Niels Jernes Vej 10 | DK-9220 Aalborg Ø | Denmark
Phone: +45 6916 4850 | e-mail: emd@emd.dk

Memo: SARAH-3 - Surface Solar Irradiation Data – Performance and Accuracy for Your Solar-PV Calculation

Authors: Morten Lybech Thøgersen (mlt@emd.dk), Marta Irena Murkowska (mim@emd.dk), Lasse Svenningsen (ls@emd.dk) & Thorkild Guldager Sørensen (tgs@emd.dk), 2024-02-20

Introducing the SARAH-3 Dataset

SARAH-3 (Surface Solar Radiation Data Set – Heliosat) - is a climate data record – recently released from EUMETSAT CM SAF [1]. SARAH-3 the third edition in a series of datasets derived from the first and second generation of MeteoSat satellites. It is released as an improvement and update to the SARAH-2 dataset [2] which is no longer updated forward in time. Spatial coverage for the SARAH datasets is up/down to 65 degrees in both longitude and latitude directions. Spatial resolution is 0.05 degrees (approximately 5.5 km) with a temporal resolution of 0.5 hour. In windPRO and energyPRO we deliver the two parameters relevant for modelling of Solar-PV – the surface incoming shortwave radiation (SIS) and surface incoming direct radiation (SID). EMD databases holds data from 1999 until present. For Solar-PV applications, the main benefits from SARAH-3 are: 1) an improved algorithm by better classification of snow-covered surfaces (previously, this could be misclassified as clouds) and, 2) using ERA5 as auxiliary data instead of ERA-Interim when using water-vapour and ozone within the modelling algorithm [3].

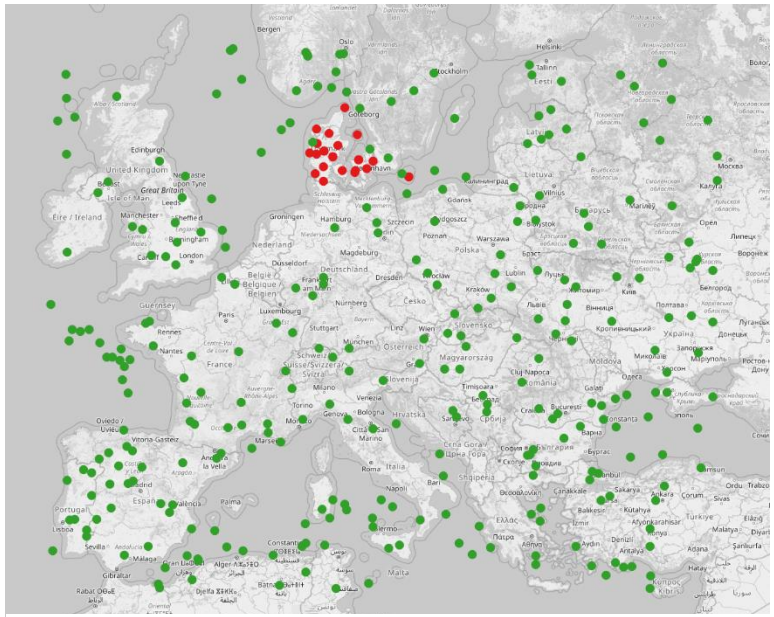

*Figure 1: Map of the evaluation locations: From DMI [11] in **red color** and by uniform random sampling in **green color**.*

Evaluation Approach

This evaluation is made in order to provide an overview of the performance and accuracy of the SARAH-3 data: It is done by comparing SARAH-3 data against in-situ data and other modelling references, i.e.: 1) ground-truth data from local meteorological stations in Denmark, 2) comparing against the global-solar-atlas [4] and, 3) against other climate datasets, including the SARAH-2 data [2]. The following approach is used, with main focus on the Global Horizontal Irradiance data (GHI). Note: GHI and SIS are compared directly here, as the SARAH-3 SIS relates to a horizontal plane.

- To compare SARAH-3 against ground truth data: Meteorological observations from 18 stations operated by the Danish Meteorological Institute (DMI) [11], see Figure 1. All stations used have at least 4 full years (and up to a maximum of 14 years) of data – and holds a 10-minute sampling rate.
- To compare SARAH-3 against other climate reference time-series datasets (200+ random locations within joint coverage area of the calculation domains):
  SARAH-2 [2], ERA5-gaussian-grid [5], ERA5 (rectangular grid) [6], CERRA Mesoscale Data [7], EMD-ConWx Mesoscale Data [8], EMD-WRF Europe + Mesoscale Data [9]
- To compare SARAH-3 against statistical reference data (atlas-data): the Global Solar Atlas [4]
- Different analysis metrics and time-scales:
  - Error (bias) as well as correlation on yearly, monthly, daily and 10-minute time-stamps
  - Two different performance metrics: Global Horizontal Irradiance (GHI) parameter and a generic solar PV-plant production (PVYIELD-YR) with optimum tilt angle for location of interest and a calculation model similar to that used in the windPRO SOLAR-PV module [10]
  - Performance metrics for additional climate datasets in order to do a direct quality evaluation and to build trust in the SARAH-3 dataset

The evaluation is made with the simple goal in mind to help building trust in the SARAH-3 dataset – and to understand strengths and weakness' when comparing to other sources.

Niels Jernes Vej 10 | DK-9220 Aalborg Ø | Denmark
Phone: +45 6916 4850 | e-mail: emd@emd.dk

Results

The Table 1 below summarize the evaluation results for the ground-truth data (18 DMI stations in Denmark). The table shows error (bias in %) of SARAH-3 and the other climate datasets, and the correlation of the monthly, daily and instantaneous (10-minute) samples. As seen below, the SARAH-3 and SARAH-2 datasets are almost identical in performance. For these Danish stations, the reanalysis dataset (ERA5) has a bias around 4%, but a variation similar to that of the SARAH datasets (around 3%). The mesoscale datasets – CERRA, EMD-WRF-Europe+ and EMD-ConWx has a significant larger bias, likely as the combination of high spatial resolution and cloud parameterization/prediction schemes are not able to correctly capture the cloud formation process within the evaluation domain.

	Error (BIAS) – Error in %						Pearson Correlation					
	GHI-YR		GHI-MONTH		PVYIELD-YR		GHI-MONTH		GHI-DAY		GHI-10MIN	
Dataset	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
SARAH-3	-1.1	3.0	-1.0	3.0	-2.0	3.4	0.998	0.001	0.988	0.003	0.948	0.008
SARAH-2	-1.2	2.9	-1.2	2.9	-2.4	3.3	0.998	0.001	0.988	0.003	0.947	0.008
ERA5 (gaussian grid)	3.8	2.7	3.8	2.7	0.2	3.2	0.997	0.001	0.976	0.004	0.915	0.007
ERA5 (rectangular grid)	4.0	3.0	4.0	3.0	3.4	3.6	0.997	0.001	0.976	0.004	0.921	0.007
CERRA	-4.0	1.7	-4.1	1.8	-6.4	3.7	0.979	0.023	0.959	0.021	0.915	0.009
EMD-WRF Europe+	23.4	2.8	23.4	2.8	29.8	3.7	0.995	0.002	0.962	0.008	0.900	0.010
EMD-ConWx	33.2	3.6	32.6	3.2	11.5	16.3	0.889	0.077	0.861	0.071	0.887	0.008

Table 1: Evaluation of SARAH-3 and other datasets against ground-truth measurements (DMI reference stations, 18 locations).

	GHI-YR: Error in %	
Dataset	Mean	Std.Dev.
SARAH-3	2.1	3.5
SARAH-2	1.7	3.5
ERA5 (gaussian grid)	2.9	4.8
ERA5 (rectangular grid)	2.9	4.6
CERRA	0.7	6.2
EMD-WRF Europe+	26.5	8.0
EMD-ConWx	28.5	7.5

Table 2: Evaluation of SARAH-3 and other datasets against GSA data (217 locations).

The Table 2 provides the comparison of the climate-datasets against the reference-dataset of GSA (Global Solar Atlas, [4]). As seen from the table, the GSA data holds a conservative estimate of the solar-irradiance of the 217 locations investigated, with GSA consistently underestimating the Solar-PV potential (here expressed through the GHI). The mesoscale datasets are, again, significantly over-predicting the resource – and in addition have also a higher uncertainty (expressed through the standard deviation of the evaluation samples). As seen from the Table 2 the performance of SARAH datasets is almost identical.

Figure 2 shows monthly data-records from 3 reference stations – Skagen, Holstebro and Nexø. The black curve shows observational data mapped onto the primary axis. The coloured curves show the difference between the evaluation dataset and the reference – mapped onto the secondary axis. All datasets have been normalized with their yearly GHI. As seen from these samples, the mesoscale-derived datasets exhibit a significant seasonal bias – with a clear trend and pattern of over-predicting and under-predicting at specific months – so when using any time-series dataset of solar irradiation – it is definitely worth looking for such seasonal biases and trends.

Conclusions

Based on time-series solar of irradiance records from the Danish Meteorological Office (DMI) and statistical data from the Global Solar Atlas (GSA), it is our conclusion that the SARAH datasets exhibit a consistent and stable performance. The SARAH datasets are our recommended first choice for PV-solar evaluations with windPRO and energyPRO, given that the PV-project is located within the spatial domain covered by these datasets. The ERA-5 derived datasets is our recommended second choice. If using solar irradiance data derived from any of the evaluated mesoscale models, these should be bias-corrected using local ground-truth data and some effort should be made to analyse and correct for seasonal bias. The performance of SARAH-3 is in good alignment with the SARAH-2 dataset, but the potential improvements within the SARAH-3 model chain - are not apparent from this study. These are presumably more visible for sites with more snow cover.

Niels Jernes Vej 10 | DK-9220 Aalborg Ø | Denmark
Phone: +45 6916 4850 | e-mail: emd@emd.dk

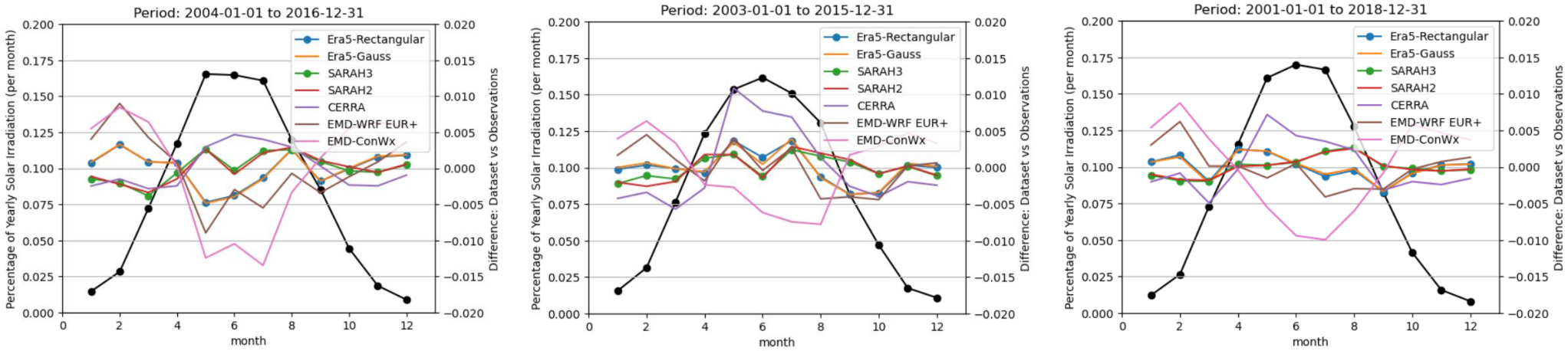

Figure 2: Evaluation of Solar Irradiation Datasets: Monthly Percentage of Yearly Solar Irradiance for Observational Data (black curve, primary axis) – and Model Deviation (coloured curves, secondary axis). Left: Station 6197 (Nexø, Bornholm). Centre: Station 6056: (Holstebro, Jutland). Right: Station 6041 (Skagen, Vendsyssel).

Project Acknowledgements

This study and the data used have been distributed through the EMD-API. The EMD-API and associated data were prepared as part of three R&D projects: The INNOWIND, RECAST and GASP projects: http://www.recastproject.dk, http://www.emd.dk/gasp and https://help.emd.dk/mediawiki/index.php/Category:InnoWind. RECAST and INNOWIND was co-funded by the Innovation Fund Denmark and GASP co-funded by EUDP, Denmark.

Data Acknowledgments and References

[1] EUMETSAT CM SAF – *SARAH-3*: https://doi.org/10.5676/EUM_SAF_CM/SARAH/V003
[2] EMD International A/S – *SARAH-2.1*: https://help.emd.dk/mediawiki/index.php/Heliosat_(SARAH)
[3] Pfeifroth & Trentmann: *Algorithm Theoretical Baseline Document - Meteosat Solar Surface Radiation and effective Cloud Albedo Climate Data Records – Heliosat – SARAH-3*, DOI: 10.5676/EUM_SAF_CM/SARAH/V003
[4] World Bank Group and ESMAP - *Global Solar Atlas* - https://globalsolaratlas.info
[5] *ERA5 Gaussian Grid*: ERA5 is being developed through the Copernicus Climate Change Service (C3S), see https://help.emd.dk/mediawiki/index.php/ERA5_Gaussian_Grid
[6] *ERA5 Rectangular Grid*: ERA5 is being developed through the Copernicus Climate Change Service (C3S), https://help.emd.dk/mediawiki/index.php/ERA5(T)_Rectangular_Grid
[7] *Copernicus Regional Reanalysis for Europe (CERRA):* CERRA is being developed through the Copernicus Climate Change Service (C3S), https://help.emd.dk/mediawiki/index.php/CERRA
[8] EMD International A/S – *EMD-ConWx Meso Data Europe*, https://help.emd.dk/mediawiki/index.php/EMD-ConWx_Meso_Data_Europe
[9] EMD International A/S – *EMD-WRF Europe+,* https://help.emd.dk/mediawiki/index.php/EMD-WRF_Europe%2B_(ERA5)
[10] *windPRO SOLAR-PV module*, https://www.emd-international.com/windpro/windpro-modules/energy-modules/solar-pv/
[11*] Danish Meteorological Institute – Open Data*: https://opendatadocs.dmi.govcloud.dk/DMIOpenData