Characterizing The Convergence And Robustness Of The Kernel Density Mapping Method For Site-Adaptation Of Global Horizontal Irradiation In Western Europe

Loïc Yezeguelian, Christophe Vernay, Thomas Carrière, Philippe Blanc

CHARACTERIZING THE CONVERGENCE AND ROBUSTNESS OF THE KERNEL DENSITY MAPPING METHOD FOR SITE-ADAPTATION OF GLOBAL HORIZONTAL IRRADIATION IN WESTERN EUROPE

Loïc Yezeguelian[a,b], Christophe Vernay[a], Dr. Thomas Carrière[a] , Prof. Philippe Blanc[b]

[a]SOLAÏS, 55 allée Pierre Ziller, 06560 Sophia Antipolis, France
[b]Centre Observation, Impacts, Energy, ARMINES / MINES ParisTech, CS 10207, 06904 Sophia Antipolis, France

ABSTRACT: The estimation of the annual solar resource at a given location is a crucial part of the financing of new photovoltaic power plants. The standard practice is to use satellite-derived databases of ground irradiance estimations. Although these databases have a long-term history that properly captures the climatological variations of a given site, they are not calibrated with ground measurements and therefore may suffer from significant biases. In this paper, we focus on Western Europe and propose to use ground measurements of global horizontal irradiation (GHI) combined with a Kernel Density Mapping algorithm to calibrate several satellite databases (Solargis, HelioClim-3 and CAMS Radiation); we then show that their biases can be reduced from roughly ±3 % down to 0.5 %. More importantly, we perform a sensitivity analysis on the minimal amount of ground measurements required to perform a proper calibration and show that 12 months is the perfect duration to achieve optimal calibration performance. Additionally, we compare the calibration performance on several different satellite databases and show that a combination of several databases can lead to lower overall uncertainty.
Keywords: GHI calibration, site-adaptation, Kernel Density Mapping, Quantile Mapping, Solargis, HelioClim-3, CAMS Radiation, bias removal, convergence, campaign duration

Table I: Notations

CDF	:	Cumulative Distribution Function
CS	:	when indexed, refers to clear-sky
GHI	:	Global Horizontal Irradiation
I	:	interpolation interval of the CDF
KDE	:	Kernel Density Estimator
KDM	:	Kernel Density Mapping
KSI	:	Kolmogorov-Smirnov Integral
LT	:	when indexed, refers to the long-term
$meas$	:	when indexed, refers to the ground measurements
MBE	:	Mean Bias Error
PDF	:	Probability Density Function
$RMSE$	:	Root Mean Square Error
sat	:	when indexed, refers to the satellite database
ST	:	when indexed, refers to the short-term measurement campaign
$\tilde{x}$	:	uncalibrated satellite point of data
x	:	calibrated satellite point of data

1 INTRODUCTION

The profitability of photovoltaic (PV) projects as well as the financial structuring strongly rely on the global performance of the PV plant which is commonly estimated with percentiles 50 % (also known as P50 in the industry) and 10 % (P90) of the yearly energy yields during its future multi-decadal lifetime. The uncertainty over such decisive assessments mainly comes from the imperfect knowledge of the local long-term irradiation (typically the Global Horizontal Irradiation – GHI). [1] The best practice for deriving typical (P50) or conservative (P90) meteorological year at a given location now consists in using satellite-based databases because of their long-term history (typically 15+ years). However, the uncertainty in the satellite GHI estimation still jeopardizes the development of large-scale PV projects for which feed-in tariffs get closer to network parity.

As for wind industry in the last decades, site-adaptation techniques are now emerging in the PV industry for correcting the GHI assessment and more precisely for removing the Mean Bias Error (MBE) of the satellite-based estimation, which is experienced whatever the modeling time step. Such a local calibration process requires local measurements of the GHI over a short-term period of several months, which in turn requires the setup of a measurement campaign at the early phase of the project. The collected data are then coupled to calibration algorithms that compare the satellite data with the measurements to model the errors. The resulting correction, in the form of a filter function, is propagated to the long-term satellite derived GHI time series.

Numerous types of calibration filters have been explored in the scientific literature. [2] Almost all these methods are mainly grouped around two mathematical procedures, namely multi-parameter regression and Quantile Mapping. The former consists in modeling the measured time-series by a combination of several relevant parameters from the satellite dataset or solar-position retrieval algorithms. Such regression can be linear or non-linear and makes use of every available data representing the instantaneous atmospheric and solar configuration of the measurement point, such as solar elevation, clearness index, etc. However, no consensus has been reached over the best combination of variables. Regression also proves to be geographically inconsistent and to strongly rely on local fluctuations of environmental parameters with post-calibration biases drastically fluctuating around zero. As for Quantile Mapping, it is rather based on a blind matching of probabilistic distributions of both satellite and measured time-series. The Cumulative Distribution Function (CDF) of the satellite irradiance data is transformed into that of the measurements covering the same time period. The mathematical transformation, taking the form of the filter, is then applied to the long-term. The benefit of this approach is that it does not depend on the geography or meteorology of the studied site. It is mathematically constructed to remove any distribution errors in the satellite data and should cancel the MBE between the long-term and the short-term without making

use of any spatial or temporal parameter other than the one it intends to correct.

While comparative studies confronting different algorithms are already available, this study intends to stress and assess the behavior of Quantile Mapping-based processes from a practical perspective. It is notably in line with the related activity of the Task 16 of the International Energy Agency program PVPS and the corresponding papers from Polo et al. [3] The main objective is to characterize the performances of the Kernel Density Mapping (KDM) method, derived from Quantile Mapping and commonly used for site-adaptation, with an innovative focus on both convergence (campaign duration) and robustness (choice of the database and spatial homogeneity). Developing an operational calibration solution raises indeed practical issues. Notably, addressing the problem of measurement campaign duration as well as selecting the best suited satellite database is mandatory prior to any site-adaptation procedure. The variability in performance caused by these choices is addressed in this paper.

To identify the performance indicators on the long-term calibration and draw reliable conclusions, experiments were conducted on sites where historical irradiance measurement data are available and with good quality. Their goal is to evidence convergence visually and numerically through a statistical approach for a number of test conditions.

2 SITE-ADAPTATION USING KDM METHODOLOGY

Providing an unbiased estimation of the GHI is the major challenge for the “solar resource” community as the size of the PV projects drastically increases (100+, even 1000 MWp) while the corresponding business plans are getting tighter. Addressing the problematic of site-adaptation is therefore mandatory to reduce uncertainty of the planned energy yield and therefore its financial risk.
The depicted calibration method consists in correcting long-term satellite-database GHI time-series using short-term horizontal pyranometer measurements. For each of the reference sites, a large number of “virtual” measurement campaigns of variable duration (ranging from one to 24 months) are simulated. Each measurement campaign leads to a specific performance (*i.e.*, long-term GHI measurement vs. calibrated GHI time-series) which is then used to study the sensitivity of the calibration process.

A calibration method highly adapted to the elimination of the MBE for a given site is KDM. Quantile Mapping removes systematic errors between two datasets by matching the CDF of the modeled time-series on its corresponding truth data counterpart. For every estimated point of data $\tilde{x}$ (in this case the GHI estimation from the satellite model to be calibrated), the associated calibrated value x is the one that should have the same image when applying the measured dataset CDF function:

$$\forall \tilde{x} \in I_{ST_{meas}} \cap I_{ST_{sat}},\ x = {CDF_{ST_{meas}}}^{-1} \circ CDF_{ST_{sat}}(\tilde{x})$$

Where $I_{ST_{meas}}$ and $I_{ST_{sat}}$ are the respective intervals modeling the observed ranges of $CDF_{ST_{meas}}$ and $CDF_{ST_{sat}}$. Index ST refers to the short term, i.e., the measurement period.

The construction of the CDF is proving to be a significant issue. As the support interval of possible GHI values is indeed continuous, that the associated CDF also is. Time-series being discrete, the easiest way to interpolate a CDF using a finite ensemble of draws is to empirically build a step function. The dataset is sorted, and the function is incremented with a constant step at each new value of GHI. Such piecewise constant function can raise inversion issues when applying ${CDF_{ST_{meas}}}^{-1}$ to subintervals where GHI points of data are scarce. In particular, the extreme high values of irradiance may thus be subject to "vertical bar" phenomena due to the non-strict monotonicity of ${CDF_{ST_{meas}}}^{-1}$. KDM methodology aims at countering these purely mathematical issues. It uses a Kernel Density Estimator (KDE) to propagate the distribution of the finite dataset of discrete GHI values in a continuous space (over the ensemble of real numbers). Instead of being considered as a probabilistic Dirac, the presence density, i.e., the Probability Density Function (PDF) of each new irradiance occurrence is expanded over $\mathbb{R}$. In the case under study, the shape of the selected kernel is a centered reduced Gaussian normalized by the number of samples N.

$$\forall s \in \mathbb{R},\quad \begin{cases} PDF(s) = \dfrac{1}{Nh}\displaystyle\sum_{k=1}^{N} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{(s-x_k)^2}{2h^2}} \\ CDF(s) = \displaystyle\int_0^s PDF(t)dt \end{cases}$$

Where h is the bandwidth of the gaussian kernel. The higher h, the smoother the PDF. Yet the bandwidth needs to be low enough to maintain sufficient curve details. In these works, $h = 3$ proved to be the best balance between smoothness and a good level of details.

It is then possible to fall back into a finite interval by truncating the CDF while removing its intercept and dividing it by the difference between the maximum and minimum value. The most relevant interval can be $I' = [0, \max(GHI_{LT_{sat}})]$, where LT refers to the long term satellite period, but it still leads to vertical bar phenomena if the maxima of the long-term satellite time-series and the measured maxima are too far apart ($CDF_{LT_{sat}}$ or its real counterpart that would have originated from the long-term pyranometer measurement campaign will reach the horizontal asymptote 1 too quickly before the other). $CDF_{ST_{meas}}$ and $CDF_{ST_{sat}}$ are therefore interpolated on interval I:

$$I = [0, \min\{argmin(|CDF_{ST_{meas}} - 0.99|), argmin(|CDF_{ST_{sat}} - 0.99|)\}].$$

It must be noted that the value 0.99 has been proposed as it generates the expected effect.

The KDM methodology can be visualized as the function modifying every date point of a time-series so that the curve of its KDE estimated CDF matches the wanted one:

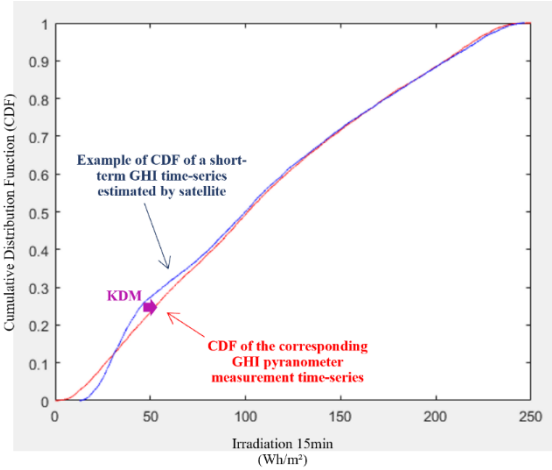


Figure 1: Visualization of the Kernel Density Mapping of a GHI time-series

It should be noted that KDM can be applied whatever the duration of the short-term or long-term time-series and whatever the database used.

3 EXPERIMENTAL PROCESS

3.1 Ground data

These works make use of seven ground-based stations located in Western Europe and belonging to the BSRN network (Baseline Surface Radiation Network) which added value mainly consists in the high-quality of the available data. [4] The one-minute GHI measurements from the ground stations over a ~12-year period (2007+) are considered as the long-term reference to be reached.

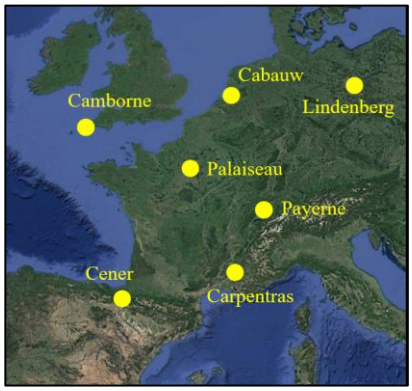


Figure 2: Location & names of the seven BSRN ground-based stations in Western Europe

Table II: Useful details for every station used. $GHI_{LT\ meas}$ is the average perceived 15-min GHI observed over the entire time series during periods when the sun is above the horizon

Name	Country	Lat(°)	Lon(°)	Elev(m)	Long-term period used	Climate	$GHI_{LT\ meas}$ (Wh/m²)
Cabauw	Netherlands	51.9711	4.9267	0.0	2007-2018	Oceanic	72.1
Camborne	UK	50.2167	-5.3167	88.0	2007-2017	Oceanic	77.3
Carpentras	France	44.0830	5.0590	100.0	2007-2018	Mediterranean	109.3
Cener	Spain	42.8160	-1.6010	471.0	2009-2019	Oceanic	97.7
Lindenberg	Germany	52.2100	14.1220	125.0	2007-2017	Continental	76.1
Palaiseau	France	48.7130	2.2080	156.0	2007-2018	Continental	85.5
Payerne	Switzerland	46.8150	6.9440	491.0	2007-2013	Continental	84.7

The heterogeneity in the seven ground-based stations, in term of climate and latitude, allows gaining confidence in the results and conclusions when addressing projects in Western Europe. Though, pyranometers are operated independently from one site to another. To equalize the data and remove any inconsistent measurement point, a quality check was applied [5]. This filtering has the effect of discarding corrupted, unmeasured, or unrepresentative pieces of data.

3.2 Satellite databases

Three satellite-based databases were used as the long-term GHI estimation to be calibrated: Solargis [6], HelioClim-3 (version 5 – HC3v5) [7] and CAMS Radiation (version 3.2 – Copernicus Atmosphere Monitoring Service) [8]. The two first ones are commercial products commonly used by the PV actors while the third one is freely provided by the EU-funded Copernicus Service.

As mentioned later, the errors made post-calibration for a given site seem to be relatively variable depending on the year of measurement, even for an identical database. Thus, it appeared to be coherent to cross-calibrate time series from the same campaign but from different databases to compensate for some of the respective biases. For that purpose, an experimental Hybrid database was also built using a regression of Solargis and HC3v5. It aims to find the best combination of several databases to enhance correlation with the short-term measurement campaign when the calibration function is mapped.

3.3 Monte-Carlo based assessment

The BSRN stations provide enough historical data to allow for experimentation based on statistical draws. To characterize the robustness of the described methodology on plausible data, it has been proposed a Monte-Carlo based variability assessment of the short-term measurement campaigns. For each of the seven sites, one iteration of the process can be summed up as follows:

- A starting date is randomly chosen.
- The corresponding short-term campaign with a specified duration in months is extracted from the

measured dataset; quality check and shadow removal are applied.

- All CDF needed are computed using the described KDM methodology.
- The Quantile Mapping function is interpolated on I and applied to the satellite long-term.
- Statistical errors are retrieved by comparing with the measurement long-term. MBE, Root Mean Square Error (RMSE) and Kolmogorov-Smirnov Integral (KSI) were chosen to assess the performance of the calibration process with respect to the target which is the actual long-term BSRN data.

This algorithm was run for n starting dates. For each of them, campaign durations were selected between one and 24 months. The statistical analysis of errors provides information on the random distribution as well as the observed median value of indicators with respect to the number of months used.
$n = 100$ was found to be enough iterations. Going over this number was not considered relevant as additional data would have great chance to overlap already existing results, bringing no further details. The experiment was conducted on the three chosen satellite databases as well as the Hybrid and respective performances were compared. Visual convergence was demonstrated using boxplot (i.e., statistical box) representation. This representation allows to easily see the distribution of the results of the n experiments for each campaign duration in months.

4 RESULTS

When experimenting the Monte-Carlo process, boxplots perfectly illustrate the expected behavior of increased confidence when working with longer measurement campaigns. Several statistical errors were monitored, the most essential being MBE followed by KSI. Since the aim of the calibration is to estimate accurately the yearly sum of irradiation, it is essential than the calibrated data is unbiased. Measures of the accuracy of the calibrated time series compared to the raw measurements (such as RMSE or MAE) are less important. These indicators provide an exhaustive view of the representativeness of the calibrated time-series obtained by demonstrating the systematic deviation and the discrepancy in terms of CDF post-calibration. RMSE is also shown in order to control that the dispersion of random errors does not increase substantially while calibrating.

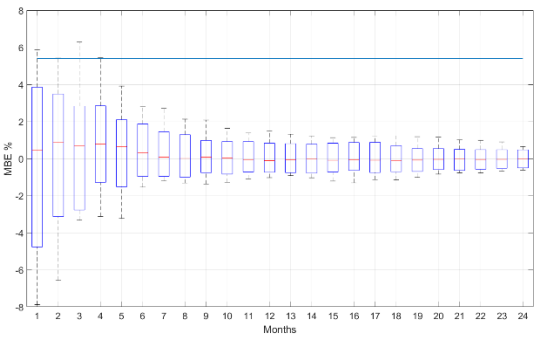


Figure 3: MBE convergence (in percent with respect to the long-term reference, Y-axis) after calibration of HC3v5 at Cener, Spain. Duration of the measurement campaign is expressed in months (X-axis). Horizontal blue line is the raw error of the satellite-based database (i.e. without calibration) while the boxplots depict the MBE spread after the calibration using campaigns with same duration (median in red, percentiles P10 and P90 – thus 80 % of the samples – as blue rectangle).

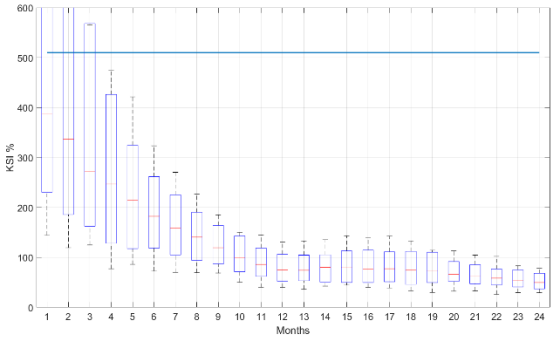


Figure 4: KSI convergence after calibration of HC3v5 at Cener, Spain. This KSI test is used to compare two statistical samples, i.e., to measure the distance between two CDF. It is possible to normalize this parameter to obtain a relative value of KSI; a value greater than 100 % indicates a significant difference between the two CDFs with a confidence level of 95 %.

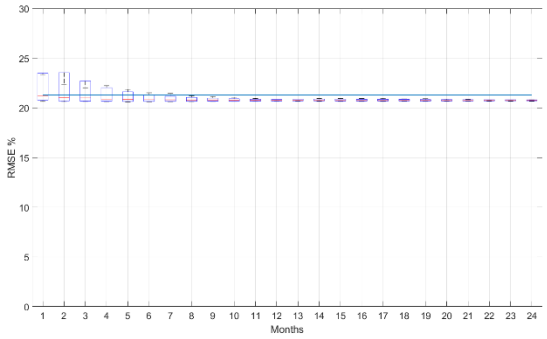


Figure 5: RMSE convergence (for 15-min time-step) after calibration of HC3v5 at Cener, Spain.

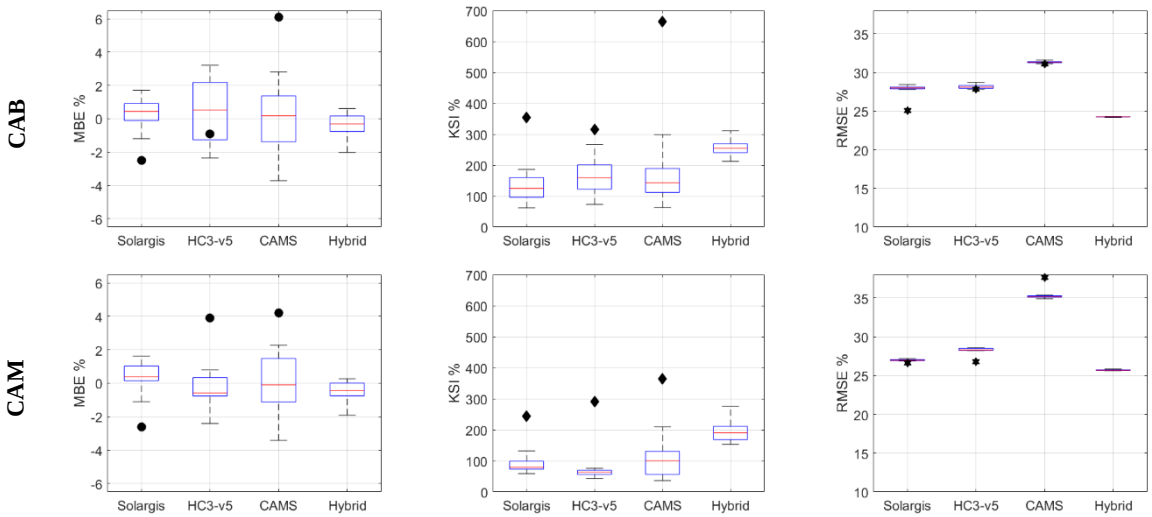
Results show that running a 12-month measurement campaign allows capturing the whole seasonal information for an efficient long-term calibration. Using more than twelve months does not bring much added value as depicted in Figures 2 and 3 (one specific site, one specific database, 100 campaigns) where MBE and KSI drastically decrease along with the campaign duration up to 12 months, then reach a plateau. A bounce-back effect can even be observed for campaigns lasting between one and two years, reflecting the fact that some seasonal effects are doubled while others are not. To reach a performance at least better than for 12 months, measuring for twice as long is a minimum, which is incompatible with the time constraints of most PV project developments.

Table 2 depicts for each of the seven locations the raw performances of the satellite-based databases along with the median of the MBE, KSI and RMSE (bold) resulting from 12-month measurement campaigns with random starting dates (100 draws). P25 and P75 are also shown

and were preferred to P10 and P90 because they are less impacted by exceptional results. Such a convergence can be observed whatever the site and the database, with different rate and convergence point. However, when focusing on the MBE, running the adapted KDM with a Hybrid database systematically improves the performances as depicted in Figure 6 where the performances for 12-month measurement campaigns are compared depending on the database.

Table III: Raw MBE, KSI and RMSE (no calibration, in bold) and median (P50) MBE, KSI and RMSE (one hundred measurement campaigns of twelve months) for the seven BSRN stations and for the three satellite-based databases (plus Hybrid). P25 and P75 are also shown to illustrate confidence.

BSRN	Database	MBE [%]				KSI [%]				RMSE [%]			
		Raw	Calib.			**Raw**	Calib.			**Raw**	Calib.		
			P25	P50	*P75*		*P25*	P50	*P75*		*P25*	P50	*P75*
CAB	Sol.	**-2.5**	*-0.1*	0.4	*0.9*	**355**	*96*	125	*160*	**25.1**	*27.9*	28.0	*28.1*
	HC3v5	**-0.9**	*-1.3*	0.5	*2.2*	**316**	*122*	160	*201*	**27.8**	*27.9*	28.1	*28.3*
	CAMS	**6.1**	*-1.4*	0.2	*1.4*	**665**	*112*	143	*189*	**31.1**	*31.2*	31.3	*31.4*
	Hybr.		*-0.8*	-0.3	*0.2*		*241*	255	*270*		*24.2*	24.3	*24.3*
CAM	Sol.	**-2.6**	*0.2*	0.4	*1.0*	**244**	*73*	80	*99*	**26.6**	*26.9*	27.0	*27.0*
	HC3v5	**3.9**	*-0.8*	-0.6	*0.3*	**291**	*56*	63	*69*	**26.8**	*28.2*	28.3	*28.5*
	CAMS	**4.2**	*-1.1*	-0.1	*1.5*	**365**	*56*	100	*131*	**37.6**	*35.1*	35.2	*35.3*
	Hybr.		*-0.8*	-0.4	*0.0*		*169*	191	*212*		*25.6*	25.7	*25.7*
CAR	Sol.	**0.4**	*-0.2*	0.3	*0.9*	**89**	*71*	88	*123*	**15.0**	*15.0*	15.0	*15.1*
	HC3v5	**3.4**	*0.2*	0.7	*1.2*	**392**	*84*	121	*153*	**15.5**	*15.2*	15.3	*15.3*
	CAMS	**1.5**	*-0.4*	0.9	*1.5*	**324**	*111*	132	*199*	**19.8**	*17.9*	18.0	*18.1*
	Hybr.		*-0.2*	0.0	*0.4*		*122*	140	*181*		*14.6*	14.6	*14.6*
CEN	Sol.	**0.8**	*-0.3*	0.1	*0.4*	**275**	*75*	94	*118*	**20.7**	*20.8*	20.9	*21*
	HC3v5	**5.4**	*-0.4*	-0.1	*0.4*	**506**	*62*	75	*94*	**21.5**	*20.7*	20.8	*20.8*
	CAMS	**2.3**	*-1.5*	0.0	*1.3*	**395**	*78*	143	*189*	**26.1**	*26.1*	26.2	*26.2*
	Hybr.		*-0.5*	-0.3	*0.0*		*220*	242	*275*		*19.9*	19.9	*20.0*
LIN	Sol.	**-2.8**	*-0.5*	0.0	*0.5*	**302**	*79*	92	*110*	**25.3**	*25.3*	25.4	*25.5*
	HC3v5	**-0.6**	*-0.8*	0.0	*0.9*	**259**	*88*	97	*124*	**26.3**	*26.5*	26.5	*26.7*
	CAMS	**2.1**	*-0.4*	0.9	*1.9*	**359**	*96*	136	*179*	**34.7**	*32.4*	32.5	*32.6*
	Hybr.		*-0.8*	-0.4	*-0.1*		*214*	235	*282*		*24.2*	24.3	*24.3*
PAL	Sol.	**-2.9**	*-0.3*	0.3	*0.8*	**324**	*80*	115	*131*	**23.5**	*23.5*	23.6	*23.6*
	HC3v5	**3.6**	*-0.4*	0.5	*0.9*	**449**	*76*	87	*103*	**23.3**	*23.3*	23.3	*23.3*
	CAMS	**4.3**	*-0.9*	0.6	*1.6*	**547**	*113*	146	*185*	**29.2**	*29.5*	29.6	*29.6*
	Hybr.		*-0.3*	0.0	*0.3*		*223*	250	*282*		*22.1*	22.1	*22.2*
PAY	Sol.	**1.1**	*-0.7*	-0.4	*0.3*	**139**	*65*	82	*96*	**22.2**	*23.3*	23.4	*23.5*
	HC3v5	**-0.4**	*-0.8*	-0.3	*0.2*	**201**	*64*	76	*94*	**25.4**	*26.4*	26.5	*26.6*
	CAMS	**5.5**	*-0.6*	-0.1	*0.3*	**384**	*54*	73	*104*	**29.8**	*28.1*	28.3	*28.4*
	Hybr.		*-0.6*	-0.4	*-0.0*		*163*	211	*249*		*22.5*	22.5	*22.6*



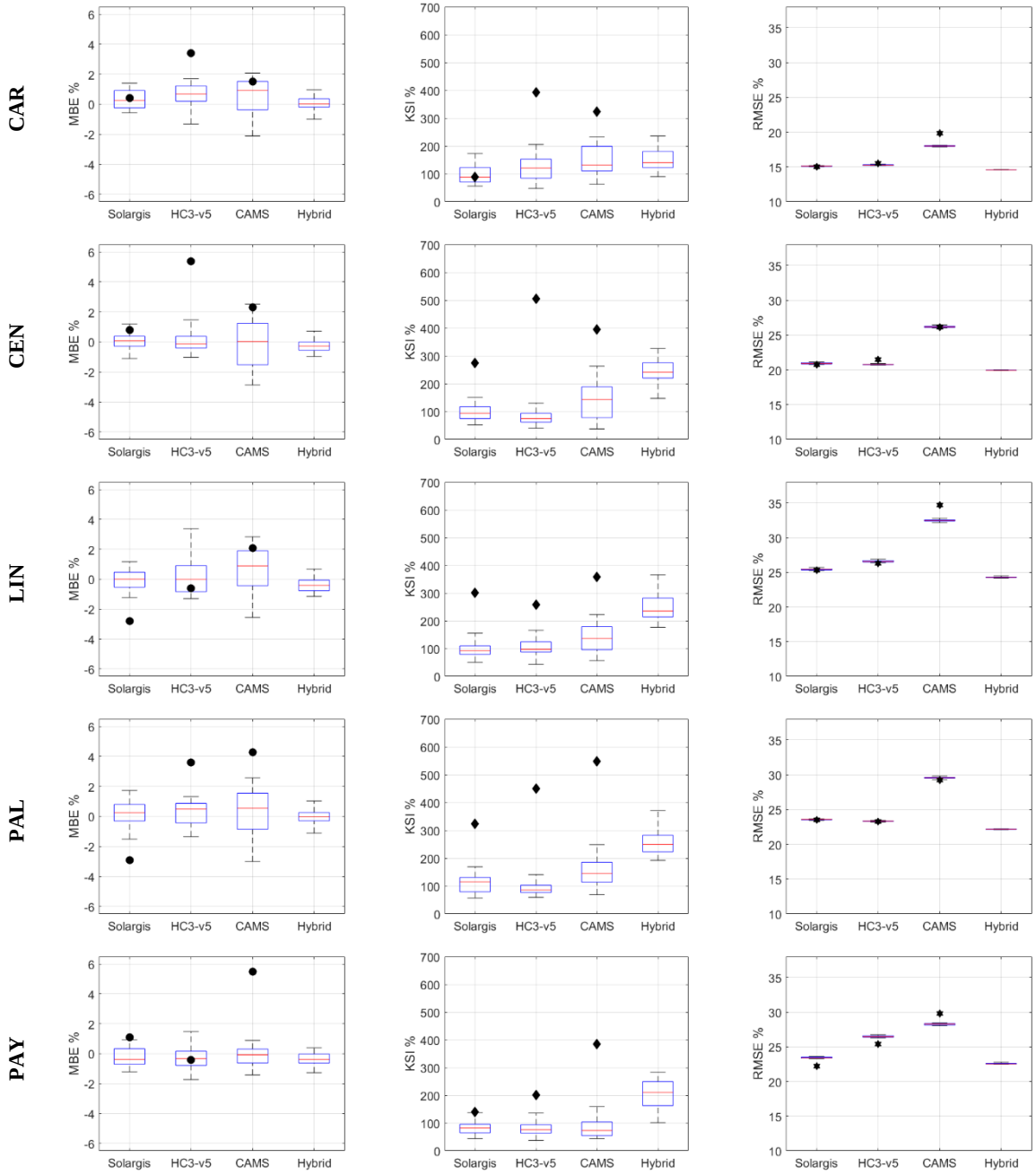


Figure 6: Comparative performances for 12-month measurement campaigns at seven BSRN locations. The black point is raw error of the satellite-based database (i.e., without calibration) while the boxplots depict indicator spread after calibration (P25 & P75).

It is then possible to see the overall behavior of each database by aggregating the data from the seven sites.

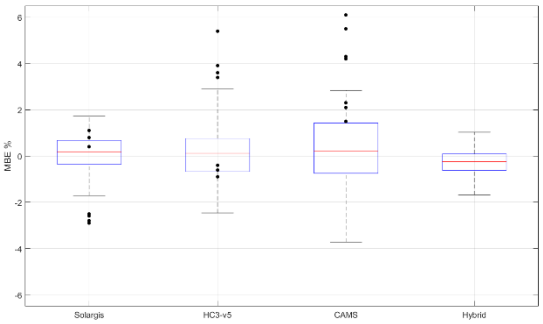


Figure 7: Overall performances for 12-month measurement campaigns at seven BSRN locations. The black points are the seven raw error of the satellite-based database (i.e., without calibration) while the boxplots depict MBE spread after calibration for the 7*100 compiled calibration processes.

The median results for the seven locations show that a 12-month measurement campaign systematically reduces the MBE when compared to the initial performance (*i.e.*, without calibration) when the initial bias is not already close to zero. Moreover, the median of the MBE systematically remains between -0.6 % and 0.9 % for all databases. Finally, the Hybrid solution systematically shows better performances with 50 % of the MBE (P25-P75 interval) remaining between -0.6 % and 0.1 % and 80% (P10-P90 interval) between -1% and 0.4%. It is however slightly negatively biased by -0.2% but this behavior can either be considered negligible or be corrected as almost constant on every given site.

When looking at Solargis results, the satellite time-series already has extremely low bias before calibration for two instances (Carpentras and Cener). Quantile Mapping displays a bias distribution which might look then worst. This cannot be considered as an issue as Solargis or Hybrid time-series systematically output very low residual post-calibration bias. In a real situation where nor raw nor calibrated bias can be computed, the use of site-adaptation guarantees to obtain reasonable remanent bias value while the uncorrected can typically range anywhere between -3.5% and 3.5% in Europe (as mentioned in both Solargis and HC3 validation reports). It is fair to state that most of calibration processes conducted in Western Europe using Solargis on any measurement campaign will lead to a global systematic error inferior to 1%. This is less true when making use of HC3v5 which can be calibrated extremely well in some cases (Cener) but poorly in other instances (Cabauw). CAMS is for its part not suitable for KDM calibration, showing too much fluctuation. No investigation has been conducted here to understand these differences between the databases. The respective models used are not visible in the resulting time-series which thus take the form of black box dataframes.

Other indicators demonstrate notable behaviors post-calibration. KSI is systematically reduced to reach values around 100% when using Solargis, HC3v5 and CAMS. This level of performance indicates that both satellite and measured CDFs are close together throughout all the interval: the two time-series are drawn from comparable random variables. Hybrid displays higher KSI on every BSRN station and is the least performing database for this one statistical error. Regarding the RMSE, it is however the one that shows the best results with levels consistently below the other three.

It is worth noticing that KDM does not have drastic impact on RMSE. This indicator indeed only highlights random errors committed by the satellite database, totally undetectable for Quantile mapping. Its monitoring still gives evidence that the method does not disperse the data, but only distorts its overall distribution keeping its original precision. Mathematically, Quantile Mapping only distorts the point cloud in the satellite vs true data scatterplot so that there is a symmetrical distribution of points around the bisector line $y = x$.

5 CONCLUSIONS

Raw bias analysis on high quality historical BSRN data proves that top-tier satellite-based databases may have systematic errors which significantly jeopardize the reliability of a yield assessment. Assuming that the systematic bias is temporally stationary, long-term satellite time-series can theoretically be perfectly compensated using calibration on short-term in-situ measurements. Yet experiments show that the performance fluctuates depending on several input features.

An effective benchmark of possible configurations allows the characterization of the calibration parameters to be favored by the surveyor in order to maximize the representativeness of the corrected irradiance data. Once the method has been set, the three main parameters that can affect site adaptation are the start date of the campaign, its duration, and the choice of the satellite database. While the first is rarely adjustable in a professional context, the other two must be considered to effectively lower the uncertainty compared to raw data. The study carried out makes it possible to rule statistically on the latter. The three major conclusions are:

- 12 months of measurement is the optimal duration. This is the minimum period to capture all seasonal effects. It is also found that an extension between 12 and 24 months does not bring any improvement and may even degrade the level of performance.
- Satellite databases are not equally suitable for calibration. Among the three tested, Solargis seems to be the most consistent due to systematically less distributed post-calibration biases on various locations. Using HelioClim-3v5 or CAMS seem to lead to higher uncertainty.
- Satellite databases can be combined after being independently calibrated in order to benefit from fluctuating performances, thus leading to better performances.

As previously pictured, a good choice of parameters can drastically decrease uncertainty on remanent bias and lead to a much higher level of confidence: calibrated time-series biases typically range between -0.5% to 0.5% while raw values range between -4% and 4%. Consequences are double: P50 yields are more representative and prone to better bankability; gap between P90 and P50 can be drastically reduced since uncertainty on the satellite database is dropped from 3% (as suggested by validation reports) to somewhere around 0.5%. PV developers could then benefit from a slightly higher debt leverage due to a P90 being closer to P50.

As a perspective, it would be interesting to extend this analysis to other calibration methods and to have a better understanding of why some campaigns work better than others.

6 REFERENCES

[1] Suri, M., Huld, T., Dunlop, E. D., Albuisson, M., Lefèvre, M., & Wald, L. (2007). Uncertainties in photovoltaic electricity yield prediction from fluctuation of solar radiation. In *22nd European Photovoltaic Solar Energy Conference*.

[2] Polo, J., Fernández-Peruchena, C., Salamalikis, V., Mazorra-Aguiar, L., Turpin, M., Martín-Pomares, L., et al. (2020). Benchmarking on improvement and site-adaptation techniques for modeled solar radiation datasets. *Solar Energy*, *201*, 469-479. https://doi.org/10.1016/j.solener.2020.03.040

[3] Polo, J., Wilbert, S., Ruiz-Arias, J. A., Meyer, R., Gueymard, C., Súri, M., Cebecauer, T. (2016). Preliminary survey on site-adaptation techniques for satellite-derived and reanalysis solar radiation datasets. *Solar Energy*, *132*, 25–37. https://doi.org/10.1016/j.solener.2016.03.001

[4] Ohmura, A., Dutton, E. G., Forgan, B., Fröhlich, C., Gilgen, H., Hegner, H., ... & Wild, M. (1998). Baseline Surface Radiation Network (BSRN/WCRP): New precision radiometry for climate research. *Bulletin of the American Meteorological Society*, *79*(10), 2115-2136.

[5] Espinar, B., Blanc, P., Wald, L., Hoyer-Klick, C., Homscheidt, M. S., & Wanderer, T. (2012). On quality control procedures for solar radiation and meteorological measures, from subhourly to monthly average time periods. *EGU General Assembly, Vienna, Austria*, 22-27.

[6] Šúri, M., Cebecauer, T., & Skoczek, A. (2011, September). SolarGIS: Solar data and online applications for PV planning and performance assessment. In *26th European photovoltaics solar energy conference*.

[7] Blanc, P., Gschwind, B., Lefèvre, M., & Wald, L. (2011). The HelioClim project: Surface solar irradiance data for climate applications. *Remote Sensing*, *3*(2), 343-361.

[8] Schroedter-Homscheidt, M., Arola, A., Killius, N., Lefèvre, M., Saboret, L., Wandji, W., et al. (2016). The Copernicus atmosphere monitoring service (CAMS) radiation service in a nutshell. *Proc. SolarPACES16*, 11-14.