CAMS Service Evolution

D4.1 – Spatially resolved and regionalized radiation uncertainty

Due date of deliverable	31/12/2024
Submission date	17/12/2024
File Name	CAMEO-D4-1-V1
Work Package /Task	Task 4.1
Organisation Responsible of Deliverable	ARMINES
Author name(s)	Yves-Marie Saint-Drenan, Jorge Lezaca with contributions from Marion Schroedter-Homscheidt
Revision number	1
Status	Issued
Dissemination Level	PUBLIC

Funded by the European Union

The CAMEO project (grant agreement No 101082125) is funded by the European Union.
Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them.

1 Executive Summary

This report describes the activities conducted in the task 4.1 entitled “Uncertainty of Solar Radiation” during the first 24 months of the project.

The initial efforts focused on the collection of in-situ solar radiation data to extend the CRS reference dataset with special focus on creating a spatially dense extension as basis for the different activities of WP4. Besides the collection itself, particular attention was paid to the formatting of the collected measurement data to meet FAIR data principles and to facilitate the data access among researchers involved in the CAMEO project and CRS teams.

A very important aspect for ensuring the relevance of the validation activities conducted in CRS and CAMEO is to be able to guarantee the quality of the reference data used. Since dense networks as evaluated in CAMEO only measure global solar irradiance as a single radiation component, it is impossible to apply the standard quality control procedure requiring the availability of global, diffuse and direct irradiances. For single component observations there is no suitable QC procedure available in the literature, Therefore, it was necessary to develop a quality control procedure specific to stations measuring only global irradiance. The proposed quality control procedure is based on the inspection of the plausibility of the measurements in comparison to sun position and a cloud-free solar radiation estimate. A tool allowing an efficient visual inspection was developed to enable efficient station handling. This procedure has been applied to measurements from more than 300 stations.

An analysis of the spatial patterns and seasonality of the error of CRS was conducted using the dense reference network as well as a few high quality BSRN reference stations located in the region of interest (France and Southwestern Germany). This analysis included a critical analysis of the reference measurements since the proposed GHI-based quality control procedure was not able to eliminate all uncertainties in the reference measurements. It was shown that most of the spatial patterns observed in monthly error statistics can be explained well by only two explanatory variables: the mean solar irradiance and the station altitude. This conclusion is important as it can be a useful information for the uncertainty quantification planned in the following tasks of WP4. It is also a precious indication for future improvements of CRS.

Table of Contents

1 Executive Summary 2
2 Introduction 4
2.1 Background 4
2.2 Scope of this deliverable 4
2.2.1 Objectives of this deliverables 4
2.2.2 Work performed in this deliverable 5
2.2.3 Deviations and counter measures 5
3 Overview of the CAMS Radiation Service (CRS) 6
4 Solar radiation data collection 7
4.1 Extension of CAMS reference radiation dataset 7
4.2 Preparation of a dataset with a high spatial density of stations 9
5 Quality control of solar radiation measurements 10
5.1 Stations measuring global, direct and diffuse irradiances 10
5.2 Stations measuring only global irradiances 11
6 Analysis of the performances of CRS 15
6.1 Methodology 15
6.2 Analysis of the spatial characteristics and seasonality of CRS error 18
6.2.1 Bias in confidentially cloud-free conditions 18
6.2.2 Standard deviation of error in confidentially cloud-free conditions 19
6.2.3 Standard deviation of error in confidentially cloudy conditions 20
6.2.4 Bias in confidentially cloudy conditions 22
6.3 Influence of the mean irradiance on CRS performances 23
6.4 Influence of the station elevation on the performances of CAMS Rad 28
6.5 Joint dependence of elevation and mean solar irradiance 29
6.6 Quantification of the influence of identified features on the spatial and seasonal patterns of CRS error 31
7 Conclusion and discussion 32

2 Introduction

2.1 Background

The provision of solar radiation estimates based on CAMS atmospheric content and cloud optical properties is an important service for the solar energy community. The integration of CAMS radiation service into the pvlib library (Holmgren et al. 2018, Anderson et al. 2023, Jensen et al. 2023) has confirmed this growing interest and significantly increased the number of users. It is therefore particularly important to meet the expectations of CAMS Radiation Service (CRS) users by improving the accuracy of radiation estimates while providing uncertainty in the estimates provided.

CRS performance is assessed in the regular CRS-EQC reports. These assessments are based on an established and recognised methodology and a validated solar radiation reference database of scientific quality. Nevertheless, the objectives of improving the CRS service and quantifying uncertainty with additional detail require in-depth analyses that go beyond the scope of regular CRS EQC activities.

One of the aims of the CAMS Service Evolution (CAMEO) project is to improve the quality of CRS by carrying out in-depth evaluations to better characterise the performance of solar radiation estimation, identify improvements and quantify its uncertainty.

2.2 Scope of this deliverable

2.2.1 Objectives of this deliverables

Identifying areas for improvement and quantifying the uncertainty of CAMS Radiation Service requires a detailed understanding of its current error patterns and properties. The regular EQC reports enable this error to be quantified with a high degree of accuracy, but the fact that the measurement stations used are sparse and very far apart from each other limits our understanding of the error. Indeed, the different stations are subject to very different meteorological situations and climatic conditions. It is often difficult to compare the evaluations made for the different stations to assess what proportion of the error is due to the specific situation of the station and what proportion represents a deficiency in the model that it would be important to correct.

Using a dense network of reference stations for the validation allows to better characterize the error structure of CRS since neighbouring stations are in a simar climate zone, are exposed to similar meteorological conditions, or have comparable orography. In addition, this makes it possible to assess changes in performance on an almost spatially continuous basis, providing very important information for understanding service improvement needs and quantifying service uncertainty.

To meet this objective, the work described in this deliverable aims to build up a dense reference radiation database, and to analyse the spatial patterns and seasonality of the CRS data.

2.2.2 Work performed in this deliverable

In this deliverable, the work as planned in the Description of Action (DoA, WP4 T4.1.1 and T.4.1.2) was performed. The efforts have been focusing on the following aspects:

- extends the CAMS ground observation database on broadband global and direct irradiances by new data sources (section 3.1)
- collection and preparation of a dense reference dataset for the need of CAMEO WP4 activities (section 3.2)
- Quality control of the data including GHI-only measuring stations and share of the collected dataset following the FAIR principle (section 4)
- Analysis of the spatial characteristics and seasonality of CRS error (sections 5.1 and 5.2)
- Preliminary investigation of explanatory variables on the spatial dependency of CRS error (sections 5.3 to 5.5)
- Evaluation of the potential of new data sources for future operation CAMS Evaluation and Quality Control (EQC) reports (section 6)

These six points will be described and discussed in the next sections.

2.2.3 Deviations and counter measures

No deviations have been encountered.

3 Overview of the CAMS Radiation Service (CRS)

The CAMS solar radiation service (CRS) provides historical values (2004 to present) of global, direct and diffuse solar irradiation, as well as of direct normal irradiation. The aim is to fulfil the needs of European and national policy developments and the requirements of (partly) commercial downstream services, e.g. for planning, monitoring, efficiency improvements, and the integration of solar energy systems into energy supply grids.

This service is based on the Heliosat-4 method, which is a fully physical model using a fast, but still accurate approximation of radiative transfer modelling and is therefore well suited for geostationary satellite retrievals. It is composed of two models based on abaci, also called look-up tables: the McClear model calculating the irradiance under cloud-free conditions and the McCloud model calculating the extinction of irradiance due to clouds. Both have been realized by using the libRadtran radiative transfer model.

The main inputs to McClear are aerosol properties, total column water vapour and ozone content as provided by the Copernicus Atmosphere Monitoring Service (CAMS) every 3 h. Cloud properties used in McCloud are derived from images of the Meteosat Second Generation (MSG) satellites (and soon from Himawari) in their 15 min temporal resolution using an adapted APOLLO-NG (AVHRR Processing scheme Over cLouds, Land and Ocean – New Generation) scheme.

A schematic structure taken from the original paper of Qu et al. (2016) summarizing the Heliosat-4 method and its McClear and McCloud components is given in Figure 1.

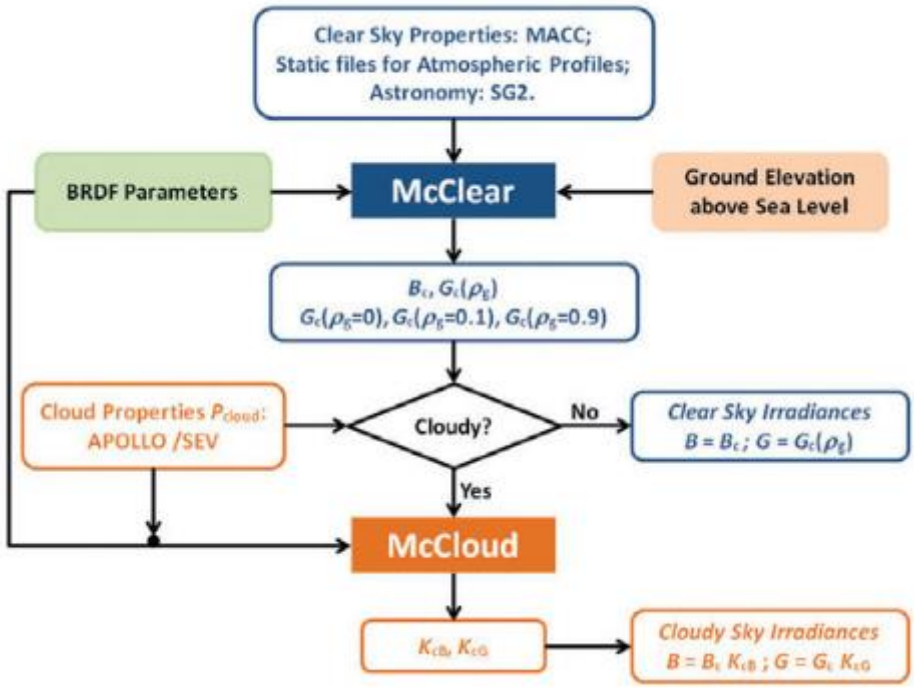

Figure 1: Schematic structure of the Heliosat-4 method used in the CAMS Radiation Service

4 Solar radiation data collection

The objective of task T4.1.1 is to extend the CAMS ground observation database on broadband global, diffuse and direct irradiances with new data sources to have a better spatial coverage of the reference data.

Two aims are followed in this task. The first objective is to extend the CAMS reference radiation dataset used for the regular CRS EQC. Indeed, the evaluation of the quality of CRS, measurements from scientific high-quality networks that are continuously updated are needed. Since the cost and maintenance constraints of high-quality stations are high, such stations are sparse and don't have the spatial density needed to evaluate spatial patterns of CRS errors at regional scale. The second objective of this task is therefore to build a dataset with a very high spatial density of stations allowing to understand the spatial patterns of the error of CAMS radiation data and identify necessary improvements of the CAMS radiation processing chain in future CAMS evolution activities.

4.1 Extension of CAMS reference radiation dataset

The collection work initiated in recent years as part of the CAMS Radiation Service (CRS) has been continued to extend the CAMS radiation reference dataset. Solar radiation measurements of the three components acquired with ISO 9060 class A instruments (secondary standard) with a time resolution of 1 min are continuously being sought. In addition, an important characteristic for inclusion in the CAMS radiation reference dataset is that the network is in operation and that the available measurements are regularly updated. The SURFRAD, SAURAN, Skynet and SOLRAD networks meet these conditions and have been included in the reference dataset as part of this task.

One-minute measurements were collected for all stations of these new networks. The collected data were converted into a Climate and Forecast (CF) compliant netCDF format, including all important metadata. The generated netCDF files have been uploaded to our Thredds data server (TDS) (http://tds.webservice-energy.org/thredds/in-situ.html) to allow the CAMS Radiation and CAMEO teams to access the data.

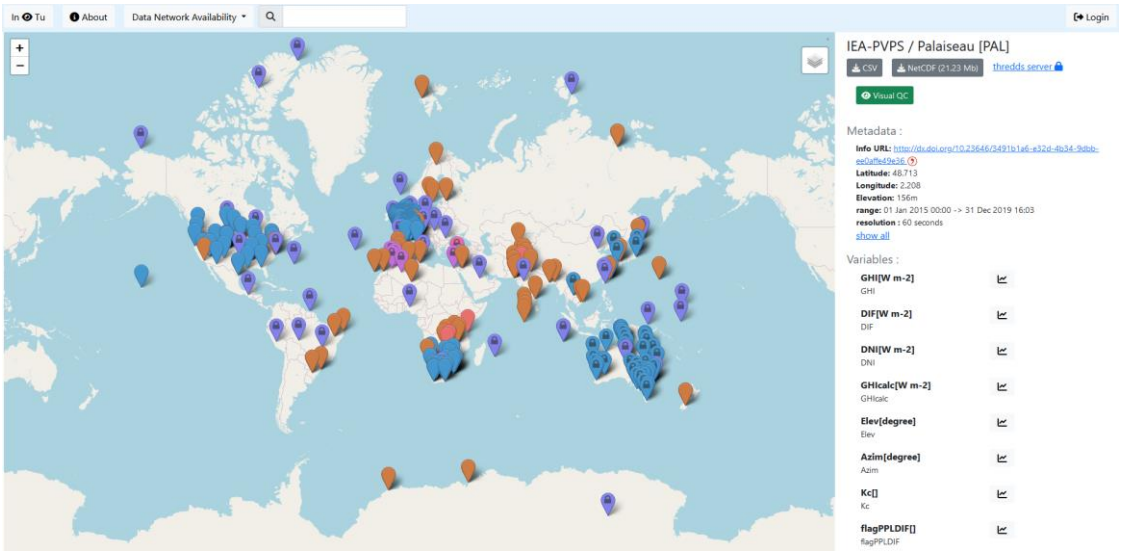

Figure 2: Screenshot of the online infrastructure illustrating the spatial distribution of the station for which solar radiation measurements are collected (https://viewer.webservice-energy.org/in-situ/#) where each colour represents a different network.

This process aims to facilitate the cooperation among researchers involved in the development, operation and validation of CAMS Radiation Service, while ensuring that our data management complies with the FAIR principles. The current content of the TDS is illustrated in Figure 2 and Figure 3 representing respectively the online access of the TDS and the spatial web-viewer coupled to the data server (https://viewer.webservice-energy.org/in-situ/#). The list of stations by networks is summarized with their most important features in Table 1.

Table 1: Overview of the stations for which solar radiation measurements are collected for the CAMS Radiation Service (white background) and the CAMEO project (light grey background for stations contribution to the EQC and dark grey background for networks with a high spatial density of stations)

Network	Number of stations	Region	Quantities measured	Period	Network in operation	Automatic update
BOM	21	Australia	G, B, D	1994-2021	Yes	No
BSRN	77	Worldwide	G, B, D	1992-2024	Yes	Yes
Ener-MENA	12	MENA region	G, B, D	2003-2024	Yes	partly
ESMAP	16	Worldwide	G, B, D	2014-2018	No	No
NREL-MIDC	12	USA	G, B, D	1992-2024	Yes	Yes
SAURAN	21	South Africa	G, B, D	2011-2024	Yes	Yes
Skynet	6	Worldwide	G, B, D	2010-2024	Yes	Yes
SOLRAD	9	USA	G, B, D	2002-2024	Yes	Yes
SURFRAD	7	USA	G, B, D	1995-2024	Yes	Yes
ISE-PVlive	40	Germany	G, G tilted	2020-2024	Yes	Yes
Météo-France	266	France	G	2015-2023	Yes	No

Figure 3: Screenshot of the Thredds data server where the collected measurements are stored and shared (http://tds.webservice-energy.org/thredds/in-situ.html)

4.2 Preparation of a dataset with a high spatial density of stations

To maximize the spatial density of the reference stations, the conditions set on the requirements for the integration of new stations to the CAMS Radiation reference dataset needed to be relaxed. Stations were sought even if they did not meet the strict requirements for instrument accuracy and maintenance or are not regularly updated. With these relaxed criteria, we have identified two networks of interest for the detailed analysis planned in CAMEO: the irradiation network operated by Météo-France, which comprises 266 stations in France, and ISE PVlive, which comprises 40 stations in south-west Germany. These two networks contain 1-min measurements but only GHI is available.

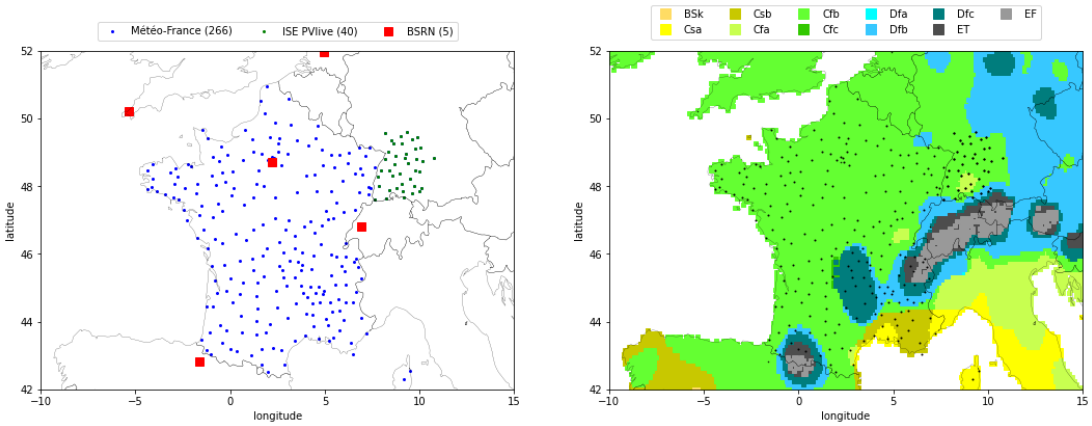

Figure 4: Spatial distribution of the Météo-France, ISE-PVLive and BSRN stations used for the analysis of the spatial and seasonal patterns of the error of CRS (left) and comparison of the location of the stations with the Köppen-Geiger climate zones (right)

The spatial distribution of the stations of the Météo-France and ISE PVlive networks is represented by blue and green squares in the left map of Figure 4. The location of BSRN stations in the domain covered by these two networks is represented by red squares. A comparison of the two networks with the BSRN stations shows that the two networks have a significantly higher spatial density, allowing to assess the spatial transition of the performances of CRS in-between BSRN stations. The locations of the stations are overlaid to a map of the Koppen Geiger climates (Peel et al., 2007) in the right map of Figure 4. It shows that many stations share a similar climate zone, with mainly difference around the mountainous regions of France (Pyrenees, the Alpes and the central massif) and the southeastern part of France along the Mediterranean coastline.

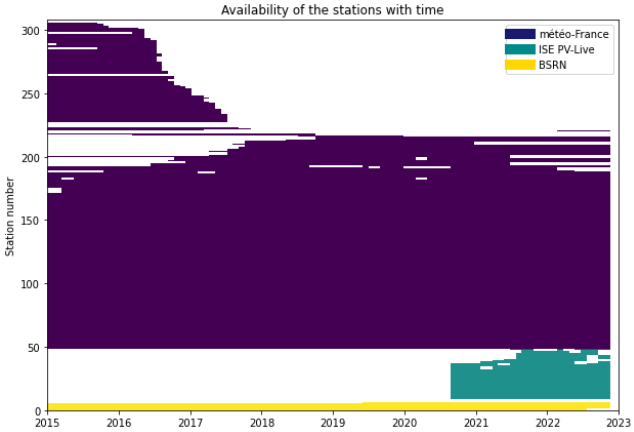

Figure 5: graphical representation of the availability of measurements of stations used in this work

The availability of the different stations is shown in Figure 5. The stations are available on different periods, as many Météo-France stations were dismantled between 2016 and 2018, and new stations were commissioned during the same period. The ISE-PVlive network started up in mid-2020 and all stations were operational by the end of 2021. In comparison the 5 BSRN stations are constantly available on the period 2015-2023. Only stations with more than two years of data have been selected for the analysis of the performance of CRS presented in the following sections. It is further assumed for simplicity that the performances of CRS are stationary over time, so that differences in the period of availability between stations can be ignored.

5 Quality control of solar radiation measurements

Detailed quality control was carried out to ensure that faulty measurements do not affect the evaluation of error and uncertainty of CRS. Standard quality control procedures were applied for stations with three components, but within CAMEO additional special tests were developed and tested to ensure the quality of measurements at stations where only global horizontal irradiance is measured. These new tests focus on detecting shadows, checking the calibration of the pyranometer and verifying the levelling of the pyranometer (see section 4.2). They may in future help to enlarge the CAMS observation database by GHI-only measuring observation networks.

5.1 Stations measuring global, direct and diffuse irradiances

Measurements of global, direct and diffuse irradiances are available at the BSRN stations used for this study. The availability of these three measures allows an exhaustive quality control (QC) of the measurements. As for the CRS EQC, we follow the BSRN QC routine initially proposed by Long and Dutton (2002). It consists of three steps that are described below.

The first step consists in verifying that measurements are comprised in a plausible range of values. Two limits are proposed by Long and Dutton (2002): the extremely rare limit (ERL) and the physically possible limits (PPL). We applied the ERL ranges that are given below:

$$-2\ W.m^{-2} \leq GHI \leq 1.2 \cdot ETN \cdot \mu^{1.2} + 50\ W.m^{-2}$$
$$-2\ W.m^{-2} \leq DIF \leq 0.75 \cdot ETN \cdot \mu^{1.2} + 30\ W.m^{-2}$$
$$-2\ W.m^{-2} \leq DNI \leq 0.95 \cdot ETN \cdot \mu^{0.2} + 10\ W.m^{-2}$$

Where GHI, DIF and DNI are the global, diffuse and direct normal irradiance measurements, μ is the cosine of the solar zenith angle and ETN is the extraterrestrial normal irradiance (solar constant corrected for the Earth's eccentricity).

The second test is aimed at verifying that the diffuse irradiance is smaller than the global horizontal irradiance, which can occur in the case of an improper calibration of the instruments or issue in the shadow ring correction (if applicable).

$$K < 1.05 \quad \text{if} \quad SZA < 75° \quad \text{and} \quad GHI > 50\ W/m^2$$
$$K < 1.1 \quad \text{if} \quad SZA \geq 75° \quad \text{and} \quad GHI > 50\ W/m^2$$

Where K=GHI/DIF is the diffuse fraction and SZA is the solar zenith angle expressed in degrees.

The last test is commonly called the closure test. It is designed to verify that direct and diffuse radiation measurements are consistent with global radiation measurements by exploiting the closure equation that links GHI, DNI and DIF:

$$GHI = DIF + \mu \cdot DNI$$

Long and Dutton (2002) proposed two different limits as a function of SZA, because the uncertainty in the ratio GHI / ($DIF + \mu \cdot DNI$) is expected to increase with SZA:

$\left|\frac{GHI}{DIF+\mu \cdot DNI} - 1\right| < 0.08$ when SZA<75° and GHI>50 W.m^{-2}

$\left|\frac{GHI}{DIF+\mu \cdot DNI} - 1\right| < 0.15$ when SZA>75° and GHI>50 W.m^{-2}

Thanks particularly to the closure test, this quality control procedure is particularly effective at identifying measurement issues and, after implementation of these tests, the measurements can be considered as reliable.

5.2 Stations measuring only global irradiances

In many stations considered for this work, only measurements of the global irradiance are available so that only the GHI range test can be applied to the data:

$$-2\ W.m^{-2} \leq GHI \leq 1.2 \cdot ETN \cdot \mu^{1.2} + 50\ W.m^{-2}$$

The required level of confidence cannot be guaranteed with this test alone. This is all the more true given that the stations in question are not maintained and calibrated with the same frequency as the BSRN stations, which further increases the uncertainty of the measurements. Solutions have therefore to be found to increase confidence in the measurements and to prevent frequent erroneous measurements from having too great an impact on the analyses. We have therefore developed a visual support to check the quality of the data, which is given in Figure 6 for measurements of GHI of the station located in Dax (Météo-France network).

The various plots included in this tool are designed to verify the plausibility of measurements from different perspectives. The different plots are annotated from 1 to 8 to facilitate the description of the tool. A time series of the measurements is first displayed (1) that allows identifying major issues as is the case in the given example in Summer 2015. A representation of the measurements as a function of the date and time of the day (2) allows to show the inherent structure of the data resulting from the rotation of the Earth and the variation of the Earth's declination. In particular, the time of sunset and sunrise can be clearly observed in this representation. The theoretical sunset and sunrise time calculated as a function of the date and station latitude and longitude are overlaid to this plot to verify that the time reference of the measurements is correct.

Global irradiance measurements are then compared to cloud-free solar irradiance data in the plot (3) where the ratio of measurements and its cloud-free counterpart are represented as a

function of time. This kind of plot is very informative as it allows detecting calibration issues (Summer 2015) or other suspicious behavior like the one that occurs between 2018 and 2019 in the example given (likely a levelling issue). An attempt has been made to quantify the tilt corresponding to the levelling issue (plot 4), but this kind of analysis is still under development and needs to be further validated.

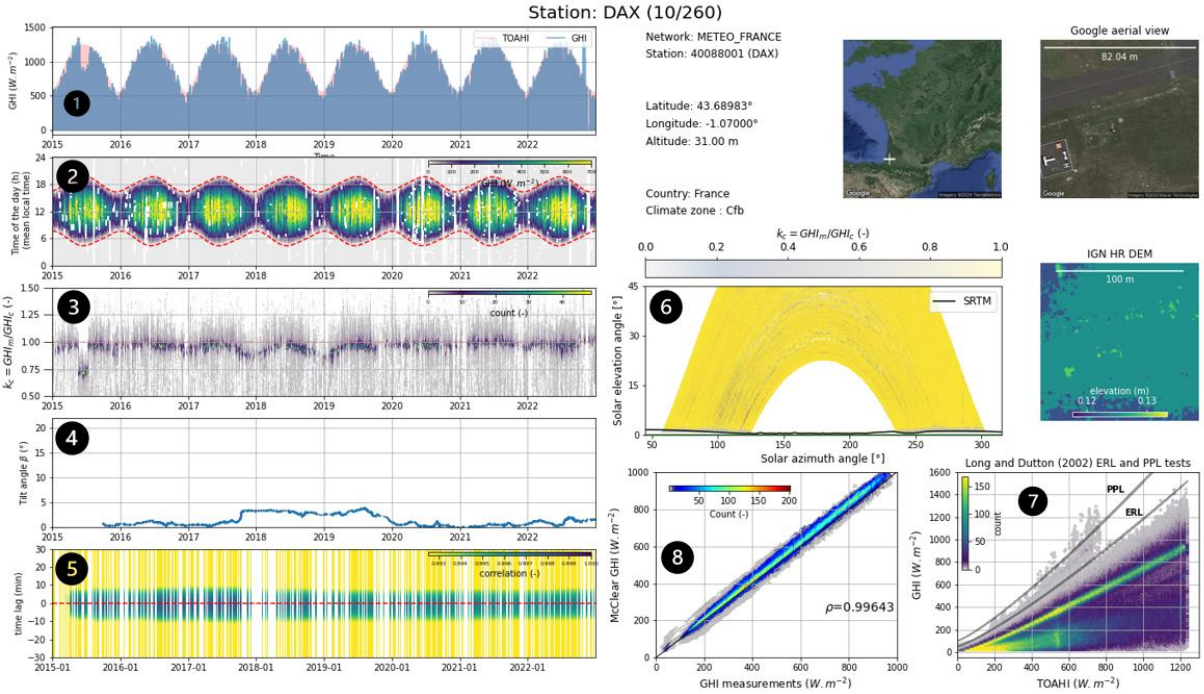

Figure 6: Quality control for the networks providing only GHI measurements

Results of a cross-correlation analysis between GHI measurements and cloud-free irradiance estimates over a sliding window is shown in the plot (5). This special analysis was aimed to understand the anomalies of the Météo-France described in the following section.

The shadow surrounding the stations are evaluated in the plot (6) where the clearsky index is represented as a function of the sun position. There was no important shadow at the Dax station, but examples of shadow profiles evaluated at different stations can be found in Figure 7. The black lines represented in the four graphs of Figure 7 are shadow profiles due to terrain orography, that have been calculated with SRTM data. All values below the SRTM horizon have been excluded from the dataset used for the analysis. Stations with a strong shadow profile have been excluded from the evaluation dataset while those with a moderate shadow have been maintained. We consider a shadow problematic if the amount of the obscured sky dome above 10° is contaminated by the shadow. In the examples given in Figure 7, the two upper cases correspond to severe shadow and have been excluded while the two lower cases have been kept.

The plots (7) in Figure 6 is represents GHI measurements as a function of the extraterrestrial horizontal irradiance (TOAHI). It is a convenient way to represent visually the upper limit ERL and PPL tests proposed by Long and Dutton (2002) introduced above. Finally, GHI measurements are compared to McClear cloud-free irradiance estimated in confidentially cloud-free conditions in the plot (8).

Such visual supports were generated for all the stations used in this work. They have been used as a basis for a visual inspection and manual flagging of all measurements.

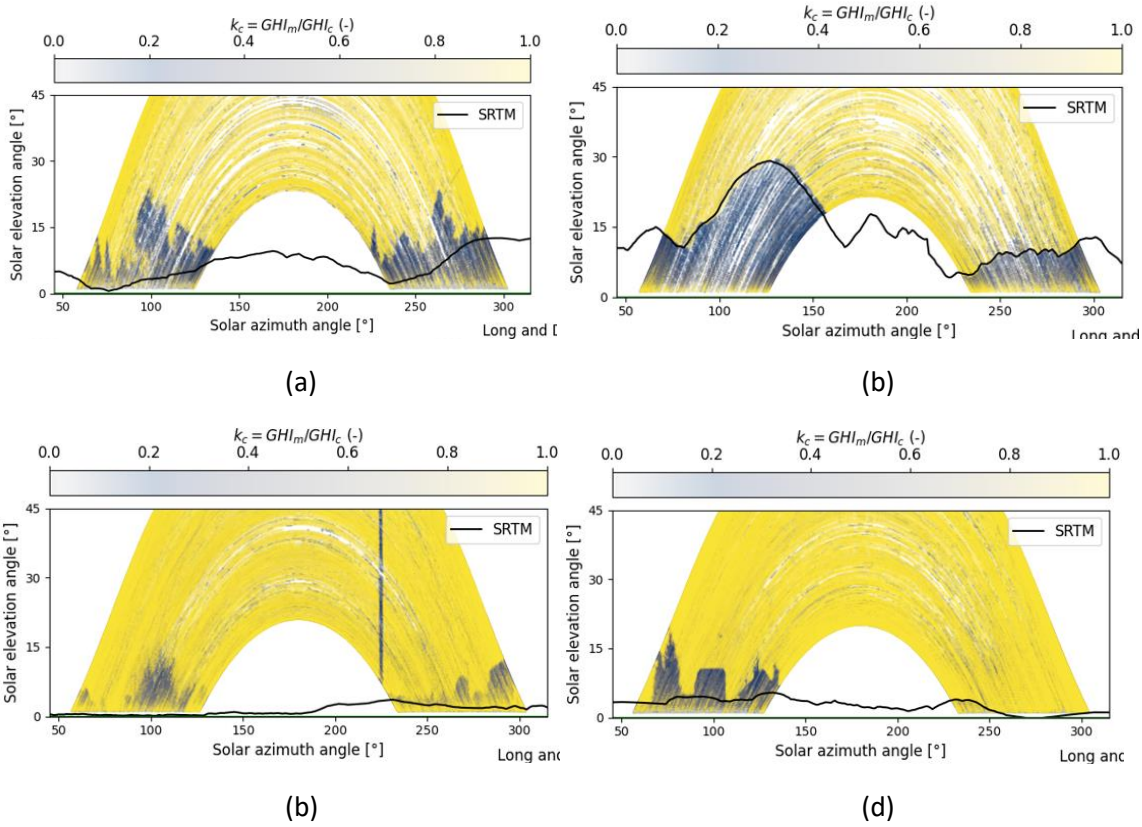

Figure 7: Examples of show profiles evaluated at different stations

To guarantee the quality of the data used in this report, we had no choice but to complete the QC by a visual inspection of the data. An important drawback of this approach is that the manual flagging of measurements is sometimes very subjective and the whole approach is hard to replicate. To alleviate these issues, some information on the manual flagging is provided below.

During the manual inspection, we excluded 33 stations from the total set of 226 stations. Among the excluded stations, 8 didn't have the minimal amount of 2 years of data, 1 had a temporal reference issue, 16 had too much shading and 1 had an obvious calibration issue. From the remaining stations, periods where an obvious issue occurred were excluded. This has been performed based on blocks of 6 month (January to June or July to December) to avoid affecting the long-term statistics due to the seasonal evolution of the solar radiation. A total number of 80 station-blocks were flagged due to levelling and calibration issues. Two examples of levelling and calibration problems are given in Figure 8 and Figure 9 respectively. In these two figures, the ratio of measurements and cloud-free irradiance is represented as a function of time and the colour of the points indicates the local density of points. Assuming that the cloud-free model is accurate, this ratio should be close to unity.

For the first station (Montauban, top graph in Figure 8), we can see that the points are close to 1 until the end of 2017, after which a seasonal profile can be observed especially in 2018 and 2019. This seasonality is no longer to be seen in 2022. This seasonal pattern is very likely resulting from a levelling issue (slightly tilted pyranometer), the probability of which was confirmed during a conversation with Météo-France employees. This behaviour is even more obvious for the second station (Le Grand Bornand bottom graph in Figure 8).

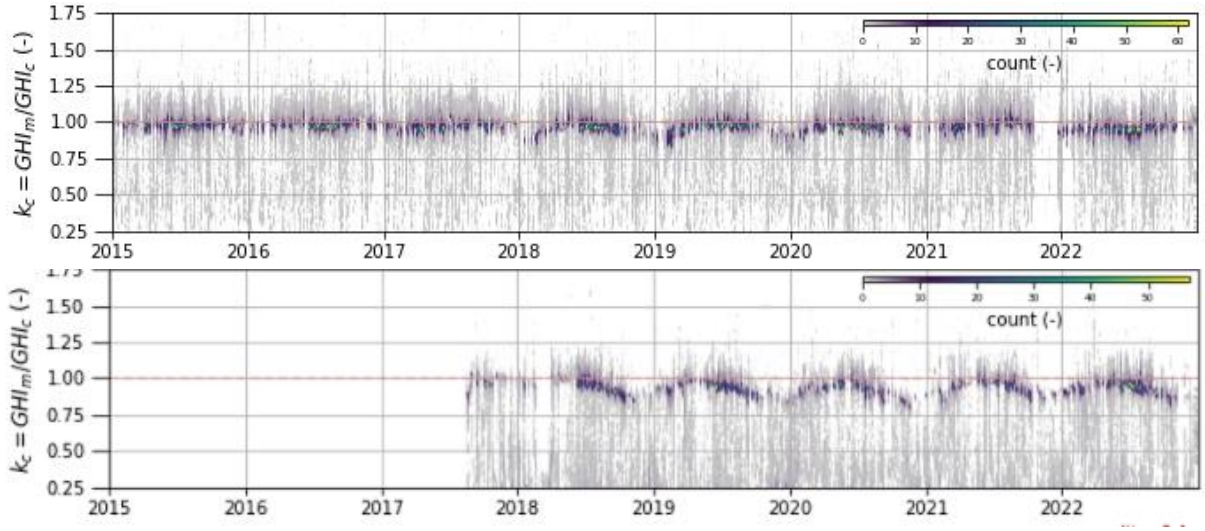

Figure 8: examples of stations with a levelling problem (top: Montauban, bottom: Le Grand-Bornand)

Calibrations issues are illustrated for two stations in Figure 9. These problems can be detected by a sudden change in the GHIm/GHIc ratio for a limited period, after which the ratio returns to unity.

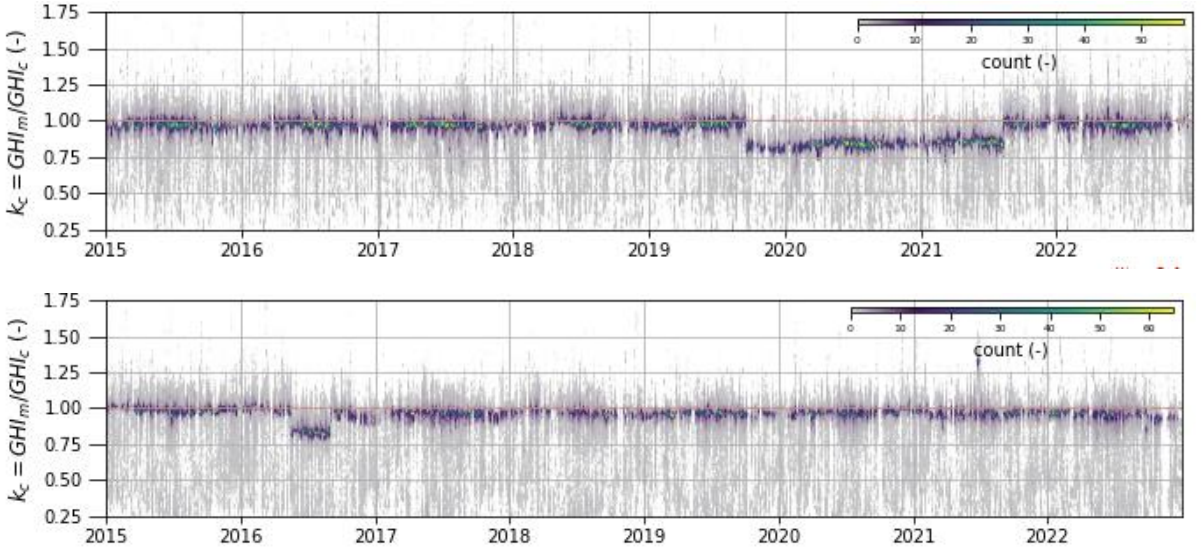

Figure 9: examples of stations with a calibration problem (top: Nice, bottom: Peone))

Finally, it should be mentioned that the manual flagging is a very subjective task. In many stations, it was not obvious whether or not there was a measurement problem. Given that too severe visual inspection can bias the result, we intentionally leave such case in the dataset.

6 Analysis of the performances of CRS

The objective of the analysis performed in this task is to identify significant spatial or seasonal structures in the error of CAMS Radiation Service that may be relevant for:

- Users of CAMS Radiation Service
- The quantification of the uncertainty of the CRS model in CAMEO
- Future improvements of APOLLO-NG and CAMS Radiation Service (CAMS service evolution)

The approach followed in CAMEO differs from the regular CRS-EQC report in that it uses an extended dataset of solar irradiance measurements with a high density of stations which, to date, have not been considered in the regular EQC owing to the fact that they do not necessarily meet scientific standards such as those in place for the BSRN network. On the other hand, the EQC needs recent and regularly updated ground observations not older than 9 months. As an example, the CRS-EQC validation report of December 2024 deals with ground measurements from the quarter March to May 2024. One of the questions this analysis must answer is whether there is any added value in integrating these large volumes of data from GHI-only observations for the regular CRS-EQC.

A potential pitfall of analysing large volumes of data is that it generates a huge amount of information that is difficult to exploit and interpret. To avoid this risk, the methodology presented in the following section aims to focus the analysis on specific aspects with a synthetic/concise approach, were the most important spatial and seasonal features of the error of CRS are identified and analysed. Consequently, our aim is not to carry out an exhaustive quantitative analysis but a targeted spatial qualitative assessment (i.e. focussing on most important features) that will serve as a basis for the quantification of the CRS error in task 4.3 entitled "*New methodologies for deriving uncertainties in CAMS radiation and deposition products relative to independent data"*. Indeed, the work planned in task 4.3 consists in developing a model of the uncertainty of CRS that will be integrated in the service. The knowledge of the variable on which CRS error depends is a very important information for developing this uncertainty model.

6.1 Methodology

The evaluation methodology used in this work follows the same principle as the one implemented in the CRS-EQC reports, with minor changes motivated and explained below.

The analysis conducted in this task is performed with version v4.6 of the CRS. This includes the version 3.6 of McClear with input parameters (aerosol optical depth, ozone, water vapor) from the latest available of cycle of CAMS IFS.

The evaluation is carried out using 15 min averages as this time resolution is used by many practitioners in the solar energy industry. This differs from CRS-EQC where the evaluation is made using hourly averages for CRS and 10 min for cloud free situations. This change with respect to the regular EQC is motivated by two objectives. Firstly, this analysis will allow us to anticipate the analysis that CRS end-user may conduct and thereby better answer future user request. Secondly, the uncertainty model (task 4.3) being developed on a 15 min basis, conducting the analysis of CRS error on this time resolution is required to ensure that the results are exploitable in this further task. For each station, the deviations are computed by subtracting 15-min averaged observations for each instant from coincident product estimations (CRS - measurements). The resulting deviations are then summarized by usual statistical quantities such as the bias and the standard deviation of the error.

In CRS-EQC reports, the following error metrics are assessed for the global, direct, and diffuse irradiances:

- Root mean square error (RMSE) and relative RMSE
- Bias and relative bias

- Correlation coefficient
- Monthly standard deviation

In addition to the quantities mentioned above, it is usual to consider other error metrics such as the mean absolute error (MAE), the R^2 or the standard deviation of the error. A calculation of all these metrics is recommended for the regular CAMS EQC in order to allow easy and convenient comparisons to all kind of metrics various users apply in their evaluation studies as well as to compare easily against available results in other studies where only a subset of metrics is published but does not add scientific information in principle. However, it would be detrimental to the objective of conciseness and synthesis of this work, and we have decided to limit the analysis to the bias and the standard deviation of the error.

The RMSE can be easily derived as root of the quadratic sum of these two quantities. In addition, these two quantities are the most important error metrics as it is usual to consider the bias as the fixable component of the error and the standard deviation as the random component of the error, which is more difficult to deal with and may require further analytic strategies.

Since two different processes of CRS are involved depending on the cloud conditions, we decided to carry out the evaluation for two different sky conditions. The first condition corresponds to confidentially cloud-free (or clear sky as widely used in the solar energy community) conditions where the measured irradiance is certainly not contaminated by clouds. In this case, the measurements are compared to the output of McClear, the cloud-free module of the CRS.

Clear sky instants are detected using the Hansen and Reno (2015) algorithm. This algorithm is based on five criteria measuring the difference between GHI measurements and a clearsky model over a sliding window. This algorithm is different from the one used for the preparation of the regular EQC report. It was chosen because it is widely used in the solar energy community, which allows us to anticipate the assessments that CRS users might make.

The second sky condition considered in the evaluation corresponds to confidentially cloudy situations where both the cloudy (McCloud) and the cloud-free (McClear) components of CRS are involved. The confidentially cloudy situations are selected for instances where the clear sky index (defined as the ratio between the measured irradiance on the ground and the clear sky irradiance) is less than 0.5.

Other cases with neither confidentially cloudy nor confidentially clear conditions are not included in the assessment in this report. Contrary to the CAMS EQC standard, we also do not report about all-sky conditions taking all instances as a whole into account in statistical metrics. The share of confidentially cloudy and cloud-free conditions varies significantly as a function of the month and the location of the season. The share of data that are neither confidentially cloudy or cloud-free ranges from 2 to 8%.

The different error metrics and cloud conditions considered in the evaluation are summarized in Table 2. To avoid any confusion, the mathematical expressions of the different error metrics are given in this table.

Table 2: Summary of the evaluation methodology implemented in this work.

			Conditions	
			Confidentially cloudy	**Confidentially cloud-free**
Subset selection			$E_{cld,mm} = \left\{ t \;\middle|\; \frac{G(t)}{G_{cs}(t)} < 0.5 \;\&\; month = mm \right\}$	$E_{cs,mm} = \{ t \mid CSD_{HansenReno}(t) = True \;\&\; month = mm \}$
Average			$\mu_{cld,mm}(G) = \sum_{t \in E_{cld,mm}} \frac{G(t)}{Card(E_{cld,mm})}$	$\mu_{cs,mm}(G) = \sum_{t \in E_{cs,mm}} \frac{G(t)}{Card(E_{cs,mm})}$
	Bias	*Absolute bias*	$\mu_{cld,mm}(\Delta G) = \sum_{t \in E_{cld,mm}} \frac{\Delta G(t)}{Card(E_{cld,mm})}$	$\mu_{cs,mm}(\Delta G) = \sum_{t \in Ecs,mm} \frac{\Delta G(t)}{Card(E_{cs,mm})}$
		Relative bias	$\frac{\mu_{cld,mm}(\Delta G)}{\mu_{cld,mm}(G)} = \frac{1}{\mu_{cld,mm}(G)} \sum_{t \in E_{cld,mm}} \frac{\Delta G(t)}{Card(E_{cld,mm})}$	$\frac{\mu_{cs,mm}(\Delta G)}{\mu_{cs,mm}(G)} = \frac{1}{\mu_{cs,mm}(G)} \sum_{t \in E_{cs,mm}} \frac{\Delta G(t)}{Card(E_{cs,mm})}$
	Standard deviation	*Absolute standard deviation*	$\sigma_{cld,mm}(\Delta G) = \sqrt{\sum_{t \in E_{cld,mm}} \frac{\left(\Delta G(t) - \mu_{cld,mm}(\Delta G)\right)^2}{Card(E_{cld,mm})}}$	$\sigma_{cs,mm}(\Delta G) = \sqrt{\sum_{t \in E_{cs,mm}} \frac{\left(\Delta G(t) - \mu_{cs,mm}(\Delta G)\right)^2}{Card(E_{cs,mm})}}$
		Relative standard deviation	$\frac{\sigma_{cld,mm}(\Delta G)}{\mu_{cld}, mm(G)} = \frac{1}{\mu_{cld,mm}(G)} \sqrt{\sum_{t \in E_{cld}} \frac{\left(\Delta G(t) - \mu_{cld,mm}(\Delta G)\right)^2}{Card(E_{cld,mm})}}$	$\frac{\sigma_{cs,mm}(\Delta G)}{\mu_{cs,mm}(G)} = \frac{1}{\mu_{cs}, mm(G)} \sqrt{\sum_{t \in E_{cs,mm}} \frac{\left(\Delta G(t) - \mu_{cs,mm}(\Delta G)\right)^2}{Card(E_{cs,mm})}}$
$\Delta G(t) = G(t) - \tilde{G}(t)$ Where: $G(t)$ is the measured solar irradiance $\tilde{G}(t)$ is the solar irradiance provided by CRS Card(E): number of elements in the set E mm is the calendar month (mm=1…12)				

6.2 Analysis of the spatial characteristics and seasonality of CRS error

In this section, our objective is to analyse the seasonality as well as spatial characteristics of the selected error metrics (bias and standard deviation of error) for confidentially cloudy and cloud-free conditions. As the magnitude of solar irradiation is known to have a significant impact on the error measurements, the absolute and relative values of the bias and standard deviation are analysed.

A common visualisation is employed for the spatial and monthly-resolved analysis for all selected error metrics and situations (Figures 6 to 13). It comprises three panels representing respectively maps of the value of interests for each month in the upper panel, seasonal evolution of the studied error metric in the medium panel and maps of the average GHI for the selected situation in the lower panel.

The middle panel allows visualizing the seasonality of the bias and standard deviation of error for all stations as well as their distribution over all stations (spread of the grey lines and violin plot). To assess whether the distribution of values for a given month is resulting from a spatially stochastic or a spatially structured error, the values represented in the seasonality plot are displayed as maps for each month in the upper panel. To facilitate the comparison of the maps from month to month, the colormap is scaled between the 5% and 95% quantiles of the displayed quantity. Since the average solar resource is expected to be an important factor for the spatial analysis of the error, maps of the mean GHI displayed in the lower panel allows a visual comparison of the spatial representation of the studied error metric with the average resource.

The graphical representation described above has been generated for the bias, the relative bias, the standard deviation of error and the relative standard deviation of error in confidentially cloud-free and confidentially cloudy conditions. This results in a set of 8 analyses that have been subjected to a detailed visual inspection by the CAMEO as well as by the CRS evolution teams. The main features resulting from this collective analysis are summarized in the following subsections.

6.2.1 Bias in confidentially cloud-free conditions

We can observe in Figure 10 that CRS has a positive monthly bias ranging from -25 to 50 $W.m^{-2}$ with majority of values between +0 to +20 in cloud-free conditions. The seasonal variation of the bias is not well pronounced. Slightly higher values of the bias are observed in the sunnier south-eastern part of France, particularly from April to October. Significantly lower bias values are found for the ISE PV-live network (northeast section of the images) but also for the BSRN stations (squares on the images). Similar observations are made for the relative bias in Figure 11, with no major differences observed between the absolute and relative bias.

The latter finding on lower bias values at BSRN and ISE PVlive with respect to Météo-France has been the subject of internal detailed investigations. While the reasons for this difference have yet to be elucidated, differences in the nature of the different measurement networks seem to provide the most plausible explanation. The instruments in the Météo-France network are maintained much less frequently than the BSRN stations. As mentioned in the previous chapter, this leads to greater uncertainty about the effect of soiling, levelling and calibration on instrument measurements, which could explain the positive bias observed. The German network is not operated with the daily maintenance protocol specific to the BSRN network, but it is a recent network that is regularly checked by the Fraunhofer ISE. We can therefore expect the bias values found for the BSRN and ISE- PVlive networks to be more reliable than those assessed with the Météo-France stations. Therefore, given their difference from the other networks, the bias values derived from the Météo-France stations should be treated with

caution, as the derived errors are potentially more influenced by the quality of the instruments than by the performance of CRS.

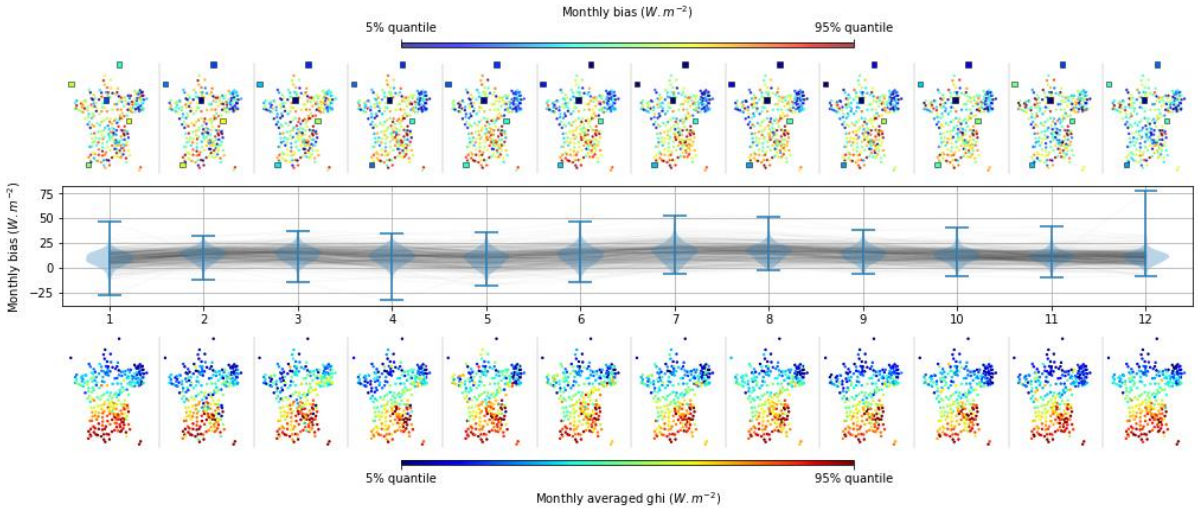

Figure 10: Monthly bias found for all stations in confidentially cloud-free conditions are represented over the year by grey lines in the middle panel. The distribution of bias values is summarized for each month by a violin plot that displays de histogram of the monthly bias of all stations as well as the 5 and 95% quantiles. For each month, maps of bias values are given in the upper panel. To facilitate the observation of the spatial patterns the range of colour values are set between the 5% and 95 % quantiles. The lower panel shows the average GHI for the situation considered.

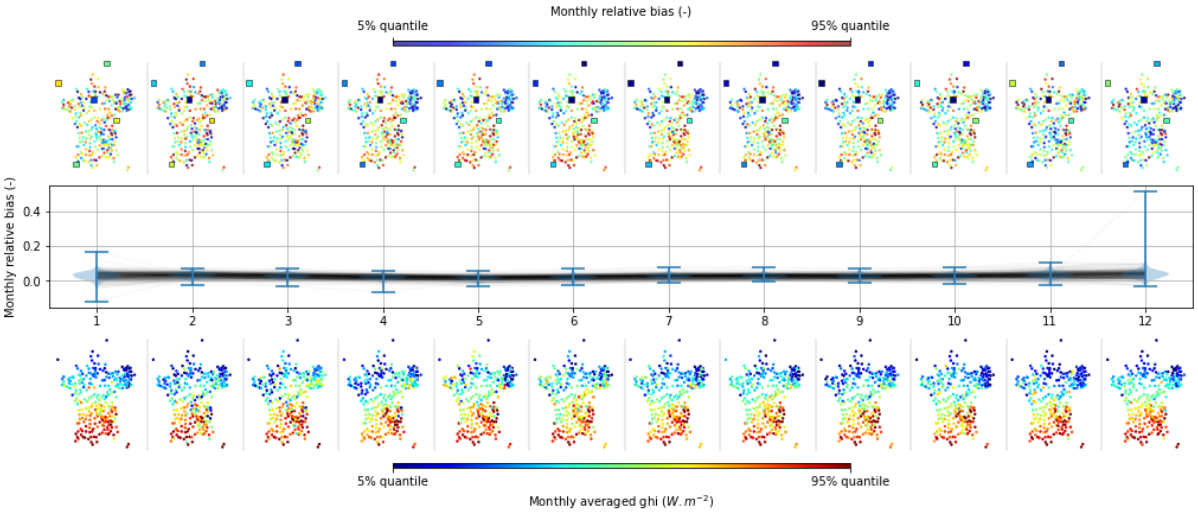

Figure 11: Same as Figure 10 for the relative bias of CRS in confidentially cloud-free conditions

6.2.2 Standard deviation of error in confidentially cloud-free conditions

The standard deviation of the error and its relative counterpart are shown in Figure 12 and Figure 13 for confidentially cloud-free conditions respectively. As with the analysis of bias in cloud-free conditions, the standard deviation of the error is greater at Météo-France stations than at BSRN and PVlive stations. This is consistent with the hypothesis that solar irradiance measurements at Météo-France stations are affected by greater uncertainty related to maintenance. As with the previous analysis, the results obtained at Météo-France stations should therefore be treated with caution.

The seasonal trend is more pronounced for the standard deviation of error than for the bias, with higher absolute values in summer than in winter, but a lower dispersion in summer than in winter. The relative values of the standard deviation of the error shown in Figure 13 increase noticeably in winter, which is partly an artefact due to the division of the standard deviation by low mean irradiance. This artefact is discussed in more detail in section 5.3.

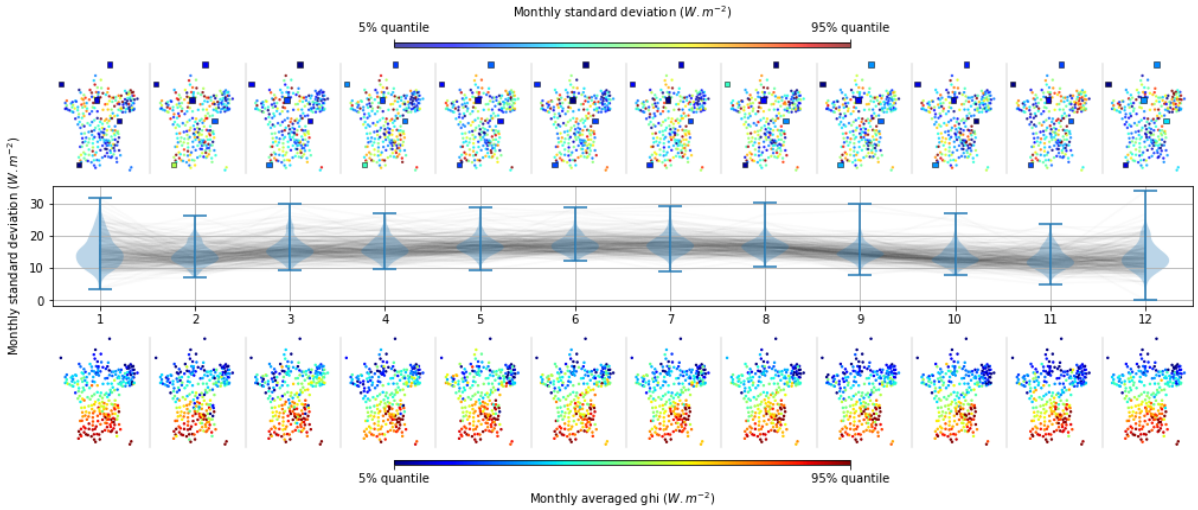

Figure 12: Same as Figure 10 for the standard deviation of the error of CRS in confidentially cloud-free conditions

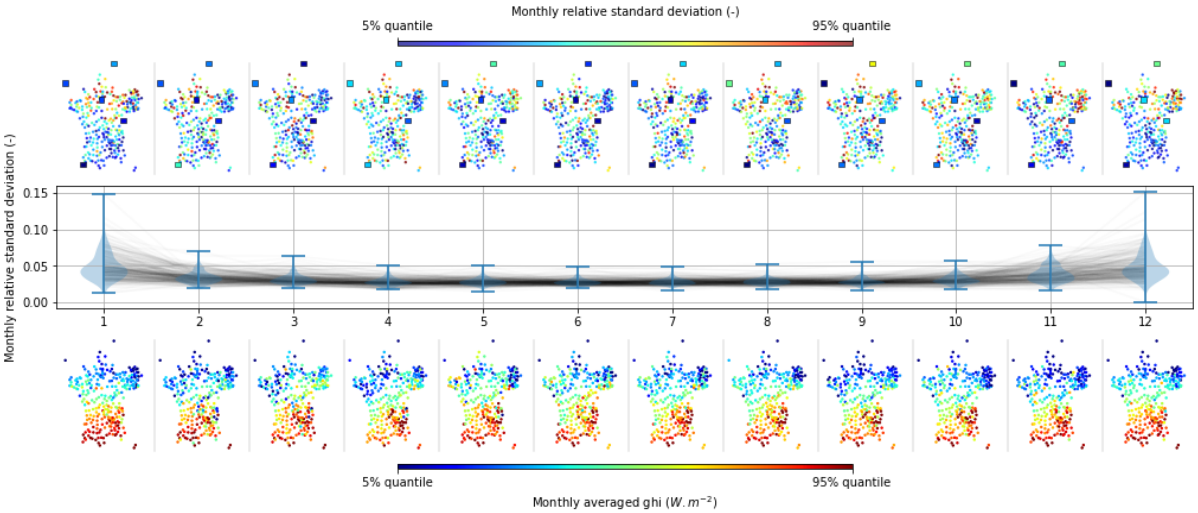

Figure 13: Same as Figure 10 for the relative standard deviation of the error of CRS in confidentially cloud-free conditions

6.2.3 Standard deviation of error in confidentially cloudy conditions

The absolute and relative values of the standard deviation of the error are shown in Figure 14 and Figure 15 for confidentially cloudy conditions. While there was a significant difference between the Météo-France measurements and the BSRN and ISE-PVLive networks in terms of CAMS Radiation performance under clear skies, the results are comparable between the different networks in cloudy condition. This can be explained by the fact that the uncertainties in the Météo-France measurements are of the same order of magnitude as the CRS errors under confidentially cloud-free conditions, but an order of magnitude smaller than the CRS

errors under confidentially cloudy conditions. This is because considering the optical properties of clouds in CRS specifically in complex cloudy conditions as heterogeneous or multi-layer clouds has limitations, resulting in an increase in error with the presence of clouds. It is therefore possible to carry out an analysis of the performance of CRS with Météo-France data in confidentially cloudy conditions, although the results obtained with Météo-France data must be interpreted with caution.

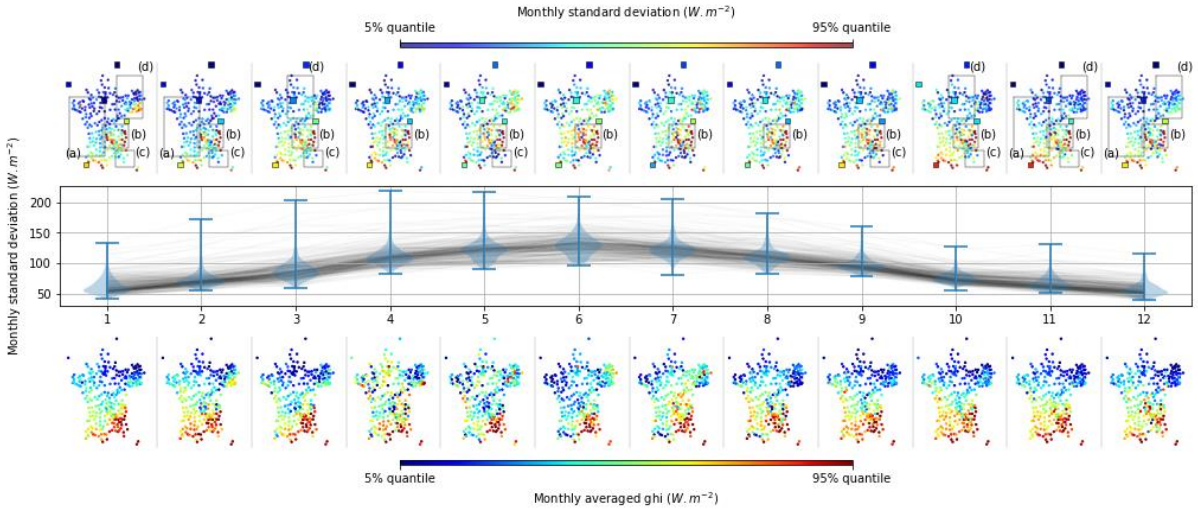

Figure 14: Same as Figure 10 for the standard deviation of error of CRS in confidentially cloudy conditions

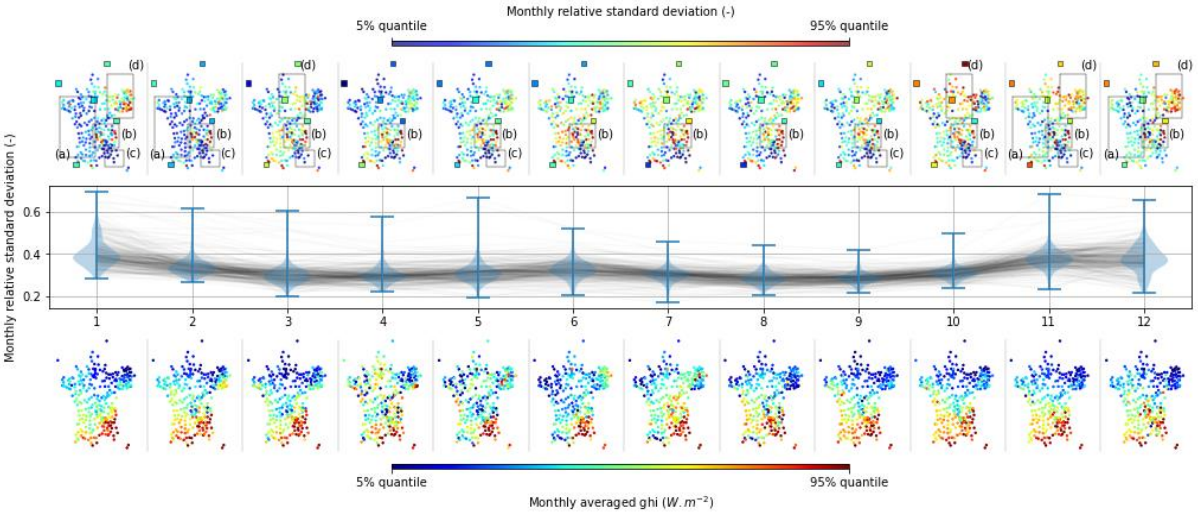

Figure 15: Same as Figure 10 for the relative standard deviation of error of CRS in confidentially cloudy conditions

The most relevant features identified in the analysis of the standard deviation of CRS in confidentially cloudy conditions are marked by rectangles in Figure 14 and Figure 15 and noted from a) to d).

In some regions marked with a), the standard deviation of the error shows a meridional trend with a higher standard deviation in the South than in the North. This dependence is most noticeable in winter, but not in summer. A comparison of these values with the GHI maps in the lower panel of Figure 14 suggests that this spatial trend is linked to mean solar irradiance. In fact, this effect is not present in the relative standard deviation maps in the upper panel of Figure 15. In this case, using the relative standard deviation removed the effect of mean irradiance on the spatial distribution of the error.

Some anomalies appear in the relative standard deviation maps while no anomalies are present in the absolute standard deviation maps. This is the case for regions marked c) and d) where the relative standard deviations are significantly higher and lower than the mean, respectively. This probably results from the problem mentioned in the previous section, where dividing an error metric by the mean solar irradiance can lead to very high (low) relative values if the mean solar irradiance is low (high). In the case shown here the error is not proportional to the mean irradiance. This is further explained and analysed in section 5.3.

In all the absolute and relative standard deviation maps presented in Figure 14 and Figure 15, a spatial pattern marked b) can be observed to the west of the Alps. This pattern has been identified as being related to orography and is analysed in more detail in section 5.4.

6.2.4 Bias in confidentially cloudy conditions

Maps of the bias of CRS in confidentially cloudy conditions shown in Figure 16 and Figure 17 have similar characteristics to those discussed in the previous section on the standard deviation of CRS in cloudy conditions. This is the case for high relative values in regions with low mean irradiance (case d) and the effect orography on the error (case b). In addition to the situations discussed above, we can observe in a meridional trend in the Southwest of France (case e). A more detailed analysis of this case, carried out in section 5.3, indicates that this spatial trend is most likely the result of the dependence of the bias on the mean irradiance.

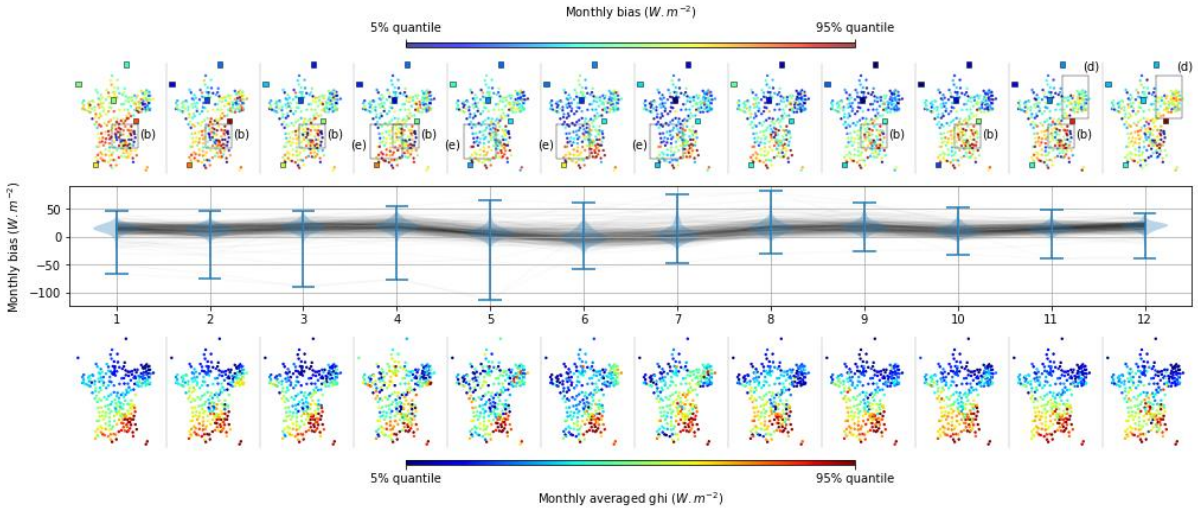

Figure 16: Same as Figure 10 for the bias of CRS in confidentially cloudy conditions

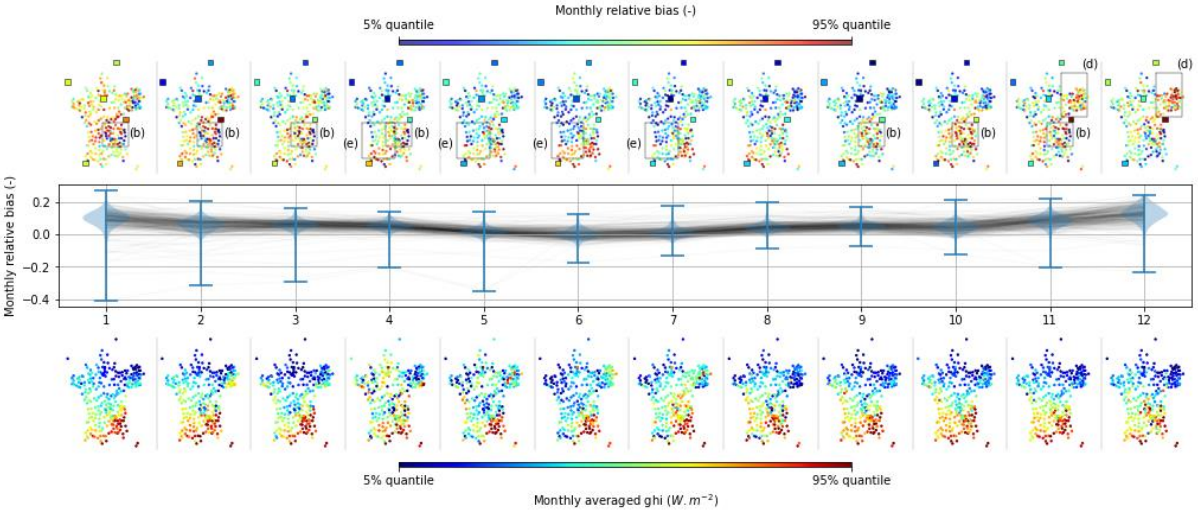

Figure 17: Same as Figure 10 for the relative bias of CRS in confidentially cloudy conditions

6.3 Influence of the mean irradiance on CRS performances

We found in the spatial and seasonal analysis of the CRS error that the different error metrics are dependent on the mean solar radiation. This dependence – illustrated in Figure 18 for the standard deviation of the error (right) and bias (middle) of CRS in February - is well-known, and it is a common practice in the solar energy community as well as in CAMS EQC reports to use relative error metrics to remove it.

However, the normalization with mean irradiation requires a linear connection between the error metric and the mean irradiation as well as a zero expected value. Otherwise, the normalization leads to artefacts. To illustrate the effect of this commonly overlooked property of irradiation validations, we further investigate the effect of mean solar irradiance on CRS error metrics in cloudy conditions for the bias and standard deviation of CRS in this section.

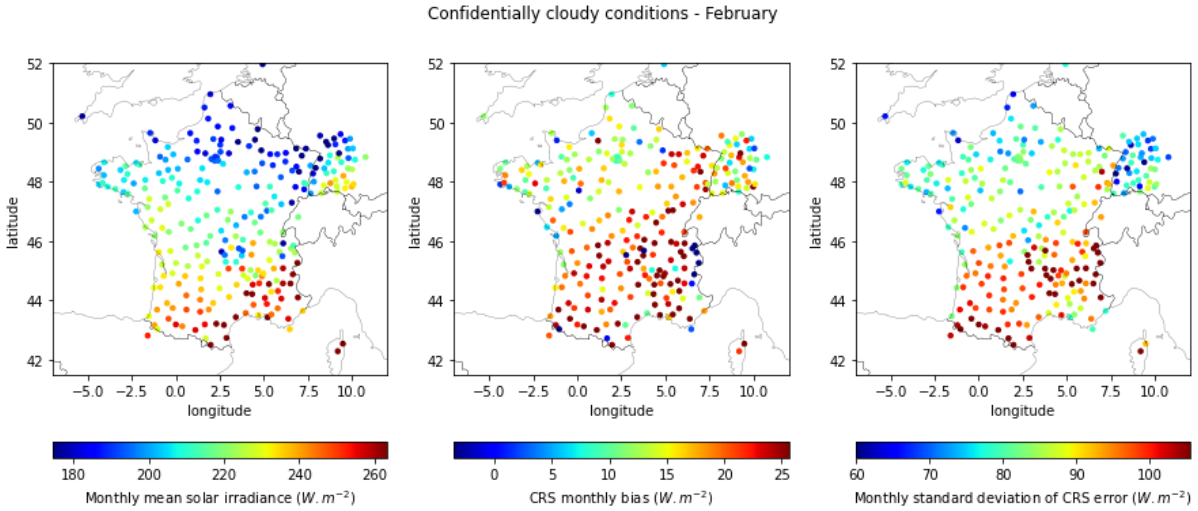

Figure 18: Maps of the mean solar irradiance and CRS standard deviation in confidentially cloudy condition in February, illustrating the dependence of CRS error metrics on the mean solar irradiance.

The absolute and relative values of the bias of CRS are shown as a function of mean solar radiation in the two graphs in Figure 19 where each point represents a station and a month, and the colour indicates the network. The bias shows no obvious dependence on mean solar irradiance. The dispersion of the points is the greatest for the Météo-France stations and a higher spread appears for all stations for mean solar irradiance values greater than 300 $W.m^{-2}$. There is no proportionality dependence between the bias and the mean irradiance. Consequently, normalising the bias by the mean irradiance results in artefacts that are clearly observable in the right plot of Figure 19. The use of relative bias should therefore be treated with caution in low GHI conditions.

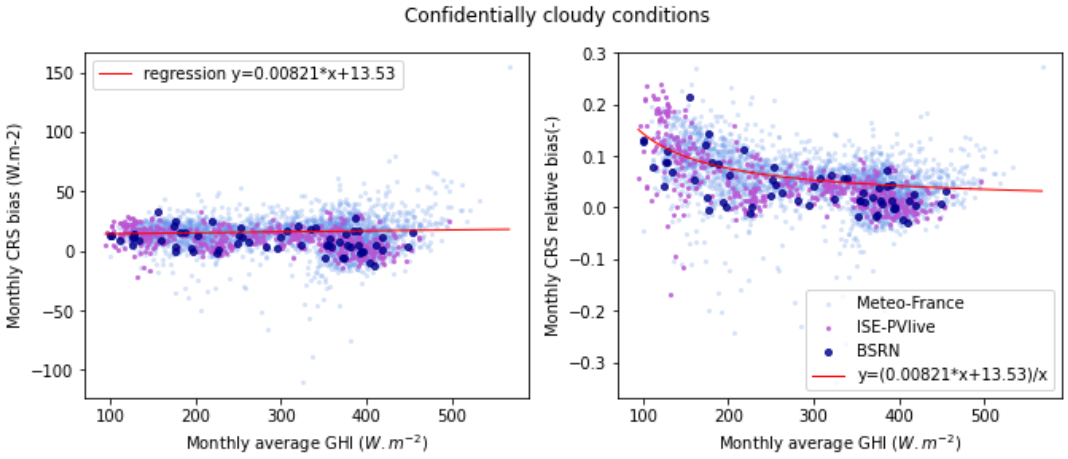

Figure 19: Scatter plot of the CRS monthly bias (absolute in the left panel and relative in the right panel) as a function of average monthly irradiance in confidentially cloudy conditions. Each point represents a month and a station, and the colours represent the different networks.

The data displayed for the absolute bias in Figure 19 are split for the different months in Figure 20. This more detailed representation shows that CRS bias varies noticeably as a function of the month of the year. The linear dependence between mean the bias and the mean irradiance is quasi-insignificant except in summer months.

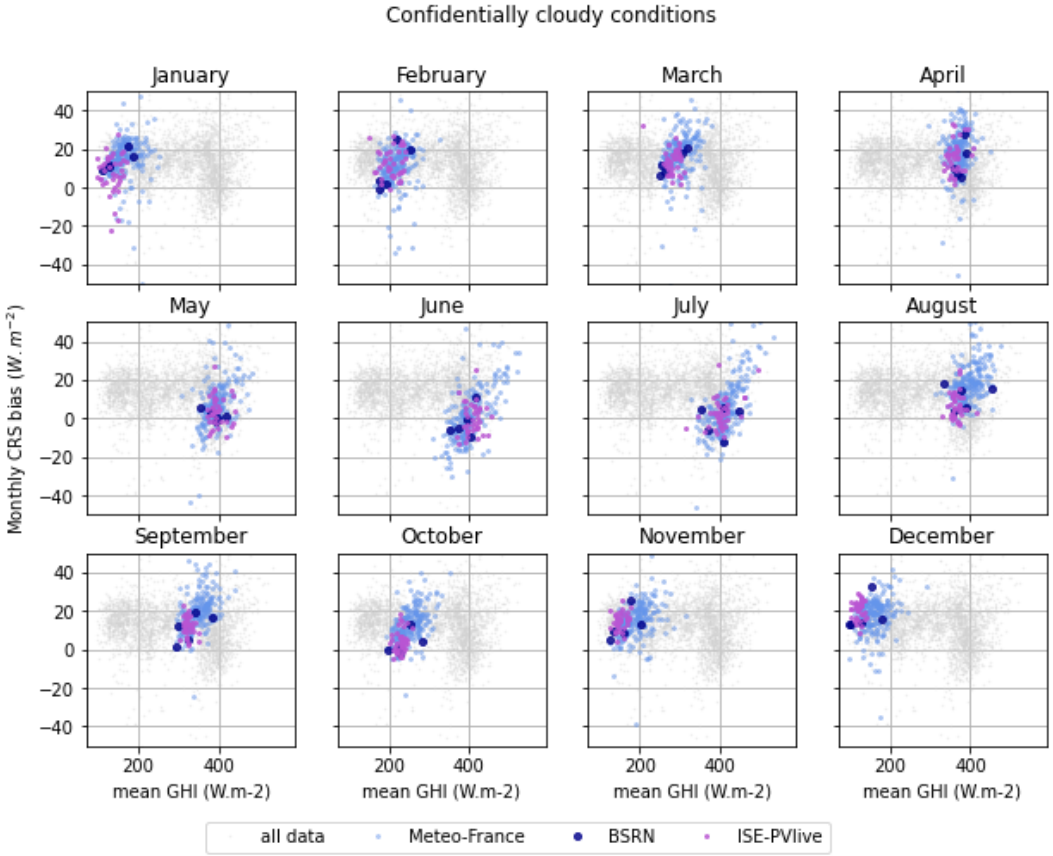

Figure 20: Same as the left plot of Figure 19 (left) split by month

The absolute and relative values of the standard deviation of CRS error are shown as a function of mean solar radiation in the two graphs in Figure 21. In this figure, each point represents a station and a month, and the colour indicates the network.

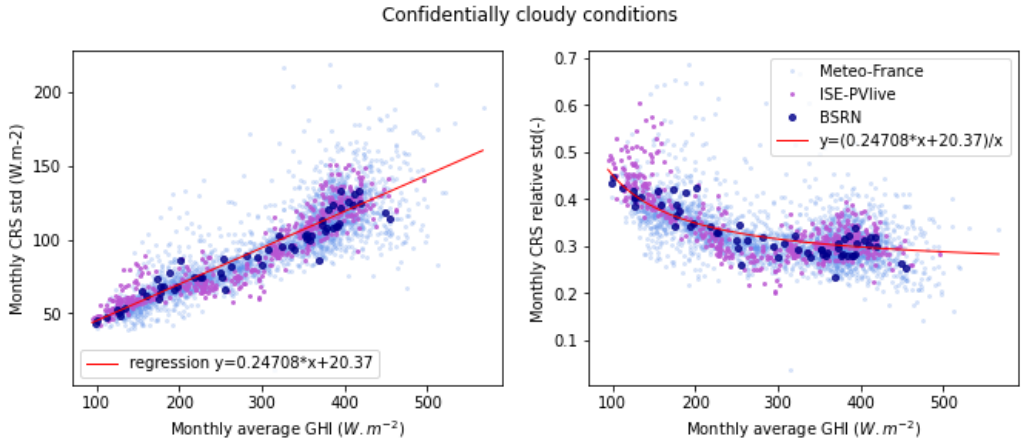

Figure 21: Scatter plot of the monthly standard deviation of the error of CRS bias (absolute in the left panel and relative in the right panel) as a function of average monthly irradiance in confidentially cloudy conditions. Each point represents a month and a station, and the colours represent the different networks.

There is a clear linear relationship between the standard deviation of the error and the average irradiance, the dependence being consistent for all networks. As this linear relationship presents an offset (20.37 $W.m^{-2}$), the hypothesis of proportionality between the standard deviation of the error and the mean irradiance does not holds (the requirement of a zero-expected value in the residuals is not met). This leads to an increase of the relative standard deviation as the mean irradiance becomes small, as show in the left panel of Figure 21.

This is the reason for the case (d) identified in the previous section, where the absolute value of the error exhibit no significant pattern while region with high relative values occur when the mean solar irradiance is low. It is thus the effect highlighted for low irradiance values in Figure 21 that was already observed in the analysis of the map (Fig. 8 and 9).

In the case (c) identified in the previous section, the absolute value of the error shows no significant pattern, while relative values are low in regions where the mean solar irradiance is high. This situation also results from the effect illustrated in Figure 21, but here lower relative values result from the fact that regions with high mean irradiance are less impacted from the above artefact than regions with a small mean irradiance.

These findings show that, similarly to the bias, the normalisation of the standard deviation of the error by the mean irradiance should therefore be interpreted with cautions.

The data represented in Figure 21 are displayed according to the month in Figure 22, which further confirms the linear dependence between the mean irradiance and the standard deviation of the error.

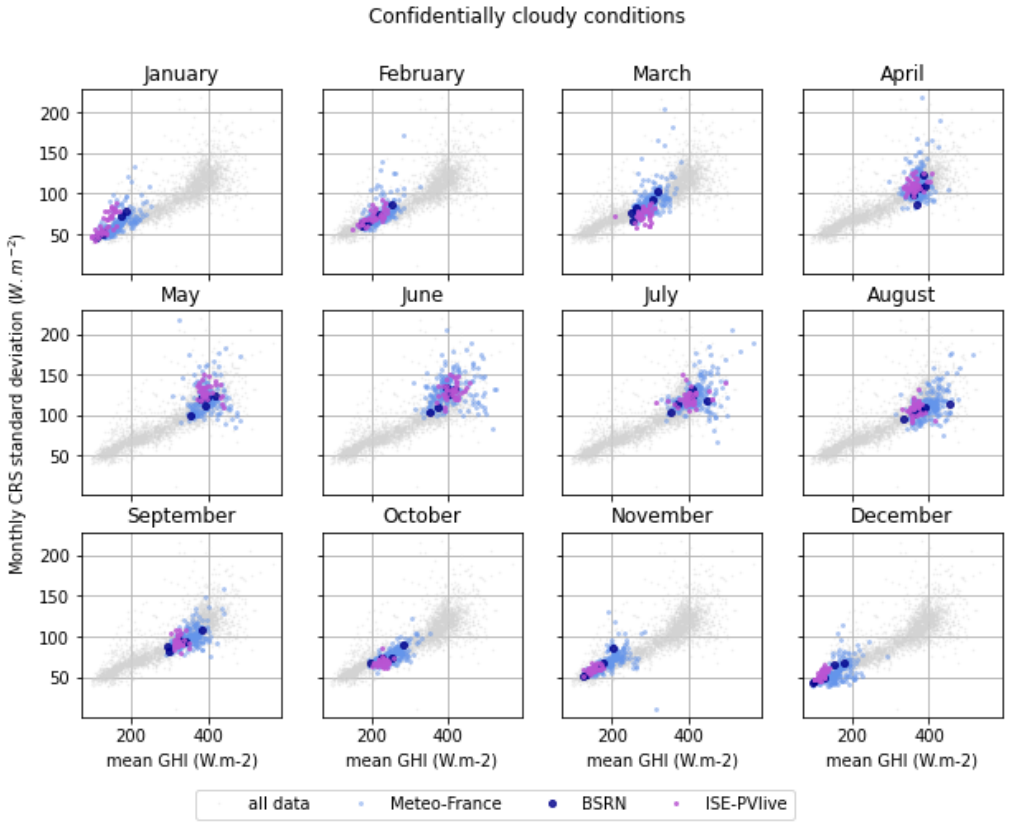

Figure 22: Same as the left plot of Figure 21 (left) split by month

We observed a spatial trend in the southwest of France, that was denoted by (e). A typical example of such a situation is shown in Figure 23. For the stations in the region of France corresponding to the case (e) identified in the previous section, the bias is plotted against the mean irradiance in Figure 24. This plot shows a significant trend between the bias and the mean irradiance. However, the bias is negative for lower irradiance and positive for larger irradiance. This explains why this pattern could not be removed using a relative bias (here again the relation of proportionality does not hold).

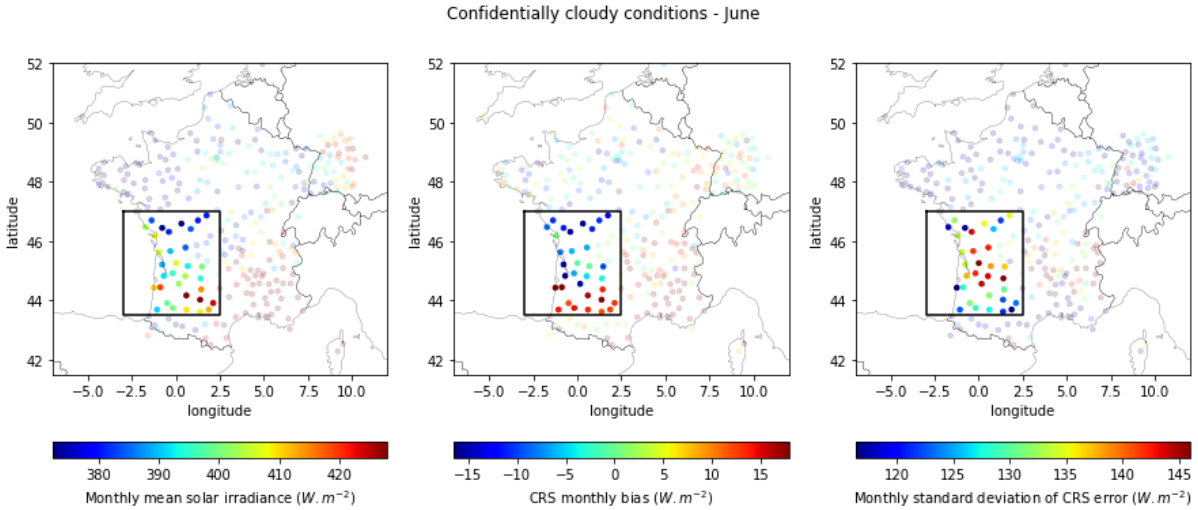

Figure 23: CRS and mean irradiance in June. The regions corresponding to the case (e) identified in the previous section are highlighted by a black rectangle.

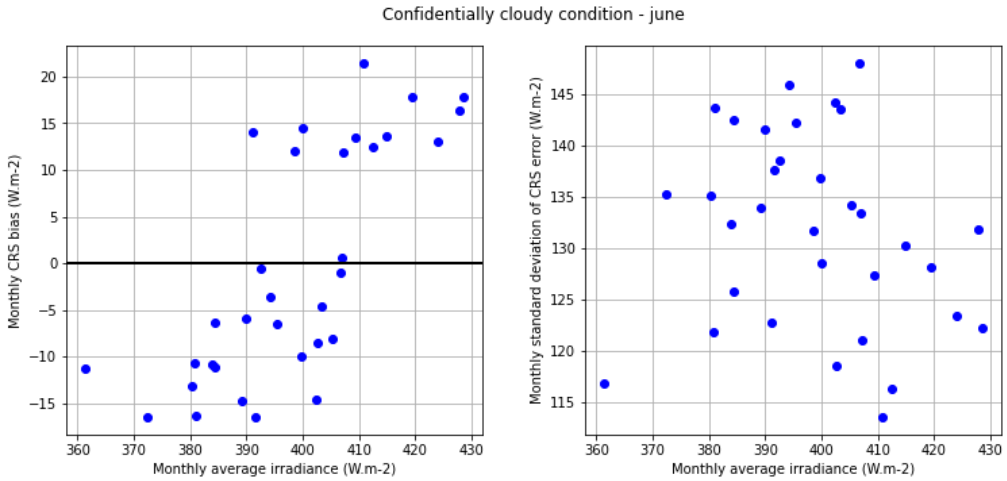

Figure 24: Plot of the bias as a function of the average irradiance for stations corresponding to the case (e) in June.

Overall, the commonly used practise to normalize error metrics by the mean observed irradiance has several drawbacks and should be used only with care. The normalized error has no physical meaning, and an analysis of the clearness index should be conducted again to get rid of the effect of the sun position while allowing a physical interpretation of the results.

6.4 Influence of the station elevation on the performances of CAMS Rad

Another observation made in the previous section is the dependence of the error of CRS on the site elevation in cloudy cases. This is illustrated in Figure 25 where a map of the site elevation (left graph) is compared to maps of the bias and standard deviation of the error of CRS in July (middle and right maps respectively).

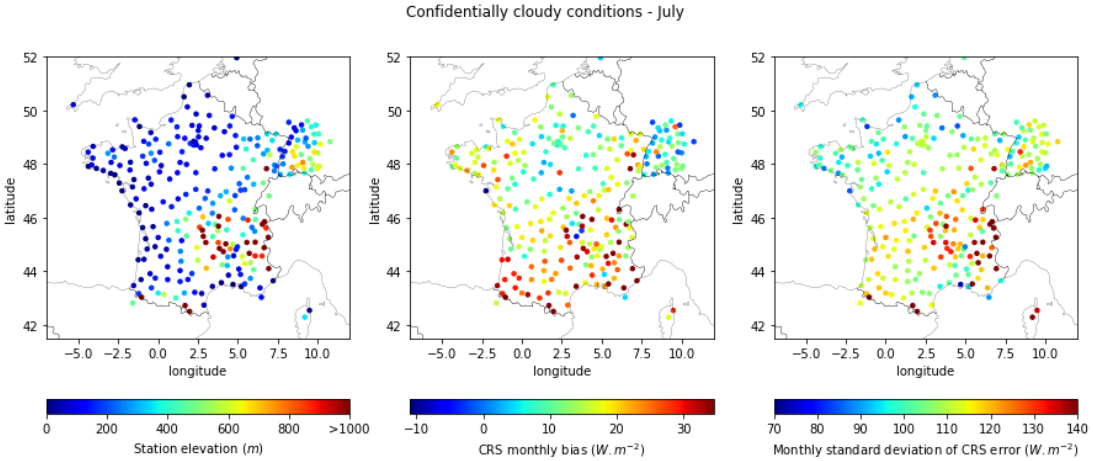

Figure 25: Maps of station elevation (left) and CRS standard deviation (right) in confidentially cloudy condition in July, illustrating the dependence of CRS error metrics on the station elevation.

We follow the same methodology as the one used in the previous section to evaluate the dependence of the error metrics on the mean solar irradiance. The bias and standard deviation of the error of CRS are represented as a function of the site elevation and by month in Figure 26 and Figure 27 respectively.

In Figure 26, the dependence between bias values and site elevation is not clear.

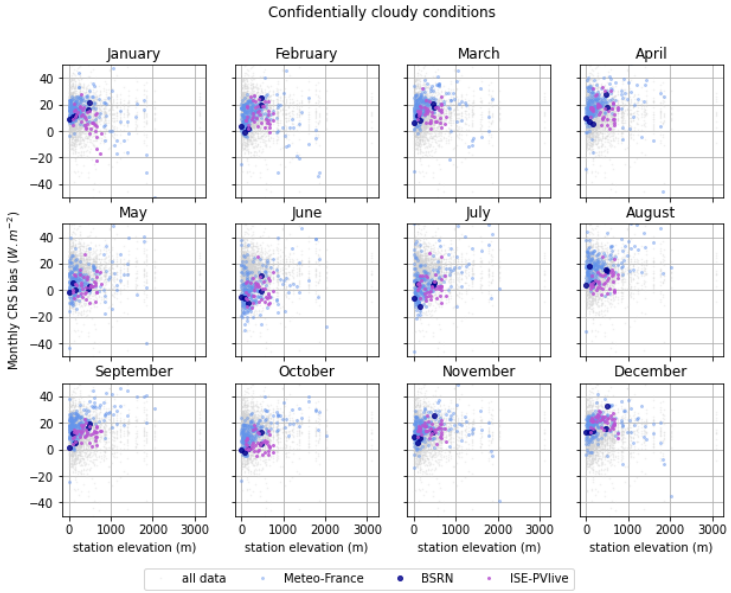

Figure 26: Scatter plot of the monthly bias of CRS as a function of site elevation in confidentially cloudy conditions for each month of the year. Each point represents a month and a station, and the colours represent the different networks.

The dependence is observed between the standard deviation of the error and the site elevation in Figure 27 clearly confirms the relevance of site elevation to explain the performance of CRS.

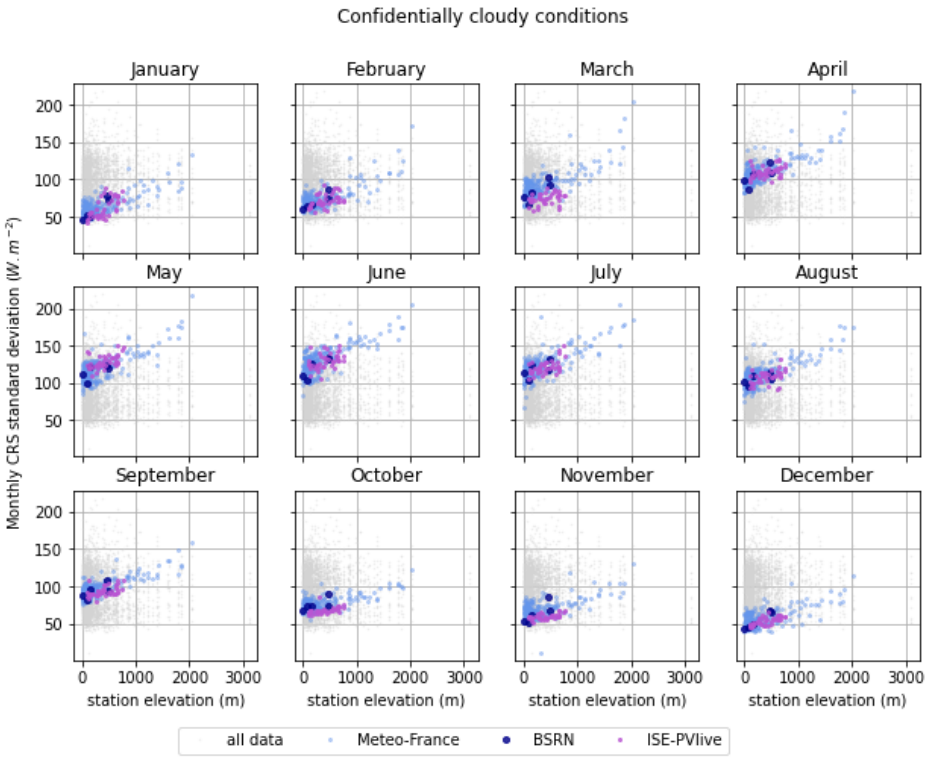

Figure 27: Scatter plot of the monthly standard deviation of CRS error as a function of site elevation in confidentially cloudy conditions for each month of the year. Each point represents a month and a station, and the colours represent the different networks.

6.5 Joint dependence of elevation and mean solar irradiance

We found in the previous section that mean solar irradiance and site elevation are two important parameters for characterizing the performance of CRS in confidentially cloudy conditions. However, these two variables are linked since we know that the solar resource is higher in elevated locations and elevated stations are located in the south east of France where the solar resource is higher. This raises the question of whether the effects of the solar resource and site elevation are not two ways of looking at the same dependency.

To answer this question, the same graphs as in Figure 26 and Figure 27 were generated with the elevation of the station represented by the colour of the scatter points. For the bias (Figure 26), as in the previous analysis, the effect of the site elevation and mean irradiance is not clear. In contrast, Figure 29Figure 27 shows that the standard deviation increases with mean irradiance and site elevation. The effects of these two variables are therefore not the result of a dependence between predictors (collinearity) but both are informative predictors for characterizing the standard deviation of the error of CRS in confidentially cloudy conditions.

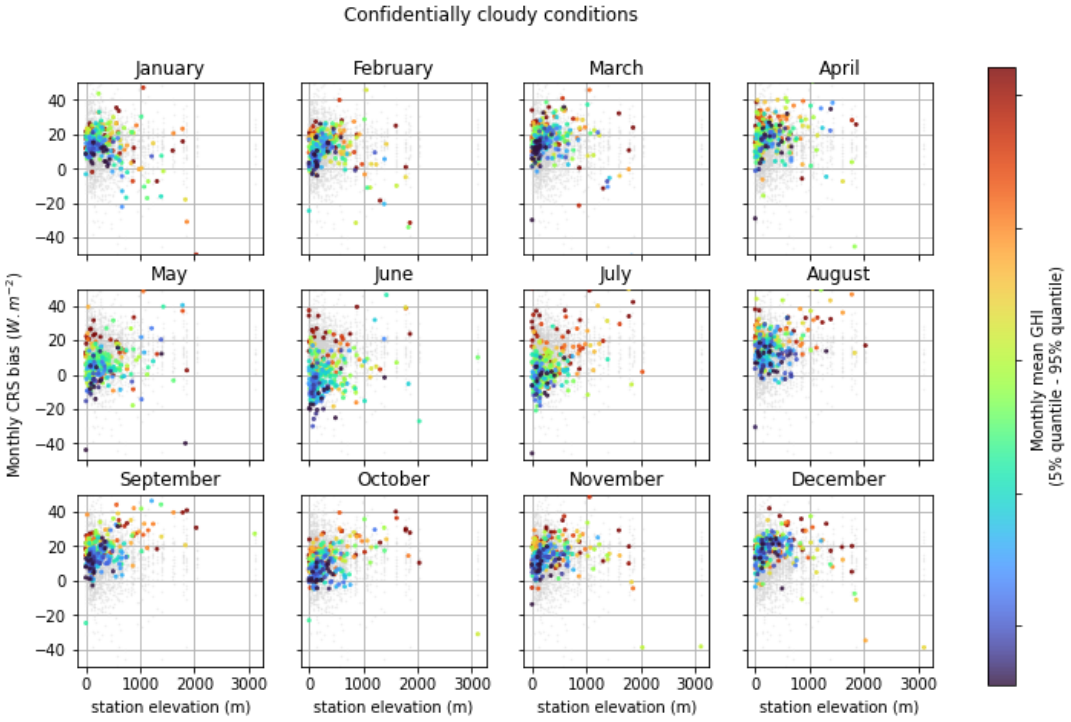

Figure 28 Same as Figure 26 where the colour of the scatter points represents the elevation of the stations

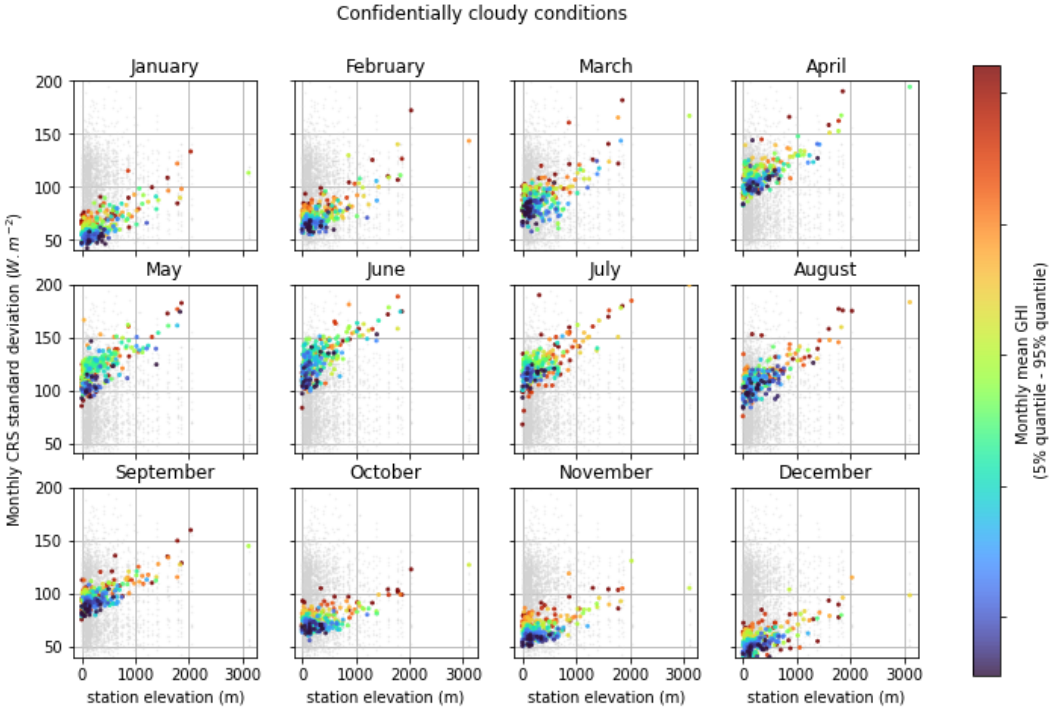

Figure 29: Same as Figure 27 where the colour of the scatter points represents the elevation of the stations

6.6 Quantification of the influence of identified features on the spatial and seasonal patterns of CRS error

In this part of the work, we studied the seasonality and spatial patterns of the error of CRS in confidentially cloudy and cloud-free conditions. We found that an import part of spatial patterns observed could be explained by the dependence of the error on mean solar irradiance and topography.

To validate and quantify this observation, we have assumed that the dependence of the bias and standard deviation of the error on station elevation and monthly mean GHI is linear and additive. In addition, we assume that the two predictors are not colinear, which is partly true since elevated region are in the south and exposed to more solar radiation than northern and less elevation locations. With these assumptions, a regression between the two error-metrics and the two explanatory variables has been made to quantify the proportion of the spatial variance for the two metrics that can be explained by these two predictors. The results are given in **Error! Reference source not found.**.

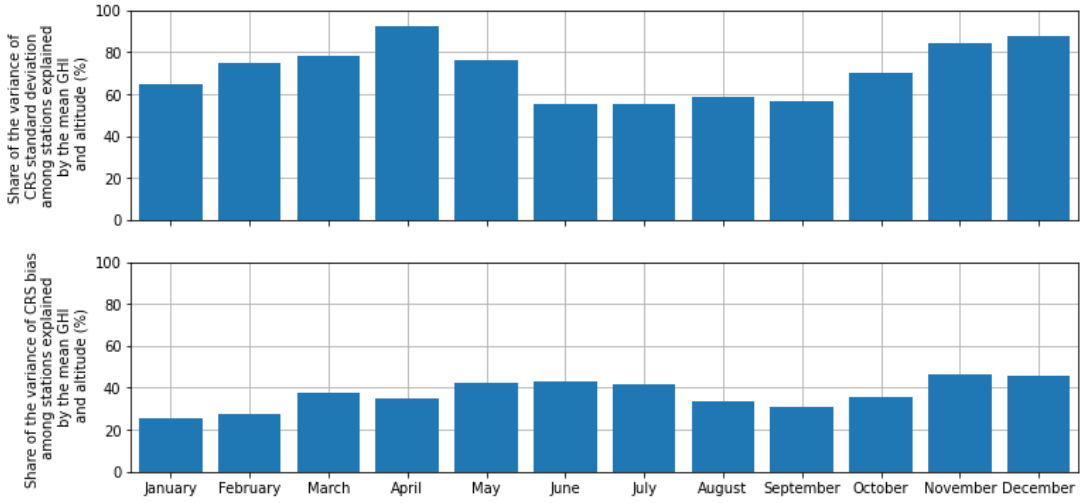

Figure 30: Results of the linear regression between the two error-metrics (standard deviation in the upper plot and bias in the lower plot) and the mean GHI and station elevation for each month.

The results illustrated in **Error! Reference source not found.** show that 50 to 90 percent of the differences among CRS standard deviation of error can be explained with the mean irradiance and the station elevation while this dependence is lower for CRS bias with values comprised between 20 and 50%. Based on this analysis, the dependence between the two selected error metrics and the station elevation and mean irradiances are analysed in more detail in section 5.3 and 5.4.

7 Conclusion and discussion

A detailed assessment of CRS error based on dense networks of solar radiation stations has been conducted in task 4.1 with the objective of identifying and characterizing spatial and seasonal patterns of CRS error whose knowledge is important for different activities of CRS and CAMEO. Such spatio-temporal patterns are important information for quantifying the uncertainty of CRS and identifying potential service improvements. It is also a valuable information for CRS users, as it improves the knowledge of the performance of CRS in comparison to regular EQC reports.

A reference dataset including global horizontal irradiation measurements with 1-min resolution has been collected using data from the Météo France pyranometric network and the German network ISE-PVlive. This dense dataset includes measurements from more than 300 stations over periods ranging from 2 to 8 years depending on the station. The use of this dataset required the development of a specific quality control procedures since no quality control matching our needs could be found in the literature for stations measuring only GHI.

The proposed quality control procedure includes different tests aimed at detecting shadow, calibration and levelling issues. In addition, a multiplot was created summarizing on a single page all relevant analysis needed for verifying the plausibility of measurements. This multiplot was generated for more than 300 stations and used for a visual inspection of the data. During this visual inspection, periods with suspicious measurements were manually excluded from the reference dataset. We expect that the quality of the dense reference network is improved thanks to this processing. However, despite the improved quality control the level of confidence in measurements from stations measuring only GHI remains lower than that from stations where global, diffuse and direct irradiance are measured, and an extensive quality control can be realized. As a result, results obtained with measurements from stations where only GHI is available should be interpreted with caution.

The evaluation showed that CRS have a larger error when compared with measurements from the Météo-France network than with other networks. The reason for this discrepancy is still unclear but it is likely related to the difference in the maintenance of the stations. This difference is particularly marked in confidentially cloud-free conditions but negligeable in confidentially cloudy conditions. The difference between sky conditions can be explained from the different magnitude of CRS error in cloudy and cloud-free conditions. This is important information with respect to the potential use of dense reference data in the EQC. We decided not to draw any conclusion on the performances of CRS using the dense reference dataset in confidently cloud-free conditions.

The analysis conducted for confidentially cloudy conditions showed that spatial patterns and a seasonality could be observed in the evaluation of CRS using the dense reference dataset. The most important patterns have been identified and subject to a detailed analysis. This showed that most pronounced patterns can be explained by a dependence of CRS error on the station altitude and mean solar irradiance. The dependence on the mean solar irradiance were found not to be systematically corrected using relative metrics but, instead, to lead to spatial artefacts. A critical analysis of the relative metrics was conducted showing that the normalization should be made with caution for low mean GHI values since the hypothesis of proportionality between error metrics and mean GHI is not always valid.

The issues found with the relative metrics raise the question on how to normalize the error metrics of CRS to get rid of the dependency on the mean solar irradiance without creating any artefact. The choice of the relative metrics has been made because it is used in the regular EQC as well as in the solar energy community even it its potential limitation are known (see e.g. Yang et al. 2020). An interesting follow-up of this analysis could consist in using clearness index instead of the relative error metrics to improve the normalization with respect to the solar irradiance.

The results presented in this report correspond to the performance of CRS v4.6 where some issues are known and currently being addressed. The validity of the analysis is therefore limited to the version of CRS operational at the time of writing. However, the dataset and methodology developed in task 4.1 represent a useful evaluation framework for longer term. It will be applied to future versions of CRS to measure the impact of version evolution on the performance of the service.

The work performed in task 4.1 is only descriptive. Since the dependence of the performance of CRS on mean solar irradiance and station elevation is an order of magnitude larger in confidentiality cloudy conditions than in confidentially cloud-free conditions (see annexes 1 and 2), this dependence should be further investigated from a modelling perspective in CRS service evolution to verify whether this effect results from issues in the coupling with clouds and, if so, identify needed model improvements.

Finally, the dependence between CRS performance and the mean solar irradiance and the station elevation can be an interesting information for the quantification of CRS uncertainty that is planned in the task 4.3 where the first findings described in this document will be further explored and quantified.

References

Anderson, K., Hansen, C., Holmgren, W., Jensen, A., Mikofski, M., and Driesse, A. "pvlib python: 2023 project update." Journal of Open Source Software, 8(92), 5994, (2023). DOI: 10.21105/joss.05994.

Holmgren, W., Hansen, C., and Mikofski, M. "pvlib python: a python package for modeling solar energy systems." Journal of Open Source Software, 3(29), 884, (2018). DOI: 10.21105/joss.00884.

Jensen, A., Anderson, K., Holmgren, W., Mikofski, M., Hansen, C., Boeman, L., Loonen, R. "pvlib iotools — Open-source Python functions for seamless access to solar irradiance data." Solar Energy, 266, 112092, (2023). DOI: 10.1016/j.solener.2023.112092.

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)

Long, C.N., Dutton, E.G., 2002. BSRN Global Network recommended QC tests, V2.0. Available online at http://epic.awi.de/30083/1/BSRN_recommended_QC_tests_V2.pdf (Last accessed 04/Nov/2024)

*Peel, M. C., Finlayson, B. L. & McMahon, T. A. Updated world map of the Köppen-Geiger climate classification. Hydrology and Earth System Sciences **11**, 1633–1644 (2007)*

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

Reno, M.J. and C.W. Hansen, "Identification of periods of clear sky irradiance in time series of GHI measurements" Renewable Energy, v90, 520-531 (2016.)

Yang D, Alessandrini S, Antonanzas J, Antonanzas-Torres F, Badescu V, Beyer HG, Blaga R, Boland J, Bright JM, Coimbra CFM, David M, Frimane Â, Gueymard CA, Hong T, Kay MJ, Killinger S, Kleissl J, Lauret P, Lorenz E, van der Meer D, Paulescu M, Perez R, Perpiñán-Lamigueiro O, Peters IM, Reikard G, Renné D, Saint-Drenan Y-M, Shuai Y, Urraca R, Verbois H, Vignola F, Voyant C, Zhang J. Verification of deterministic solar forecasts. Sol Energy 2020;210:20–37. http://dx.doi.org/10.1016/j.solener.2020.04.019, URL https://www.sciencedirect.com/science/article/pii/S0038092X20303947, Special Issue on Grid Integration.

Annex 1: CRS error metrics as a function of mean solar irradiance in confidentially cloudy and cloudless conditions

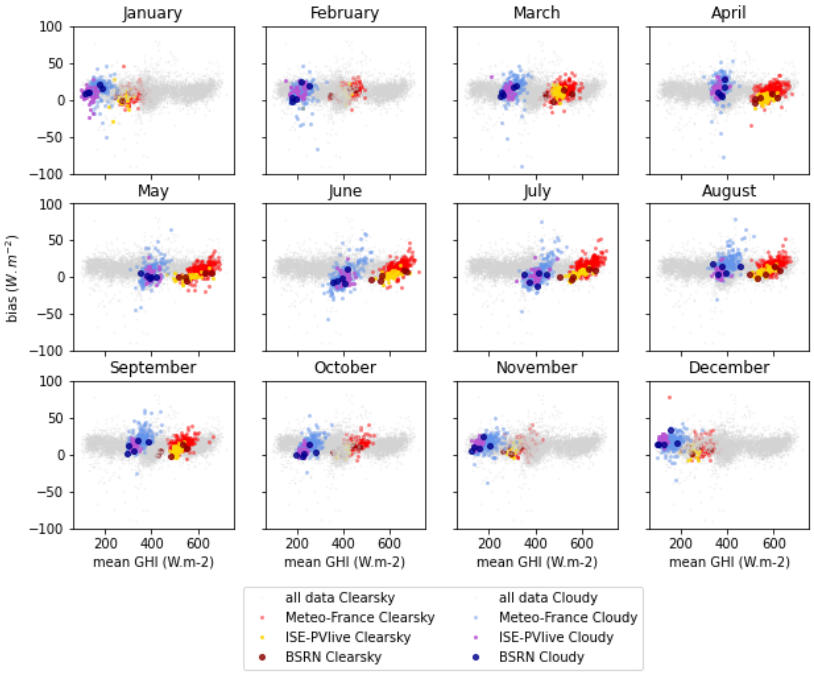

Figure 31: Scatter plot of the monthly bias of CAMS Radiation as a function of average monthly irradiance in confidently cloudy and cloud-free conditions. Each point represents a month and a station, and the colours represent the different networks.

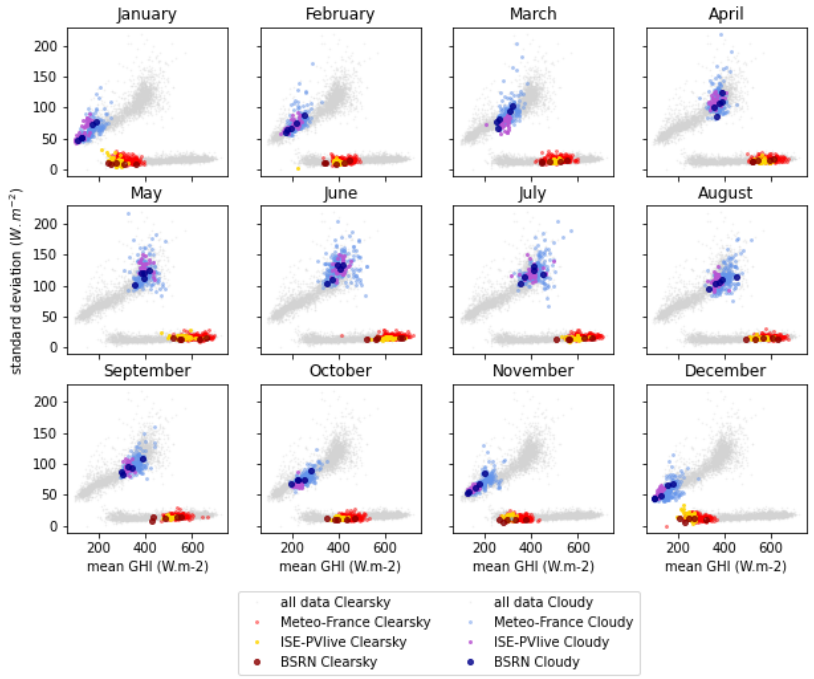

Figure 32: Scatter plot of the monthly standard deviation of the error of CAMS Radiation as a function of average monthly irradiance in confidently cloudy and cloud-free conditions. Each point represents a month and a station, and the colours represent the different networks.

Annex 2: CRS error metrics as a function of station elevation in confidentially cloudy and cloudless conditions

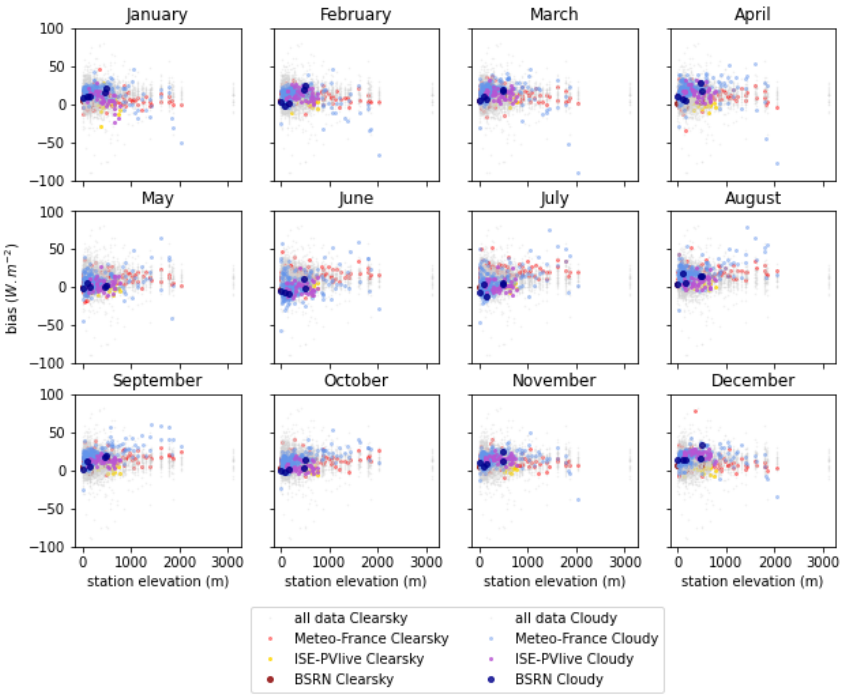

Figure 33: Scatter plot of the monthly bias of CAMS Radiation as a function of station elevation in confidently cloudy and cloud-free conditions. Each point represents a month and a station, and the colours represent the different networks.

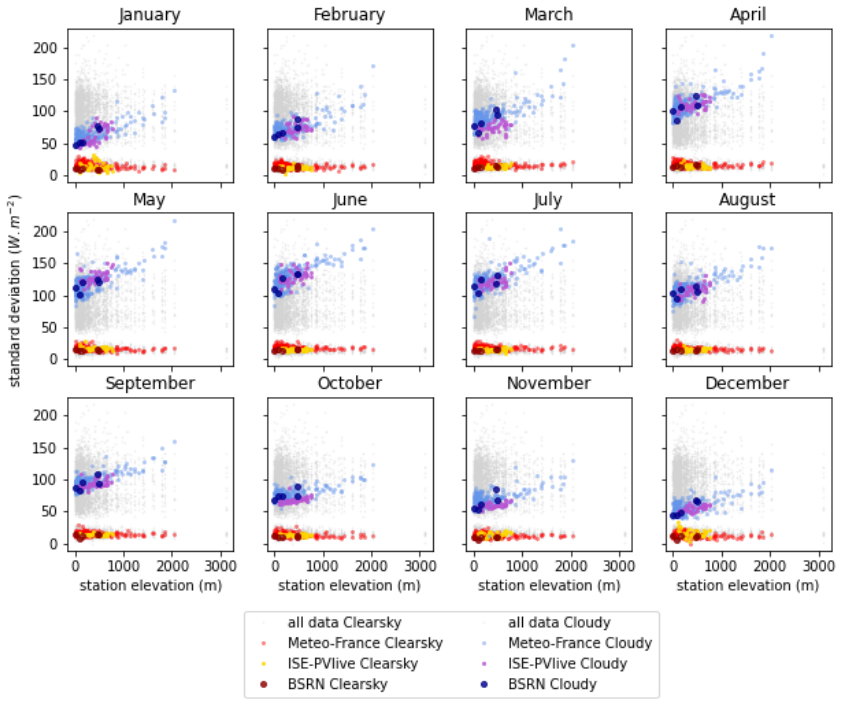

Figure 34: Scatter plot of the monthly standard deviation of the error of CAMS Radiation as a function of station elevation in confidently cloudy and cloud-free conditions. Each point represents a month and a station, and the colours represent the different networks.

Document History

Version	Author(s)	Date	Changes
0.1	Yves-Marie Saint-Drenan Jorge Lezaca	28 November 2024	Initial version
0.2	Yves-Marie Saint-Drenan Jorge Lezaca	03 December	Revised version including comments from Faiza Azam, and Mireille Lefevre
1.0	Yves-Marie Saint-Drenan Jorge Lezaca	16 December	Final version including comments from CAMEO internal reviewers

Internal Review History

Internal Reviewers	Date	Comments
Vincent Huijnen + Miró van der Worp KNMI, Yana Karol GRASP	Dec 2024	