Wide-area sky view factor analysis and Fourier-based decomposition model for optimizing irradiance sensors allocation in European solar photovoltaic farms

A software tool

Alcañiz, Alba; van Kouwen, Mathijs I.; Isabella, Olindo; Ziar, Hesan

DOI
10.1016/j.solener.2024.113139

Publication date
2025

Document Version
Final published version

Published in
Solar Energy

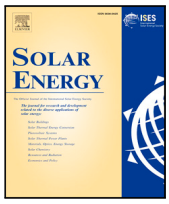
Citation (APA)
Alcañiz, A., van Kouwen, M. I., Isabella, O., & Ziar, H. (2025). Wide-area sky view factor analysis and Fourier-based decomposition model for optimizing irradiance sensors allocation in European solar photovoltaic farms: A software tool. *Solar Energy*, *286*, Article 113139. https://doi.org/10.1016/j.solener.2024.113139

Solar Energy 286 (2025) 113139

Contents lists available at ScienceDirect

Solar Energy

journal homepage: www.elsevier.com/locate/solener



Wide-area sky view factor analysis and Fourier-based decomposition model for optimizing irradiance sensors allocation in European solar photovoltaic farms: A software tool

Alba Alcañiz *, Mathijs I. van Kouwen, Olindo Isabella, Hesan Ziar

Photovoltaic materials and devices group, Sustainable Energy Technology department, Faculty Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, The Netherlands

ARTICLE INFO

Keywords:
Pyranometer allocation
Solar farm
Software tool
Irradiance map
Irradiance decomposition
Sky view factor

ABSTRACT

Solar farm installers generally struggle with the allocation of irradiance sensors throughout the plant area, which are essential for monitoring purposes. Despite the existence of the International Electrotechnical Commission guidelines for photovoltaic (PV) plant monitoring, no specific guidance is provided when it comes to allocating sensors. This can be especially problematic for solar farms in hilly terrain. In this work, a software tool is built to allocate horizontal and in-plane irradiance sensors. Additionally, advice on the optimum number of sensors and the prevented error is provided based on the layout of the farm. The methodology consists of calculating the irradiance at every point of the solar farm area and finding the one closest to the average. This average is computed differently depending on the sensor type and monitoring purpose. A modification of the BRL irradiance decomposition model is also proposed to reduce the bias of the original model. The software has been applied to two case studies of existing solar farms in hilly areas in Greece and Germany, showing its applicability for real case scenarios in different climates and geological landscapes. The runtime of the software tool is mainly a function of solar farm size and the land morphology of its location.

This methodology has been only developed for monofacial fixed-tilted PV farms.

1. Introduction

With the increasing need for clean energy sources, more photovoltaic (PV) farms are being installed worldwide. These large-scale installations prioritize the use of precise irradiance data. Therefore, they typically measure the irradiance onsite by using both pyranometers and working-class reference cells. The difference in operational principles among these devices can result in noteworthy variations in irradiance [1–3]. Moreover, these instruments have an inherent uncertainty. The measurement uncertainty of pyranometers is around 1%, although in the field it can increase to 1.3 to 1.7% in economically relevant hours of the day [4]. Accurate irradiance measurements are key for the precise assessment of PV farm projects [2] and these deviations can be considerably detrimental for some purposes.

The type of irradiance measured and the time resolution depends on the purpose of these measurements [5–7]. One can distinguish between global horizontal irradiance (GHI) and plane-of-array (POA) irradiance G_{POA} measurements. While the former is placed horizontally, the latter is at the same inclination as the PV modules. Horizontal sensors are employed to connect with satellite data or for forecasting purposes when measuring at high temporal resolutions. Forecasting can also be performed with high-resolution POA sensors, which are also employed for power loss analysis (PLA) on less than a daily basis and performance ratio (PR) calculations on an annual basis. This latter variable can be significantly affected by irradiance deviations [1], highlighting the need for highly accurate irradiance measurements.

To reduce the uncertainty, the standard 61724-1:2021 from the International Electrotechnical Commission (IEC) "serves as guidance for monitoring system choices" [8]. This standard requires class A pyranometer measurements of in-plane irradiance and GHI throughout the plant. The standard also provides the number of sensors that should be placed horizontally and in-plane depending on the PV plant layout. Guidelines are also provided regarding the location of these pyranometers:

- The location shall be chosen as representative.
- Shading on the sensor shall be avoided. If it occurs, it may only be within a half hour from sunrise and sunset and it shall be documented.

* Corresponding author.
E-mail address: A.AlcanizMoya@tudelft.nl (A. Alcañiz).

https://doi.org/10.1016/j.solener.2024.113139
Received 20 March 2024; Received in revised form 4 July 2024; Accepted 21 November 2024
Available online 6 December 2024

Table of symbols :

Symbol	Definition	Unit
α	Ground albedo	–
d_n	Day of the year	–
D_n	Normalized day of the year	–
$\Delta\epsilon$	Prevented measurement deviation	W h/(m^{2}y)
DHI	Diffuse Horizontal Irradiance	W/m^2
DNI	Direct Normal Irradiance	W/m^2
ϵ_d	Measurement deviation for minimizing daily bias	Wh/(m^2 y)
ϵ_h	Measurement deviation for minimizing hourly bias	Wh/(m^2 y)
ϵ_y	Measurement deviation for minimizing yearly bias	Wh/(m^2 y)
G_{hor}	Horizontal irradiance considering obstructions	W/m^2
G_{POA}	Plane of array irradiance	W/m^2
G_{ref}	Reference irradiance	W/m^2
γ	Angle of incidence of the PV modules	°
GHI	Global Horizontal Irradiance	W/m^2
k_d	Diffuse fraction	–
N_{gp}	Number of PV module groups in the farm	–
N_{pl}	Number of planes in the farm	–
N_s	Number of irradiance sensors	–
nMBE	Normalized mean bias error	–
nRMSE	Normalized root mean squared error	–
R_d	Diffuse transposition factor	–
rPMD	Relative prevented measurement deviation	–
SVF	Sky View Factor	–
SVF	Weighted average SVF based on the number of modules per group	–

- GHI sensors should be leveled to within 0.5°. G_{POA} sensors should be aligned with the intended plane within 0.5° tilt and 1° azimuth angles.

These IEC guidelines can be however confusing due to the absence of clarification on what constitutes a *representative* location. A meticulous allocation approach is crucial when planning the monitoring infrastructure, especially in PV farms with spatial differences in irradiance due to altitude variations or horizon obstructions [9]. The usage of few sensors or incorrectly allocated sensors can lead to, among others, incorrect evaluations and fault assessments [5,6,10].

In the literature, some methods for sensor allocation have already been proposed. In [9], an optimization algorithm was presented based on minimizing the error induced by different cloud conditions. However, the proposed method was developed only for a flat rectangular PV plant with an unobstructed horizon. The results reported in [4], show that the optimal sensor placement is away from the edges of the PV farm and at roughly 25% and/or 75% across the module. These results, however, were obtained for an unobstructed tracked bifacial PV farm. Another allocation algorithm used an interpolated irradiance map [11] which was obtained from the sensor data using the Kriging interpolation technique [12]. However, the methodology requires sensor data information hence it cannot be used before the sensors are installed.

A workaround would be using satellite data, making ground-based irradiance measurements redundant. However, satellite irradiance data currently offers low spatial resolution and is less accurate than ground-based measurements [13,14]. Therefore, most PV power plant commissioners prefer using on-site irradiance sensors [15]. In conclusion, no universal and scalable algorithms exist for irradiance sensor allocation.

This work aims to fill this gap by developing a software tool that provides the optimal sensor allocation for PV farms considering their elevation profile. The irradiance is calculated at all the possible sensor locations of the PV farm, and the most *representative* location is proposed based on an error minimization process. The error definition depends on the purpose of irradiance measurement. Additionally, advice on the optimal number of sensors is provided based on the layout of the farm.

This work is structured as follows. The methodology of the software is explained in detail in Section 2. The databases required for the software are presented in Section 3. The main results consisting of the software overview and its application to two case studies are presented in Section 4. These outcomes are discussed in Section 5 before concluding in Section 6.

2. Methodology

This section explains the methodology employed to determine the optimal sensor allocation. It assumes that the user knows the location and layout of the future PV plant. The user can or cannot know how the sensors will be used and the number of sensors to be placed.

So far, this methodology has been developed for monofacial fixed-tilted PV farms. Nevertheless, a similar framework applies to the increasingly widespread bifacial and/or tracking PV installations.

The strategy of the methodology consists of calculating the irradiance at each point of the PV plant area and selecting the most representative points as those with values closest to the average. In particular, it consists of the following steps, which are graphically presented in Fig. 1:

1. Determine the sensor distribution amongst groups and planes.
2. Import Digital Surface Model (DSM) data.
3. Perform Sky View Factor (SVF) calculations for relevant plane-of-arrays and interpolate grid.
4. Import GHI and albedo α.
5. Generate a G_{POA} map for the relevant plane-of-arrays using an improved Boland–Ridley–Lauret (BRL) decomposition model and the Perez transposition model [16].
6. Determine the reference irradiance G_{ref} and calculate the error map for every sensor.
7. Suggest the best sensor locations.
8. Create output report including figures and tables.

In the following, each of these steps is thoroughly explained. However, before diving in, some assumptions need to be taken:

- Pyranometers are placed within the PV plant area, defined here as the terrain to which the PV farm operator has access. Within that terrain, all obstacles such as trees or water streams are disregarded.
- POA sensors are always placed maximally 10 m away from the PV group for which they measure. This step expands the boundaries

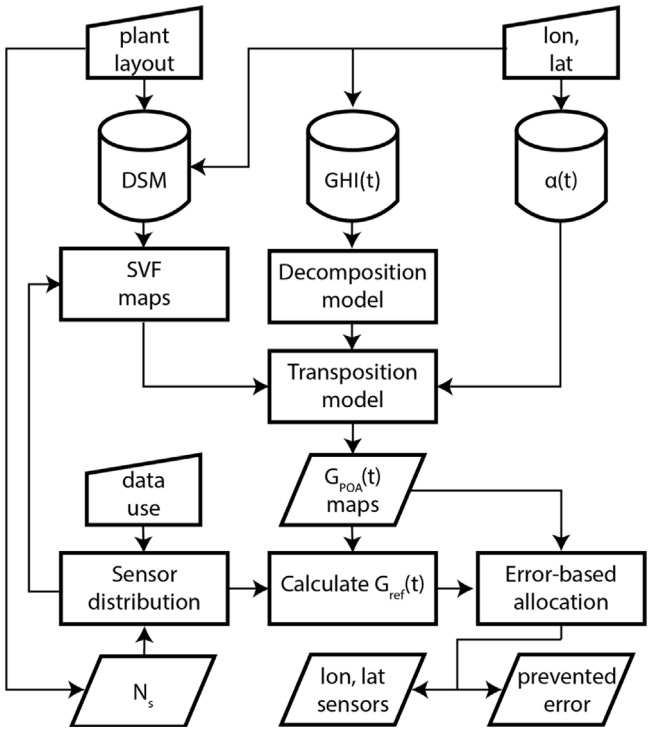


Fig. 1. Flowchart of the methodology followed in this work. The trapezoids are user inputs, the cylinders denote imported data, the rectangles are predefined algorithms, and the parallelograms denote (intermediate) outputs.

of the PV group to increase accessibility with the assumption that no significant alterations occur in this additional terrain.

PV group is defined here as a cluster of modules with the same characteristics. To ensure that the sensors capture all the variability across the farm, the user-defined PV groups should be selected by grouping modules that are expected to perform similarly. For instance, by separating modules that could be subjected to different shading levels or that are connected to different inverters.

The difference between PV group and PV plant area is visible in Fig. 14 for one of the case studies (marked with blue and red dots and labeled as plant area and panel area, respectively).

- PV modules are evenly distributed within user-input PV groups.
- GHI and its decomposed components are spatially constant in the whole PV plant (justified by the use of hourly averaged data).
- For computational efficiency, the SVF calculations for the PV panels and the pyranometers are done at 1.5 m height, approximately the middle point of modules from the ground [10].
- Albedo is spatially constant but temporally variant.

Section 2.1 starts by specifying the required inputs. The creation of sky view factor and irradiance maps is explained in Sections 2.2 and 2.3, respectively. The core of the methodology, the error-based minimization problem, is presented in Section 2.4. Section 2.5 elaborates on the tactic to find the optimal number of sensors for each PV farm. Finally, outputs and performance indicators are described in Section 2.6.

2.1. Inputs

The user inputs are:

1. Purpose of the irradiance sensors, which influences the sensor type and its optimal spatial distribution. The defined purposes are [5–7]:
   (a) PR: Yearly calculation of performance ratio using G_{POA}.
   (b) PLA, daily: Daily power loss analysis with G_{POA}.
   (c) PLA, real: Power loss analysis in real-time (hourly) with G_{POA}.
   (d) POA, for: Forecasting models using hourly G_{POA}.
   (e) Hor, for: Forecasting models using hourly G_{hor}.
   (f) Sat: Connection with satellite data via GHI.

   where G_{hor} is the horizontal irradiance considering horizon obstructions, that is the irradiance measured by horizontally placed pyranometers. It is different from satellite readings of GHI that are based on a free horizon.

   When selecting multiple data usages, the user is prompted to rank them so the software tool prioritizes purposes in case of conflicts.
2. PV farm location.
3. Location, tilt, and azimuth of modules, clustered into groups of equal inclination.
4. PV plant capacity.
5. Number of sensors, if known.

2.2. Sky view factor

After defining the inputs, the sky view factor (SVF) is computed, a parameter necessary to estimate the incident irradiance. The SVF is defined as the portion of the sky visible from the PV module. It ranges from 0 to 1 and is affected by the inclination of the module and altitude differences in the terrain or horizon obstructions. Digital Surface Model (DSM) data is employed to consider these terrain differences. This data provides the height for equidistant points across an area. The spatial resolution of the SVF map is therefore determined by that of the DSM data. Linear interpolation can be employed to increase the spatial resolution and with it the number of possible sensor locations. The SVF is computed for the entire PV plant area obtaining an SVF map.

The calculation of the SVF is based on the work of Keijzer [17] and further implemented in the work of de Jong [18]. The sky hemisphere is divided into altitude bands and azimuth slices. The maximum obstructed altitude angle is determined for every azimuth slice. Only DSM grid points at a distance lower than a certain radius of evaluation are considered.

The number of altitude bands is already optimized in [17] to be 90, but the azimuth slices and radius of evaluation are optimized here by using data from the 3797 European solar parks that are registered in the Global Power Plant Database (GPPD) [19] (see Fig. 2(a)). Fig. 2 shows part of the results of the optimization process as a trade-off between accuracy and computational resources. 30m-resolution DSM data from Sentinel Hub is employed in this analysis [20].

The incremental increase in SVF is plotted against the number of azimuth slices for all locations in Fig. 2(b). The results in the figure are in line with what is expected: a higher amount of azimuth slices leads to a more accurate and higher SVF. In this figure, one can also observe how most existing solar plants are in areas without large local height differences since most PV plants have an SVF of 0.97 or higher, and the mean SVF is close to 0.999. 1080 azimuth slices are selected as optimal, given that the SVF barely changes after this point. This corresponds to an azimuth resolution of 0.3°.

Using this optimized value, the SVF was computed for varying radii of evaluation on four planes with different tilts. Fig. 2(c) illustrates the mean change in SVF compared to the value at 250 m for the horizontal plane. As expected, the SVF reduces with increased radius. A too-small radius overestimates the SVF due to the underestimation of horizon obstruction. The SVF decreases the most in the 500 m to 750 m range, as most PV plants have few obstructions in their direct surroundings. At large radii, only large hills, mountains, or high-rise buildings will have a significant impact on the SVF. Similar behaviors were observed for other tilted planes and orientations. Based on this analysis, an optimum value of 2000 m was selected. It should be noted that these results are obtained for the imported DSM data thus they may not apply to data with different resolutions.

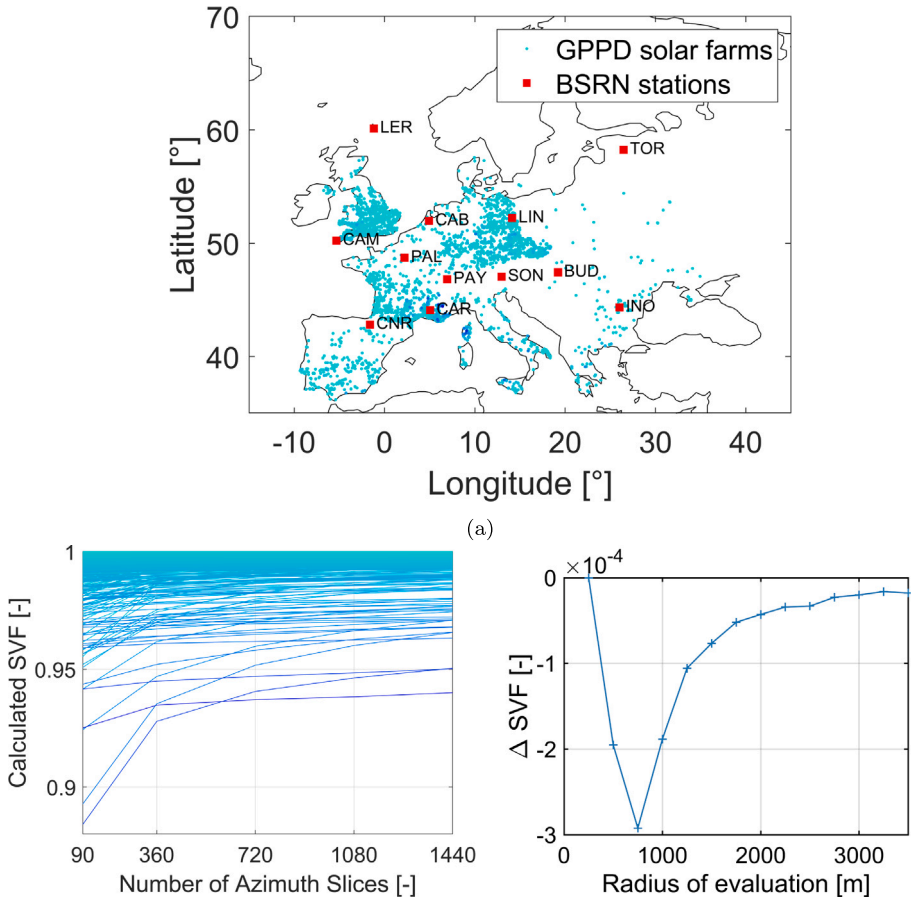


Fig. 2. Optimization of SVF parameters for a horizontal plane for the GPPD European PV farms. (a) Location of the GPPD solar farms in cyan and of the Baseline Surface Radiation Network (BSRN) meteorological stations in red squares with the corresponding labels. (b) Effect of increasing the azimuth slices. Each colored line corresponds to a PV farm of the matching color as depicted on the map. (c) Mean change in SVF for increasing radii of evaluation compared to the SVF obtained with a radius of 250 m.

2.3. Irradiance maps

Once the SVF is available, the POA irradiance (G_{POA}) can be calculated. This represents the irradiance received by the modules. Several inputs are required, including GHI, Diffuse Horizontal Irradiance (DHI), Direct Normal Irradiance (DNI), the ground albedo α, the PV module angles, and the solar angles. For definitions of these parameters, the reader is referred to [21].

At every time step, the solar angles can be calculated from the plant location [21]. The albedo is imported as explained in Section 3.3. GHI can be imported as well (see Section 3.2), but obtaining DHI and DNI measurements is challenging due to the need for costly equipment. However, decomposition models can be employed to calculate DHI from GHI. Subsequently, DNI can be derived using the closure equation [22].

The decomposition step is typically the main source of uncertainty in G_{POA} evaluation [23]. Therefore, insight into the best-performing decomposition models is valuable. Despite the several comparisons between decomposition models in Europe available in the literature [24–32], no model stands out.

Therefore, a comparison of the most common decomposition models is performed in Appendix A. Among all the models tested, BRL proves to be the overall best performer, exhibiting superior results in terms of normalized root mean squared error (nRMSE) and an average bias. The bias exhibits a seasonal pattern consistent across all models, with an overprediction of the diffuse component in autumn and an underprediction in spring. This seasonal normalized mean bias error (nMBE) trend is mitigated in this work for the BRL model through a Fourier-mediated adjustment, resulting in the BRL-MvK model, a BRL tailored for European climates.

Once GHI, DHI, and DNI are obtained, G_{POA} can be computed. This irradiance comprises three components: diffuse, direct, and ground-reflected. Whereas the latter two can be calculated deterministically using equations 18.18, 18.20, and 18.22 from [21], the diffuse component is determined using a diffuse transposition factor R_d [22]. Different sky diffuse models exist for calculating this diffuse fraction. Comparison studies based in Europe [33–43] show that the Perez model [16,44] is superior most often.

The time-varying $G_{POA}(t)$ is calculated for each location of the plant over two years. Since only two years are selected, the influence of solar cycles, which have an 11-year periodicity, is not considered. Nevertheless, their impact on yearly incident irradiance is minor and can be disregarded. Since the $G_{POA}(t)$ depends on the PV module angles, this process is performed for every PV group or plane. When horizontal sensors are to be placed, the process is repeated considering a 0° tilt.

2.4. Error-based minimization problem

Once the irradiance maps are ready, the pyranometers can be allocated. As mentioned in the introduction, the objective is to find the most *representative* location. At that location, the irradiance time series reflects the irradiance incident on the group of PV modules to which the sensor is assigned. A representative time series can be defined as a series in which the values are close to the reference time series at each instant. Since each module in the group is subjected to a different irradiance profile and all modules have the same weight, the reference time series is defined as the mean irradiance.

Using the time-dependent irradiance map, the reference irradiance $G_{ref}(t)$ is constructed by averaging the irradiance for each PV group at every instant. That value is then compared to the calculated $G_{POA}(t)$ for every potential pyranometer location l in the PV plant area $l \in L$. L is the number of possible locations in which a sensor can be placed. $G_{POA,l}$ is the POA irradiance at a location l. The location with the minimal error is chosen as the most representative one. The error $\epsilon_{p,l}$ is calculated as the sum of the absolute differences for every time instant, see Eq. (1). $p \in [h, d, y]$ as the error can be calculated on an hourly, daily, or yearly basis. The time resolution of the error is determined by the monitoring purpose, as the minimization problem is defined by the period $t \in T$ over which no bias should occur.

$$\arg\min_{l \in L} \left(\epsilon_{p,l}\right) = \arg\min_{l \in L} \left(\sum_{t \in T} |G_{POA,l}(t) - G_{ref}(t)| \right) \tag{1}$$

G_{ref} varies depending on the sensor orientation and PV plant distribution. The following subsections explain the procedure for each possibility.

Single group

First, the objective is to place an in-plane pyranometer representative of a group of PV modules that reside in a single area with identical tilt and orientation (same plane). Assuming equal distribution across the PV group, G_{ref} is the average across all irradiance time series within the PV group, Eq. (2).

$$G_{ref}(t) = \frac{1}{L} \sum_{l=1}^{L} G_{POA,l}(t) \tag{2}$$

Multiple groups, same plane

In this second case, all PV modules have identical tilt and orientation (same plane), but they are not clustered, therefore belonging to multiple groups of modules. G_{ref} is defined as the average of the irradiance of each group weighted by the number of modules in each group. Eq. (3) expresses this weighted average, where N_{gp} are the number of groups in the PV farm, N_n are the number of modules in each group n, and L_n is the number of locations in which a sensor can be placed in group n.

$$G_{ref}(t) = \frac{1}{\sum_{n=1}^{N_{gp}} N_n \cdot \sum_{n=1}^{N_{gp}} L_n} \sum_{n=1}^{N_{gp}} \sum_{l=1}^{L_n} G_{POA,l}(t) \cdot N_n \tag{3}$$

Since the spatial resolution of the imported data may not coincide with the distance between PV modules, the possible sensor locations L_n need to be considered. The schematic in Fig. 3 can help understand this. A PV plant is represented by three groups of modules ($N_{gp} = 3$), all with the same orientation. Sensor locations are marked by colored dots: red for group 1, yellow for group 2, pastel pink for group 3, and gray for boundaries. Considering this, $N_1 = 9$, $N_2 = 5$, and $N_3 = 12$, but $L_1 = 4$, $L_2 = 3$, and $L_3 = 8$. In the software tool, some of these boundaries would also be considered as pyranometer locations (as long as they are within 10 m of a PV module) but they are discarded for ease of understanding in this example.

Horizontal sensors

For the horizontal sensors, the following adaptations occur in the minimization procedure:

- The SVF grid covers the whole plant area.
- G_{POA} is replaced by G_{hor} for a forecasting sensor and GHI for a satellite connection sensor. This distinction is because local differences should be included only for forecasting purposes.
- The hourly bias is always minimized.

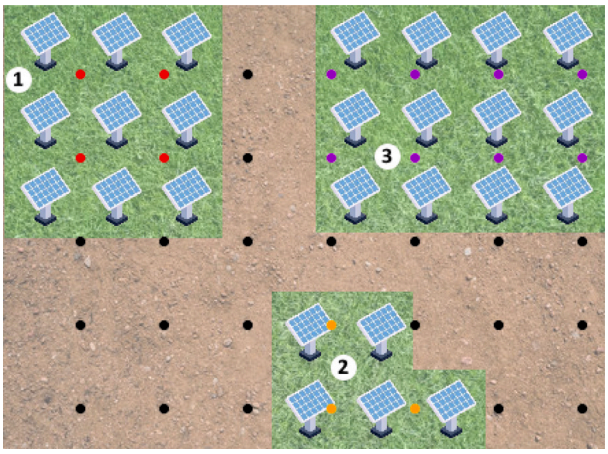


Fig. 3. Example of a PV farm sketch with 3 groups of PV modules and matrix of sensors' locations to help understand the calculation of G_{ref}.

Multiple sensors

When multiple sensors within the same plane and group are to be placed, the minimization procedure changes:

1. The grid points that are within the lowest 10% of the error map are identified.
2. Using k-means clustering [45], k clusters are formed, where k is equal to the number of to-be-placed pyranometers.
3. The minimization is performed for each cluster.

If simply the k best locations are chosen, all locations will be next to or near the global minimum and thus spatially very close. k-means clustering ensures that the pyranometer locations are spatially spread.

2.5. Number of sensors

If the user indicates no number of sensors, advice is generated by the software considering no (financial) limitations. The rules are that when at least one in-plane data usage is chosen, a sensor is advised for every plane and an additional sensor per PV group. One sensor is advised for each use of horizontal irradiance measurements. Fig. 4 graphically summarizes this decision process.

The number of sensors advised by the software tool is from now on referred to as the *ideal number of sensors*. The eventual distribution of the ideal number of sensors N_s when all future purposes are chosen is:

- $N_{s,POA,pl} = N_{pl}$
- $N_{s,POA,gp} = N_{gp}$
- $N_{s,hfor} = 1$
- $N_{s,sat} = 1$

Where N_{pl} stands for the number of planes in the PV plant (defined as the unique combinations of tilt and orientation), $N_{s,hfor}$ denotes the number of horizontal sensors for forecasting, and $N_{s,sat}$ denotes the sensor for satellite connection.

When the user inputs a number of sensors that deviate from the ideal, the following rules, implemented via nested if-else loops, ensure the optimal distribution:

1. When $N_s = 1$, and both in-plane and horizontal sensors are required based on the data purposes, the in-plane sensor gets priority if $N_{pl} = 1$. Otherwise, $N_{s,hfor} = 1$.
2. When $N_s = 2$, and more in-plane and horizontal sensors are required to accomplish all intended data purposes, $N_{s,POA} = 1$ and $N_{s,hfor} = 1$, as long as it does not violate another rule.

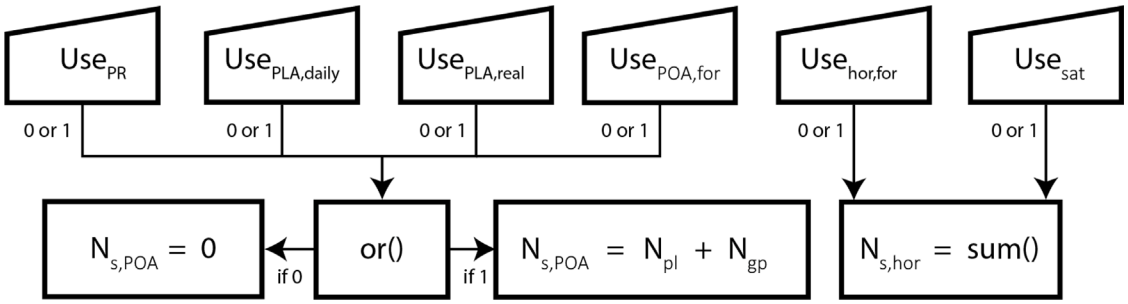


Fig. 4. Flowchart illustrating how the advice for the number of horizontal and in-plane sensors follows from the foreseen data collection purposes as input by the user. The trapezoids denote binary variables that indicate whether a certain data purpose is entered by the user. The subscripts in the figure denote the six data-gathering purposes.

3. When $N_s > 2$ and one or two horizontal sensors are to be placed: $N_{s,hfor} > 0$.
4. Distribution of in-plane sensors follows:
   - $N_{s,POA,pl} = N_{pl}$ or 0 (all planes or no planes have a sensor).
   - $N_{s,POA,gp} = N_{s,POA} - N_{s,POA,pl}$.
   - In-plane group sensors are distributed based on the group size.
5. When conflicts occur, the user-input priorities determine how the pyranometers will be placed.

2.6. Performance indicators

Next to providing allocation advice, the software tool intends to show the benefits of its use via performance indicators. The main used performance indicator is the prevented measurement deviation (PMD) $\Delta\epsilon_p$ (see Eq. (4)). It indicates the accuracy improvement of the allocated sensor compared to the worst-case location, thus it can be interpreted as the prevented error made. It can be defined for every time resolution, with $p \in [y, d, h]$.

$$\Delta\epsilon_p = \epsilon_{p,l_w} - \epsilon_{p,l_b} \tag{4}$$

where l_w and l_b are the worst-case and the advised sensor location, respectively. These values are directly retrievable from the error map.

Another performance indicator is the relative PMD:

$$rPMD_p = \frac{H \cdot \Delta\epsilon_p}{H' \cdot \sum_{h=1}^{H} G_{ref}(h)} \cdot 100\% \tag{5}$$

With $H' = 8760$ and H the number of hourly time steps employed, as the methodology is commonly performed for two years of data.

3. Databases

This section presents the databases needed for the software tool. Section 3.1 justifies the selection of the digital surface model data. Section 3.2 briefly presents the retrieval of the GHI satellite data, and Section 3.3 explains the choice and preparation for the time-dependent albedo data.

3.1. Digital surface model

The digital surface model (DSM) data is used in this work to calculate the SVF. To access DSM data in the entirety of Europe and to prevent bulk downloads of the height data, an online tool is chosen that combines globally available DSM data: Sentinel Hub [20]. Sentinel Hub is a big data satellite imagery service. Users can quickly and easily access satellite data from complete archives using an application programming interface (API). The Sentinel Hub service offers at least 30 m resolution DSM data for Europe.

3.2. Global horizontal irradiance

Since the possibility of using an API and having a high resolution was a priority in this work, the PV-GIS database was selected for importing historical measurements of GHI [46]. Using the BSRN ground-based measurements [47], the PV-GIS imported satellite-based measurements are validated for the year 2016. The RMSE is in the range of 90 to 140 W/m^2 for all stations. The source of this inaccuracy likely arises from the cumulative error propagation throughout all the modeling steps. For this work, however, ensuring unbiased data is more crucial than historical accuracy. The average absolute nMBE is only 3% and that is considered good enough.

3.3. Time-dependent albedo

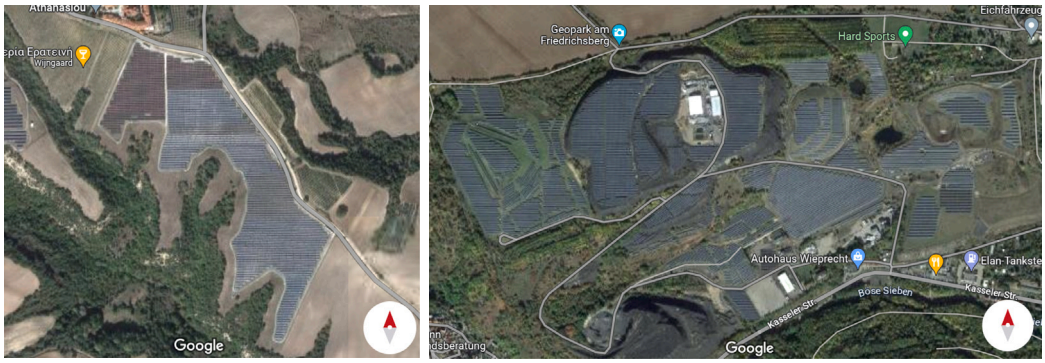
The effective ground albedo α fluctuates during the day due to the reflectivity of the ground being dependent on the incident angle of radiation [48,49], and throughout the year due to vegetation changes, snow coverage, and other factors [49,50]. When albedometers are unavailable, satellite data can be used to determine the ground albedo [51,52]. Nevertheless, the spatial resolution of this data is too low to be used on a local scale. The available albedo databases with a spatial resolution of up to 1 km sacrifice temporal resolution, but their spatial resolution is still insufficient for local-scale applications [53].

Considering this, satellite data is used as a source of time-dependent ground albedo. A resolution of one hour is needed, as the data will be combined with the hourly GHI. The only database with hourly albedo in Europe is the NASA POWER database [54]. The spatial resolution of these albedo measurements is only 1° × 1°, which is around 110 km × 110 km. Although this is not a high spatial resolution, using a spatially inaccurate time-dependent albedo is considered closer to reality than assuming a constant albedo.

Many stochastic climatic processes induce a high variability in the yearly albedo pattern. Therefore, as recommended by the literature [55, 56], hourly albedo data from 2017 until 2021 are averaged to create one sample year of data.

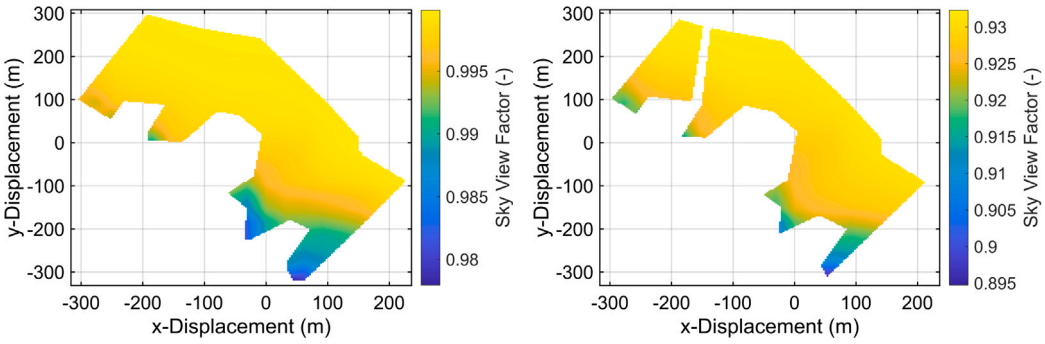
4. Results

The developed methodology was implemented in Matlab in the form of a software tool. Its code is available as supplemental material for this article. The software is presented in Appendix B and a video is supplied as additional material to showcase the usage of the tool. Its application in two case studies is explored in Section 4.1. A simplification of the current version of the software is explained in Section 4.2.



(a) (b)

Fig. 5. Satellite images of the solar park (a) in Kolindros (Greece) and (b) in Eisleben (Germany) used as case studies. *Source:* Images from [58].



(a) (b)

Fig. 6. Surface plots of the SVF map for the horizontal (a) and in-plane (b) sensors of the Kolindros case study. Please note that the gradient scales are tuned for each subfigure.

4.1. Case studies

This section presents the use of the tool on two existing PV plants located in hilly terrain as case studies. For each of them, the resulting pyranometer allocations are presented, together with all preliminary results that have led to an informed decision on the most representative locations.

The 6 MW_p solar park in Kolindros, Greece, is situated close to the mountain range where the mountain Olympos is. As seen in Fig. 5(a), it is a solar park with only two groups of PV modules, oriented south. The tilt is estimated to be 30° which is the optimal tilt for south-oriented panels at that location [57].

The Solar Park Krughütte is a 29.1 MW_p plant in Eisleben, Germany. It consists of small groups of PV modules south- and southeast-oriented, as seen in Fig. 5(b). The tilt is estimated to be 33°, which is optimal for that location [57].

4.1.1. Solar park in Kolindros

The tool starts by calculating the SVF for all grid points within the possible sensor allocation area. Since the SVF depends on the tilt and orientation, SVF maps are created for the horizontal plane and the 30°-inclined and south-oriented plane. The results are shown in Fig. 6. The SVF for horizontal planes and tilted planes decreases towards the southern part of the PV plant. This is a direct result of the local height differences in the PV plant, as the altitude is relatively low in the south compared to the northwest corner. The tilted SVF is lower than the horizontal one, as seen by the range of values of the color bars. For horizontal planes, the SVF ranges between 0.98 and 0.995, while for tilted planes, it varies from 0.895 to 0.93. The maximum difference in SVF across the PV plant for the tilted plane is 0.035. Although a small number, it can lead to noticeable G_{POA} differences over the year.

Using the SVF maps, G_{POA} maps are determined with the imported GHI time series for the years 2018–2019. For every grid point, the error is calculated. The error definition and time resolution depend on the intended data usage, as explained in Section 2.4. Fig. 7 displays the hourly error maps for horizontal sensors for forecasting and connection with satellite purposes.

As seen in Fig. 7(a), the error for the sensor used for horizontal forecasting is minimal at a band of grid points that crosses diagonally the PV plant (indicated by the red dashed line in the figure). Thus, the most representative locations are clustered around this line. Fig. 7(b) shows the error map for sensors that will measure close to GHI. Therefore, the optimal point is found at locations where the horizontal SVF is closest to unity.

Moving on now to the error map for in-plane sensors in Fig. 8, the lowest error values are once again concentrated within a band that diagonally traverses the PV plant from northwest to southeast. The group sensors are placed using error maps for all separate groups, created using the reference irradiance of that group.

The software tool was tested for the Kolindros case without inputting a predetermined number of sensors. The data usage input by the user was prioritized as follows:

1. Daily power loss analysis
2. Forecasting using G_{POA}
3. Forecasting using G_{hor}
4. Connection to satellite data

The resulting monitoring infrastructure, therefore, needed in-plane sensors, a horizontal sensor for measuring G_{hor}, and a horizontal sensor for measuring GHI. Three in-plane sensors were advised because there are two groups and one plane in the PV plant. In conclusion: $N_s = 5$. The in-plane sensors were optimized using the yearly error since PR

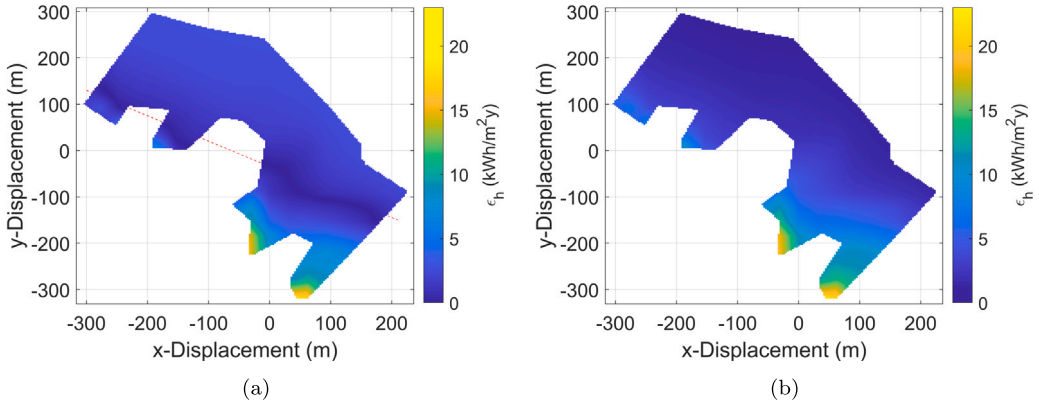


Fig. 7. Surface plots of the hourly error map for the horizontal sensors to be placed for (a) forecasting purposes and (b) satellite connection purposes in the Kolindros case study. The red dashed line in (a) is included to help the eye identifying the band of points with minimal error.

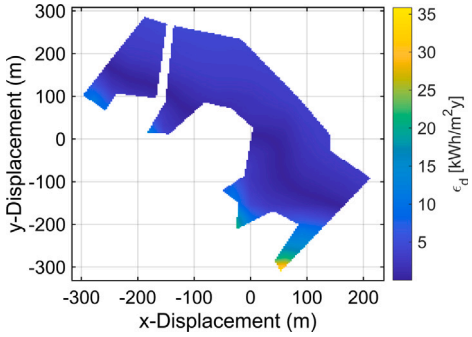


Fig. 8. Surface plots of the daily error map for in-plane sensors in the Kolindros PV plant.

calculations have the highest priority amongst the G_{POA} uses. The locations of the sensors are shown in Fig. 9. These results are in line with what is expected from the error and SVF maps. For instance, the horizontal satellite sensor is positioned where neighboring obstacles are minimal (highest SVF) while the horizontal forecasting sensor is allocated in the middle of the farm, where SVF is a good representative of the whole farm.

The hourly measurement deviations of the sensors (Eq. (4)) are 0.2 Wh/m^{2}y for the G_{hor} sensor, 3.2 Wh/m^{2}y and 4.5e−4 Wh/m^{2}y for the group sensors, and 0.1 Wh/m^{2}y for the plane sensor. This is a negligible measurement deviation which demonstrates the efficacy of the allocation algorithm for these sensors. For the GHI sensor, the measurement deviation is 1.1 kWh/m^{2}y. This is higher than for the other sensors because the SVF at that location is not exactly unity, leading to a deviation of GHI. This measurement deviation, however, is still small compared to the annual incident irradiance of 3.2 MWh/m^{2}y, which is the irradiance measured by a hypothetical horizontal sensor placed on top of the PV farm with no horizon obstacles. The average relative measurement deviation over all allocated sensors is 0.8%. Although the measurement errors throughout the year are small, the instantaneous relative measurement error can be up to three times larger. The largest errors occur at hours with large DHI and low albedo.

When a predefined number of sensors lower than five was input, the same locations for a part of the sensors in Fig. 9 were found:

- $N_s = 4$: same results without the left group sensor (due to this group being the smallest).
- $N_s = 3$: same results without the group sensors.
- $N_s = 2$: in-plane sensor and forecasting sensor (based on priority).
- $N_s = 1$: only the in-plane sensor (based on priority).

For $N_s > 5$, the additional sensors are used as extra group sensors, distributed amongst the two groups. The largest group always has an equal or larger number of sensors than the smaller group. In Fig. 10, $N_s = 8$ was used to observe the effect of the k-means clustering. As expected, the group sensors are well distributed over the module area. The fact that all minima in the error map are clustered around a line leads to all cluster minima, and thus group sensors, are positioned around that line as well.

4.1.2. Solar park Krughütte in Eisleben

For the second case study, since there are two distinct PV module orientations, the SVF maps for the southeast- and south-oriented planes are shown separately in Fig. 11. The SVF pattern results directly from the DSM. The western (left) part of the PV plant is on a hill, up to 50 m higher than the eastern part. This is reflected by a high SVF in the west and a low SVF in the east. The maximum SVF differences seen within planes are 0.02, significantly lower than at the Kolindros plant.

The resulting hourly error maps for horizontal sensors in the Eisleben plant are seen in Fig. 12. The maximum measurement deviation is 8 to 10 kWh/m^{2}y, over two times lower than in the Kolindros case. This results from the lower SVF difference.

The error maps for allocating the southeast- and south-oriented in-plane sensors at the Solar Park Krughütte are shown in Fig. 13. One can observe a high correlation with the SVF maps shown in Fig. 11.

The assumed user input for prioritized data purposes is:

1. Real-time fault detection
2. Forecasting using G_{POA}
3. Forecasting using G_{hor}
4. Connection to satellite data

Due to these data purposes, the resulting monitoring infrastructure needed in-plane sensors, a horizontal sensor for G_{hor}, and a horizontal sensor measuring GHI. Fourteen in-plane sensors were advised since there are twelve groups and two planes in the PV plant. Overall, $N_s = 16$. To have sixteen pyranometers in the PV plant might be unnecessary and too expensive. Some groups can easily be combined into one larger group with only one in-plane sensor. However, the software tool is currently only capable of using panel groups as defined by the user. The in-plane sensors are optimized using the hourly error since the real-time PLA has the highest priority amongst the G_{POA} uses.

In Fig. 14, the advised sensor allocation in Eisleben is visualized. The horizontal satellite sensor is at the maximum SVF location, the far west of the solar park. The horizontal forecasting sensor is in the middle of the plant as expected from Fig. 12(b), and the plane sensors are also in line with the error maps in Fig. 13. Two group sensors appear as one due to the close proximity of their group perimeters.

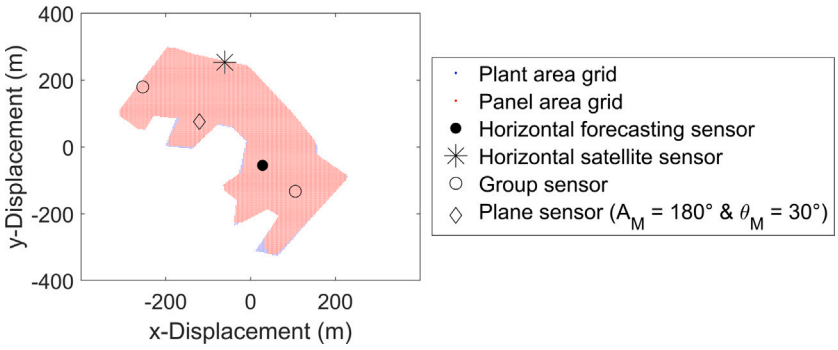


Fig. 9. Allocation of five sensors as advised by the software tool for the Kolindros case study.

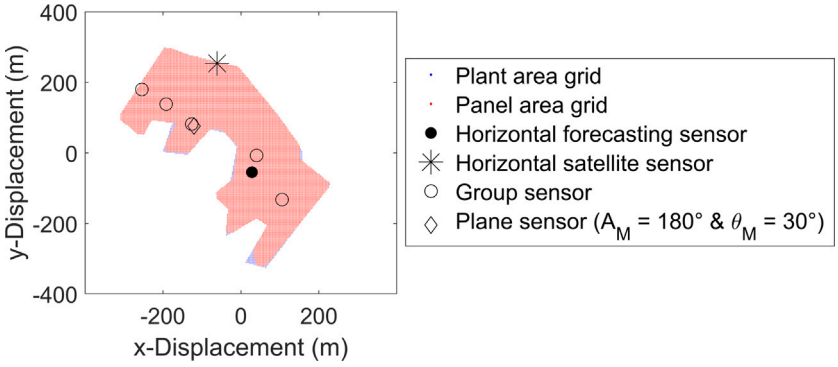


Fig. 10. Map of the allocation of eight sensors for the Kolindros case study.

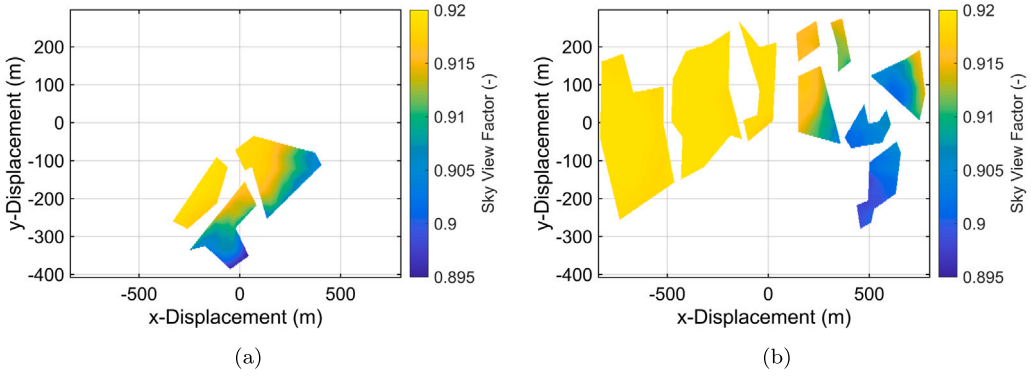


Fig. 11. Surface plots of the SVF map for the (a) southeast-oriented and (b) south-oriented planes of the solar park Krughütte.

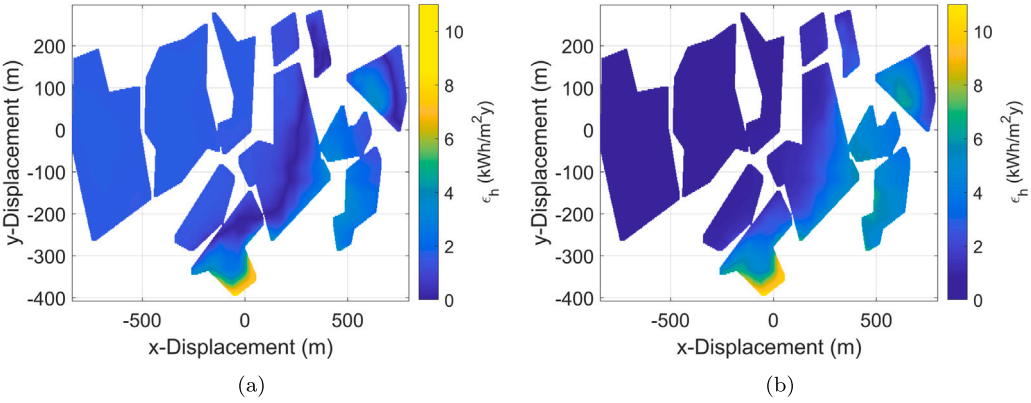


Fig. 12. Surface plots of the hourly error map for horizontal sensors to be placed for (a) forecasting purposes and (b) satellite connection purposes in the Eisleben PV plant.

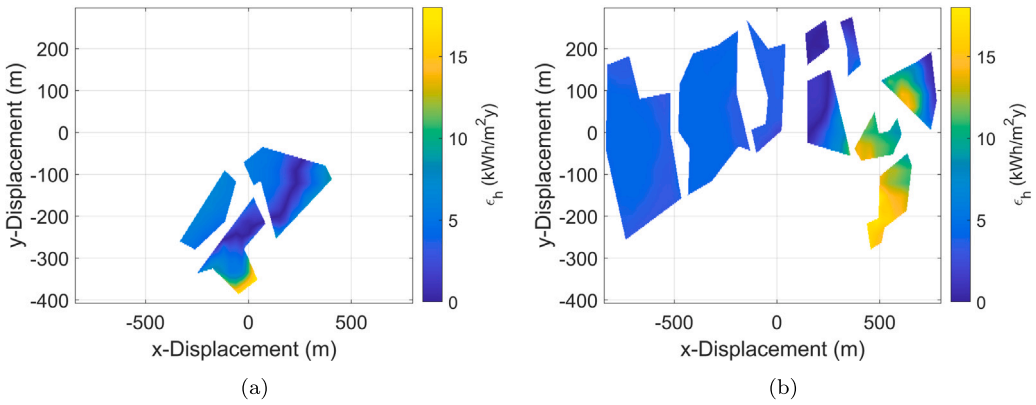


Fig. 13. Surface plots of the hourly error map for in-plane sensors of the (a) southeast-oriented and (b) south-oriented planes in the Eisleben PV plant.

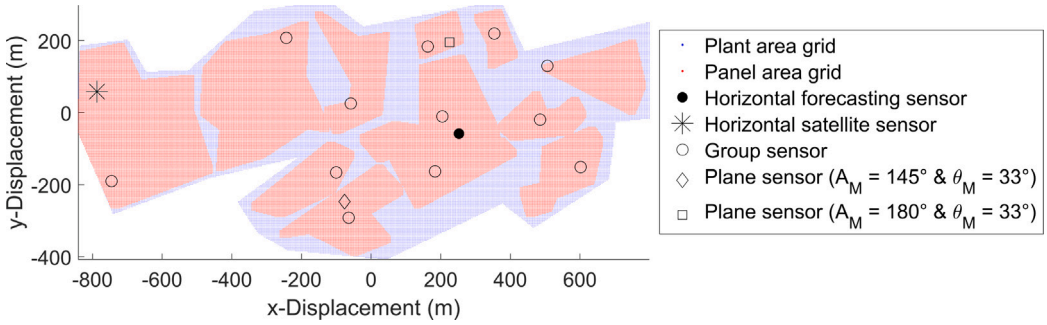


Fig. 14. Allocation of sixteen sensors as advised by the software tool for the Eisleben case study.

The average relative prevented measurement bias of in-plane sensors for this case study is 0.3%. This is lower than the Kolindros case, as expected from the smaller local SVF differences in the Eisleben case.

After analyzing the two case studies, one can compare the results with the IEC 61724 recommendations. Considering the size of the PV farms, their guidelines suggest at least two horizontal irradiance sensors and two in-plane ones. This recommendation is followed for the farm in Kolindros, where the software suggests 5 sensors, but not for the Krughütte farm where the software suggests 16 sensors. As already mentioned, 16 sensors may be unnecessary and financially unfeasible. However, having only 4 may be inaccurate considering that two distinct orientations are present. That can lead to inaccuracies affecting the operations of the PV farm.

4.2. Simplification

The sensor allocation was found to have a high dependency on the SVF map. The horizontal satellite sensor is allocated at the point with the highest SVF, and the other sensors are all allocated at the contour lines of the SVF.

It has been proven mathematically in Appendix C, by exploring the definition of the terms involved, that the minimization problem of all the error functions reduces to Eq. (6). In the equation, $\breve{SVF}$ is the average SVF weighted by the number of modules per group. Thus when the measurement deviations are minimized over all the locations, the difference between the SVF of locations and the weighed SVF is minimized. This equality means that the software tool does not have to distinguish between the time resolution of the errors when minimizing. This explains that all minimum errors are located along a contour line with the value of the weighted average SVF on the map.

$$\arg\min_{l\in L}\left(\epsilon_{h,l}\right)=\arg\min_{l\in L}\left(\epsilon_{d,l}\right)=\arg\min_{l\in L}\left(\epsilon_{y,l}\right)=\min_{l\in L}|\Delta SVF_l| \qquad (6)$$
$$\Delta SVF_l = SVF_l - \breve{SVF}$$

Therefore, the whole allocation algorithm can be done based on solely the SVF maps. For all planes and groups, $\breve{SVF}$ can be calculated, and the location at which the SVF is closest to this value is the most representative. Since the horizontal sensor used for connection to satellite data is placed at the highest SVF, the SVF map is also sufficient here.

These simplifications are based on the assumptions that the GHI, DHI, and DNI time series are location-independent, at least within the geographical scope of the installation. Additionally, the found simplification is based on the assumption that the albedo is spatially constant. This is valid for the current input possibilities of the software tool. Moreover, the simplification is valid only when the angle of incidence amongst all modules is constant. That is true because, in the current version of the software tool, all modules in a group are bound to be equally tilted and oriented. When discontinuously tilted groups are included as input options, the simplification no longer holds, and the G_{POA} and error maps should be included again. Finally, the behind-the-horizon effect is neglected in the current algorithm. This implies that the solar radiation when the sun is below the horizon is ignored, such as during sunrise or sunset. If this is included in the software tool, an extra location-dependent variable is added, making the simplification invalid.

5. Discussion

The reported average relative prevented measurement bias for the two presented case studies is below 1%. This value is lower than the uncertainty of pyranometers, which can be 1.3 to 1.7% in economically relevant hours of the day [4,59]. Therefore, one could question the significance of this work. However, one also needs to consider that the accuracy of the employed data, especially the DSM data, can underestimate the bias. If more accurate terrain data was obtained, for instance from a scanning drone [60], the height differences could be more precisely captured and the detailed shading profiles could be

included in the code. The former may also impact the variation in tilt due to the terrain of the modules, assumed to be constant in the case studies. The latter could be used to distinguish between the irradiance profiles of the modules located at the edges of the plant compared to those at the center. All of these aspects would impact the prevented error.

The geographical scope of this work has been limited to Europe. However, most of the methodology can be extended to other geographical regions by employing databases covering those regions. The only point that requires special attention is the choice of the irradiance decomposition model. Considering the importance of this step in the methodology, an analysis similar to the one performed in Appendix A should be undertaken before extending the software tool to new regions.

Related to this latter point, the accuracy of the newly presented BRL-MvK model should be tested for other locations. It would be interesting to observe whether the seasonal variation of the bias reported in this work is present in regions outside of Europe.

Finally, the software tool should be upgraded to cover the increasingly common bifacial and tracked PV systems. To achieve this, special attention should be given to the estimation of the irradiance.

Out of the two, tracked PV systems may be the simplest to implement. The yearly energy yield should be calculated considering the angles of the tracker instead of the irradiance on a fixed plane. If the tracking mechanism is analytically implemented, this extension can be relatively straightforward. More consideration is required for trackers that adjust the angles in real-time to maximize the overall irradiance harvesting.

Bifacial systems may require an additional sensor to measure the rear irradiance, therefore a sensor facing the ground. The guidelines provided in the IEC 61724 [8] could be used as a starting point when adding this feature to the tool. Additionally, an accurate modeling of the albedo should be included, that is not only time-dependent but also location-dependent. However, the lack of available albedo data may currently hinder this implementation. Efforts in the literature to estimate the albedo from satellite data may help overcome this limitation [61]. If the albedo is finally made location-dependent throughout the farm, the algorithm simplification presented in Section 4.2 would no longer hold.

6. Conclusion

In this work, an irradiance sensor allocation algorithm for Europe has been proposed and implemented in the form of a software tool. The strategy consists of finding representative sensor locations following IEC guidelines. Additionally, the software proposes the optimum number of sensors in a PV farm based on the IEC guidelines and the purpose of data usage and provides metrics to show the benefits of using the developed software. In the process of the software tool creation, several steps were optimized. By employing data from existing farms, the optimum parameters for the sky view factor calculations using digital surface elevation data are found to be a radius of 2 km and an azimuth resolution of 0.3°. Considering the high bias of the irradiance decomposition models of the literature, a custom model has been proposed which is based on the Perez model. The software was applied to two case studies of existing PV farms in hilly terrain. Although the method has been currently developed for PV farms composed of monofacial fixed PV modules with similar orientation (i.e. no East-West configuration), the methodology can be extended to bifacial and/or tracking PV installations.

CRediT authorship contribution statement

Alba Alcañiz: Writing – original draft, Visualization, Software, Methodology. **Mathijs I. van Kouwen:** Visualization, Software, Methodology. **Olindo Isabella:** Writing – review & editing, Supervision. **Hesan Ziar:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research leading to these results has received funding from the Horizon 2020 Program, under Grant Agreement 952957, Trust-PV project. The authors would like to thank Alex Katsikogiannis for his assistance in obtaining the data from the solar park in Greece.

Appendix A. Decomposition model comparison

Since no decomposition model stands out in the literature, a comparison was conducted to determine the most suitable one. The analysis utilized all 12 European meteorological stations from the Baseline Surface Radiation Network (BSRN) [47] (see Fig. 2(a) for their location). The stations provide ground-based measurements of GHI, DHI, and DNI, enabling the comparison of modeled values with measured ones. BSRN combines data from numerous previously independent radiation measurement stations globally. Given the diverse sources, no standardized quality checks are applied. Thus, preliminary quality assessments, including the conversion to hourly resolution, are conducted on all irradiance data. All hourly data points where the solar zenith exceeds 85° are discarded to prevent anomalies that occur at low solar altitude angles. This zenith limit was used for preparing the BSRN data and the irradiance data of the software tool.

This study focuses on six widely used models: Erbs [62], DISC [63], Reindl2 [64,65], BRL [66], Every [67], and Engerer [68,69]. For every station, the hourly time series $\widehat{\mathrm{DHI}}$ was calculated for each model using the GHI of that station, and compared to the BSRN-measured DHI. All these decomposition models use the clearness index to estimate the diffuse fraction, which is related to the time of day. Some models also included other time-dependent variables such as the sun altitude in Reindl2, the zenith angle in DISC, and apparent solar time in BRL, Every, and Engerer. These latter three furthermore included continuity in the model by including daily clearness index and persistence. The performance indicators are the normalized root mean squared error (nRMSE, Eq. (7)) and the normalized mean bias error (nMBE, Eq. (8)).

$$\mathrm{nRMSE} = \frac{H}{\sum_{h=1}^{H} \mathrm{DHI}} \sqrt{\frac{\sum_{h=1}^{H} (\widehat{\mathrm{DHI}} - \mathrm{DHI})^2}{H}} \tag{7}$$

$$\mathrm{nMBE} = \frac{H}{\sum_{h=1}^{H} \mathrm{DHI}} \frac{\sum_{h=1}^{H} (\widehat{\mathrm{DHI}} - \mathrm{DHI})}{H} \tag{8}$$

where H is the number of hourly time instants. The results were analyzed to identify the best decomposition model for hourly irradiance data in Europe, see Fig. 15.

The nRMSE is notably higher for the BUD and SON stations compared to the others. This can be attributed to the relatively poor data quality of these stations. The BRL model outperforms all other models in all stations except for the SON station, which had shown data limitations due to the high altitude. The simple Erbs model outperforms the bivariate models Reindl2 and DISC, the two least accurate models. Therefore, increasing the complexity of Erbs, resulting in Reindl2 and DISC, did not improve its performance. The Every and Engerer models are adaptations of the BRL model. However, their nRMSE is higher than the original model. For the Every model, the decreased performance could be due to the model being parameterized in Australia. The Engerer model also showed decreased performance compared to the BRL model, possibly due to the model being designed for minute data originally.

Regarding the nMBE of the analysis, no model outperforms the others at all stations. However, on average, the Engerer model performs

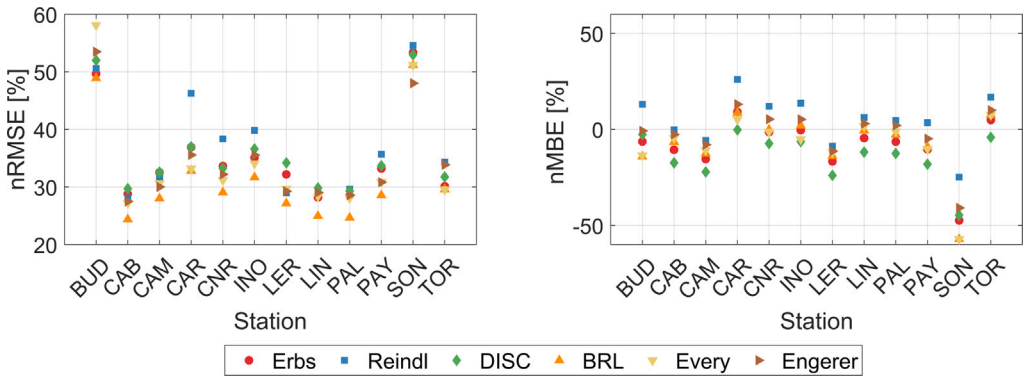


Fig. 15. nRMSE and nMBE of six decomposition models, shown for the twelve different datasets retrieved using the BSRN database. Refer to Fig. 2(a) for the location of the BSRN station employed.

best, with an average absolute nMBE of 8.9% compared to the worst-performing 14% for the DISC model and close to 11% for the remaining models. Nevertheless, Engerer outperforms BRL in only seven out of twelve stations. In contrast to the nRMSE results, Every and Engerer models slightly surpass the preceding BRL model. Furthermore, the Erbs outperforms its succeeding models DISC and Reindl2. All models highly underestimate the diffuse fraction at the SON station, which could be again due to the extremely high altitude of the station. This raises a general concern for using decomposition models at high altitudes.

The performance analysis concludes that the Engerer model performs best based on average nMBE. However, the BRL model is considered the overall best model, consequently outperforming other models based on nRMSE and performing averagely based on the nMBE. In this study, these results are assumed to apply to satellite irradiance data. However, a note should be made that this conclusion is only robust when hourly ground-based input irradiance is used.

In addition to examining annual metrics, seasonal variations were investigated. The daily errors of all stations are averaged, leading to an average nRMSE or nMBE for each day. When exploring the nRMSE results (of which no figure is included), most models have their highest accuracy in late autumn and early winter. However, no clear pattern has been found since the Reindl2 and DISC models show different seasonal behavior, and the similarities in the BRL, Every, and Engerer model patterns probably emerge from their relation.

More interesting results are found when exploring the seasonal pattern in nMBE shown in Fig. 16. All models overpredict the diffuse component in autumn and underpredict it in spring. This bias pattern also emerges when the same analysis is done at a station level (not shown). It is hard to explain this behavior since most seasonal fluctuations have extremes in winter and summer. Despite conducting a comprehensive analysis of the seasonality of variables, the observed nMBE pattern could not be explained.

While the origin of this seasonal bias remains unknown, compensating for it can improve existing models. A correction factor can be applied to a model with high nMBE to create a much better fit. However, this approach does not apply to a model with a high nRMSE. Based on the results, opting for the BRL model is recommended. To compensate for the seasonal bias, a multiplication factor is introduced to the BRL model:

$$\widehat{\mathrm{DHI}}_{improved} = \widehat{\mathrm{DHI}} \frac{1}{\mathrm{nMBE}_{fit}} \tag{9}$$

By simply including the multiplication factor, the diffuse fraction k_d (ratio of irradiance that is diffusely reflected) may exceed unity. This is mitigated within the new model by capping all values of k_d at 1.

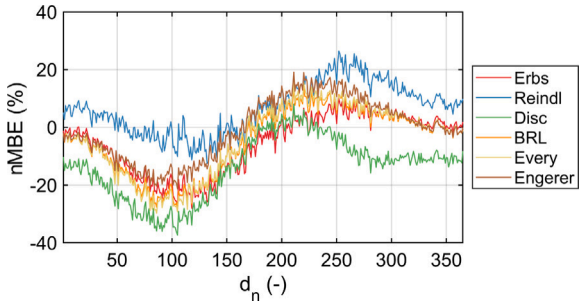


Fig. 16. Average daily nMBE of six decomposition models for a dataset containing irradiance data of twelve European BSRN stations. A seasonal bias pattern occurs with negative bias in spring and positive bias in autumn.

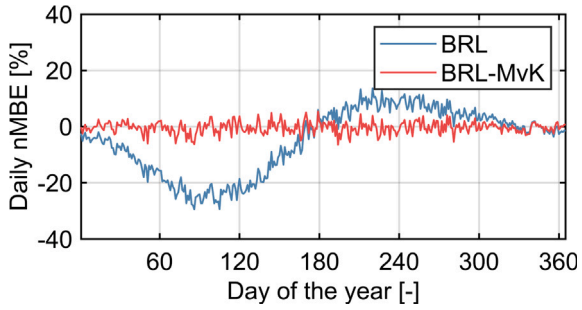


Fig. 17. Daily nMBE averaged for all the BSRN stations of the BRL and BRL-MvK model throughout a year.

The daily average nMBE curve for the BRL model was used to determine the multiplication factor. A Fourier expansion with two terms and a period of 365 days was used.

$$\mathrm{nMBE}_{fit} = a_0 + a_1 \cos(D_n) + b_1 \sin(D_n) + a_2 \cos(2 \cdot D_n) + b_2 \sin(2 \cdot D_n)$$
$$D_n = \frac{d_n \cdot 2\pi}{365} \tag{10}$$

With $a_0 = 1.037$, $a_1 = 0.02052$, $a_2 = -0.02688$, $b_1 = 0.1069$, $b_2 = -0.002202$, and d_n the day number in a calendar year. Using Eq. (10), the seasonal nMBE pattern is mitigated using the compensating factor as seen in Fig. 17. The new bias-compensated model is named the BRL-MvK model.

The improvement of the model is also observed with the annual nMBE of each station. The annual nMBE is shifted towards a more positive bias for all stations (of which no figure is included), decreasing

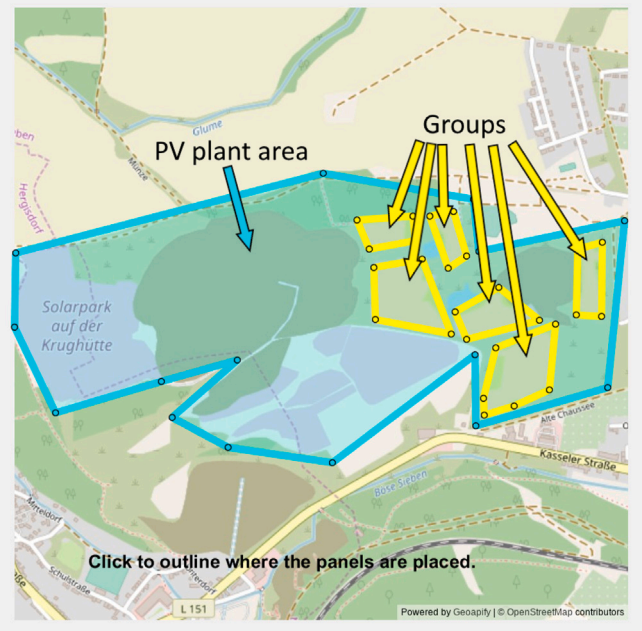


Fig. 18. GUI where the user is asked to draw the PV group areas (yellow) for the Solar Park Krughütte in Eisleben, Germany. The already drawn blue area is the PV plant area. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

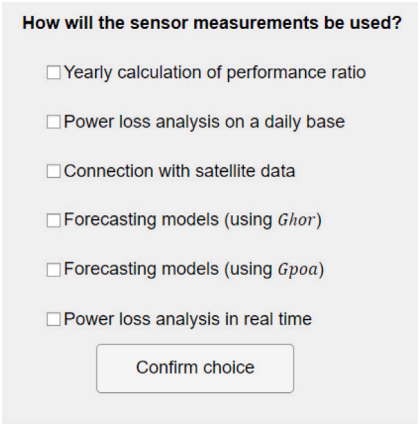


Fig. 19. GUI element where the user inputs the purpose of the data for future use.

Table B.1
Runtime of the different steps of the software tool for the two case studies.

Runtime [min]	Kolindros	Eisleben
Distribute sensors	0.0	0.0
Import data	0.4	0.4
Calculate SVF	4.5	19.9
Create irradiance maps	0.0	0.0
Create error maps	0.1	0.5
Suggest locations	0.0	0.0
Create report	0.6	1.0
Total	5.6	21.8

the inter-station average of the absolute value of the annual nMBE values from 11.2% to 10.6%. This shift results in improved model performance for most (formerly) negative-biased stations and a decreased performance for all formerly positive-biased stations. Kindly note that Fig. 17 refers to the daily biased error averaged for all BSRN stations while 10.6% refers to the mean absolute of the nMBE of the 12 stations. The nRMSE of the BRL-MvK model outperforms the BRL model in all stations and the mean is improved from 31.8% to 30.5%.

Appendix B. Software tool overview

The software tool initiates by prompting the user to input location parameters:

- What is the central longitude [°] and latitude [°] of the PV plant (with 4 decimal points of accuracy)?
- A square image is downloaded. How many km should the sides be to include the full PV plant area?
- What is the capacity of the PV plant?

After which a prompt is generated confirming the input location of the user using a map tile retrieved via Geoapify [70].

When the location is deemed correct, the perimeters of the PV plant and the PV group area are defined. The user is asked to draw polygons, as seen in Fig. 18. The blue area indicates the PV plant area and the yellow areas are the PV groups area, which is provided per group. Before being prompted to draw the areas, the user is asked whether all PV modules in the plant are uniformly tilted and oriented. If so, the following questions are asked:

- What is the azimuth orientation (deg) of the PV panels (E = 90°, S = 180°, W = 270°)?
- What is the tilt (deg) of the PV panels (flat = 0°)?

If the modules are not uniformly tilted and oriented, the two latter questions are asked for every group after drawing it.

Before defining the yellow group areas, the user is asked if all groups have an equal number of solar panels. If so, the N_{gp} of each group is set to equal values. Otherwise, they are asked to input the number of modules after each drawn PV group.

Afterward, the user is prompted to input the data purposes using the GUI element shown in Fig. 19. If more goals are checked in the GUI element, the user is asked to input the relative priority of each purpose.

The last question is about the number of sensors to be allocated. If the user does not input a predefined number of sensors, advice is given.

During the calculations, the user is updated on the progress and the undertaken steps. The code is profiled to give insight into how time-intensive the software tool is and what the speed-determining modules are. Table B.1 provides the time profiling for the two case studies presented in Section 4.1. For the two case studies presented, the software takes around 5.6 min for the 6 MW_p solar park in Kolindros (Greece), and 21.8 min for the 29.1 MW_p solar park in Eisleben (Germany). Although in general the higher the capacity, the longer the computational time, the capacity is not the sole indicator of computational time demand. The total runtime of the software tool mostly depends on the SVF calculations, for which the higher the local height differences, the more time-intensive the computations become. The creation of the output report and the data import are also time-intensive steps.

Appendix C. Algorithm simplification

The allocation algorithm is based on the minimization problem in Eq. (1). Independently of the time resolution of the error, the measurement deviations take the absolute difference between $G_{POA,l}$ and G_{ref}. For clarity, the definition of $G_{POA,l}$ is shown in Eq. (11).

$$G_{POA,l} = SVF_l \cdot R_d \cdot DHI + (1 - SVF_l) \cdot \alpha \cdot GHI + DNI \cdot \cos(\gamma) \quad (11)$$

where γ is the angle of incidence onto the PV modules. $G_{POA,l}$, R_d, DHI, DNI, GHI, and α are time series but the time dependency has been removed from this equation for ease of understanding.

The reference irradiance is the weighted average of the G_{POA} of a certain area, as shown in Eq. (3). Since the reference irradiance is

location-independent, it can be rewritten including the definition of $G_{POA,l}$:

$$G_{ref} = \breve{SVF} \cdot R_d \cdot DHI + (1 - \breve{SVF}) \cdot \alpha \cdot GHI + DNI \cdot \cos(\gamma) \tag{12}$$

where $\breve{SVF}$ is the weighted average of the SVF of all grid points inside the PV group for which the sensor measures irradiance, Eq. (13).

$$\breve{SVF} = \frac{1}{\sum_{n=1}^{N_{gp}} N_n \cdot \sum_{n=1}^{N_{gp}} L_n} \sum_{n=1}^{N_{gp}} \sum_{l=1}^{L_n} SVF_l \cdot N_n \tag{13}$$

The simplification of G_{ref} only holds because DHI, DNI, GHI, R_d, α, and γ are all (assumed to be) spatially independent. Using this new notation of G_{ref}, the following derivation can be done:

$$\begin{aligned} |G_{POA,l} - G_{ref}| &= |SVF_l \cdot R_d \cdot DHI + (1 - SVF_l) \cdot \alpha \cdot GHI \\ &\quad - \breve{SVF} \cdot R_d \cdot DHI - (1 - \breve{SVF}) \cdot \alpha \cdot GHI| \end{aligned} \tag{14}$$

$$= |(SVF_l - \breve{SVF}) \cdot (R_d \cdot DHI - \alpha \cdot GHI)| \tag{15}$$

Eq. (15) holds for sums over all data points of in-plane and reference irradiance time series of length T. The minimization problem for the three time resolutions of the measurement deviation can be rewritten using $\Delta SVF_l = SVF_l - \breve{SVF}$:

$$\begin{aligned} \arg\min_{l \in L}(\epsilon_{h,l}) &= \arg\min_{l \in L} \sum_{h=1}^{H} |\Delta SVF_l \cdot [R_d(h) \cdot DHI(h) \\ &\quad - \alpha(h) \cdot GHI(h)]| \cdot \frac{H'}{H} \end{aligned} \tag{16}$$

$$\begin{aligned} \arg\min_{l \in L}(\epsilon_{d,l}) &= \arg\min_{l \in L} \sum_{d=1}^{D} |\Delta SVF_l \sum_{hd=1}^{24} [R_d(hd) \cdot DHI(hd) \\ &\quad - \alpha(hd) \cdot GHI(hd)]| \cdot \frac{H'}{24 \cdot D} \end{aligned} \tag{17}$$

$$\begin{aligned} \arg\min_{l \in L}(\epsilon_{y,l}) &= \arg\min_{l \in L} \sum_{y=1}^{Y} |\Delta SVF_l \sum_{hy=1}^{8760} [R_d(hy) \cdot DHI(hy) \\ &\quad - \alpha(hy) \cdot GHI(hy)]| \cdot \frac{1}{Y} \end{aligned} \tag{18}$$

where $h \in H$, $d \in D$, and $y \in Y$ are the hour, day, and year of each time step, respectively. The last term normalizes to a yearly value, with $H' = 8760$ being the hours per year.

All terms that are both inside the absolute signs and location independent cancel out in the minimization problem. The fractions at the end of Eqs. (16), (17) and (18) cancel out because they are necessarily a positive constant. Therefore, all minimization problems are reduced to Eq. (6).

This simplification does not imply that the error maps are equal. Differences occur when the time series inside the absolute brackets from Eq. (16)–(18) have both negative and positive terms. This leads to measurement deviations canceling out when different periods are chosen to minimize bias. However, that condition rarely occurs, being met in only 0.2% and 0.14% of the data points in the Kolindros and Eisleben case studies, respectively. This simplification has been implemented for these two case studies, and the difference in results is negligibly small.

Appendix D. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.solener.2024.113139. We provide additional resources to enhance the understanding and applicability of the techniques presented in this article. A comprehensive video demonstrates the functionality and usage of the developed software tool. Additionally, the complete source code used to develop the software tool is made available with the objective of promoting collaboration and continuation of this work.

References

[1] N.H. Reich, B. Mueller, A. Armbruster, W.G. Van Sark, K. Kiefer, C. Reise, Performance ratio revisited: is PR> 90% realistic? Prog. Photovolt., Res. Appl. 20 (6) (2012) 717–726.

[2] M. Rivera, C. Reise, Silicon sensors vs. Pyranometers–review of deviations and conversion of measured values, in: 37th European PV Solar Energy Conference and Exhibition, Vol. 7, 2020, p. 11.

[3] M. Mussard, H.N. Riise, M.S. Wiig, S. Rønneberg, S.E. Foss, Influence of temperature coefficient and sensor choice on PV system performance, IEEE J. Photovolt. (2023).

[4] M. Korevaar, D. Nitzel, Simulation of POA front irradiance sensor mounting position, in: European PVPMC 8, 2023.

[5] K. Lambers, Asset manager at energy21, 2023, Private communication.

[6] J. Hommel, Bakker en co, 2023, Private communication.

[7] H. Buikema, Asset manager at Powerfield Netherlands B.V., 2023, Private communication.

[8] IEC, Photovoltaic System Performance - Part 1: Monitoring (IEC 61724-1:2021), Standard, International Electrotechnical Commission, 2021, URL https://webstore.iec.ch/publication/65561.

[9] M. Pau, N. Locci, C. Muscas, A tool to define the position and the number of irradiance sensors in large PV plants, in: 2014 IEEE International Energy Conference, ENERGYCON, IEEE, 2014, pp. 374–379.

[10] G. Oviedo, BayWa r.e., 2023, Private communication.

[11] L.L. Jiang, R. Srivatsan, D.L. Maskell, Optimal irradiance sensor placement for photovoltaic systems using mutual information based greedy algorithm in Gaussian process, in: 2017 IEEE 44th Photovoltaic Specialist Conference, PVSC, IEEE, 2017, pp. 120–126.

[12] A. Di Piazza, M.C. Di Piazza, G. Vitale, A kriging-based partial shading analysis in a large photovoltaic field for energy forecast, in: International Conference on Renewable Energies and Power Quality, ICREPQ'09 Valencia, Spain, 2009.

[13] M. Ernst, A. Thomson, I. Haedrich, A. Blakers, Comparison of ground-based and satellite-based irradiance data for photovoltaic yield estimation, Energy Procedia 92 (2016) 546–553.

[14] A. Harsarapama, M. Korevaar, O. Isabella, M. Zeman, Accuracy of irradiance measurement for a PV park versus the number of sensors, in: Optics for Solar Energy, Optica Publishing Group, 2017, RTh2B–2.

[15] A. Miller, B. Lumby, Utility scale solar power plants: a guide for developers and investors, in: Guidelines Book Written for IFC, World Bank Group, 2012.

[16] R. Perez, P. Ineichen, R. Seals, J. Michalsky, R. Stewart, Modeling daylight availability and irradiance components from direct and global irradiance, Solar Energy 44 (5) (1990) 271–289.

[17] M. Keijzer, A Multi-Surface Reflected Irradiance Model for Pyranometer Corrections and PV Yield Calculations in Complex Urban Geometries, Delft University of Technology, the Netherlands, 2019, URL http://resolver.tudelft.nl/uuid:d1a6baf0-7721-441e-bc2c-3a05bdaea89d.

[18] D. de Jong, H. Ziar, Photovoltaic potential of the dutch inland shipping fleet: An experimentally validated method to simulate the power series from vessel-integrated photovoltaics, Solar RRL 7 (8) (2023) 2200642.

[19] Global Energy Observatory, Google, KTH Royal Institute of Technology in Stockholm, Enipedia, World Resources Institute, Global power plant database, 2018, URL https://earthengine.google.com/, Published on Resource Watch and Google Earth Engine; http://resourcewatch.org/.

[20] Sinergise Laboratory for geographical information systems, Sentinel-Hub, 2023, URL https://www.sentinel-hub.com/.

[21] A. Smets, K. Jäger, O. Isabella, R. Van Swaaij, M. Zeman, Solar Energy: The Physics and Engineering of Photovoltaic Conversion, Technologies and Systems, Bloomsbury Publishing, 2016.

[22] D. Yang, Solar radiation on inclined surfaces: Corrections and benchmarks, Sol. Energy 136 (2016) 288–302.

[23] D. Yang, Estimating 1-min beam and diffuse irradiance from the global irradiance: A review and an extensive worldwide comparison of latest separation models at 126 stations, Renew. Sustain. Energy Rev. 159 (2022) 112195.

[24] L. Laiti, L. Giovannini, D. Zardi, G. Belluardo, D. Moser, Estimating hourly beam and diffuse solar radiation in an alpine valley: A critical assessment of decomposition models, Atmosphere 9 (4) (2018) 117.

[25] C. Bertrand, G. Vanderveken, M. Journée, Evaluation of decomposition models of various complexity to estimate the direct solar irradiance over Belgium, Renew. Energy 74 (2015) 618–626.

[26] S. Dervishi, A. Mahdavi, Comparison of models for the derivation of diffuse fraction of global irradiance data for Vienna, Austria, in: Proceedings of the Building Simulation, 2011.

[27] J.D. Mondol, Y.G. Yohanis, B. Norton, Solar radiation modelling for the simulation of photovoltaic systems, Renew. Energy 33 (5) (2008) 1109–1120.

[28] F. Lanini, Division of Global Radiation into Direct Radiation and Diffuse Radiation, University of Bern, 2010.

[29] A. Pérez-Burgos, M. Díez-Mediavilla, C. Alonso-Tristán, M. Dieste-Velasco, Retrieval of monthly average hourly values of direct and diffuse solar irradiance from measurements of global radiation in Spain, J. Renew. Sustain. Energy 10 (2) (2018).

[30] S. Mishra, Selection Map for PV Module Installation Based on Shading Tolerability and Temperature Coefficient, Delft University of Technology, 2018.
[31] E. Paulescu, R. Blaga, Regression models for hourly diffuse solar radiation, Sol. Energy 125 (2016) 111–124.
[32] D.C. Martinsen, M.B. Øgaard, M.M. Nygârd, M.S. Wiig, M. Di Sabatino, E.S. Marstein, H.N. Riise, Poster: Evaluation of irradiance and transposition models for a nordic climate, in: 8th World Conference on Photovoltaic Energy Conversion, 2022.
[33] C. Demain, M. Journée, C. Bertrand, Evaluation of different models to estimate the global solar radiation on inclined surfaces, Renew. Energy 50 (2013) 710–721.
[34] G. Notton, P. Poggi, C. Cristofari, Predicting hourly solar irradiations on inclined surfaces based on the horizontal measurements: Performances of the association of well-known mathematical models, Energy Convers. Manage. 47 (13–14) (2006) 1816–1829.
[35] M. Horváth, T. Csoknyai, Evaluation of solar energy calculation methods for 45 inclined, south facing surface, Energy Procedia 78 (2015) 465–470.
[36] M. Cucumo, A. De Rosa, V. Ferraro, D. Kaliakatsos, V. Marinelli, Experimental testing of models for the estimation of hourly solar radiation on vertical surfaces at Arcavacata di Rende, Sol. Energy 81 (5) (2007) 692–695.
[37] A. Padovan, D. Del Col, Measurement and modeling of solar irradiance components on horizontal and tilted planes, Sol. Energy 84 (12) (2010) 2068–2084.
[38] J.D. Mondol, Y.G. Yohanis, B. Norton, Solar radiation modelling for the simulation of photovoltaic systems, Renew. Energy 33 (5) (2008) 1109–1120.
[39] A. Jamodkar, Energy Yield Prediction of Solar Powered E-bike Charging Station, Delft University of Technology, 2017, URL https://repository.tudelft.nl/islandora/object/uuid:2abcd917-7298-4ae5-8c0a-392bb31a1ef2/datastream/OBJ/download.
[40] C. Vasar, O. Prostean, G. Prostean, Evaluating solar radiation on a tilted surfaces-a study case in Timis (Romania), in: IOP Conference Series: Materials Science and Engineering, vol. 106, IOP Publishing, 2016, 012026.
[41] R. Posadillo, R.L. Luque, Evaluation of the performance of three diffuse hourly irradiation models on tilted surfaces according to the utilizability concept, Energy Convers. Manage. 50 (9) (2009) 2324–2330.
[42] P.G. Loutzenhiser, H. Manz, C. Felsmann, P. Strachan, T. Frank, G. Maxwell, Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation, Sol. Energy 81 (2) (2007) 254–267.
[43] P. Ineichen, Global Irradiance on Tilted and Oriented Planes: Model Validations, University of Geneva, 2011.
[44] R. Perez, R. Seals, P. Ineichen, R. Stewart, D. Menicucci, A new simplified version of the perez diffuse irradiance model for tilted surfaces, Solar Energy 39 (3) (1987) 221–231.
[45] S. Lloyd, Least squares quantization in PCM, IEEE Trans. Inf. Theory 28 (2) (1982) 129–137.
[46] T. Huld, R. Müller, A. Gambardella, A new solar radiation database for estimating PV performance in Europe and Africa, Solar Energy 86 (6) (2012) 1803–1815.
[47] A. Driemel, J. Augustine, K. Behrens, S. Colle, C. Cox, E. Cuevas-Agulló, F.M. Denn, T. Duprat, M. Fukuda, H. Grobe, et al., Baseline surface radiation network (BSRN): structure and data description (1992–2017), Earth Syst. Sci. Data 10 (3) (2018) 1491–1501.
[48] F. Yang, K. Mitchell, Y.-T. Hou, Y. Dai, X. Zeng, Z. Wang, X.-Z. Liang, Dependence of land surface albedo on solar zenith angle: Observations and model parameterization, J. Appl. Meteorol. Climatol. 47 (11) (2008) 2963–2982.
[49] R.E. Dickinson, Land surface processes and climate—Surface albedos and energy balance, in: Advances in Geophysics, vol. 25, Elsevier, 1983, pp. 305–353.
[50] H. Ziar, F.F. Sönmez, O. Isabella, M. Zeman, A comprehensive albedo model for solar energy applications: Geometric spectral albedo, Appl. Energy 255 (2019) 113867.
[51] W. Lucht, C.B. Schaaf, A.H. Strahler, An algorithm for the retrieval of albedo from space using semiempirical BRDF models, IEEE Trans. Geosci. Remote Sens. 38 (2) (2000) 977–998.
[52] Y. Qu, Q. Liu, S. Liang, L. Wang, N. Liu, S. Liu, Direct-estimation algorithm for mapping daily land-surface broadband albedo from MODIS data, IEEE Trans. Geosci. Remote Sens. 52 (2) (2013) 907–919.
[53] S. Marathe, Effect of Urban Development on Local Albedo, Delft University of Technology, 2022, URL https://repository.tudelft.nl/islandora/object/uuid%3A7c13032f-1758-4cf1-b9df-ec823ab5b3f7.
[54] A.H. Sparks, nasapower: a NASA POWER global meteorology, surface solar energy and climatology data client for R, J. Open Source Softw. 3 (30) (2018) 1035.
[55] X. Lin, J. Wen, Q. Liu, D. You, S. Wu, D. Hao, Q. Xiao, Z. Zhang, Z. Zhang, Spatiotemporal variability of land surface albedo over the Tibet Plateau from 2001 to 2019, Remote Sens. 12 (7) (2020) 1188.
[56] E. Palle, P. Goode, P. Montañés-Rodríguez, A. Shumko, B. Gonzalez-Merino, C.M. Lombilla, F. Jimenez-Ibarra, S. Shumko, E. Sanroma, A. Hulist, et al., Earth's albedo variations 1998–2014 as measured from ground-based earthshine observations, Geophys. Res. Lett. 43 (9) (2016) 4531–4538.
[57] Y.-M. Saint-Drenan, L. Wald, T. Ranchin, L. Dubus, A. Troccoli, An approach for the estimation of the aggregated photovoltaic power generated in several European countries from meteorological data, Adv. Sci. Res. 15 (2018) 51–62.
[58] Google maps, 2023, URL https://www.google.com/maps. (Accessed: 22 April 2023).
[59] Kipp & Zonen, Suncertainty: the Kipp & Zonen pyranometer uncertainty calculator, 2016, URL https://www.kippzonen.com/News/743/Suncertainty-the-Kipp-Zonen-pyranometer-uncertainty-calculator. Visited on 4/7/2024.
[60] I. Horvath, A. Alcañiz, F. Segata, J. Lemmens, T. Hall, W. Vanheusden, Automated PV Digital Twin-Based Yield Simulation Framework, Tech. Rep., Trust-PV, 2022.
[61] C. Schaaf, Z. Wang, MCD43c3 MODIS/Terra+ aqua BRDF/Albedo albedo daily L3 global 0.05 deg CMG. NASA EOSDIS land processes DAAC, 2015.
[62] D. Erbs, S. Klein, J. Duffie, Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation, Solar Energy 28 (4) (1982) 293–302.
[63] E.L. Maxwell, A Quasi-Physical Model for Converting Hourly Global Horizontal to Direct Normal Insolation, Tech. Rep., Solar Energy Research Inst., Golden, CO (USA), 1987.
[64] D.T. Reindl, W.A. Beckman, J.A. Duffie, Diffuse fraction correlations, Solar Energy 45 (1) (1990) 1–7.
[65] D.T. Reindl, W.A. Beckman, J.A. Duffie, Evaluation of hourly tilted surface radiation models, Solar Energy 45 (1) (1990) 9–17.
[66] B. Ridley, J. Boland, P. Lauret, Modelling of diffuse solar fraction with multiple predictors, Renew. Energy 35 (2) (2010) 478–483.
[67] J.P. Every, L. Li, D.G. Dorrell, Köppen-Geiger climate classification adjustment of the BRL diffuse irradiation model for Australian locations, Renew. Energy 147 (2020) 2453–2469.
[68] N. Engerer, Minute resolution estimates of the diffuse fraction of global irradiance for southeastern Australia, Sol. Energy 116 (2015) 215–237.
[69] J.M. Bright, N.A. Engerer, Engerer2: Global re-parameterisation, update, and validation of an irradiance separation model at different temporal resolutions, J. Renew. Sustain. Energy 11 (3) (2019).
[70] Geoapify GmbH, Geoapify, URL https://www.geoapify.com/.