# Evaluating CM-SAF Solar Radiation CLARA-A1 and CLARA-A2 Datasets in Scandinavia

Bilal Babar, Rune Graversen and Tobias Boström

*Energy and Climate group, Department of Physics and Technology, The Arctic University - University of Tromsø, Norway*

**Abstract**

Estimating/retrieving solar radiation through satellite-based remote sensing provides larger spatial coverage compared to other methods. Accurate estimates of incoming solar radiation is important when planning new solar energy installations. In addition, these estimates are also used in climate studies. Geostationary satellites are ideal for estimating solar radiation but cannot be used for high latitudes because of an unfavourable viewing angle; however, polar-orbiting satellites provide an alternative. **CL**oud, **A**lbedo **RA**diation edition 2 (CLARA-A2) is the latest retrieval product of cloud properties, surface albedo and surface solar radiation by Satellite Application Facility on Climate Monitoring (CM-SAF) based on Advance Very High Resolution Radiometer (AVHRR) observations from polar orbiting satellites. This data set covers the whole earth and provides daily and monthly averages. In this study, we have evaluated the CLARA-A2 data set and the previous version CLARA-A1 to *in-situ* high-quality observations from specific locations in Scandinavia, with a focus on solar radiation at high latitudes. The results show that both datasets perform within the target accuracies of CM-SAF, although the new data points, which were previously not available in CLARA-A1 due to snow-cover and cloud differentiation, have high deviations. Nevertheless, yearly average energy estimates are more accurate in CLARA-A2 because of these new points. For Swedish locations, mean absolute deviation (MAD) of 8.1 $\mathrm{Wm}^{-2}$ and 8.7 $\mathrm{Wm}^{-2}$ for CLARA-A1 and A2 respectively were calculated for updated values. Similarly, for Norwegian locations MAD of 8 $\mathrm{Wm}^{-2}$ and 8.9 $\mathrm{Wm}^{-2}$ were calculated for

*Preprint submitted to Elsevier* *March 13, 2018*

*Email address:* bilal.babar@uit.no (Bilal Babar, Rune Graversen and Tobias Boström)

CLARA-A1 and A2. Overall, for all locations MAD lies at 8.1 $Wm^{-2}$ and 8.8 $Wm^{-2}$ for CLARA-A1 and A2, respectively. CLARA A2 has more temporal data points than CLARA A1, however, the MAD of the new data points that were not available in CLARA-A1 are 15.2 $Wm^{-2}$ and 17.7 $Wm^{-2}$ for Swedish and Norwegian sites, respectively.

*Keywords:* CLARA A1 and A2, Scandinavia, ECMWF, Arctic, solar radiation estimation, polar orbiting satellites

---

## 1. Introduction

The surface radiation budget at the Earth plays a central role in climate monitoring and analysis of different meteorological parameters. Recent studies such as (Stroeve et al., 2014; Arndt and Nicolaus, 2014) make use of the surface radiation fluxes to indicate changing atmospheric and environmental conditions. In addition, surface radiation averages are used in the planning phase of the feasibility of solar energy conversion installations such as solar thermal or photovoltaic systems. Feasibility studies are important for choosing the optimal energy mix, as evident from the recent global status report by Renewable Energy Policy Network for the 21$^{st}$ Century (Ren21, 2017). The increase in the solar energy deployment in the past few years makes such datasets even more important for feasibility studies of future installations. In the Arctic regions there has been a growing interest in the use of clean and renewable energy sources, but the lack of reliable solar data hinders the socio-political decision-making processes. The focus of this paper is on validation and discussion of the improvements and shortcomings of the second edition of **CL**oud, **A**lbedo **RA**diation (CLARA) dataset for high latitude areas of Norway and Sweden. The retrieval quality of both data sets is tested against *in-situ* observations from locations at varying latitudes. In addition, these sites have different topography, especially in the Norwegian part.

Large solar power plants require preliminary data such as potential site locations and area-specific designs. The potential of a location is needed on a

monthly and annual basis (Stoffel et al., 2010). The designs may vary, for example at high latitude locations, single or dual axis tracking increases the output energy by approximately 50% (Huld et al., 2010; Good et al., 2011). In addition, inter-annual variability of solar energy is used as a measure of change in received levels of radiation through a certain period to find uncertainties in the energy production at the locations where the solar energy units are planned (Kariuki and Sato, 2018). Long time series usually of the magnitude of multi-decadal order of solar radiation are analyzed in the preplanning of power plants (Meyer et al., 2006). In most cases satellite-based databases or climate models are used to simulate solar-radiation parameters on a longer term, as these are usually not available from *in-situ* ground measuring stations. A common belief is that active solar energy production at high latitudes is not feasible since often the solar energy potential is underestimated. It is often neglected that the cold climate can be beneficial for solar energy harvesting as the efficiency of silicon solar cells increase at low temperatures (Skoplaki and Palyvos, 2009), and the presence of snow covers reflect solar radiation thereby boosting the output power. However, there are some challenges with solar energy at high latitudes such as a large seasonal variation in solar insolation, and a mismatch with the users demands. In this paper we focus on the challenge of accessing accurate solar irradiation data at high latitudes.

Various specialized databases are available for surface radiation estimation, including, European Solar Radiation Atlas (ESRA), solar data (SoDa), Satel-Light, Meteonorm, Photovoltaic Geographical Information System (PVGIS) etc (Dunlop et al., 2006). However, most datasets are based on geostationary satellites and therefore do not provide coverage above 60-65 degrees latitude. Others that use different satellite assimilation techniques take very few ground measuring stations into account, and thus cannot be considered as accurate for high latitudes. For locations above 60 degrees, retrieval methods based on observations from polar-orbiting satellites provide a solution, since these are shown to result in more accurate estimates than those obtained based on other remote sensing methods or empirical model estimation technique (Pinker and Laszlo,

1992; Besharat et al., 2013). As shown by Polo et al. (2016), satellite estimation of solar radiation has considerably improved and it is the second best option after the ground measurement methods. The Satellite Application Facility on Climate Monitoring (CM-SAF) provides multiple climate data records for cloud detection, albedo and surface radiation. CLARA data sets are one such product that can be used at high latitude locations because of its global coverage.

The most accurate *in-situ* instrument for recording global horizontal irradiance (GHI) is a pyranometer (Iqbal, 2012). In high-latitude Arctic regions, there are few meteorological stations and only a subset of these record solar radiation. The large distances between measurement hinder the exploitation of new sites for solar energy based on *in-situ* observations. Alternatively, solar radiation maps based on polar orbiting satellites can be used at these locations.

Some previous studies including Riihelä et al. (2015) and Urraca et al. (2017) have performed error statistics on the estimation of CLARA-A1 and CLARA-A2. In (Riihelä et al., 2015), authors performed an extensive evaluation of CLARA-A1 and SARAH-A1 over Sweden and Finland, while in (Urraca et al., 2017) a few sites from Norway were included. The novelty of this work lies in the comparison of the 2 datasets on Norway and Sweden over a larger number of sites and years. Moreover, the strength and weakness of the datasets are analysed in depth.

This paper is organised as follows. Section 2 describes the sites used in the study and the sources of *in-situ* measurements. Section 3 describes methods used to process the data and the statistical evaluations performed. Section 4 presents the result and a discussion on these results. Section 5 concludes this work.

## 2. Sites

The locations used in this study are at different latitudes in Norway and Sweden. The reason for this is that the performance of Cloud, Albedo Radiation (CLARA) datasets can be assessed by taking into account that at higher

latitudes there are more images provided by polar orbiting satellites (14 per day at poles). Coordinates of the locations, altitude and terrain information are provided in table 1. The *in-situ* data used to validate both data sets are acquired from two different sources. For Norway, the data are from Norsk institutt for bioøkonomi (NIBIO), and for the Swedish locations, the data are from the database of Sveriges meteorologiska och hydrologiska institut (SMHI). Both databases contain average hourly measurement by Kipp and Zonen CPM11 or CMP13 pyranometers. The equipment is regularly maintained and datasets are quality controlled by the respective organizations. In case of SMHI, Baseline Surface Radiation Network (BSRN) routines by (Long and Dutton, 2010) are used for quality assurance. Missing or erroneous data are corrected by using meteorological variables described by (Davies and McKay, 1989). The network was upgraded in 2006-2007 and the average ratio between old and new measurements was found to be 0.997. More detail on the upgrade is given by (Carlund, 2011). NIBIO calibrates the equipment once every year and had a major overhaul in 2013. The equipment is inspected and maintained on daily or weekly basis (http://lmt.bioforsk.no/about). In this study, an additional quality check of the on-site observations was performed, and any data flagged for low quality were discarded. In addition, NIBIO measurements having more than 10% of hourly missing values in a year were discarded (see appendix for details about the years not included in the study).

## 3. Method

### *3.1. Data Source*

CLARA edition 2 (CLARA-A2) by CM-SAF is the latest edition of CLARA datasets and was released in December 2016. The solar radiation estimates for CLARA are derived from the Advance Very High Resolution Radiometer (AVHRR) sensors on board METOP and NOAA polar orbiting satellites. The dataset is available for a 34 year period from 1$^{st}$ January 1982 to 31$^{st}$ December 2015, which is an extension of 6 years relative to the previous edition. The

Table 1: Information on the location, altitude and land cover type of the sites included in the study

| **Norway** | **Latitude** | **Longitude** | **Altitude (m)** | **Land Cover Type** |
|---|---|---|---|---|
| Tromsø | 69.65 | 18.9 | 12 | Island |
| Pasvik | 69.45 | 30.04 | 27 | Lakes/forest |
| Sortland | 68.6 | 15.28 | 14 | Coastal/fjords |
| Vågønes | 67.28 | 14.45 | 26 | Forest/Coastal |
| Tjøtta | 65.83 | 12.43 | 10 | Coastal/archipelago |
| Oslo | 60.12 | 11.3 | 162 | Rural/agricultural |
| Særheim | 58.76 | 5.65 | 90 | Inland/rural/agricultural |
| Lyngdal | 58.13 | 7.04 | 4 | Urban/Fjords/near coastal |
| | | | | |
| **Sweden** | **Latitude** | **Longitude** | **Altitude (m)** | **Land Cover Type** |
| Kiruna | 67.83 | 20.43 | 408 | Sparse forest |
| Luleå | 65.55 | 22.13 | 17 | Coastal |
| Umeå | 63.82 | 20.25 | 10 | Near coastal |
| Borlange | 60.48 | 15.43 | 140 | Urban/forest |
| Stockholm | 59.35 | 18.07 | 30 | Coastal |
| Göteborg | 57.70 | 12.00 | 5 | Coastal |
| Lund | 55.71 | 13.21 | 73 | Urban |

dataset covers the whole globe with a spatial resolution of 0.25x0.25 degrees on a regular lat-lon grid, which translates to 27.8 km at the equator. Average Surface In-coming Shortwave radiation (SIS) values are available for daily and monthly time resolutions. Instantaneous AVHRR images are processed to derive a spatio-temporal averaged dataset, consisting of cloud cover, surface albedo and surface-radiation products. The second edition is an improvement over the first edition because of the upgraded retrieval method and 6 years of additional data.

CLARA-A2 uses aerosol information, vertical integrated vapor and ozone, along with the surface albedo product to estimate incoming solar radiation, (Jörg Trentmann and Team, 2016). Estimation of surface albedo is a challenging task, which includes calculating top-of-the-atmosphere reflectance, classification of snow covered pixels, radiometric and geolocation topography correction, land use classification etc. (Kati Anttila and Team, 2016). In the case of high-latitude complex topography, a number of these methods are used to calculate the surface albedo including topography correction and classification of snow covered pixels. The viewing and illumination geometry at the satellite sensor becomes complex at low sun elevation. Such conditions increase the bidirectional surface reflectance thereby making the estimation process more complex (Kati Anttila and Team, 2016). This aspect will be further discussed in later sections. Figure 1 shows the CLARA-A2 yearly-averaged incoming solar radiation for 2009 on a horizontal surface.

Certain limitations exist in CLARA-A2; one of the main limitation is the availability of AVHRR observations. For calculating the daily averages, at least 20 observations are needed within a day and in each grid cell. In case of less than 20 images, the daily average field in question is filled with a value of -999 $Wm^{-2}$ that represents a missing value. For a given grid cell, at least 20 days of observations is required to produce the monthly averages for SIS for a given grid cell. In case of availability of less than 20 days, the field is filled with a missing value.

A shortcoming of the dataset is the low number of satellites in the 1980s and the early 1990s, and for this reason only the period from 1995 and beyond

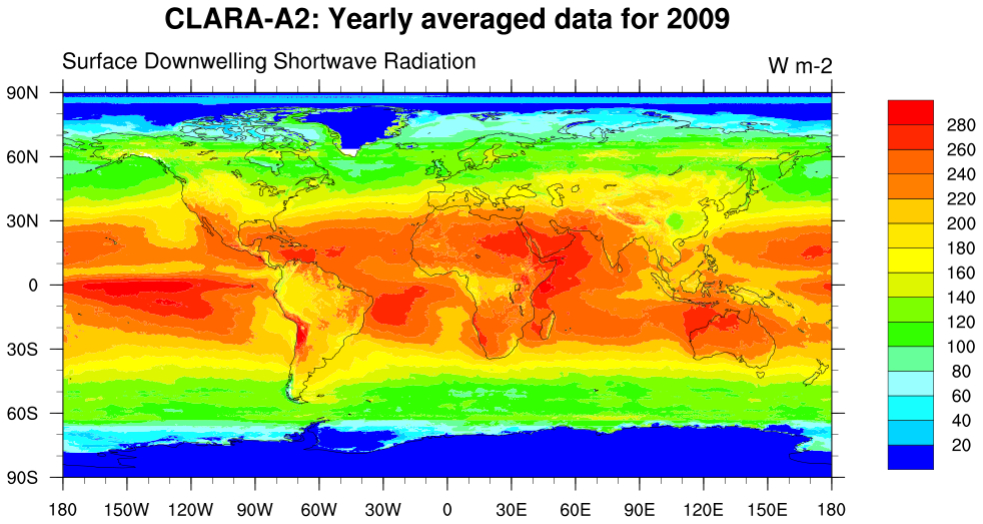


Figure 1: CLARA-A2 yearly averaged solar irradiation data for 2009 on a horizontal surface.

is considered in this study. Another shortcoming includes the orbital drift of the satellites that results in different local observation times, which changes the observation conditions. Over Greenland the data quality was found to be insufficient to fulfil the threshold accuracy requirements, therefore, the southern tip of Greenland appears to be white which shows the area having missing values.

The major improvements in the latest CLARA edition on grid cell are from the cleaning and homogenizing of the basic level-1 AVHRR radiance data and the use of Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) cloud information. In the second edition, the cloud screening ability near poles is enhanced. Especially cloud detection over snow-covers is optimized and false cloud detection is reduced by using CALIOP cloud mask and CALIOP estimated cloud-optical thickness (Karlsson et al., 2017). A new dynamic aerosol optical depth (AOD) is used in CLARA-A2 surface albedo (SAL) calculations, which was previously set at a constant value of 0.1 (Kati Anttila and Team, 2016). Moreover, the new edition uses wind speed in addition to sun zenith angles in SAL calculations (Kati Anttila and Team, 2016). Digital elevation

model used in this study is from NOAA (National Centers for Environmental Information). The snow depth data used to show the average snow depth of the areas in the analysis was obtained from ERA-Interim reanalysis (Dee et al., 2011).

### *3.2. Data processing*

The ground-measured data used in this study are hourly averaged global horizontal irradiation. Refer to section 2 for more details. The data from the SMHI database are quality controlled and flagged. From this dataset, sites flagged for bad quality were not used in the comparison. The NIBIO database is also quality controlled but not flagged. For Norway, hourly data for any year with large data gaps (10% or more of hourly values) were discarded. Missing values in this dataset were replaced by linear interpolation without taking diurnal solar elevation variation into account. For both NIBIO and SMHI, secondary standard pyranometers are used to record but these quality equipment have errors even when well-maintained and serviced. CMP11 Kipp and Zonen pyranometer have a flux measurement error of 2-5%. For monthly values lower uncertainty of 2% is expected in summer periods and 5% is expected in winter period (Wang et al., 2012). These uncertainties set an upper limit to the evaluation accuracy when estimates are compared with ground measured data (Riihelä et al., 2015).

Both CLARA datasets provide data of daily and monthly averages with a spatial resolution of 0.25x0.25 degrees (27.8 km x 27.8 km at the equator). Instead of fetching data for the closest grid point from the site locations, inverse distance weighted interpolation was used to calculated radiation values at precisely the site locations. Whenever the surrounding four grid points have more than 1 missing value for a certain time; the interpolation was replaced by a missing value of -999 $\mathrm{Wm}^{-2}$. By using this method, a slight improvement was observed in the overall deviations.

### *3.3. Statistical Evaluation of Estimations*

Different statistical measures are used to evaluate the model deviations. The most widely used measure is the Root Mean Squared Deviation (RMSD). As

an additional measure the BIAS or mean bias deviation (MBD) is used in the evaluation. Using MBD gives an insight in the general trends of under or over estimations. Mean absolute deviation (MAD) is also used for the evaluations of datasets. Because of the absolute values used in this measure, the negative and positive deviations do not cancel out each other as in the MBD. This is a good measure to compare different models as the one with smaller MAD will be the more reliable for estimations (Last et al., 2001).

## 4. Results and discussion

Table 2 shows the results of the statistical evaluation performed over the period of 1995 to 2009 over Sweden and Norway. The evaluations are arranged in decreasing latitudes in the tables. For most of the sites, CLARA-A2 provides lower RMSD values for daily means, but for monthly means, CLARA-A1 performs better or very similar to CLARA-A2.

In terms of biases, CLARA-A1 performs better at most of the sites. At some locations though the opposite pattern is found, but overall the Swedish locations show an overestimation and the Norwegian locations an underestimation. In a previous work by Riihelä et al. (2015), a similar overestimation was reported for CLARA-A1 in Sweden. The frequency of observations of the satellite also contributes to the errors, where 20 images are used to estimate daily and monthly averages, while the available frequency of ground observations is once every hour.

For both data sets, the threshold, target and the optimal accuracy is 15, 10 and 8 $\mathrm{Wm}^{-2}$ respectively, for monthly averages and 30, 25 and 20 $\mathrm{Wm}^{-2}$ for daily averages as described in (Karlsson et al., 2012; Jörg Trentmann and Team, 2012) and (Karlsson et al., 2017; Jörg Trentmann and Team, 2016), respectively. The MAD in table 2 indicates that all the results are well within these specified thresholds, and most of the sites show an optimal accuracy of 8 and 20 $\mathrm{Wm}^{-2}$ for monthly and daily averages, respectively. For Norwegian locations, monthly MAD of 8 $\mathrm{Wm}^{-2}$ was recorded for CLARA-A1 while for CLARA-A2 it was 8.9

$Wm^{-2}$ and for Swedish locations, monthly MAD was 8.1 $Wm^{-2}$ for CLARA-A1 and 8.7 $Wm^{-2}$ for CLARA-A2 $Wm^{-2}$. The overall MAD for CLARA-A1 and A2 for daily averages were 20.05 $Wm^{-2}$ and 15.65 $Wm^{-2}$ and for monthly averages 8.06 $Wm^{-2}$ and 8.82 $Wm^{-2}$, which is also within the limits of CM-SAF. For most of the sites the daily accuracies are improved in the later CLARA edition relative to the former, while CLARA-A1 performs better on monthly accuracies for most of the sites. Furthermore, CLARA-A2 has more monthly and daily mean data points than CLARA-A1, especially at higher latitudes as shown by the Hovmöller diagram in the figure 2. Higher latitudes have more snow covers, which are estimated more frequently in CLARA-A2. The availability of the datasets will be elaborated further in the subsequent sections.

Polar orbiting satellites follow a sun synchronous orbit in which the temporal resolution of sensing increases with latitude. About 14 daily observations are recorded close to the poles per satellite swath, whereas only two observations are available close to the equator (Karlsson et al., 2017). At latitudes below 65 degrees the number of images captured by polar orbiting satellites is not high enough to obtain the daily means when the day length is short, while the availability rises again above 65 degrees because of the overlapping of the satellite swath. At even higher latitude, the coverage is larger but the main challenge at such high latitudes is the snow covered surfaces (Urraca et al., 2017). In this study, the Norwegian locations have snow covers in addition to a very complex terrain including a high number of fjords and mountains (see figure 5). It is highly likely that satellite retrieval estimation methods deteriorate on mountain regions because the spatial resolution of incident light on satellite sensor is not high enough to compensate for the complex terrain, while sudden changes in weather conditions due to mountains are not compensated for with low sensing frequency as in the case of polar orbiting satellites.

Table 2: CLARA-A1 and CLARA-A2 monthly averaged comparison results from 1995 to 2009. The deviations are represented by root mean square deviation (RMSD), mean bias deviation (MBD) and mean absolute deviation (MAD). Numbers in parenthesis are the results for daily mean values. The table shows the results for Norway and Sweden seperately along with results from all sites.

| | RMSD($Wm^{-2}$) | | MBD($Wm^{-2}$) | | MAD($Wm^{-2}$) | |
|---|---|---|---|---|---|---|
| Location | A1 | A2 | A1 | A2 | A1 | A2 |
| Norwegian Locations | | | | | | |
| Tromsø | 18 (46) | 16 (24) | 3.4 (4.3) | -4 (-3) | 4.2 (10.4) | 8.7 (12) |
| Pasvik | 11 (36) | 16 (22) | 1 (2.1) | -2.9 (-2) | 3.3 (8.6) | 6.2 (8.8) |
| Sortland | 11 (21) | 18 (24) | -3.7 (-2.8) | -11.3 (-10.7) | 4.4 (7.6) | 11.5 (14.3) |
| Vågønes | 13 (35) | 11 (17) | 1.3 (2.8) | -2 (-1) | 4.3 (9.9) | 5.4 (9.6) |
| Tjøtta | 8 (33) | 7 (16) | 2.2 (3.6) | -1.3 (-0.3) | 3.7 (10.7) | 4.2 (8.4) |
| Oslo | 9 (33) | 10 (18) | -2.3 (-0.6) | -3.7 (-2.3) | 4.1 (12.5) | 5.8 (10.4) |
| Særheim | 7 (31) | 7 (16) | 1.2 (2.7) | -1.9 (-0.3) | 4.3 (13.8) | 4.4 (9.5) |
| Lyngdal | 12 (24) | 20 (34) | -2.7 (-1.7) | -7.6 (-6.6) | 6.4 (11.6) | 9.5 (13.9) |
| All Norwegian locations | 11 (34) | 14.2 (24.9) | -0.1 (1.9) | -5.6 (-4.1) | 8 (18.7) | 8.9 (13.5) |
| Swedish Locations | | | | | | |
| Kiruna | 8 (29) | 18 (24) | -0.5 (0.8) | -0.5 (0.8) | 2.6 (7.5) | 7.8 (11.1) |
| Luleå | 9 (27) | 9 (16) | 1.2 (2.7) | -0.8 (0.1) | 3.5 (8.4) | 4.3 (7.4) |
| Umeå | 8 (27) | 11 (17) | 0.5 (2.5) | -4 (-2.6) | 3.7 (8.9) | 6.8 (9.3) |
| Borlange | 9 (27) | 9 (17) | -1 (0.7) | -3.6 (-2.1) | 4 (10.8) | 9.4 (5.5) |
| Stockholm | 8 (28) | 9 (18) | 2.4 (4.6) | 3.2 (1.7) | 4.7 (12.5) | 5.5 (9.9) |
| Göteborg | 7 (25) | 7 (16) | 1.9 (3.6) | 0.8 (2.3) | 4.5 (12.4) | 4.5 (9.4) |
| Lund | 9 (25) | 8 (17) | -2.1 (-0.9) | -1.8 (-0.1) | 4.9 (11) | 5.2 (10.4) |
| All Swedish Locations | 11.7 (41.6) | 13 (46.5) | 0.5 (2.9) | -2.5 (-1.1) | 8.1 (21) | 8.7 (17.4) |
| **All Locations** | **11.4 (38.7)** | **13.5 (38.5)** | **0.2 (2.5)** | **-3.8 (-2.4)** | **8 (20)** | **8.8 (15.6)** |

This study is conducted on mountainous regions with snow covers, which not only introduces random errors but also negative biases. Furthermore, because the satellite estimation methods use the visible spectrum channels for the detection of clouds, the sensors cannot differentiate between clouds and snow cover, which further contributes to increasing the errors (Urraca et al., 2017). However, 0.6 and 0.8 µm channels are used separately in order to detect snow covers and calculating the albedo (Kati Anttila and Team, 2016). Albedos for snow are high in the near ultra-violet and visible spectrum, but it starts dropping drastically in the near infra-red region between 0.8 and 1.5 µm (Wiscombe and Warren, 1980). Most of the high latitude sites in this study have snow cover for a large part of the year. Which implies a further increase of errors in the datasets. Although the new dataset have more coverage over snow-covers, which was previously not available in CLARA-A1, but such new values have large errors. These large errors are likely due to the differentiation between snow and cloud covers (see figure 5).

### *4.1. Inter-annual stability*

As discussed earlier, inter-annual stability of a dataset provides insight into the uncertainties associated with the energy production of solar energy plants. Areas where typical ground measuring equipment are not available can take advantage of datasets provided by CM-SAF. Therefore, such datasets should be consistent throughout the periods of investigation. In figure 3 the box plot of MBD of both CLARA-A1 and A2 datasets are shown. It can be seen from the figure that the CLARA-A2 dataset has lower median bias than the CLARA-A1 dataset, with median values being closer to the zero bias. The CLARA-A2 dataset has more extreme minimum values, compared to CLARA-A1, while the maximum values are in most cases better in the CLARA-A2 dataset. Moreover, the $25^{\text{th}}$ and $75^{\text{th}}$ percentile values in CLARA-A2 data set lies approximately around -2 and 2 $\text{Wm}^{-2}$, while in CLARA-A1 these values are approximately around 0 and 4 $\text{Wm}^{-2}$. These results show that the newer edition of CLARA has more stability in terms of biases over the years included in the study period.

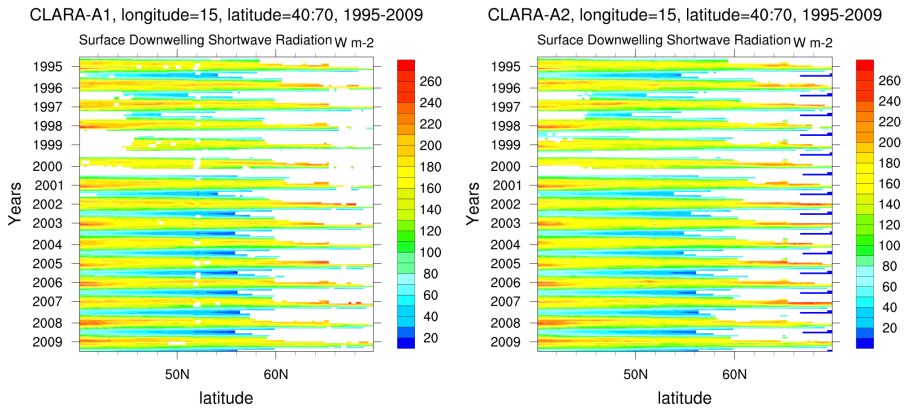


Figure 2: Hovmöller plots for CLARA-A1 and A2 datasets for the included years in the study. The plots are centered at 10 degrees longitude and span from 40 to 70 degrees latitude.

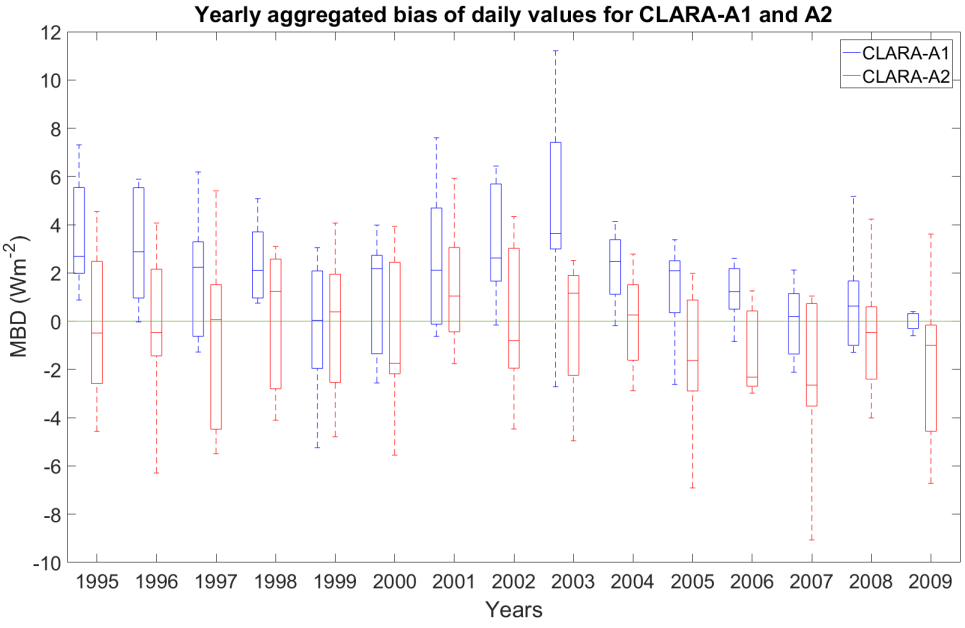


Figure 3: Box plot showing the inter annual stability of CLARA-A1 and A2. The stability is shown in terms of mean bias deviation. $25^{\text{th}}$ and $75^{\text{th}}$ percentile values are shown by the length of the box

## 4.2. Data availability

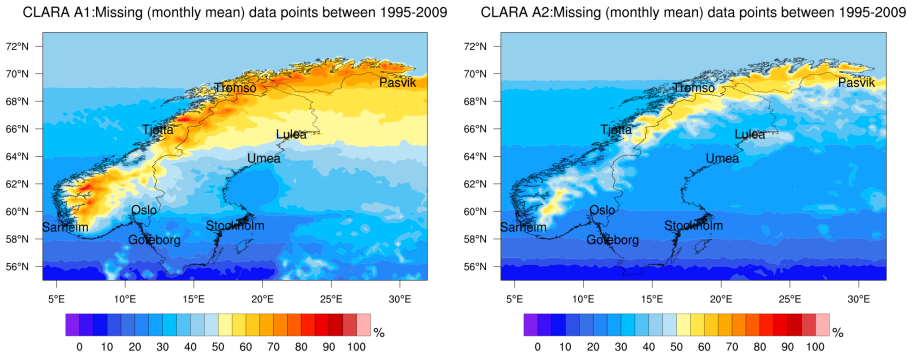


Figure 4: Percentage of monthly averaged data missing values in the datasets. Figure on the left shows the missing points in CLARA-A1 dataset between 1995 and 2009. Figure on the right shows the missing points in CLARA-A2 for the same period

One of the improvements of CLARA-A2 is the differentiation of snow-covered surfaces from cloud covers in the surface albedo calculations. Both CLARA datasets do not provide coverage over snow-covered surfaces (Riihelä et al., 2015; Karlsson et al., 2017) and such time periods are filled with missing values. Nevertheless, because of the improvement in surface albedo calculations, CLARA-A2 provide more data points than CLARA-A1. The additional data points in CLARA-A2 are mostly from the snow-cover time periods, hence there is not much improvement in the overall skills. In most cases, there is a higher degree of deviation at such locations, which further increase the deviations as a whole. As shown in figure 4, CLARA-A1 has roughly between 50 and 80% missing values in Norway and around 40 to 60% missing values in Sweden. In comparison CLARA-A2 has approximately 30 to 60% missing data in Norway and 20 to 50% missing data in Sweden. This further explains the results in Table 2, where CLARA-A1 performs better than CLARA-A2 and that the skills for the Swedish locations are better than those at the Norwegian locations. The complex topography of Norwegian locations along with a high percent-

age of snow covers at these areas have resulted in inaccurate estimations that previously were replaced by missing values and thus not taken into account in statistical evaluations. Figure 5 below shows the average snow depth in the study period between 1995 and 2009 along with a digital elevation model of the study area. By comparing figure 5 with the maps in figure 4, it can be seen that in CLARA-A1 snow-covers correspond to missing values.

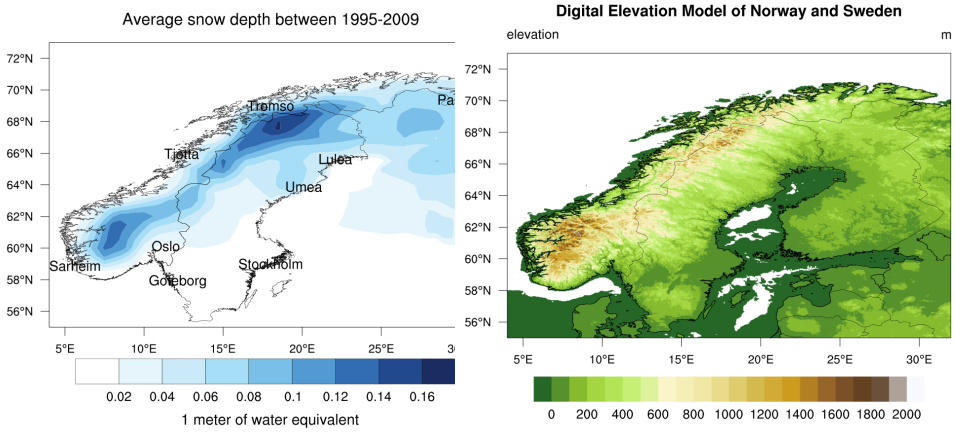


Figure 5: Average snow depth between 1995 to 2009 from ERA-Interim and topography. Larger snow depth occurs at complex terrains, and most missing data points lie in such regions.

Similarly, in CLARA-A2 there are less missing values on snow covered grid points, but still the highest amount of missing data are found on the higher snow-depth grid points and high elevation locations.

### *4.3. Seasonal variations in the datasets*

To further investigate the datasets, seasonal variation of both datasets were calculated. Data from 1995 to 2009 were divided into quarterly datasets by assigning the months from February to April to the $1^{st}$ quarter, May to July to the $2^{nd}$ quarter, August to October to the $3^{rd}$ quarter and November to January to the $4^{th}$ quarter. In this manner, we could separate the darker and

snow covered periods from the summer months.

Figure 6 illustrate the quarterly frequency of missing data in the CLARA-A1 data set and illustrates the increase in the availability of data points in the new edition compared with the previous edition. It further illustrates that due to the fact that most of the northern parts of Norway and Sweden has snow-covers, most of the missing data point in CLARA-A2 lie in these regions. The availability has increased in these northern location in CLARA-A2 when compared to CLARA-A1, though not so much in the high snow-depth mountain regions (see figure 5). The highest amount of missing values lie in the February to April months when the polar night has ended and the snow is melting.

Table 3 gives the seasonal deviations of the two datasets. It can be seen that in the $1^{st}$ and $2^{nd}$ quarter, CLARA-A2 provides more valid data points than does CLARA-A1 (see also figure 6). Missing data or no valid value at grid points means that these months are not taken into account when making any of the calculations in the study. When compared to the snow-depth map in figure 5, the regions of missing values lie approximately on the areas having higher snow-depth and complex topography. The $1^{st}$ and $4^{th}$ quarters have special conditions, where the $1^{st}$ quarter has low sun-elevation angles and the $4^{th}$ quarter includes the polar-night period. Moreover, the $1^{st}$ and $3^{rd}$ quarter have similar and opposite sun elevation angles (in the $1^{st}$ quarter the solar elevation increases while in the $3^{rd}$ quarter it decreases) but the $1^{st}$ quarter has more snow-cover than the $3^{rd}$ quarter. It also shows that in the $1^{st}$ quarter both the MBD and MAD are larger in CLARA-A2 than CLARA-A1. Low RMSD values are observed below 60 degrees in Swedish locations but not in Norwegian locations. The MBD or bias is mostly negative for CLARA-A2, with high values for Norway than for Sweden. However, due to the unavailability of data in some high latitude locations it was not possible to calculate the deviations. In the $2^{nd}$ quarter, CLARA-A1 has better RMSD measures until around north of 60 degrees after which CLARA-A2 either starts improving or provides similar values as CLARA-A1 (except for Pasvik, Sortland and Kiruna). Similarly, CLARA-A1 again provides better MBD and MAD values. In the $3^{rd}$ and $4^{th}$

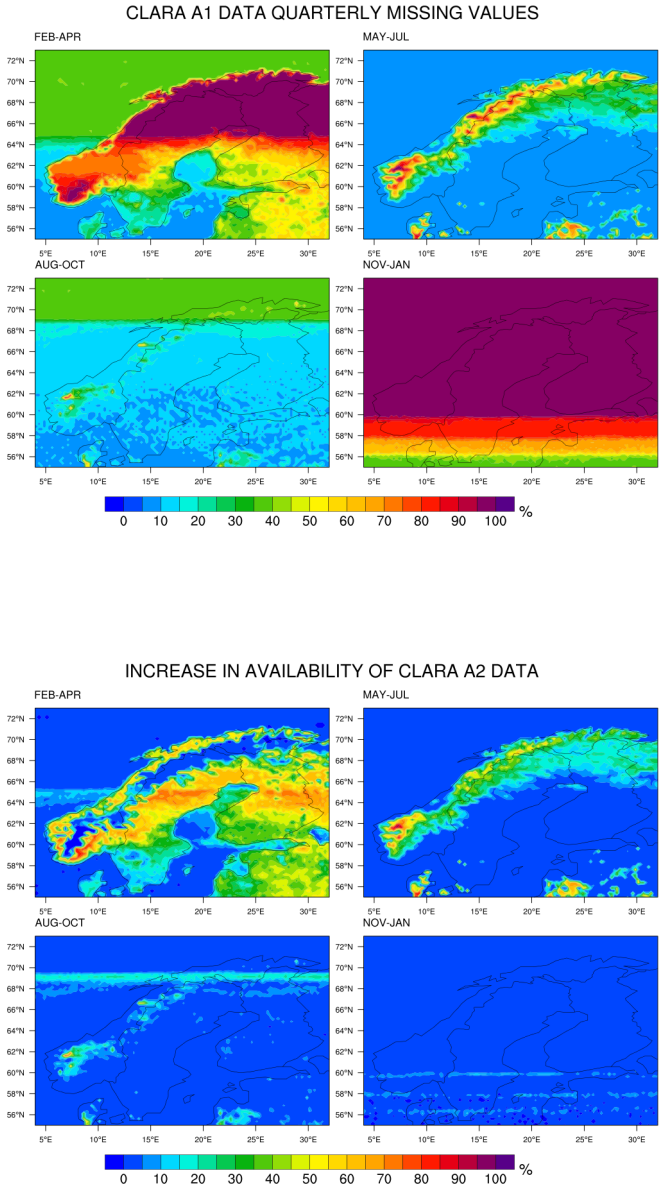


Figure 6: The top figure shows the percentage of monthly missing data in CLARA-A1 in each quarter. The lower figure shows the percentage increase in the availability of CLARA-A2 dataset in each quarter. The highest increase is in the areas that have complex topography in addition to snow covers.

Table 3: Quarterly deviations for CLARA-A1 and CLARA-A2 datasets. The table shows the seasonal variation in the biases of both datasets. Monthly average values for the years included in the study were divided into four quarters that are denoted by Q. CLARA A1 and A2 datasets are denoted by A1 and A2, respectively.

| Norway/Quarter | | RMSD($Wm^{-2}$) | | MBD($Wm^{-2}$) | | MAD($Wm^{-2}$) | | Correlation | |
|---|---|---|---|---|---|---|---|---|---|
| | | A1 | A2 | A1 | A2 | A1 | A2 | A1 | A2 |
| Tromsø | Q1 | - | 11.6 | - | -5 | - | 5 | - | 0.9 |
| | Q2 | 26 | 25.2 | 5.6 | -11.3 | 5.9 | 19.8 | 0.60 | 0.9 |
| | Q3 | 14.6 | 11.4 | 7.8 | 1.2 | 11 | 8.7 | 0.98 | 0.9 |
| | Q4 | - | 2.3 | - | -1.2 | - | 1.2 | - | - |
| Pasvik | Q1 | - | 3.2 | - | -0.6 | - | 0.6 | - | - |
| | Q2 | 13.9 | 27.5 | 4.3 | -5.9 | 4.8 | 13 | 0.95 | 0.7 |
| | Q3 | 10.2 | 11.5 | 0 | -3.5 | 8.4 | 9.7 | 0.98 | 0.9 |
| | Q4 | - | 3.7 | - | -1.4 | - | 1.4 | - | - |
| Sortland | Q1 | 22.3 | 22.9 | -1.6 | -11.6 | 1.6 | 11.6 | - | 0.97 |
| | Q2 | 12 | 22.6 | -5.1 | -17.9 | 6.7 | 18.4 | 0.9 | 0.97 |
| | Q3 | 10 | 13.7 | -7.6 | -11.9 | 8.5 | 12 | 0.9 | 0.99 |
| | Q4 | 4.7 | 9.6 | -0.5 | -3.7 | 0.6 | 4 | 0.9 | 0.88 |
| Vågønes | Q1 | 5.1 | 8 | 0.2 | -4 | 0.2 | 4 | - | 0.99 |
| | Q2 | 12 | 8.4 | 6.9 | 3 | 6.9 | 6.8 | 0.9 | 0.99 |
| | Q3 | 15 | 15.5 | -1.4 | -5.3 | 8.6 | 8.8 | 0.9 | 0.96 |
| | Q4 | 9.4 | 7.2 | -0.4 | -2 | 1.4 | 2.1 | 0.6 | 0.94 |
| Tjøtta | Q1 | 9.4 | 7.8 | 0.6 | -1.9 | 0.6 | 2.2 | - | 0.89 |
| | Q2 | 10.8 | 7.7 | 6.9 | 2.1 | 8.5 | 6.6 | 0.9 | 0.98 |
| | Q3 | 6.5 | 7.8 | 1 | -3.9 | 5 | 6.4 | 0.9 | 0.99 |
| | Q4 | 3 | 3.4 | 0.3 | -1.5 | 0.6 | 1.5 | 0.9 | 0.99 |
| Oslo | Q1 | 10.7 | 29.4 | -1.2 | -12.5 | 1.2 | 12.5 | 0.97 | 0.87 |
| | Q2 | 21.5 | 20.6 | -6 | -3.4 | 10 | 8.3 | 0.88 | 0.89 |
| | Q3 | 12.4 | 11.7 | -5.3 | -3.8 | 9.1 | 8.5 | 0.97 | 0.97 |
| | Q4 | 6.1 | 9.3 | -1 | -2.5 | 1.5 | 2.5 | 0.93 | 0.93 |
| Særheim | Q1 | 5.7 | 6.7 | 1.5 | -3 | 2.9 | 3.3 | 0.99 | 0.98 |
| | Q2 | 6.8 | 5.8 | 3.3 | 1.7 | 5.6 | 4.5 | 0.99 | 0.99 |
| | Q3 | 7.9 | 9 | -0.3 | -3.6 | 7 | 7.3 | 0.99 | 0.99 |
| | Q4 | 3.7 | 5.9 | 0.4 | -2.5 | 1.5 | 2.6 | 0.98 | 0.99 |
| Lyngdal | Q1 | 10.2 | 34.5 | -0.5 | -10.8 | 2.9 | 10.8 | 0.97 | 0.66 |
| | Q2 | 12.5 | 13.8 | -1.2 | -4.4 | 9.7 | 10.9 | 0.96 | 0.96 |
| | Q3 | 14.4 | 16.6 | -8.2 | -11.2 | 10.2 | 11.9 | 0.97 | 0.98 |
| | Q4 | 8.3 | 11.5 | -1.1 | -4.2 | 2.8 | 4.2 | 0.90 | 0.90 |

| Sweden/Quarter | | RMSD($Wm^{-2}$) | | MBD($Wm^{-2}$) | | MAD($Wm^{-2}$) | | Correlation | |
|---|---|---|---|---|---|---|---|---|---|
| | | A1 | A2 | A1 | A2 | A1 | A2 | A1 | A2 |
| Kiruna | Q1 | - | 15 | - | -4.3 | - | 4.3 | - | 1.00 |
| | Q2 | 8.3 | 29.6 | 0.9 | -12.6 | 3.1 | 18.3 | 0.94 | 0.84 |
| | Q3 | 8 | 8.7 | -3.1 | -3.3 | 6.4 | 6.8 | 0.99 | 0.99 |
| | Q4 | 3.1 | 4 | 0.2 | -1.7 | 0.7 | 1.7 | 0.86 | 0.99 |
| Luleå | Q1 | - | - | - | - | - | - | - | - |
| | Q2 | 12.5 | 12.1 | 6 | 2.6 | 7 | 7.6 | 0.94 | 0.96 |
| | Q3 | 7.9 | 8.8 | -1.3 | -3.5 | 6.3 | 7.3 | 0.99 | 0.99 |
| | Q4 | 3.7 | 4.5 | -0.1 | -2.3 | 0.9 | 2.2 | 0.76 | 0.99 |
| Umeå | Q1 | 2.3 | 13.9 | 0.3 | -6.4 | 0.4 | 6.4 | 0.84 | 0.99 |
| | Q2 | 9.3 | 11.6 | 4.8 | -1.5 | 6 | 9.3 | 0.97 | 0.98 |
| | Q3 | 9.3 | 10 | -2.7 | -5.2 | 7.5 | 8.6 | 0.99 | 0.99 |
| | Q4 | 3.3 | 5.4 | -0.3 | -3 | 0.8 | 3 | 0.85 | 0.99 |
| Borlange | Q1 | 4.2 | 11 | -0.8 | -6.5 | 0.9 | 6.5 | 1.00 | 0.99 |
| | Q2 | 7.7 | 6.4 | 0.2 | -1 | 6 | 5.1 | 0.98 | 0.99 |
| | Q3 | 9.7 | 9.6 | -4.3 | -5.6 | 7.3 | 8.2 | 0.98 | 0.99 |
| | Q4 | 10 | 9.7 | 0.8 | -1.2 | 2 | 2.3 | 0.61 | 0.68 |
| Stockholm | Q1 | 14.1 | 13.5 | 1.2 | -2 | 5.2 | 6.1 | 0.88 | 0.90 |
| | Q2 | 22 | 23 | 5 | 7 | 18.2 | 19.3 | 0.81 | 0.82 |
| | Q3 | 29.9 | 31.4 | -1.2 | 0.2 | 22.9 | 23.2 | 0.81 | 0.80 |
| | Q4 | 10 | 11.1 | 0.6 | -2 | 3.9 | 4.2 | 0.84 | 0.83 |
| Göteborg | Q1 | 5.5 | 3.5 | 2.5 | -0.9 | 3 | 1.8 | 1.00 | 1.00 |
| | Q2 | 9 | 8.9 | 5.8 | 6.4 | 7.5 | 7.7 | 0.99 | 0.99 |
| | Q3 | 7.5 | 7.6 | -0.9 | 0.01 | 6.1 | 6.2 | 0.99 | 0.99 |
| | Q4 | 4 | 6.2 | 0.2 | -2.4 | 1.5 | 2.5 | 0.98 | 0.98 |
| Lund | Q1 | 5 | 6.6 | 1.8 | -2.8 | 2.7 | 4 | 1.00 | 0.99 |
| | Q2 | 8.5 | 7.5 | -1.5 | 2.9 | 4.9 | 6.2 | 0.98 | 0.99 |
| | Q3 | 12.6 | 9.1 | -8.8 | -4 | 9.4 | 7.1 | 0.98 | 0.99 |
| | Q4 | 4.6 | 6 | -0.03 | -3.3 | 2.5 | 3.4 | 0.99 | 0.99 |

quarters, all the measures are either similar in both the datasets or slightly worse in CLARA-A2 for both Norwegian and Swedish location. Based on the observations it can be said that although CLARA-A2 has more coverage over snow-covered areas it still provides large deviations at high latitude locations.

### *4.4. Analysis of the new and updated monthly average values in CLARA-A2*

By comparing CLARA-A1 and A2, it can be seen that there are two major changes in the availability of data. First, there are fewer missing values in A2 and secondly, the adjacent grid point values are also updated in CLARA-A2 due to the use of different methods of estimation. This section provides an evaluation of the new and updated monthly means estimations separately. The values marked with "New" are the values which were not available in CLARA-A1 (marked as a missing values) but that are available in CLARA-A2. The values marked with "Updated" are those values which were available in CLARA-A1 but these got updated because of the use of new algorithms. In this way we could separately analyse the improvement of CLARA-A2. Table 4 shows the RMSD, MBD, MAD and the number of new values in CLARA-A2. For the newly added added data points in CLARA-A2 the MAD target accuracies for all locations are above the limits (17.7 $\mathrm{Wm}^{-2}$ for Norway and 15.2 $\mathrm{Wm}^{-2}$ for Sweden). Individually for both Sweden and Norway, the updated values are very similar and within the target (8.3 $\mathrm{Wm}^{-2}$ for both Norway and Sweden). Table 4 also shows the overall accuracies of both datasets for all Norwegian and Swedish locations. Overall accuracies for both datasets also are within the limits.

Furthermore, the new values in CLARA-A2 have a constant negative bias that shows the underestimation in these values. The cause for this underestimation can be attributed to the inaccurate detection of snow-covers. The RMSD section of the table shows that the new values have very high deviations for high-latitude locations in both countries; nevertheless, the updated values have relatively low RMSD because of the upgraded retrieval method and absence of snow-covers.

Table 4: Analysis of the new and updated solar radiation values in CLARA-A2 for Norwegian and Swedish locations. The column marked with New are the values which were not available in CLARA-A1 (shown in last column, No. of new values), while the updated values are the ones which were available in CLARA-A1 but were updated in CLARA-A2

| Norwegian Location | RMSD ($Wm^{-2}$) | | MBD($Wm^{-2}$) | | MAD($Wm^{-2}$) | | No. of new values |
|---|---|---|---|---|---|---|---|
| | New | Update | New | Update | New | Update | |
| Tromsø | 25 | 14 | -4.7 | 0.9 | 5 | 3.2 | 20 |
| Pasvik | 44 | 12 | -2.3 | -0.1 | 2.3 | 3.4 | 8 |
| Sortland | 30 | 15 | -4.4 | -6 | 4.5 | 6.1 | 18 |
| Vågønes | 9 | 13 | -1.1 | -0.7 | 1.3 | 3.9 | 23 |
| Tjøtta | 7 | 8 | -0.7 | -0.5 | 0.7 | 3.4 | 14 |
| Oslo | 16 | 8 | -2 | -1.7 | 2 | 3.8 | 15 |
| Særheim | 9 | 7 | -0.3 | -1.6 | 0.3 | 4.2 | 4 |
| Lyngdal | 46 | 14 | -2 | -5.6 | 2 | 7.4 | 12 |
| ALL SITES | 25.9 | 11.5 | -17 | -3.7 | 17.7 | 8.3 | 114 (12%) |

| | | | |
|---|---|---|---|
| CLARA-A1(All Included) | 11 | -0.1 | 8 |
| CLARA-A2(All included) | 14.2 | -5.6 | 8.9 |

| Swedish Location | RMSD ($Wm^{-2}$) | | MBD($Wm^{-2}$) | | MAD($Wm^{-2}$) | | No. of new values |
|---|---|---|---|---|---|---|---|
| | New | Update | New | Update | New | Update | |
| Kiruna | 37 | 8 | -4.7 | -0.6 | 4.7 | 2.9 | 26 |
| Luleå | 17 | 9 | -0.5 | -0.1 | 0.5 | 3.6 | 6 |
| Umeå | 18 | 8 | -2.4 | -1.3 | 2.4 | 4.1 | 27 |
| Borlange | 12 | 8 | -1.1 | -2.4 | 1.1 | 4.4 | 18 |
| Stockholm | 6 | 23 | -0.2 | 2 | 0.2 | 13 | 8 |
| Göteborg | 6 | 7 | -0.1 | 0.9 | 0.1 | 4.4 | 6 |
| Lund | 9 | 7 | -0.6 | -1.1 | 1.1 | 4.1 | 30 |
| ALL SITES | 20.6 | 11.9 | -14.5 | -0.7 | 15.2 | 8.3 | 121 (9.6%) |

| | | | |
|---|---|---|---|
| CLARA-A1 (All included) | 11.7 | 0.5 | 8.1 |
| CLARA-A2 (All included) | 13 | -2.5 | 8.7 |

### *4.5. Analysis of annual energy estimates*

The total annual energy estimate at a site is an important parameter for planning purposes. In addition to daily and monthly averages that are used in the inter-annual stability for energy production, annual energy averages give an insight into the total energy that can be harvested at potential site locations. Table 5 shows the RMSD, MBD and MAD of yearly averaged hourly solar irradiances of CLARA-A1 and A2. In this analysis, CLARA-A2 performs considerably better than CLARA-A1 in all areas. Moreover, average annual energy is also listed for both CLARA datasets and *in-situ* values. For calculating yearly energy values, mean hourly values from ground-measured data and mean daily values from CLARA datasets were used. By comparing the energy potential estimates it can be seen that CLARA-A2 provides better estimates than CLARA-A1. The energy estimates are better in CLARA-A2 due to the fact that it provides more data points than CLARA-A1. Fewer data points in the time series means that the energy estimates for CLARA-A1 results in lower estimates than both CLARA-A2 and ground observed data.

The energy estimates provided in table 5 are for the yearly solar radiation received on a horizontal plane per area averaged over the study period. At high latitude locations, the elevations of the sun are often very low and consequently the horizontal solar density decreases. The difference between high and low latitude locations is considerably less when looking at an optimally inclined or a tracking surface.

## 5. Conclusion

In this work, we evaluated two datasets derived from polar orbiting satellites. CLARA-A2, the newer version of the CM-SAF polar orbiting satellite-based database, is derived with a procedure including improvements in cloud cover and snow cover distinction; hence, there are more data points taken into account in the new dataset. Still, missing values exist in the new dataset due to lack of differentiation between clouds and snow covers. However, the newer

Table 5: This table shows annual average solar radiations error analysis for CLARA-A1 and A2 for Norwegian and Swedish locations in terms of RMSD, MBD and MAD. The portion of the table labelled as Power is expressed in $Wm^{-2}$. The right side of the table shows the annual average energy estimates of CLARA-A1, A2 and ground-observed data expressed in $kWhm^{-2}y$.

| Norwegian Locations | Power | | | | | | Energy | | |
|---|---|---|---|---|---|---|---|---|---|
| | RMSD ($Wm^{-2}$) | | MBD ($Wm^{-2}$) | | MAD ($Wm^{-2}$) | | A1(avg) $kWhm^{-2}y$ | A2(avg) $kWhm^{-2}y$ | Obs(avg) $kWhm^{-2}y$ |
| | A1 | A2 | A1 | A2 | A1 | A2 | | | |
| Tromsø | 69.7 | 9.8 | 68.7 | 7.7 | 68.7 | 7.7 | 469.1 | 643.7 | 687.4 |
| Pasvik | 65.7 | 12 | 65.3 | 9.3 | 65.3 | 9.8 | 497.4 | 544.6 | 718.2 |
| Sortland | 50.7 | 4.7 | 48.8 | 2.1 | 48.8 | 3.7 | 600.3 | 664.8 | 780.4 |
| Vågønes | 53.7 | 13.8 | 53.1 | 12.9 | 53.1 | 12.9 | 600 | 724 | 733.9 |
| Tjøtta | 57.5 | 25.6 | 56.8 | 25.3 | 56.8 | 25.3 | 698.9 | 749.9 | 768.2 |
| Oslo | 48.2 | 31.3 | 47.2 | 30.2 | 47.2 | 30.2 | 827.5 | 902.4 | 948.7 |
| Særheim | 29.7 | 21.6 | 28.8 | 21.3 | 28.8 | 21.3 | 913.8 | 901.7 | 921.7 |
| Lyngdal | 31.9 | 21.7 | 29.7 | 17.2 | 29.7 | 18.7 | 915.7 | 939.8 | 1032.9 |

| Swedish Locations | Power | | | | | | Energy | | |
|---|---|---|---|---|---|---|---|---|---|
| | RMSD ($Wm^{-2}$) | | MBD ($Wm^{-2}$) | | MAD ($Wm^{-2}$) | | A1(avg) $kWhm^{-2}y$ | A2(avg) $kWhm^{-2}y$ | Obs(avg) $kWhm^{-2}y$ |
| | A1 | A2 | A1 | A2 | A1 | A2 | | | |
| Kiruna | 48.6 | 9.1 | 47.5 | 8.4 | 47.5 | 8.4 | 525 | 654.7 | 804.5 |
| Luleå | 62.3 | 34.5 | 61.5 | 34.3 | 61.5 | 34.3 | 704.3 | 728.1 | 895.8 |
| Umeå | 51 | 18.9 | 48.6 | 17.4 | 48.6 | 17.4 | 777.2 | 860.4 | 916.7 |
| Borlange | 43.7 | 29.8 | 42.7 | 28.9 | 42.7 | 28.9 | 846.7 | 893.3 | 937.2 |
| Stockholm | 38.3 | 32.8 | 36.6 | 30.4 | 36.6 | 30.4 | 984.5 | 998 | 993.4 |
| Göteborg | 32.1 | 26.9 | 30.3 | 24.9 | 30.3 | 24.9 | 968.3 | 966.5 | 969.6 |
| Lund | 18.8 | 17.4 | 4.9 | 9 | 13.4 | 11.8 | 791.1 | 1013 | 1034.7 |

edition does not considerably improve the estimates for Northern Scandinavia. The evaluation metrics used in the study provides an insight into the performance of these datasets. CLARA-A2 is observed to provide underestimation at most locations, while CLARA-A1 provides more positive biases. This underestimation can be associated with the snow and cloud detection and the difficulties to differentiate between the two, which hopefully will be further improved in CLARA-A3, the next edition of this dataset that is planned to be launched in 2020. The CLARA-A2 dataset has less intra-annual variability than CLARA-A1, and along with the spatiotemporal resolution, it provides a more reliable dataset for areas below 60 degrees latitude. For the magnitude of errors presented in this study, consideration should be given to the complex topography especially in the case of Norwegian sites. Table 2 shows that MBD and MAD values are predominantly higher at Norwegian location. However, at most locations the target monthly average accuracies of 9 $\mathrm{Wm}^{-2}$ for CLARA-A2 and 10 $\mathrm{Wm}^{-2}$ for CLARA-A1 are achieved, along with daily average accuracies of 18 $\mathrm{Wm}^{-2}$ for CLARA-A2 and 20 $\mathrm{Wm}^{-2}$ for CLARA-A1. A quarterly deviation analysis shows that due to the complex topography and snow cover in Norwegian locations, CLARA-A2 does not provide more accurate estimates than CLARA-A1. Analysis on the new data points of CLARA-A2, that were previously not available, shows that these new values have very high deviations. Nevertheless, yearly energy estimates of CLARA-A1 are predominantly lower than CLARA-A2 estimates since there are simply more data points in CLARA-A2. To conclude, even if CLARA-A2 has a higher negative bias than CLARA-A1 at the specific common data points, CLARA-A2 still has more accurate yearly energy estimates because it has more data points than CLARA-A1.

## Appendix

Years within the studying period of 1995 to 2009 not included in this work.

Table 6: Detail of years not included in the study for each location.

| | |
|---|---|
| Kiruna | N.A |
| Luleå | N.A |
| Umeå | N.A |
| Borlange | N.A |
| Stockholm | 1998 |
| Göteborg | N.A |
| Lund | N.A |
| Tromsø | 1995,1996,2000,2001,2002,2006,2007,2008 |
| Pasvik | 1995,1996,2006,2007 |
| Sortland | 1995,1996,1997,2000,2003,2007 |
| Vågønes | 1995,1996,1997,2007 |
| Tjøtta | 1995,1996,1997,2006,2007 |
| Oslo | 1995,1996,1997,1998,2006,2007 |
| Særheim | 1995,1996,2000,2006,2007 |
| Lyngdal | 1995,1996,2003 |


## Acknowledgements

The authors would like to thank Jörg Trentmann and CM-SAF for providing support and assistance regarding the datasets. We would also like to thank SMHI and NIBIO for providing ground measured solar radiation data. This work is supported by Troms county and industry development fund under the project title,"Renewable energy in the arctic - academy and business in a joint effort" RDA12/46.


## References


J. Stroeve, T. Markus, L. Boisvert, J. Miller, A. Barrett, Geophysical Research Letters 41 (2014) 1216–1225.

S. Arndt, M. Nicolaus, The Cryosphere 8 (2014) 2219–2233.

R. Ren21, Renewable Energy Policy Network for the 21st Century, Paris, France (2017).

T. Stoffel, D. Renné, D. Myers, S. Wilcox, M. Sengupta, R. George, C. Turchi, Concentrating solar power: best practices handbook for the collection and use of solar resource data (CSP), Technical Report, National Renewable Energy Laboratory (NREL), Golden, CO., 2010.

T. Huld, T. Cebecauer, M. Šúri, E. D. Dunlop, Progress in Photovoltaics: Research and Applications 18 (2010) 183–194.

C. Good, H. Persson, Ø. Kleven, M. Norton, T. Boström, in: Presentert på" 26th European Photovoltaic Solar Energy Conference and Exhibition.

B. W. Kariuki, T. Sato, Renewable Energy 116 (2018) 88–96.

R. Meyer, S. Lohmann, C. Schillings, C. Hoyer, Solar energy resource management for electricity generation from local level to global scale (2006) 55–68.

E. Skoplaki, J. A. Palyvos, Solar energy 83 (2009) 614–624.

E. D. Dunlop, L. Wald, M. Suri, Solar Energy Resource Management for Electricity Generation from Local Level to Global Scale., Nova Science Publishers Inc., 2006.

R. Pinker, I. Laszlo, Journal of Applied Meteorology 31 (1992) 194–211.

F. Besharat, A. A. Dehghan, A. R. Faghih, Renewable and Sustainable Energy Reviews 21 (2013) 798–821.

J. Polo, S. Wilbert, J. A. Ruiz-Arias, R. Meyer, C. Gueymard, M. Suri, L. Martín, T. Mieslinger, P. Blanc, I. Grant, et al., Solar Energy 132 (2016) 25–37.

M. Iqbal, An introduction to solar radiation, Elsevier, 2012.

A. Riihelä, T. Carlund, J. Trentmann, R. Müller, A. V. Lindfors, Remote Sensing 7 (2015) 6663–6682.

R. Urraca, A. M. Gracia-Amillo, E. Koubli, T. Huld, J. Trentmann, A. Riihelä, A. V. Lindfors, D. Palmer, R. Gottschalg, F. Antonanzas-Torres, Remote sensing of environment 199 (2017) 171–186.

C. N. Long, E. G. Dutton (2010).

J. Davies, D. McKay, Solar Energy 43 (1989) 153–168.

T. Carlund, Upgrade of SMHI's meteorological radiation network 2006-2007: Effects on direct and global solar radiation, SMHI, 2011.

S. K. Jörg Trentmann, C. Team (2016).

A. R. T. M. Kati Anttila, Emmihenna Jääskelinen, C. Team (2016).

K.-G. Karlsson, K. Anttila, J. Trentmann, M. Stengel, J. F. Meirink, A. Devasthale, T. Hanschmann, S. Kothe, E. Jääskeläinen, J. Sedlar, et al., Atmospheric Chemistry and Physics 17 (2017) 5809–5828.

D. P. Dee, S. Uppala, A. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Balmaseda, G. Balsamo, P. Bauer, et al., Quarterly Journal of the royal meteorological society 137 (2011) 553–597.

K. Wang, J. Augustine, R. E. Dickinson, Journal of Geophysical Research: Atmospheres 117 (2012).

A dictionary of physics, 2015.

J. M. Last, J. H. Abramson, G. D. Freidman, A dictionary of epidemiology, volume 4, Oxford University Press New York, 2001.

K. Karlsson, A. Riihelä, R. Müller, J. Meirink, J. Sedlar, M. Stengel, M. Lockhoff, J. Trentmann, F. Kaspar, R. Hollmann, et al., Satell. Appl. Facility Clim. Monit. http://dx. doi. org/10.5676/EUM_SAF_CM/CLARA_AVHRR 1 (2012).

R. M. Jörg Trentmann, C. Team (2012).

S. K. Jörg Trentmann, C. Team (2016).

W. J. Wiscombe, S. G. Warren, Journal of the Atmospheric Sciences 37 (1980) 2712–2733.