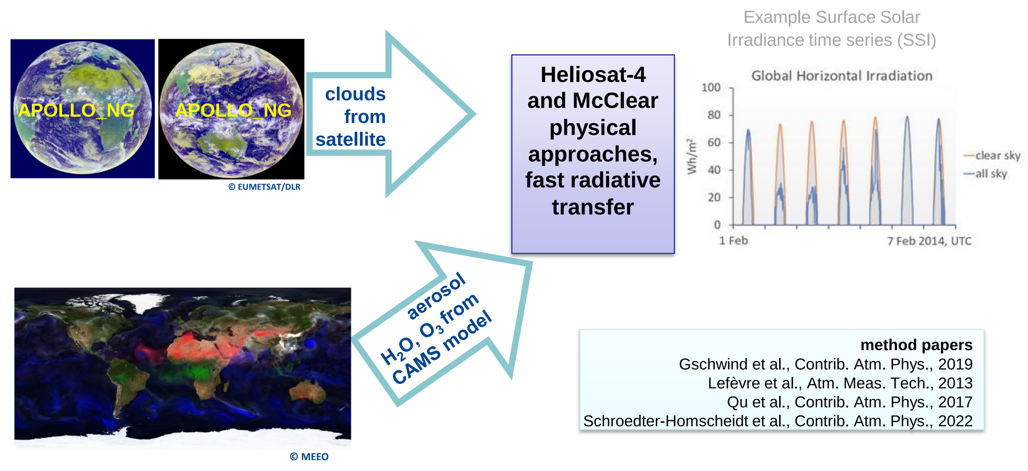
CAMS radiation service for solar energy: Exploring the error space with data-driven and spatially resolved methods and service evolution

Faiza Azam[1], Jorge Lezaca[1], Marion Schroedter-Homscheidt[1], Antti Arola[2], Antti Lipponen[2], Laurent Saboret[3], Mireille Lefevre[4], Yves-Marie Saint-Drenan[4]
faiza.azam@dlr.de
[1] German Aerospace Center (DLR), Institute of Networked Energy Systems, Oldenburg, Germany, [2] Finnish Meteorological Institute (FMI), Kuopio, Finland, [3] Vaisala, [4] MINES Paris, PSL University, Sophia-Antipolis, France

Copernicus Atmospheric Monitoring service (CAMS) solar radiation services (CRS)



Primary product: SSI Time Series

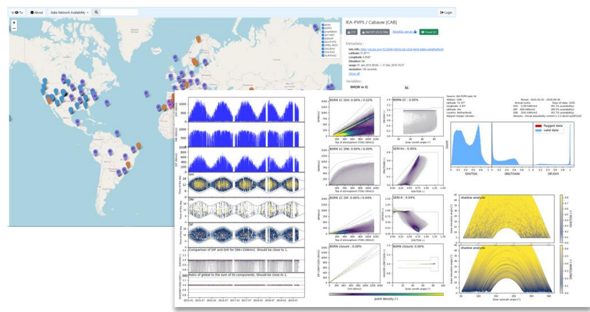
- global, diffuse, direct and direct normal irradiance
- 2004 onwards in MSG & 2016 onwards for HIMAWARI filed of view (FOV)
- 1 min, 15 min, 1 hour, 1 day, 1 month resolution
- interactive access on CAMS ADS and SODA portal
- OGC script access possible or via open source library PVlib
- access to all input data in expert mode (1 min)

Pre-calculated SSI gridded data

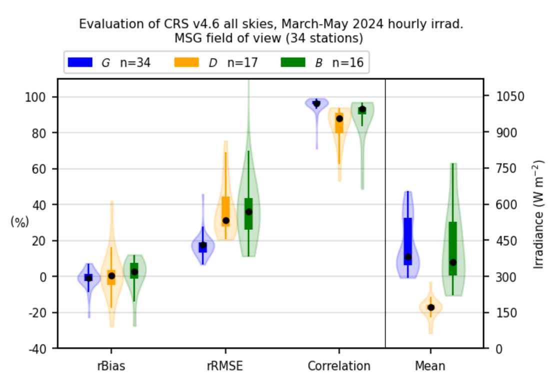
- global, diffuse, direct and direct normal irradiance
- 15 min temporal resolution selected
- 2005-2023 in MSG & 2016-2023 in HIMAWARI FOV
- interactive access on CAMS ADS
- 0.1° spatial grid selected

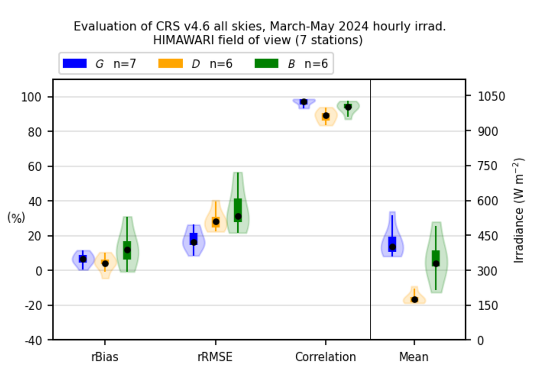
CRS routine evaluations and Quality control (EQC)

- Provision of ground based SSI measurements collected on global scale to support CRS evaluations
- Consolidated metadata; harmonized netCDF structure following Climate and Forecast (CF) conventions and the FAIR (Findable, Accessible, Interoperable, Reproducible) principle
- Automatized visual Quality Control (QC)
- Python library libinsitu for data processing
- Data access through THREDDs data server https://viewer.webservice-energy.org/in-situ/



- CRS irradiance vs ground observations evaluations: Use of high quality radiation networks
- Application of CAMS expert QC on ground observations
- Standardize error metrics (MBE, RMSE, …) on cloudless / cloudy skies: station-wise and aggregated comparisons
- Quarterly updates (3 months evaluation with 6 months delay)
- Publicly available https://atmosphere.copernicus.eu/supplementary-services





CAMEO: the detailed assessment of the error space of CRS

Questions to be addressed in CAMEO

1. Are the aggregated error metrics provided to the users sufficient for expert applications?
2. Can we assess the main error sources of the CRS irradiance estimates?
3. Could we give more pixel-wise (non aggregated) uncertainty indicators for irradiance estimations to the users?

SHAP analysis as a rational for the selection of uncertainty model input parameters

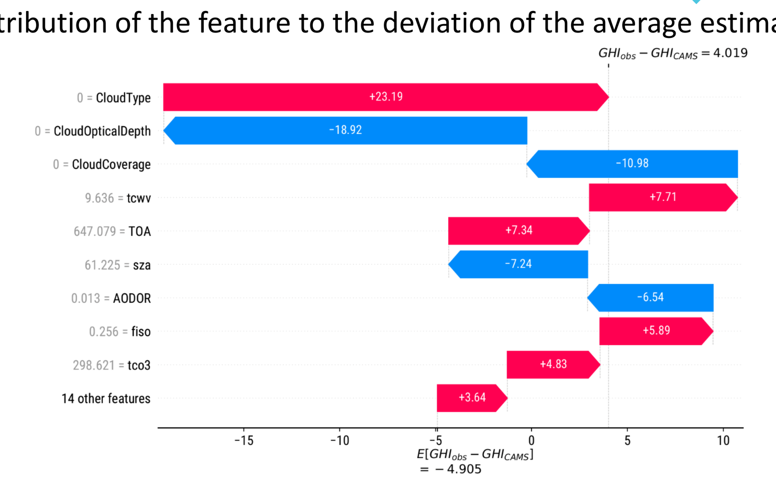
SHapley Additive exPlanations (SHAP) analysis:

Method (game theory) to understand the contributions of single model parameters/features on the total error of the CRS irradiance estimates

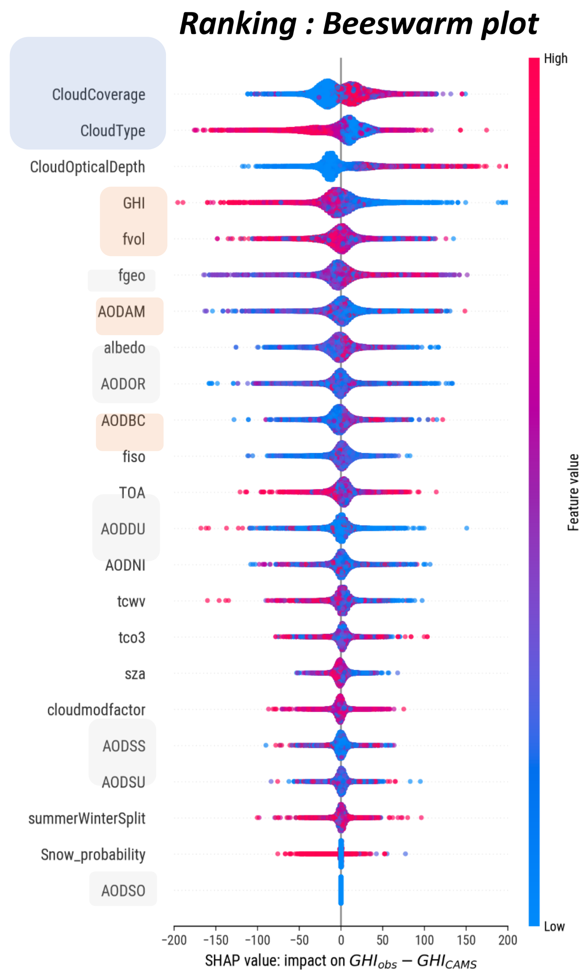
$$f(X) = GHI_{cams}$$
$$e(X) = GHI_{obs} - f(X)$$

$X \rightarrow$ $XGboost$ (overfitted) $\rightarrow e(X)$

SHAP value:
Contribution of the feature to the deviation of the average estimate

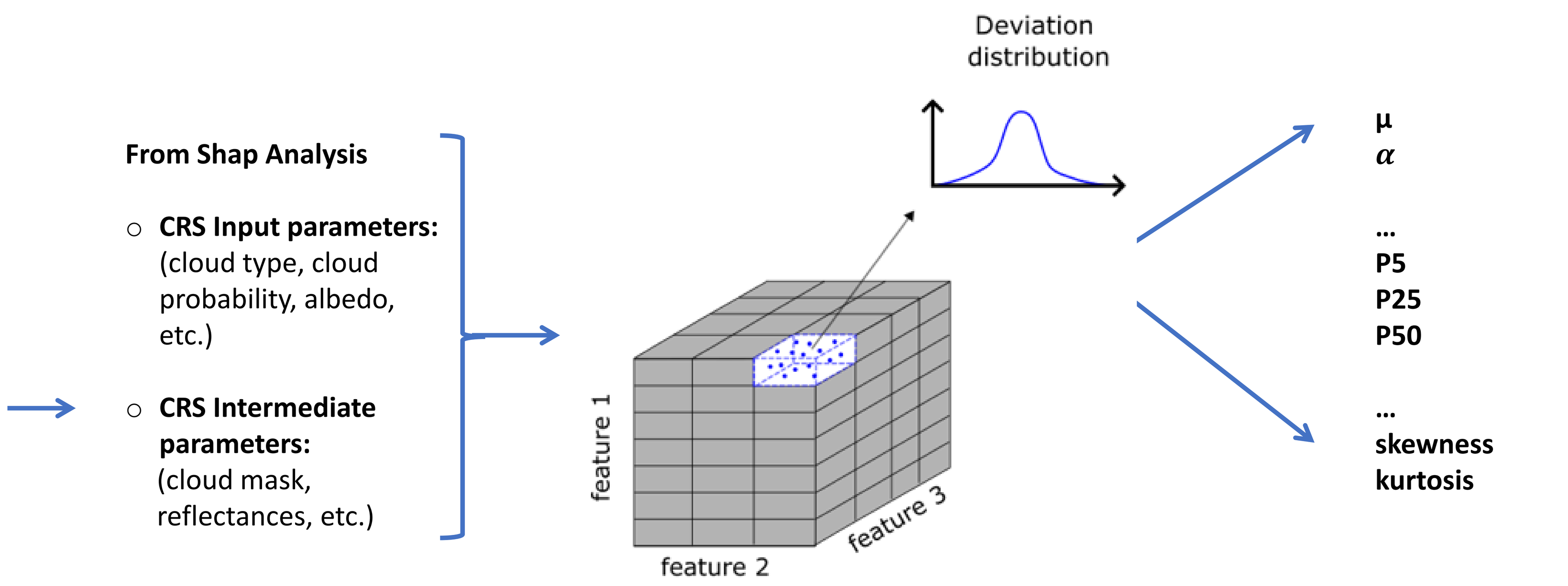


Ranking : Beeswarm plot



Pixel-wise error model development

Principle of the look up table (LUT) and neural network (NN) approaches:



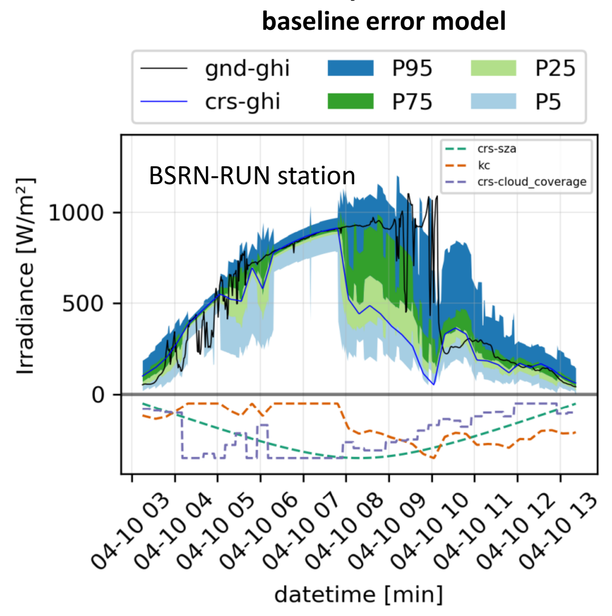
Methodology :

- **Training** : find a probability distribution for every bin in our multidimensional feature space (56 internal cams parameters)
- **Inference** : for the specific location, estimate the uncertainty from the probability distribution that corresponds to its specific feature combination (time dependent)
- **Approaches** : LUT approach / NN approach optimizing moments of a parametric distribution (MBE (μ), STDE (α), kurtosis, skewness)

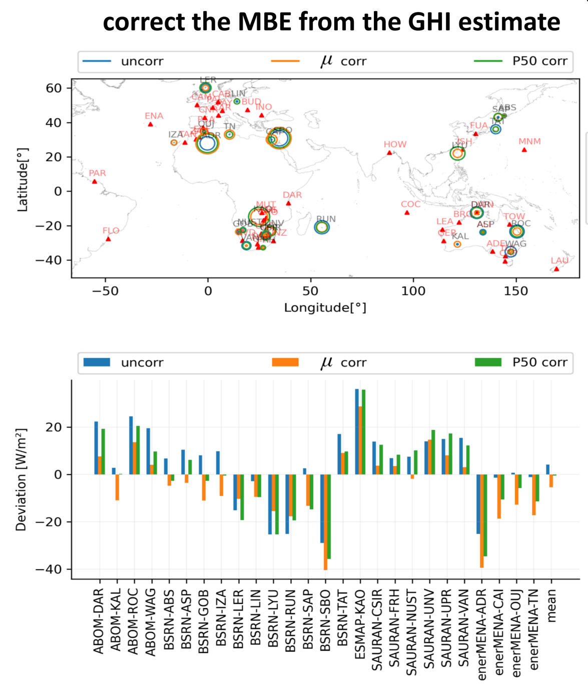
Localized error model 1: uncertainty inference based on LUT

- LUTs created using three parameters:
  - cloud probability: {from 0% to 100% every 5%} -> (20 ranges)
  - Kc = GHI/GHIc: {from 0 to 1.4 every 0.1} -> (14 ranges)
  - CRS-solar zenith angle (crs-sza): {from 0° to 80° every 2°} -> (40 ranges)
- Percentiles used in order to achieve a probabilistic-based inference of the CRS errors {from P0 to P100 every 1 percentile} + {MBE, MAE, STDE, RMSE} -> (105 metrics)
- Multiple years are used for training

GHI uncertainty inference from the baseline error model



Assessment of the model to deterministically correct the MBE from the GHI estimate



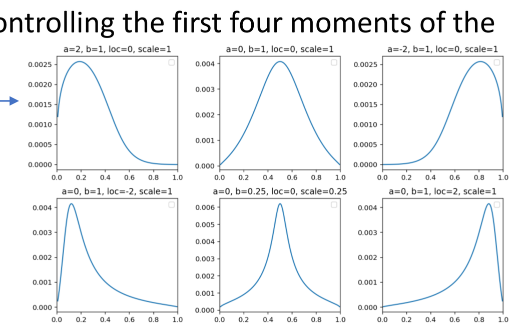
Red triangles: training sites (40)
Colored circles: inference sites (26)
Size of the circle: deviation

Blue: uncorrected estimate
Orange: μ-corrected estimate (using the MBE of the bin)
Green: P50-corrected estimate (using the median of the bin)

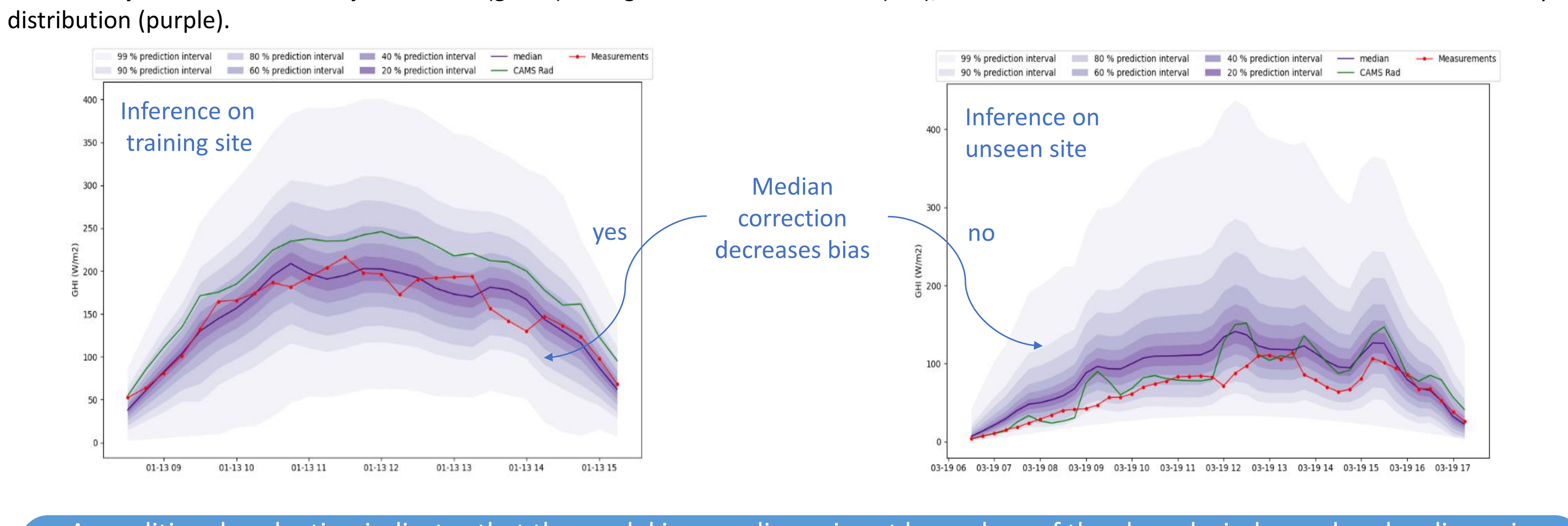
- Ground observation values well contained inside the 90% confidence interval (P5 - P95)
- The spread of the confidence intervals correlates well with the local variability situation, i.e., narrow intervals on clear and overcast situations and wider intervals on variable situations.
- Potential as a bias correction methodology

Localized error model # 2: uncertainty inference based on Deep learning

- Uncertainty modelled by a Johnson SU parametric distribution with four parameters controlling the first four moments of the distribution (mean, variance, skewness, and kurtosis)
- Six key variables for the calculation of SSI in CRS are selected as predictors:
  - clear sky index calculated with CRS estimates (Kc)
  - cloud type
  - cloud cover
  - cloud optical depth
  - solar zenith angle (SZA)
  - solar azimuth angle (SAA)
- 4 stations and 2 years are used for training
- The four parameters are predictedas a function of six predictors using a neural network (implementation in teserflow probablility



Uncertainty inference for 2 days: CRS-GHI (green) and ground measurements (red), CRS-GHI-corrected with the median of the uncertainty distribution (purple).



- A conditional evaluation indicates that the model is over-dispersive at low values of the clear sky index and under-dispersive at high values of the clear sky index for cases of unseen sites. This lack of spatial generalisation can be attributed to an overtraining issue.
- The model is well calibrated when testing and training data (not shown here) are same but a significant decrease in accuracy is observed when the model is applied to stations that were not used for training.

References

- **Qu et al.,** Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method, Contrib. Atm. Sci., 2017
- **Schroedter-Homscheidt et al.,** Surface solar irradiation retrieval from MSG/SEVIRI based on APOLLO Next Generation and HELIOSAT-4 methods,Contrib. Atm. Sci.,/Meteorol. Z. Vol. 31 No. 6 (2022), p. 455 – 476, DOI: 10.1127/metz/2022/1132
- **Lefèvre et al.,** McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions, AMT, 2013
- **Gschwind et al.,** Improving the McClear model estimating the downwelling solar radiation at ground level in cloud-free conditions – McClear-v3, Contrib. Atm. Sci./Meteorol. Z., 2019
- **User's Guide and validation reports** at https://atmosphere.copernicus.eu/supplementary-services

Acknowledgements

- The CRS routine evaluations are carried out under CAMS2_73_bis Project under CAMS program, which is part of the EU Copernicus Program and implemented by ECMWF on behalf of the EU.
- 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.
- We are thankful to ground station operators as the Baseline Surface Radiation Network (BSRN), Enermena network and Bureau of Meteorology (BOM) network for the ground measurements of SSI

ON BEHALF OF ECMWF FOR THE EUROPEAN COMMISSION