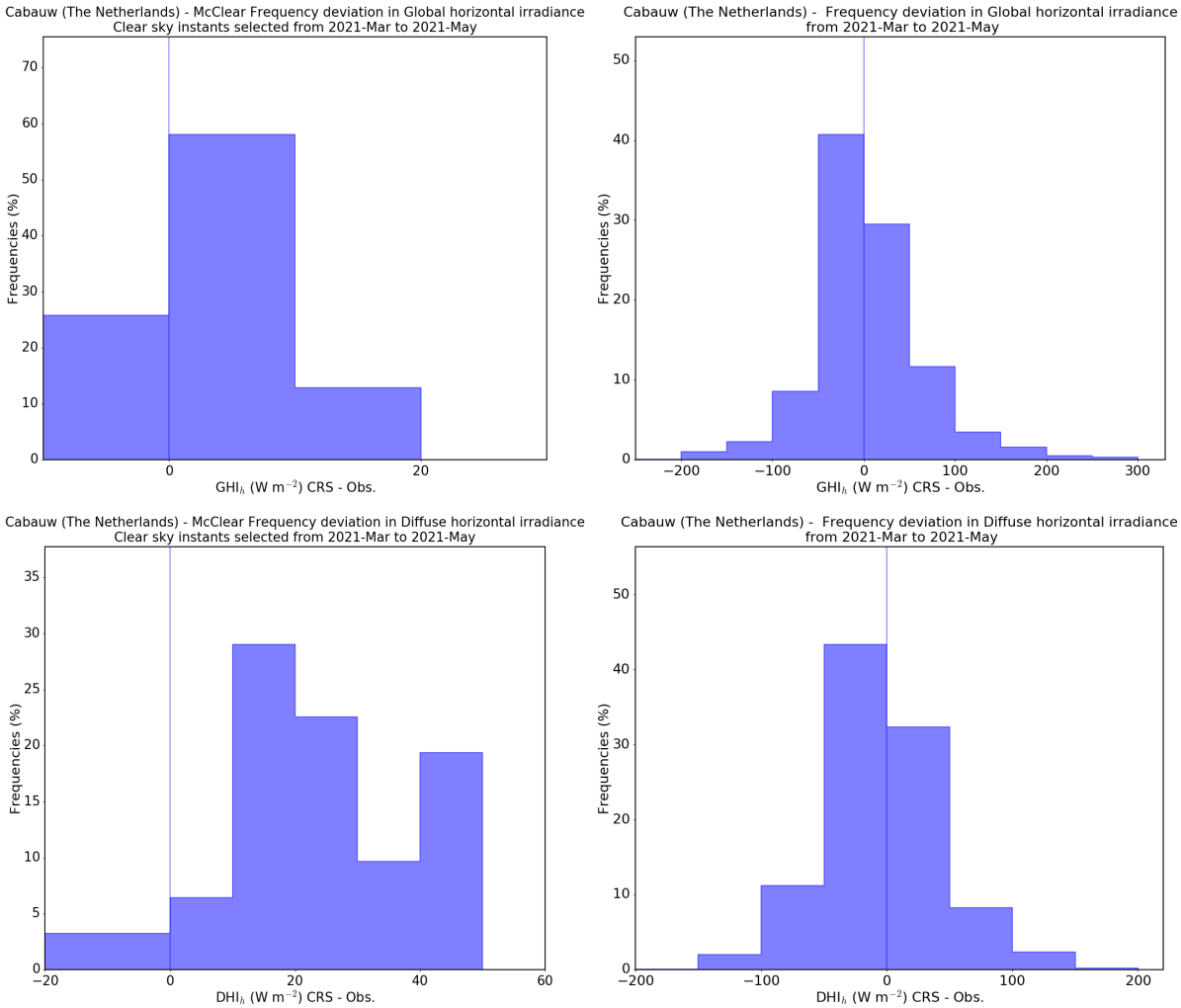



*Figure 1. 2-D histogram between ground measurements (station) and the CRSv4.0 product for Hourly Mean of Irradiance. Left: McClear. Right: CAMS-RAD.*

The histogram of the deviations, or as below the frequency distribution of the deviations, indicates the spreading of the deviations and their asymmetry with respect to the bias. Ideally, frequency should be 100% for deviation equal to 0. The more compact the frequency distribution of the deviations, the better.



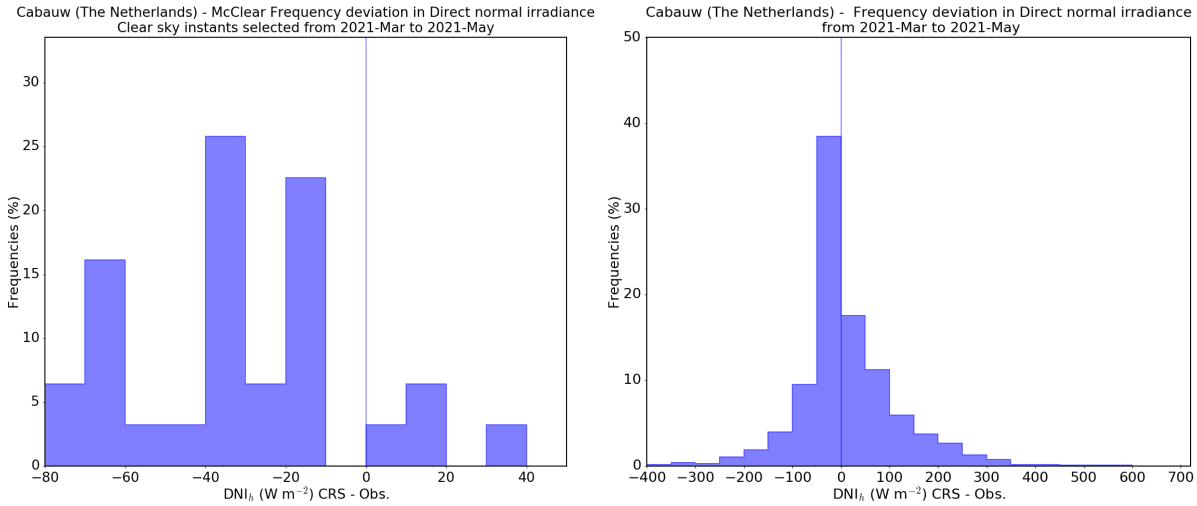



*Figure 2. Frequency distribution of the deviations (CRSv4.0 - measurements). Left: McClear. Right: CAMS-RAD.*

## III. Comparison of histograms

The graphs above deal with comparisons of measurements and CRSv4.0 values on a one-to-one basis: for each pair of coincident measurement and CRSv4.0 estimate, a deviation is computed and the resulting set of deviations is analysed.

This section deals with the statistical representativeness of the measurements by CRSv4.0. The frequency distributions of the measurements at station (red line) and the estimates (blue line) are computed and compared. A frequency distribution (histogram) shows how Hourly Mean of Irradiance values are distributed over the whole range of values. Ideally, the blue line should be superimposed onto the red one. If the blue line is above the red one for a given sub-range of values, it means that CRSv4.0 produces these values too frequently. Conversely, if the blue line is below the red one, CRSv4.0 does not produce values in this sub-range frequently enough.

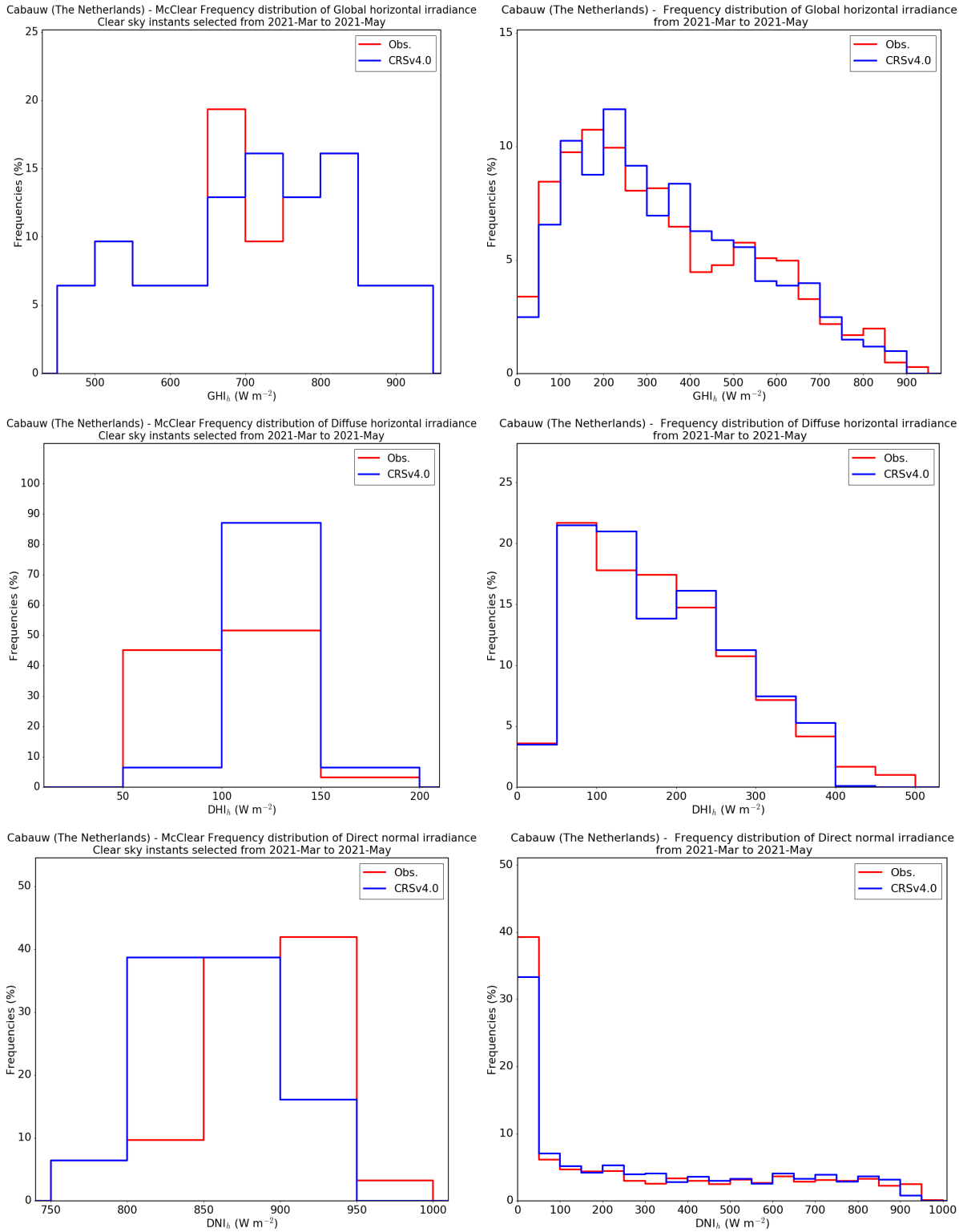
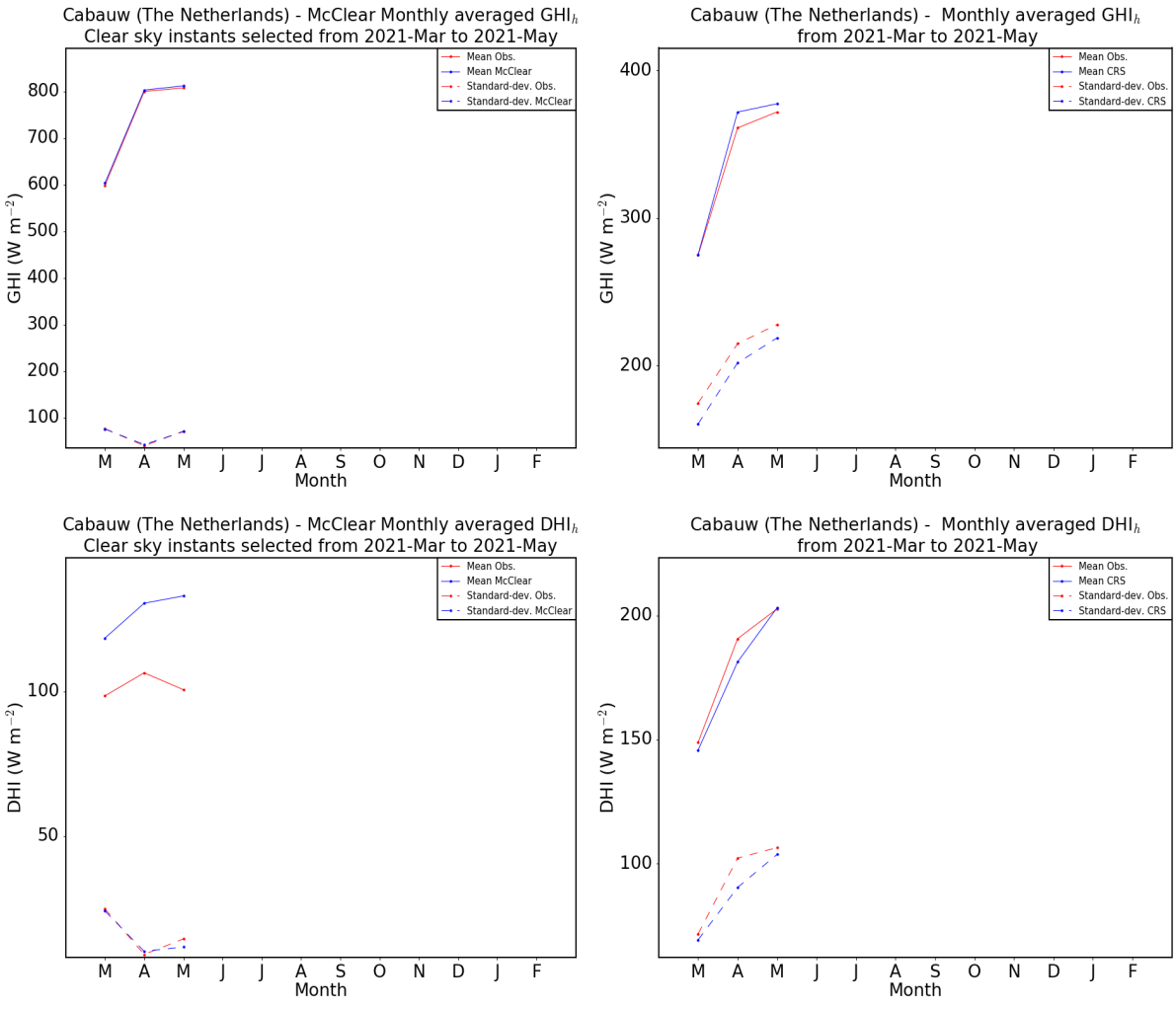



*Figure 3. Frequency distributions of the measurements station (red line) and CRSv4.0 (blue line) for Hourly Mean of Irradiance. Left: McClear. Right: CAMS-RAD.*

# IV. Comparison of monthly means and standard deviations

For each calendar month (i.e., Jan, Feb, Mar...) in the selected period, all measurements kept for validation and the coincident CRSv4.0 estimates were averaged to yield the monthly means of Hourly Mean of Irradiance and the standard deviations. The standard-deviation is an indicator of the variability of the radiation within a month. In the following graph, monthly means are shown with diamonds and standard deviations as crosses. Red color is for measurements and blue color for CRSv4.0. The closer the blue symbols (CRSv4.0) to the red ones (measurements), the better. A difference between red dot (measurements) and blue diamond (CRSv4.0) for a given month denotes a systematic error for this month: underestimation if the blue diamond is below the red dot, overestimation otherwise. For a given month, a blue cross above the red one means that CRSv4.0 produces too much variability for this month. Conversely, CRSv4.0 does not contain enough variability in the opposite case.



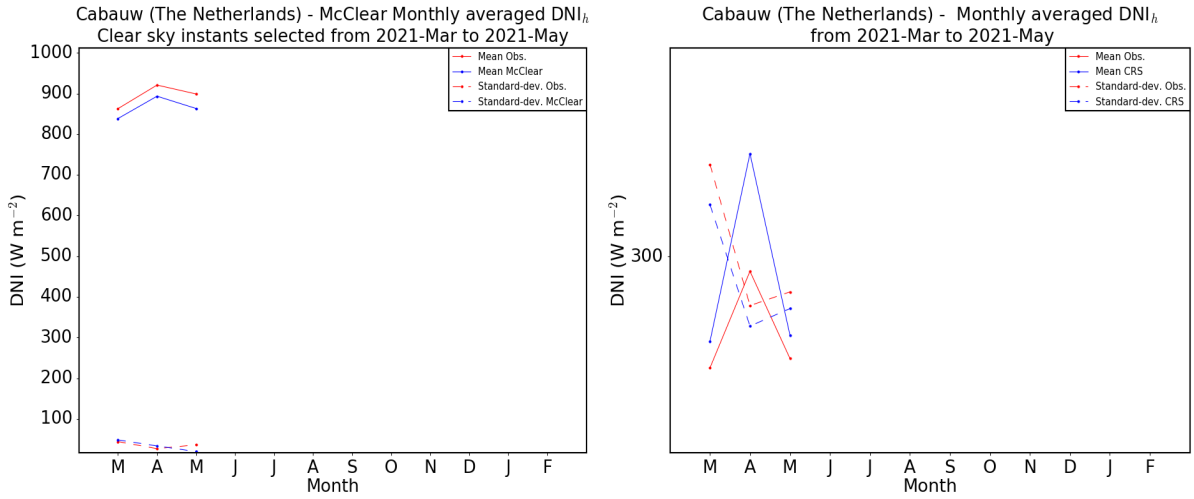



*Figure 4. Monthly means of Hourly Mean of Irradiance measurements at station (red dots) and CRSv4.0 (blue diamonds), and monthly standard-deviation of measurements (red crosses) and CRSv4.0 (blue crosses). Left: McClear. Right: CAMS-RAD.*

# V. Performances in clearness index

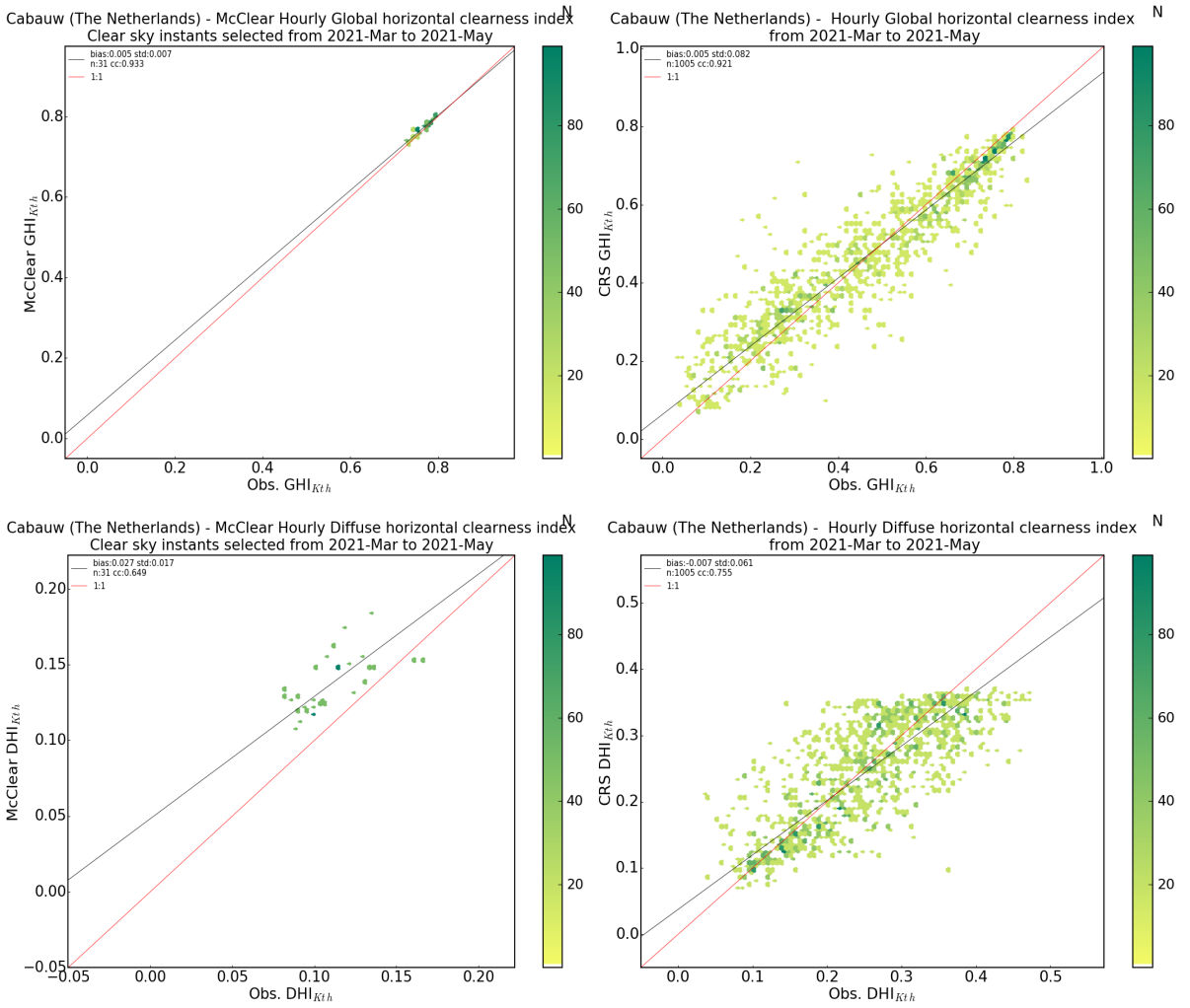
## V.1. Summary of performances

*Summary of the performance of the CRSv4.0 product for Hourly Mean of Clearness Index at Cabauw*

| | McClear | | | CAMS-RAD | | | |
|---|---|---|---|---|---|---|---|
| | **Global** | **Diffuse** | **Direct Normal** | **Global** | **Diffuse** | **Direct Normal** | **Unit** |
| Mean of measurements at station kept for validation | 0.768 | 0.110 | 0.657 | 0.457 | 0.251 | 0.204 | |
| Number of data pairs kept for validation | 31 | 31 | 31 | 1005 | 1005 | 1006 | |
| Percentage of data pairs kept relative to the number of data >0 in the period | 100 | 100 | 100 | 100 | 100 | 100 | |
| Percentage of data pairs kept relative to the number of valid instants in the period | 3 | 3 | 3 | 96 | 96 | 96 | % |
| Bias (positive means overestimation; ideal value is 0) | 0.005 | 0.027 | -0.021 | 0.005 | -0.007 | 0.012 | |
| Bias relative to the mean of measurements | 1 | 25 | -3 | 1 | -3 | 6 | % |
| RMSE (ideal value is 0) | 0.009 | 0.032 | 0.031 | 0.082 | 0.061 | 0.078 | |
| RMSE relative to the mean of measurements | 1 | 15 | 3 | 18 | 24 | 38 | % |
| Standard deviation (ideal value is 0) | 0.007 | 0.017 | 0.023 | 0.082 | 0.061 | 0.077 | |
| Relative standard deviation | 1 | 29 | 5 | 18 | 24 | 38 | % |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Correlation coefficient (ideal value is 1) | 0.933 | 0.649 | 0.813 | 0.921 | 0.755 | 0.938 | |

## V.2. 2-D histograms (scatter density plots) - Comparison of histograms



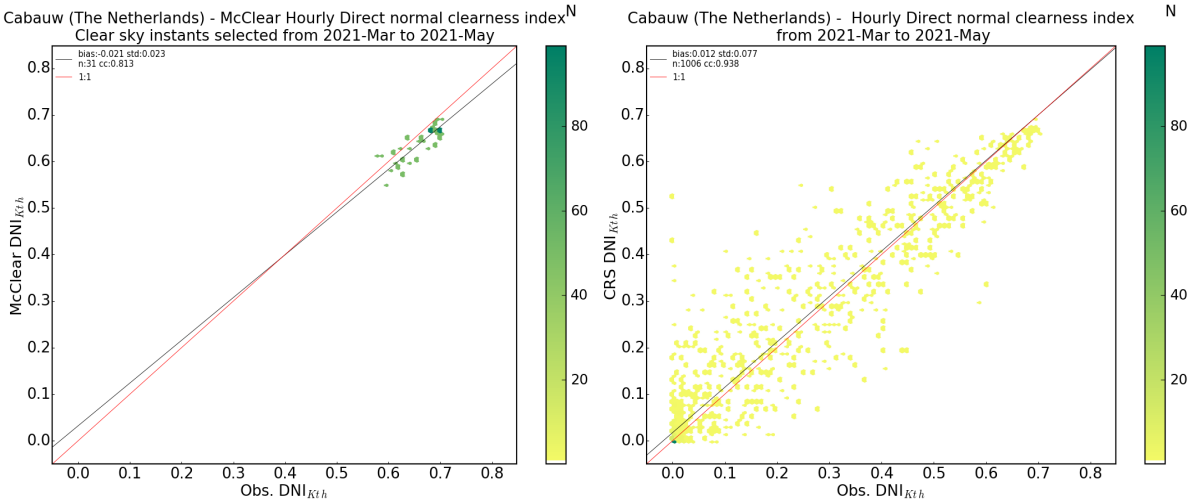
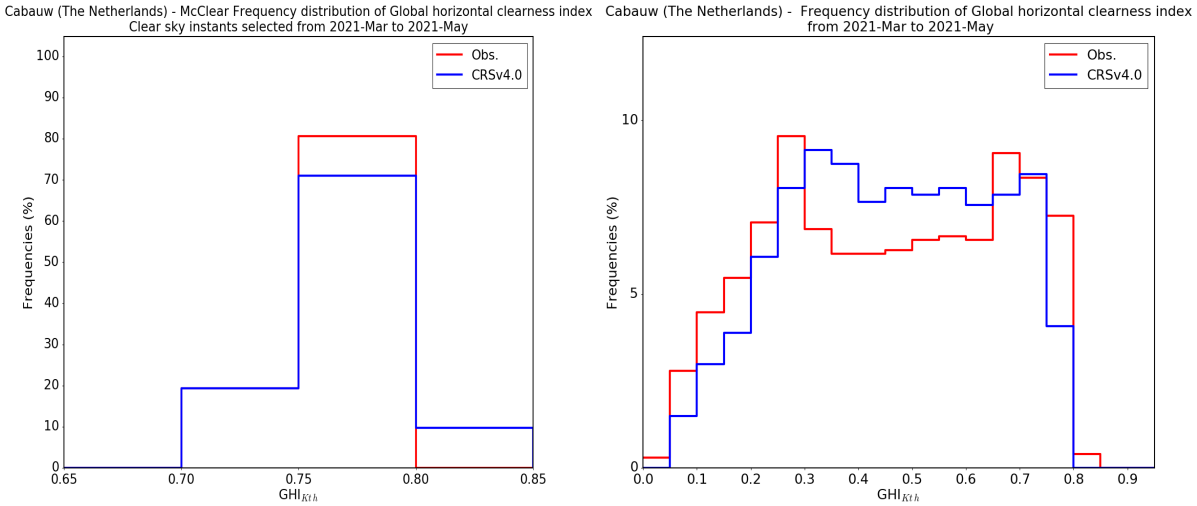



*Figure 5. 2-D histogram between ground measurements (station) and the CRSv4.0 product for Hourly Mean of Clearness Index. Left: McClear. Right: CAMS-RAD.*



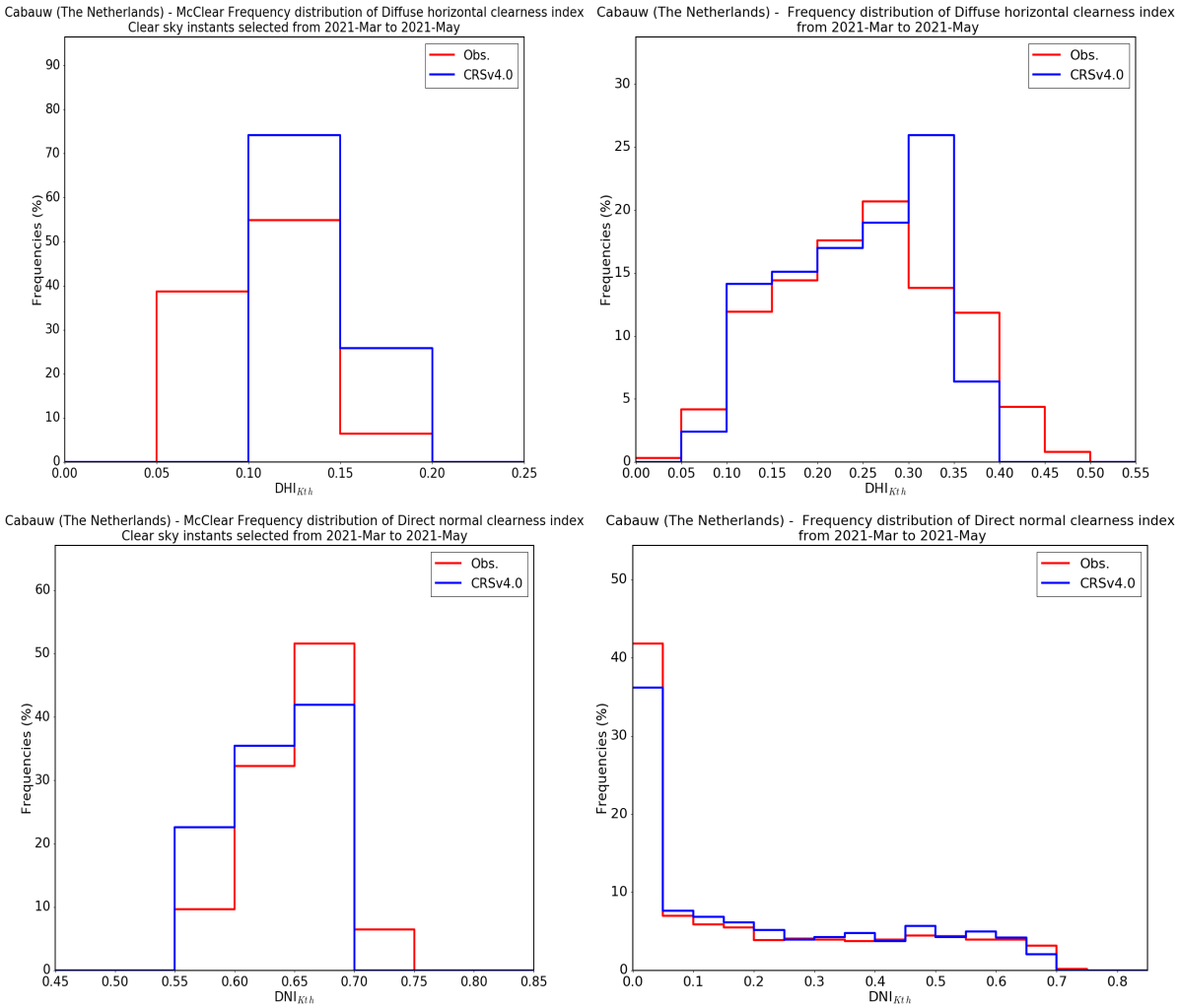



*Figure 6. Frequency distributions of the measurements station (red line) and CRSv4.0 (blue line) for Hourly Mean of Clearness Index. Left: McClear. Right: CAMS-RAD.*

Copernicus Atmosphere Monitoring Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: info@copernicus-atmosphere.eu