Reply to anonymous Referee #1

We would like to thank the anonymous Referee #1 for his/her detailed and constructive review. The comments and suggestions were very helpful and significantly contributed to improving the clarity, robustness, and overall quality of the manuscript.

In the following, analytical replies are provided to each of the reviewer's comments. Reviewer's comments are written in bold font. Line numbers, when provided, refer to the new version with track changes.

The manuscript "PV power modelling using solar radiation from ground-based measurements and CAMS: Assessing the diffuse component related uncertainties leveraging the Global Solar Energy Estimator (GSEE)" by Nikolaos Papadimitriou et al. focuses on the exploitation of the impact of the partitioning of the global horizontal irradiance (GHI) in its direct and diffuse components on the PV power production by simulations with the widely used Global Solar Energy Estimator (GSEE) model. The solar irradiance, air temperature, aerosol optical depth input sources are the BSRN and AERONET measurements from 5 sites in Europe, North Atlantic Ocean and Sahara desert and the CAMS model. The diffuse fraction (DF), being rarely measured except in a limited number of sites, is estimated within the GSEE through the logistic Boland-Ridley-Lauret (BRL) model, based on the clearness index as the main parameter. The effects of different cloud cover and aerosol optical depth on DF, and the corresponding impact on the simulation of power production is evaluated for fixed and 2-axis tracking PV systems based on c-Si and CdTe technology. The effects on different timescales are also explored. Finally, an assessment of the financial impacts deriving from evaluating the DF of desert dust from the BRL model compared to the measured GHI components, for a hypothetic PV solar farm around Tamanrasset is provided.

The main conclusion is that the best agreement in DF estimation with BRL model and ground-based measurements is for cloud-free and very load to moderate aerosol, i.e. for the simplest atmospheric conditions to be modelled. The worst situation for a reliable power production estimation is for partially cloudy skies. In sites impacted by high dust load, the BRL underestimated the DF and the power generation is overestimated.

We would like to thank the reviewer for his/her time and comments.

General comments

While it is clear that the GSEE is widely used and the BRL model is optimized for both the Northern and Southern Hemispheres, I suspect it is not the best model for analyzing the effects of aerosols on the GHI partition, as the clearness index is mostly influenced by cloudiness and the reader is not informed about how the aerosols are accounted for. The authors should address this aspect, which is a key point in the development of the work.

as well as discuss similar works, if any, dealing with this topic.

Despite the large amount of calculations done in this work, the results are not valued by an adequate discussion, both in qualitative and quantitative terms. As a general comment I suggest quantifying the results in the text (Results and Conclusions sections) and not leaving the values only in the supplementary material.

The authors do not cite in the Introduction any previous paper dealing with the estimation of PV power generation under different cloudiness and/or aerosol load conditions, nor compare any of their results with previous works. If a similar work is not found in literature, this aspect, which increases the importance of this study, should be emphasized both in the introduction and in the conclusions.

Overall, I recommend a major revision of the key points before publication.

Reply

We acknowledge the limitations of the BRL model, especially regarding the effects of aerosols on GHI partition. However, it constitutes a key submodule of the GSEE library, as it is integrated within the internal processing chain of the climate data interface, a feature that plays a central role in the applicability of GSEE to climate-driven assessments. Some text has been added in the Introduction to clarify this (lines 84-86)

Moreover, following an extensive review of the relevant literature indicates that, despite the large number of studies discussing the role of aerosols and clouds on the amount of and the distribution of the solar irradiance that reaches the Earth's surface (Fountoulakis et al., 2021; Kosmopoulos et al., 2018; Papachristopoulou et al., 2022; Amiridis et al., 2024; Kosmopoulos et al., 2017; Calastrini et al., 2024; Kouklaki et al., 2023), there is a lack of studies addressing reliability of PV power simulations under diverse atmospheric conditions due to inaccuracies in the representation of the diffuse component in PV power models. Furthermore, we were not able to find any study examining the reliability of CAMS radiation data for PV power potential assessments. Some relevant text has been added both the Introduction and the Conclusions (lines 98-102 and lines 586-589)

In addition, the manuscript has been revised to include a more detailed quantitative discussion of the evaluation is Sections 3.3 – 3.5, and some tables have been transferred from the Supplement to the main text.

REFERENCES

Amiridis, V., Kazadzis, S., Gkikas, A., Voudouri, K. A., Kouklaki, D., Koukouli, M.-E., Garane, K., Georgoulias, A. K., Solomos, S., Varlas, G., Kampouri, A., Founda, D., Psiloglou, B. E., Katsafados, P., Papachristopoulou, K., Fountoulakis, I., Raptis, P.-I., Georgiou, T., Gialitaki, A., ... Zerefos, C. (2024). Natural aerosols, gaseous precursors and their impacts in Greece: A review from the remote sensing perspective. Atmosphere, 15(7), 753. https://doi.org/10.3390/atmos15070753

Calastrini, F., Messeri, G., & Orlandi, A. (2024). Long-range mineral dust transport events in Mediterranean countries. Air, 2(4), 444–467. https://doi.org/10.3390/air2040026

Kosmopoulos, P. G., Kazadzis, S., Taylor, M., Athanasopoulou, E., Speyer, O., Raptis, P. I., Marinou, E., Proestakis, E., Solomos, S., Gerasopoulos, E., Amiridis, V., Bais, A., & Kontoes, C. (2017). Dust impact on surface solar irradiance assessed with model simulations, satellite observations and ground-based measurements. Atmospheric Measurement Techniques, 10(7), 2435–2453. https://doi.org/10.5194/amt-10-2435-2017

Kosmopoulos, P., Kazadzis, S., El-Askary, H., Taylor, M., Gkikas, A., Proestakis, E., Kontoes, C., & El-Khayat, M. (2018). Earth-Observation-based estimation and forecasting of particulate matter impact on solar energy in Egypt. *Remote Sensing*, *10*(12), 1870. https://doi.org/10.3390/rs10121870

Kouklaki, D., Kazadzis, S., Raptis, I.-P., Papachristopoulou, K., Fountoulakis, I., & Eleftheratos, K. (2023). Photovoltaic spectral responsivity and efficiency under different aerosol conditions. *Energies*, *16*(18), 6644. https://doi.org/10.3390/en16186644

Papachristopoulou, K., Fountoulakis, I., Gkikas, A., Kosmopoulos, P. G., Nastos, P. T., Hatzaki, M., & Kazadzis, S. (2022). 15-year analysis of direct effects of total and dust aerosols in solar radiation/energy over the Mediterranean Basin. Remote Sensing, 14(7), 1535. https://doi.org/10.3390/rs14071535

Specific comments

Lines 67-72: since the use of an empirical model constitutes a relevant part of the work, I suggest to detail a little the description of these models and in particular the

description of the BRL model, for example by mentioning here the variables that are used to derive the DF.

Reply

Some text added in order to highlight the innovative formulation besides the BRL model (lines 70-81)

Line 71: some information about the BRL model should be provided, since the model is widely used in the paper. Is this the only model for DF estimation incorporated in the GSEE?

Reply

GSEE includes only the BRL model. Although the user in its single usage (modelling a PV plant with hourly time-series data in a specified project location) has the option to use other diffuse fraction models (such as those included in pvlib), the climate data interface tool is designed for deploying exclusively the BRL implementation.

Some text added to clarify this (lines 127-131).

Table 3 with the libRadtran input parameters is somehow unclear. The SZA input is "with step 90°": what do the authors mean? In addition, is the wavelength dependence of the surface albedo, SSA and gg accounted for? Finally, I suggest "integrated water vapor" instead of "water vapor".

Reply

The reference to SZA input was a transcription error that was introduced during the manuscript revision. As the BRL model requires hourly input data at exact hourly timestamps for at least one full day, we used the libRadtran option to set as input datetime accompanied with the coordinates instead of SZA directly. After, in the analysis we computed the corresponding SZA values. In section 3.2, we stated that we chose the summer solstice as a representative day with sufficient number daylight hours.

Regarding the wavelength dependence of SSA and gg is not accounted for in the present analysis. In the present, this dependence is not explicitly accounted for, as the objective is not a fully spectrally resolved radiative transfer analysis, but rather to investigate the differences associated with some representative SSA values.

We accepted also the suggestion for adding the word Integrated before Water Vapor.

Lines 290-295 and Figure 2: the comparison of the DF from the BSRN measurements and the BRL model is tricky. Do the differences for SZA>60° increase because of the difficulty of the model in estimating the DF for high SZAs? The bottom of Figure 2 shows that this happens for cloud-free conditions above 70° (not 60° SZA as said in the text) and not for all sites: Izana do not show the SZA dependence. It appears also for the cleanest conditions, i.e. for AOD500≤0.05.

In my opinion this deserves a little bit more investigation, instead of simply limiting the comparison to SZAs below 60°, as in Figure 3.

Reply

We agree that the shift becomes observable within the range of 60-70 degrees. However, the specific point at which SZA starts to affect cannot be identified with precision. Therefore, a value of 60 degrees adopted as reference limit for practical reasons related to solar energy production applications, as above this value the low absolute irradiance levels contribute less to total energy yield. Moreover, we investigated this in section 3.2 in the sensitivity, where under clear sky conditions the shift arising closer to 70 degrees.

In Izana, there is the influence of altitude, where the levels of the diffuse irradiance are significantly lower.

The difficulty of the model in estimating the diffuse fraction for high SZA values may arise from the symmetry of a typical daily profile of the diffuse fraction and the hourly clearness, as the model requires full-day input data for hourly clearness. This is confirmed also by the sensitivity analysis, as the model has similar behavior.

Some clarifying text has been added.

Line 326: under overcast conditions the BRL DF takes a range of values, approximately from 0.6 to 1, while the BSRN DF is close to 1. This means that even for homogeneous sky conditions, isotropic radiation the BRL model is not capable of providing reliable DF estimates. The authors could also refer to the 3D variability of cloud properties, whose effect cannot be accounted for by a model like BRL.

Reply

The vast majority of cases where the BRL diffuse fraction is below 0.8 while the observed is close to 1 correspond to periods involving rapid transitions between partly cloudy and overcast skies, occurring either during the hour itself or immediately before or after it. These discrepancies can be mainly attributed to limitations of the DNI-based characterization methodology for cloudiness.

Moreover, the presence of aerosols can amplify these discrepancies. Even under cloudy conditions, aerosol may coexist and contribute, despite the difficulty to be measured and quantified.

Furthermore, a limited number of cases identified during intense dust events at Tamanrasset and Izana, where the reduction of DNI was so pronounced that the applied DNI-based criterion classified these conditions as overcast.

However, we did not further investigate these inconsistencies, as the energy production levels during such periods are very low.

Some clarifying text has been added (lines 327-337).

Section 3.2: the dependence on latitude and altitude is not discussed, although presented in Figure 5.

The results show that for totally scattering aerosols (SSA=1) the BRL model underestimates the DF, while for absorbing aerosols (SSA=0.7) is overestimates the DF.

The authors may briefly discuss the results of the sensitivity analysis in terms of how the BRL logistic model treats the aerosol-radiation interactions. The model does not explicitly include the aerosol optical properties, but incorporates their effects into the clearness index, together with those exerted by the clouds, which are by no means larger. In this section, instead, the authors examine in detail how aerosols influence the partitioning of GHI into direct and diffuse components, and the results obtained with the BRL model are strongly biased, as expected. This is also a consequence of how the model was conceived, in particular of the data with which the relationship between the DF and the geometric, meteorological and atmospheric variables were determined.

Reply

Some clarifying text added to discuss the effect of altitude and altitude, and to emphasize on the range of the discrepancies as an outcome of the effect of aerosol with different optical properties.

Section 3.3: the authors should comment the results presenting the quantification of the differences in power production derived from using only GHI or GHI and DHI as GSEE model input. Not all sites and all atmospheric conditions have to be considered, but at least for two sites with different characteristics, such as Izana and Lindenberg, for fixed and 2-axis tracking systems.

Reply

Some text has been added, as well as Tables with the computed indices for Carpentras and Tamanrasset (as representative locations) have been transferred from the supplement to the manuscript.

Lines 492-494: the data gaps are in the GHI data and/or in the irradiance components? Can the author suggest a reason for these gaps? Does using a less stringent condition on the number of days per month (for example 15 days) allow for a less fragmented annual evolution?

Reply

The data gaps occur mainly close to the solar noon during the summer months and arise from removing data through the quality control checks applied in the BSRN. Rejection of the data during the BSRN QC procedure is possibly related to operational issues at the station. Even though the data gaps are in most cases less than 2-3 hours, they may affect the BRL performance throughout these days. Thus, these days have been removed from the analysis.

Lines 495-501: I find it very useful to present the differences also in percentage, referred to the energy production obtained using GHI and DHI as a reference.

Reply

A Figure showing the percentage differences added in the supplement. Some discussion relative to the new figure has been added in the manuscript (lines 499-501).

Section 3.5: more details about the CAMS data selection are needed. For example: the authors say that the CAMS solar radiation data are adapted to the investigated sites, but how is this done? By interpolation, by considering the nearest CAMS grid point? Moreover, the Izana site is excluded from the analysis because the altitude of the site is

not directly comparable to the model vertical grid. For this reason, it is useful to know the CAMS 3D spatial grid and to add it in Section 2.3.

Another missing point is the quantification of the irradiance differences on power production, a qualitative discussion is not sufficient.

Finally, the authors should refer to previous papers, if any, dealing with the use of CAMS irradiance data for modelling PV potential power production.

Reply

For the data selection we used the "CAMS solar radiation time-series" product, where the user sets as input the coordinates as well as the altitude of a specific location, and then the output data are offered in ASCII format. Therefore, the interpolation methodology is integrated to the CAMS product algorithm. Some text added to clarify this in section 2.3.

Some text has also been added to discuss quantitively the validation as well as to highlight the novelty of this work, as we could not find studies directly examining CAMS performance for PV power modelling.

Conclusions: should summarize the main results and report some numerical data, which otherwise are only relegated in the supplementary material. For example, the results may be evaluated for tow or three sites with different characteristics in terms of latitude and aerosol regimes. I suggest reporting the impact on the power generation not only as absolute values, but also as percent, to facilitate understanding. In addition, even if the results pertaining to the panels with CdTe technology are not different from those obtained for c-Si panels, they should be briefly recalled.

When discussing the assessment of the financial impacts of the desert dust effects at Tamanrasset, the authors should clearly address that the analysis considers the differences in diffuse function derived from the measurements and calculated by the model. The sentence in lines 474-476 "Therefore, site assessments that do not account for the impact of desert dust aerosols may overestimate financial performance…." may be misleading, as it may be interpreted as the assumption that desert dust is not accounted for in the model simulations. I suggest reformulating the sentence.

Reply

We tried to improve the conclusions section as recommended by the reviewer.

Technical corrections

We thank the reviewer for the detailed comments and suggestions. All minor editorial, typographical, and wording-related comments (e.g., punctuation, wording clarifications, terminology corrections) have been carefully addressed and incorporated into the revised manuscript.

Responses to the comments requiring further clarification are provided below.

Line 142: can the authors quantify how much is the uncertainty on using a fixed surface albedo of 0.3?

Reply

The default GSEE value of 0.3 can be considered as representative for most types of surfaces (e.g., the surface albedo for urban landscapes, desert, and green grass is usually between 0.2 and 0.4). There are of course darker surfaces (e.g., forests, asphalt).

In the context of past studies (see reply to reviewer 1 in Papachristopoulou et al., 2024: https://doi.org/10.5194/amt-2023-110-AC1), we have investigated the sensitivity of the GHI irradiance to the surface albedo under various conditions (see the Figures below). For SZAs below 75° and under clear sky conditions, changing surface albedo by 10% changes the GHI by less than 1% (i.e., the difference is within the uncertainty range of the ground-based measurements). Under cloudy conditions the % differences are larger (i.e., ~ 5% for a 10% change in surface albedo for Cloud Optical Thickness of 12), but under such conditions the amount of PV power potential is small. Since very high surface albedo values are rare at the latitudes where the study is focused, and differences from the default value are generally smaller than 0.2, we decided to use the default surface albedo value. The manuscript has been also updated with this information (see lines 142-145)

Line 159: is this reference correctly cited?

Reply

Yes, the reference is correctly cited. It corresponds to an online source from the official GSEE model website, which has been cited accordingly in the revised manuscript.

Table 1: is it necessary? In my opinion the text description is exhaustive.

Reply

We have retained Table 1, as it provides a concise overview of the input parameters used in the GSEE model for the purposes of this study, complementing the textual description and improving clarity for the reader.

Line 198: is it "daily" or "hourly"?

Reply

It is daily, as this refers to the pre-processing of the data that used as input to the climate interface

Lines 256-257: I suggest to include a sentence on how the solar radiation is estimated in CAMS and the 3D spatial resolution, an important information for the CAMS data selection operated for the analysis described in Section 3.5.

Reply

Additional text has been added to clarify that the CAMS solar radiation time-series product is used. The description now briefly outlines how the data are retrieved, including the use of location coordinates and station altitude as inputs (lines 261-263).

Line 367: maybe "ground altitude"?

Reply

This was a typographical error, which has been corrected in the revised manuscript. The correct value is 0.1 km, rather than "0,1".

Line 598: express the power overestimation also in percent number and the site where this is observed (should be Tamanrasset). Is the 49.2 Wh/kWp/hour on hourly value? I could not find this number in the supplementary material tables.

Reply

This issue was caused by a typographical error and the inadvertent omission of several rows from the table, which have now been corrected in the revised manuscript. The value of 49.2

Wh/kWp per hour refers to an hourly estimate for the Tamanrasset site and is now correctly reported in Table 4. The relative mean bias error (rMBE) is expressed percentage terms.

Line 613: Table 4 is not present.

Reply

Table 4 had been removed during a previous revision and transferred to the Supplementary Material; however, the reference in the main text was not updated accordingly. In the revised manuscript, Table 4 has been reinstated in the main text and the reference has been corrected.

Figure 6: bottom graphs. Why the data produced with only GHI for Izana seem to be cut around 900 Wh/kW$_p$ per hour?

Reply

The PV systems considered in this study have a nominal capacity of 1 kWp. The PV model applies a default system loss factor of 10%. This effectively limits the maximum achievable power output to approximately 90% of the nominal capacity (i.e., around 900 W/kWp). This effect becomes apparent at the Izaña site due to its low latitude combined with its specific geographical and atmospheric conditions, which lead to high irradiance levels. As a result, the simulated PV output appears capped around 900 Wh/kWp per hour when only GHI is used.

Table S2: the metrics for high and very high aerosol load are missing.

Reply

The issue was due to missing rows in the previous version of Table S2. The table has been corrected and is now included as Table 4 in the revised manuscript, containing all metrics for low, high, and very high aerosol load conditions.

<u>Reply to anonymous Referee #2</u>

We would like to thank the anonymous Referee #2 for the careful evaluation of the manuscript and the helpful comment and suggestions. The reviewer's remarks were valuable in improving the presentation, consistency, and technical clarity of the study, and helped us to address several points that required further clarification.

Detailed responses to the reviewer's comments are provided below. The reviewer's comments are reported in bold font, and line numbers, when provided, refer to the revised manuscript with track changes enabled.

This manuscript presents a comprehensive assessment of the impact of uncertainty in the diffuse component of solar radiation on the prediction of photovoltaic energy production using the Global Solar Energy Estimator (GSEE) and different data sources. This analysis includes categorization based on sky and aerosol conditions. The manuscript is well-organized and well-written, and the topic is worthy of study. However, there are some aspects that still need clarification or improvement.

We would like to thank the reviewer for his/her time and comments.

The specific comments are as follows:

1. Throughout the document, the existence of a reflected component of solar radiation on a tilted surface is not explicitly mentioned. It is understood that the authors include this component in the "diffuse" radiation, although, usually, in much literature, diffuse (sky) and reflected components are treated separately. Furthermore, the document does not specify which transposition model is used to determine solar irradiance on the inclined surface (in the same way that the separation model included in GSEE for estimating the diffuse fraction is mentioned). Several transposition models exist that treat diffuse sky irradiance in different ways, from the simplest, which considers it isotropic, to the more elaborate ones that separate the sky into different regions (background diffuse, circumsolar region, horizon brightening). This is important because fixed systems with different tilt angles and two-axis tracking systems are studied in different locations, and specifying which transposition model was used will allow for a better discussion and understanding of the results.

Reply

The GSEE package includes the submodule “trigon”, which contains a set of functions for computing the in-plane irradiance. These functions are based on trigonometric formulations, that account of the surface albedo, thereby including the ground-reflected component of solar radiation.

The main submodule of the GSEE library used to simulate the electric output of a PV panel requires as input the separated irradiance data (GHI, and diffuse fraction). Subsequently, it internally calls the trigon transposition model to compute the plane-of-array irradiance, which is then used to estimate the PV power output.

In contrast to BRL submodule – where apart from the climate data interface, alternative diffuse fraction models can be selected by the user in single-site application – the transposition model in GSEE is fixed and cannot be modified by the user.

Some text has been added (lines 133-137)

2. Line 193 states that the data were "resampled to hourly”, but this is unclear and needs clarification. Is an hourly average calculated? Refer also to Line 260 where hourly values are mentioned. The BRL model is based on irradiance values integrated over a one-hour period (Ridley et al., 2010). Please clarify. Regarding the AERONET data, how were the values treated given that the raw data does not have a defined periodicity?

Reply

The resampling to hourly resolutions refers to the calculation of hourly mean values. The implementation of the BRL model included within the GSEE requires full-days timeseries of the hourly clearness index. For this purpose, the mean GHI of each hour is divided with the TOA on horizontal plane, which calculated as follows: Solar_Constant (Day_of_Year) * cos(SZA), where SZA is evaluated at the midpoint of the hour.

Regarding the AERONET data, although the original measurements do not have a fixed temporal periodicity, the data are resampled by computing hourly mean values from all available observations within each hour. This procedure ensures temporal consistency with the hourly irradiance and BRL model inputs.

The calculation of hourly mean values is clarified in the revised text.

3. In Line 211: How the values of $0 < SV < 1$ for partly cloudy conditions are determined? This is not clear. Also, if using a value of SV between 0 and 1 to

characterize the intra-hour conditions then the value of irradiance must be hourly mean values (see also comment 2). Please clarify this aspect and justify the use of SV compared to the use of other indicators, such as the clear-sky index (not the clearness index), defined as the ratio of actual irradiance to irradiance for clear sky conditions from a suitable model, for example. This also relates to what is said in Line 290, where the use of cameras is suggested to overcome the issue of small and scattered clouds within the sky dome that enhances the diffuse component while not blocking the direct normal irradiance. There are other effective methods for identifying sky conditions. Please comment and clarify.

Reply

Representing the effect of cloudiness was challenging, as it requires the deployment of several observations. However, the DNI-based formulation aims to provide an indicative measure of the intra-hour cloudiness conditions. Alternative approaches, such as the clear-sky index or the Cloud Modification Factor require estimates of the clear sky GHI, which also introduce uncertainty. Some text added in lines 202 to 208.

4. In Line 220, regarding the AERONET data, please clarify "... data were resampled at hourly intervals ..." in view of comment 2 above.

Reply

Some clarifying text has been added.

5. In lines 234 and 247, there appear to be some typos (an extra "c" and "my", respectively).

Reply

All the typos have been removed.

6. In Figure 3: the represented data is only for SZA < 60, correct? Please, confirm this and mention it in the figure caption.

Reply

Yes, we confirm that the figures is for SZA < 60. The figure caption revised to include this information.

7. In Figure 8 and associated text, the reason why using a 30-days moving average is not clear. Is a centered moving average used? Please clarify, also regarding the data gaps which, up to this point in the document, were not evident (may be including in Table 2 this information will help).

Reply

A centered 30-day moving average is used in Figure 8. The purpose of applying the moving average is to reduce short-term variability and highlight the underlying temporal behavior of the analyzed quantities, facilitating a clearer comparison of trends.

The data gaps occur mainly around solar noon and arise from measurements removed during the application of the BSRN quality control checks. Although these gaps are, in most cases, shorter than 2–3 hours, they may affect the BRL model performance for the corresponding days. This information has now been clarified in the revised text (lines 189-192).

PV power modelling using solar radiation from ground-based measurements and CAMS: Assessing the diffuse component related uncertainties leveraging the Global Solar Energy Estimator (GSEE)

Nikolaos Papadimitriou[1,2], Ilias Fountoulakis[2], Antonis Gkikas[2], Kyriakoula Papachristopoulou[3], Andreas Kazantzidis[1], Stelios Kazadzis[3], Stefan Pfenninger[4], John Kapsomenakis[2], Kostas Eleftheratos[5,6], Athanassios A. Argiriou[1], Lionel Doppler[7], and Christos S. Zerefos[2,6,8,9]

[1]Department of Physics, University of Patras, 26500 Patras, Greece

[2]Research Centre for Atmospheric Physics and Climatology, Academy of Athens, 11521 Athens, Greece

[3]Physikalish-Meteorologisches Observatorium Davos, World Radiation Center (PMOD/WRC), 7260 Davos, Switzerland

[4]Faculty of Technology, Policy, and Management (TPM), Delft University of Technology, 2628 BX Delft, the Netherlands

[5]Department of Geology and Geoenvironment, National and Kapodistrian University of Athens, 15784 Athens, Greece

[6]Biomedical Research Foundation, Academy of Athens, 11527 Athens, Greece

[7]Deutscher Wetterdienst, Meteorologisches Observatorium Lindenberg – Richard Assman Observatorium (DWD, MOL-RAO), 15848 Lindenberg (Tauche), Germany

[8]Navarino Environmental Observatory (N.E.O.), 24001 Messinia, Greece

[9]Mariolopoulos-Kanaginis Foundation for the Environmental Sciences, 10675 Athens, Greece

Corresponding author: Nikolaos Papadimitriou (npapadimitriou@academyofathens.gr, Vasilissis Sofias 79, 11521 Athens, Greece)

Abstract

Accurate PV power production modelling requires precise knowledge of the distribution of solar irradiance among its direct and diffuse components. Since this information is rarely available, this

requirement can be addressed through the use of diffuse fraction models. In this study, we try to quantify the errors in PV modelling when measurements of the diffuse solar irradiance are not available. For this purpose, we use total and diffuse solar irradiance data obtained from ground-based measurements of BSRN to simulate the PV electric output using GSEE. We have chosen five sites in Europe and North Africa, with different prevailing conditions, where BSRN measurements are available. GSEE incorporates an implementation of the Boland-Ridley-Lauret (BRL) diffuse fraction model, along with a Climate Data Interface that enables simulations across different time scales. We evaluate the capability of BRL in providing accurate estimations of the diffuse fraction under diverse atmospheric conditions, with particular attention on the presence of clouds and aerosols and assess the extent to which its associated errors propagate to energy production modelling. Furthermore, we compare GSEE outputs when using CAMS radiation time-series as input instead of ground-based measurements, to quantify the impact of the CAMS radiation product uncertainties in PV modelling.

Keywords

Solar energy modelling; CAMS radiation; PV power modelling; aerosol; dust; solar radiation

1. Introduction

Decarbonizing the power sector in a sustainable manner is pivotal in the effort to mitigate climate change (Edenhofer et al., 2011; Owusu & Asumadu-Sarkodie, 2016; IPCC, 2023) and the large-scale deployment of Solar Energy offers significant prospects toward this objective (Kakran et al., 2024). The available solar energy is a variable source, fluctuating across different timescales with a unique solar-resource profile over individual locations (McMahan et al., 2013). Therefore, accurate solar energy forecasting and resource assessment is crucial for minimizing the risk in selecting project location, designing the appropriate solar-energy conversion technology, and integrating new sources of solar based power generation into the electricity grid (Stoffel, 2013), while short-term, intra-hour forecasts are critical for power plant operations, grid-balancing, real-time unit dispatching, automatic generation control, and trading (Pedro et al., 2017).

~~For practical reasons, it is critical to extend~~ Extending solar irradiance forecasting to ~~encompass methods linked to solar-based power generation.~~derive PV power forecasts is essential in solar energy applications. PV power modelling can be ~~derived~~achieved through the following additional steps to solar irradiance forecasting: (i) decomposing Global Horizontal Irradiance (GHI) into Diffuse

Horizontal Irradiance (DHI) and Direct Normal Irradiance (DNI~~)~~); (ii) calculating the plane-of-array irradiance incident on the surface of PV planes, whether static or mounted on a solar tracking system, and (iii) simulating the PV power production primarily based on the in-plane irradiance (Blanc et al., 2017).

The scarcity of concurrent measurements of both solar irradiance components, coupled with the complexity of their theoretical computation, has driven the development of numerous empirical models for estimating the diffuse fraction (ratio of the diffuse-to-global solar radiation). A seminal contribution in this area was made by Liu and Jordan (1960), who established a correlation between the diffuse fraction and the clearness or cloudiness index (ratio of the global-to-extraterrestrial radiation). These models predominantly rely on the clearness index as the principal predictor. They are generally classified into single-predictor models and multi-predictor models, with the latter incorporating additional astronomical variables for enhanced precision (Paulescu & Blaga, 2019). Typically, these models are expressed as polynomial equations, ranging from the 1st to the 4th degree, that link the diffuse fraction to the clearness index ~~(Jacovides et al., 2006).~~ $DF = f(clearness\ index, * params)$ (Jacovides et al., 2006). Boland et al. (2001) proposed the use of a logistic function instead of linear or simple nonlinear functions of the clearness index. Ridley et al. (2010) developed a multiple-predictor logistic model, known as the Boland-Ridley-Lauret (BRL), which combines simplicity and reliable performance across both the Northern and Southern Hemispheres. The BRL model extends Boland’s approach by adopting the hourly clearness index as the principal predictor and introducing the following additional parameters: apparent solar time, daily clearness index, solar altitude, and a measure of the persistence of global radiation level. In the implementation of the BRL included in the GSEE, the users set as input only the hourly clearness. Moreover, this implementation adopts the updated parameters proposed by Lauret et al. (2013), which derived using data from nine worldwide locations covering a variety of climates and environments across Europe, Africa, Australia and Asia. While the existing models consider all-sky conditions, in solar energy modelling it is critical to focus on cloud-free skies, where energy production is maximized. Under such conditions, aerosols become the primary parameter influencing the distribution of solar irradiance among its components. (e.g., Blaga et al., 2024). Specifically, the BRL model accounts for aerosols indirectly through the clearness index, which is indicative of the overall atmospheric attenuation of solar radiation.

~~Regions~~In regions dominated by abundant sunshine, such as the Mediterranean and Middle East, which are favorable for solar based power generation, the attenuation of solar irradiance is strongly influenced by aerosols, and particularly desert dust aerosols. Several studies highlighted the impact of desert dust aerosol in the downwelling solar irradiance and the energy production in these regions (Fountoulakis et al., 2021; Papachristopoulou et al., 2022~~).~~; Kosmopoulos et al., 2018; Kouklaki et al., 2023). The significance of considering the effect of aerosols in short-term solar irradiance forecasting and nowcasting is emphasized by Kazantzidis et al. (2017), Raptis et al. (2023) and Papachristopoulou et al. (2024).

The Global Solar Energy Estimator (GSEE; Pfenninger & Staffell, 2016) is a widely used open access model for simulating PV power output, designed for rapid calculations and ease of use~~,~~. It comes with an implementation of the BRL diffuse fraction model (Ridley et al., 2010; Lauret et al., 2013).

While PV power modelling is essential for linking solar resources to energy production, the existing literature does not adequately address its reliability under diverse atmospheric conditions. To the best of our knowledge, the existing literature does not include studies that explicitly address the uncertainties in PV energy production modeling associated with the partitioning of solar radiation into its direct and diffuse components at the model input. In this study, we supply GSEE with input data from ground-based measurements as well as from the Copernicus Atmospheric Monitoring Service (CAMS), aiming to investigate differences in PV power output simulations, which arise from providing only GHI as input radiation data. At the outset, we focus on evaluating the reliability of BRL under diverse atmospheric conditions, with particular attention to the dependence of its accuracy on the presence of clouds and aerosols. To further explore this, we conduct a sensitivity analysis using radiative transfer model (RTM) simulations under cloud-free skies. Following these analyses, we assess the extent to which the associated uncertainties in the estimation of the diffuse fraction spread to the power generation over hourly intervals. This step involves simulating PV plants with varying configurations.

GSEE is also effective for analyzing trends and variability in solar based power generation through its climate interface submodule (e.g., Hou et al., 2021~~).~~), where the BRL model is integrated within the internal processing chain The accuracy of the climate interface in estimating the total daily PV power output is also evaluated in this study.

2. Data and Methodology

2.1 Global Solar Energy Estimator (GSEE)

The modelling of the PV power output is conducted using the version 0.3.1 of GSEE (Pfenninger & Staffell, 2016). The model features functions for simulating a complete PV system, incorporating characteristics and specifications such as location, installed capacity, technology, tracking (fixed, 1-axis, 2-axis), tilt angle, and orientation.

The user provides as input time-series data of solar radiation, and optionally, ambient air temperature and surface albedo. Specifically, the model requires GHI and, when available, the Diffuse Fraction. If the diffuse component is not provided, the provided implementation of the BRL diffuse fraction model (Ridley et al., 2010; Lauret et al., 2013) is employed to estimate it, relying only on time-series of the hourly clearness index and the geographical coordinates. While in the single-site application of the GSEE model with hourly time resolution the user has the option to adjust the input and select alternative diffuse fraction models implemented by external libraries, e.g., pvlib (Anderson et al., 2013), the climate data interface automatically invokes the BRL model as part of the internal processing workflow. GSEE utilizes the provided information for the distribution of the irradiance components and applies trigonometric calculations to determine the total solar irradiance incident on the panel's inclined plane. More precisely, for the plane-of -array irradiance calculation a GSEE includes the submodule "trigon" (transposition model), which is based on trigonometric formulations, that account of the surface albedo, thereby including the ground-reflected component of solar radiation. However, the transposition model is integrated within the GSEE internal algorithms, so it cannot be modified by the user.

After solar irradiance the most significant parameter regarding energy production is air temperature (e.g., Dubey et al., 2013). If temperature is not provided by the user, the model assumes a default value of 20 °C. In this study, temperature was used as input only in the simulations with BSRN data, as it is provided alongside ~~actinometric~~radiation measurements. A surface albedo value of 0.3 considered by default from the model, introduces some uncertainty in our simulations~~.~~, which however is estimated to be small. Under cloudless conditions, a 10% difference in surface albedo changes the GHI by ~1% for SZA < 75°. Differences are larger under cloudy conditions (~ 10% difference in GHI for a 10% difference in surface albedo). Nevertheless, surface albedo at the selected sites is generally low and relatively invariant throughout the year (even at the most northern site of Lindenberg there is only a limited number of days with increased surface albedo due to snow cover).

The available options for the panel type are crystalline silicon (c-Si) and Cadmium Telluride (CdTe), where the power output is modeled based on the relative PV performance model described by Huld et al. (2010). For fixed panels, a built-in latitude dependent function for the optimal tilt is also included.

Moreover, GSEE includes a Climate Data Interface submodule that enables the processing of gridded climate datasets, with varying temporal resolutions, ranging from hourly to annual. Within the context of this submodule, the use of BRL serves as part of the resampling and upsampling processes applied to input climate datasets with daily resolution. For processing data with lower-than-daily resolutions, it incorporates the use of Probability Density Functions (PDFs), which describe the probability with which a day with a certain amount of radiation occurs within a month (Renewables Ninja, n.d.). This methodology accounts for the non-linear distribution of mean monthly radiation across individual days, ensuring a more representative temporal disaggregation. The processes applied to the mean daily irradiance are described in detail in Section 3.4.

For the purposes of this study, we simulated solar plants with capacity of 1 kWp, and for both available technologies. The simulations with c-Si technology, considered as default by the model, are presented ~~detailed~~ in detail the following sections. The results of the simulations with CdTe technology are provided in the supplement, and are not thoroughly discussed, since they are very similar to the results for the c-Si technology. Regarding the mounting approach, the solar plants were either static and oriented to the south or equipped with a 2-axis solar tracking system. In the case of fixed panels, we selected the optimal tilt angle relying on the latitude dependent built-in function.

The input parameters defining the characteristics of the simulated PV plants are summarized in Table 1.

Table 1. Input parameters defining the characteristics of the simulated PV plants

Capacity	**Mounting Approach**		**Technology**	
1 kWp	Fixed	2-axis tracking	c-Si	CdTe

	Orientation: south	Tilt Angle: f(latitude) built-in function for optimal tilt			

2.2 Ground-based measurements

We supplied GSEE with ground-based irradiance as well as ambient temperature measurements collected from five stations of the Baseline Surface Radiation Network (BSRN; Driemel et al., 2018). Moreover, information about aerosols was retrieved from co-located stations of the Aerosol Robotic Network (AERONET; Holben et al., 1998; Dubovik et al., 2000).

Information for the stations utilized for this study is summarized in Table 2, and their geographical location is depicted in Figure 1.

Table 2. Detailed information about the location of the ground-based stations used in this study.

STATION	Latitude [° N]	Longitude [° E]	Elevation [m]
Carpentras (CAR)	44.08	5.06	100
Cener (CNR)	42.82	-1.60	471
Izaña (IZA)	28.31	-16.50	2373
Lindenberg (LIN)	52.21	14.12	125
Tamanrasset (TAM)	22.79	5.53	1385

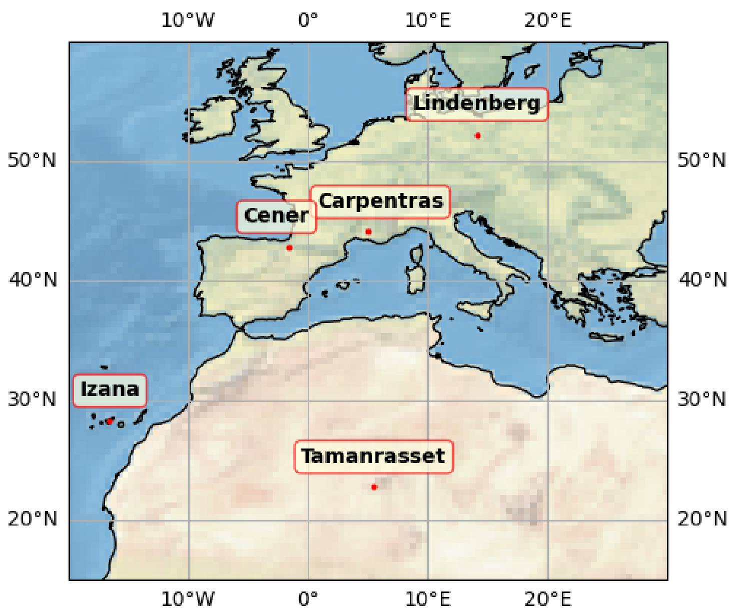

Figure 1. Locations of the BSRN and co-located AERONET stations that are used in the current study

BSRN station-to-archive files were accessed and manipulated using the SolarData v1.1 R package (Yang, 2019), and the BSRN-recommended quality check (QC) tests (Long & Dutton, 2010) applied to the collected data. Some data gaps arose due to measurements removed during the QC procedure. Although these data gaps are, in most cases, shorter than 2-3 hours, they may affect the BRL performance throughout the corresponding days. Consequently, days affected by such data gaps excluded from the analysis. We retrieved data for 2017, with 1-minute temporal resolution. We used GHI, DHI, and Temperature as inputs to the GSEE model. Initially, the data were resampled to hourly~~, and~~ and mean hourly values of GHI and DHI are calculated. Then, the simulations were conducted using either GHI and DHI, or only GHI along with the deployment of BRL. The input to BRL consists of hourly clearness index, derived by dividing GHI measurements with the solar radiation

Formatted: Font color: Text 1

incident on a horizontal plane at the Top of the Atmosphere (TOA) above the examined location. Subsequently, the 1-min timeseries resampled also to a daily resolution and transformed into three-dimensional arrays, $GHI = f(time, lat, lon)$, where the spatial dimensions of each dataset corresponded to a unique point defined by the coordinates of the associated station. Simulations with the daily time-resolved dataset were performed using the Climate Data Interface.

~~Measurements~~Representing cloudiness is a challenging task that requires several observations. For this purpose, aiming to obtain an indicative measure of the intra-hour cloudiness conditions we adopted the following formulation. Specifically, measurements of Direct Normal Irradiance (DNI) were utilized to obtain information for cloudiness relying on the conditions stated by WMO (2021), according to which sunshine duration is the total period where DNI exceeds $120\ W/m^2$. Alternative approaches such as the Cloud Modification Factor, require estimates of the clear sky irradiance, which introduces additional uncertainty. For the purpose of this analysis, we introduced a solar visibility (SV) parameter. Specifically, we assigned the value 0 when sun was obscured and the value 1 when visible. Aiming to describe the mean intra-hour cloudiness conditions, we considered the sky as cloud-free, cloudy, and partly cloudy based on the mean SV for the entire corresponding hour as follows:

$$\langle SV \rangle_{hour} : \begin{cases} 1 & cloud - free \\ \in (0,1) & partly\ cloudy \\ 0 & cloudy \end{cases}$$

For aerosol information, we accessed the AERONET Version 3 (V3) (Giles et al., 2019) and retrieved level 2.0 data (from direct sun measurements) for Aerosol Optical Depth at 500nm (AOD_{500}), which serves as a representative measure of the aerosol load; Ångström Exponent between 440 and 870 nm wavelengths ($AE_{440-870}$), where values near 0 correspond to coarse dust particles and values around 2 to fine (e.g., smoke) particles (Dubovik et al., 2002); and Fine Mode Fraction at 500nm (FMF_{500}) obtained from the Spectral Deconvolution Algorithm (SDA) retrievals, to distinguish aerosol into fine and coarse mode. The data were resampled at hourly intervals and a mean hourly value calculated. After, the hourly mean values divided into clusters ~~regarding AO~~based on AOD_{500}, reflecting different levels of aerosol load and allowing us to quantify their impact on solar energy production. To investigate the impact related exclusively to aerosols, we included only hours with cloud-free sky conditions. The clusters are defined in detail as follows:

- $AOD_{500} \leq 0.05$: Low aerosol load

- $0.05 < AOD_{500} \leq 0.15$: Moderate aerosol load
- $0.15 < AOD_{500} \leq 0.3$: High aerosol load
- $AOD_{500} > 0.3$: Very high aerosol load

To evaluate the performance of the Climate Interface over daily intervals, we defined the sunny (cloudless) days using the condition: $\langle SV \rangle_{day} \geq 0.9$. Next, to characterize the average aerosol conditions on sunny days, we applied the following classification:

- $\langle AOD_{500} \rangle_{day} \leq 0.05$: very-low aerosol
- $\langle AOD_{500} \rangle_{day} > 0.05$: aerosol-laden

Detailed comparisons of the energy production over hourly and daily integrals ~~c~~ under the various predefined sky conditions are provided in the supplement through evaluation metrics.

The selected locations have quite different atmospheric conditions regarding cloudiness and aerosols. Additionally, they vary in altitude. A brief overview of the prevailing conditions derived from the ground-based data is provided on the supplement. Regarding cloudiness, it is notable that in Lindenberg the sky is generally overcast, whereas in southern locations sunshine dominates. In terms of aerosols, very high aerosol loads occur more frequently in Tamanrasset. As for aerosol type, there is considerable variation among the examined locations: Carpentras, Cener, and Lindenberg are primarily influenced by fine mode aerosols, while Tamanrasset and Izaña are mostly affected by coarse mode aerosols.

For investigating the impact of desert dust aerosol in solar based power generation, Tamanrasset serves as a representative and exceptional case because it is in a region with important sources of Saharan dust aerosols (Faid et al., 2012). Meanwhile, Izaña, located in subtropical North Atlantic, is a high mountain station within the free troposphere, affected ~~my~~ mineral dust when the Saharan Air Layer top exceeds the station height, especially through August to October (Toledano et al., 2018; Cuevas et al., 2018). Due to its high altitude, Izaña avoids contamination from local or regional sources (Barreto et al. 2022). The Canary Islands, where Izaña is located, are influenced by extreme dust events that cause a significant decrease in PV power generation (Canadillas-Ramallo et al., 2021). In South Europe, which is also affected by the transport of Saharan dust across the Mediterranean, aerosol types exhibit a mixture as a result of simultaneous local pollution and low concentration of mineral dust (Logothetis et al., 2020).

2.3 Copernicus Atmospheric Monitoring Service (CAMS)

We retrieved data from the CAMS radiation service (Schroedter-Homscheidt et al., 2022; Qu et al., 2017), from the solar radiation time-series product (CAMS, 2020). The CAMS solar radiation service provides historical estimates for global solar radiation, along with its components, from 2004 to present. These values are provided with a frequency as fine as 1-minute. In this study, we used the hourly time-series of GHI and DHI for all-sky conditions, setting the input coordinates to match the locations of the BSRN stations. The solar radiation time-series product (CAMS, 2020) performs interpolations integrated in its internal algorithm and provides time-series for the coordinates and the altitude of a single-site location. We compared the solar energy production derived from the use of CAMS data with that derived from the use of ground-based measurements from BSRN.

2.4 Radiative Transfer Model (RTM)

We performed Radiative Transfer (RT) simulations aiming to further assess the uncertainties in estimating the diffuse fraction arising from the effect of aerosols. The simulations were conducted using libRadtran (Emde et al., 2016; Mayer & Kylling, 2005), a widely used software package, allowing the computation of radiances, irradiances, and actinic fluxes. A sensitivity analysis was performed by comparing the diffuse irradiance calculated from libRadtran with the estimations of BRL. This analysis examines the dependence of the aerosol-related discrepancy as function of Solar Zenith Angle (SZA) and latitude, considering the effect of parameters such as surface albedo and altitude.

To conduct aerosol parameterizations, we considered the default aerosol extinction profile (Shettle, 1989) and set asymmetry factor (gg) to 0.7, while varying the Single Scattering Albedo (SSA) and the Ångström Exponent (AE), and defining AOD_{500} by adjusting the value of the parameter-b in Ångström's law (Ångström, 1929) as follows:

$$\tau_\lambda = b \cdot \lambda^{-a} \rightarrow AOD_{500} = b \cdot (0.5\ \mu m)^{-AE}$$

The standard aerosol profiles (Anderson et al., 1986) were used for all sites. According to Fountoulakis et al. (2022), using a more accurate vertical distribution of aerosols in the troposphere would have a negligible effect in the GHI and DHI at the Earth's surface.

Table 3 illustrates the libRadtran settings used in this study.

Table 3. LibRadtran inputs

Parameter	Input
Atmospheric profile	Mid-latitude summer (April-September)/mid-latitude winter (October - March) (Anderson et al., 1986)
Extraterrestrial spectrum	(Kato et al. 1999)
~~SZA~~Datetime	~~with step 90°~~date and time input accompanied by project location coordinates
Altitude	0.1/2 km
Surface albedo	0.2 / 0.8
Number of streams	6
RT solver	sdisort (Buras et al., 2011)
AE	0 – 2 with step 1
SSA	0.7, 0.9, 1.0
gg	0.7
TOC (Total Ozone Column)	300 DU
Integrated Water ~~vapor~~Vapor	15 mm

3. Results

3.1 Performance verification of the BRL diffuse fraction model

The performance of BRL was evaluated by comparing the actual diffuse fraction, obtained directly from resampled to hourly BSRN ground-based measurements, with that derived using BRL. ~~Initially~~As a first step, to isolate the influence of SZA from that associated with the atmospheric conditions, the difference in diffuse fraction (DF) between the observed and the one estimated using BRL as a function of SZA is presented in Figure 2. The atmospheric conditions are represented separately for both all-sky and cloud-free sky conditions and are grouped into clusters, as outlined in Section 2.2. The patterns reflecting the differences under the distinct sky conditions indicate an additional dependency on SZA, which becomes apparent approximately ~~beyond 60°.~~at SZA between 60° and 70°. In most cases, there is an almost constant displacement with respect to y=0 below 60°, as well as a ~~negative trend~~change in behavior when SZA exceeds this value. Izaña presents a special case, as the station is located at a very high altitude~~, with adjacent clouds occasionally being situated at a~~. At such high altitudes the contribution of the diffuse component to the total irradiance is significantly smaller relative to lower ~~elevation than the station itself. As a result~~altitude sites, which seems to be captured more accurately by BRL at high SZAs. We must also note that (i) at Izaña, the actual diffuse irradiance ~~experiences~~may experience an additional enhancement due to the contribution of ~~these~~ adjacent lower-lying clouds – an effect that is not accounted for in the diffuse fraction model, and (ii) during dust events the site is usually inside – and not under – the dust layer, which results in more complex interactions between dust and solar radiation relative to lower altitude sites. Defining an exact limit (for the lower altitude sites), where the behavior is changing, is challenging; therefore, 60° was selected for practical energy-related applications, focusing on periods with meaningful energy contribution, and is supported by the sensitivity analysis (Section 3.2) under clear-sky conditions. Concerning the same grouped atmospheric conditions, Figure 3 illustrates the comparison between the observed and the estimated diffuse fraction for $SZA \leq 60°$. This approach allows us to examine BRL performance after eliminating the influence of SZA, thereby providing a more comprehensive view of its reliability.

Formatted: English (United Kingdom)

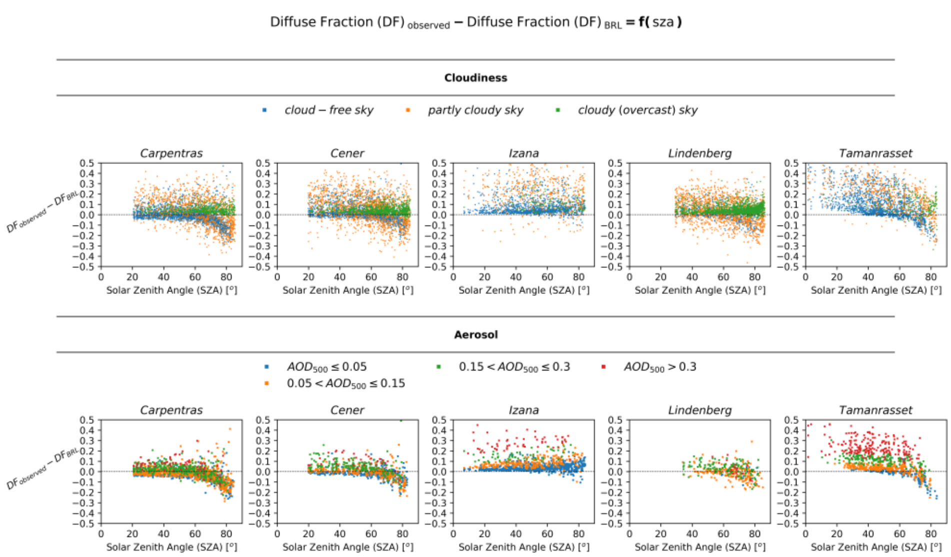

Figure 2. Difference between the ~~observed and the~~ diffuse fraction estimated by the ground-based measurements and by using the BRL ~~diffuse fraction~~model as a function of SZA under diverse

atmospheric conditions: (top) classification with respect to cloudiness and (bottom) classification with respect to aerosol optical depth

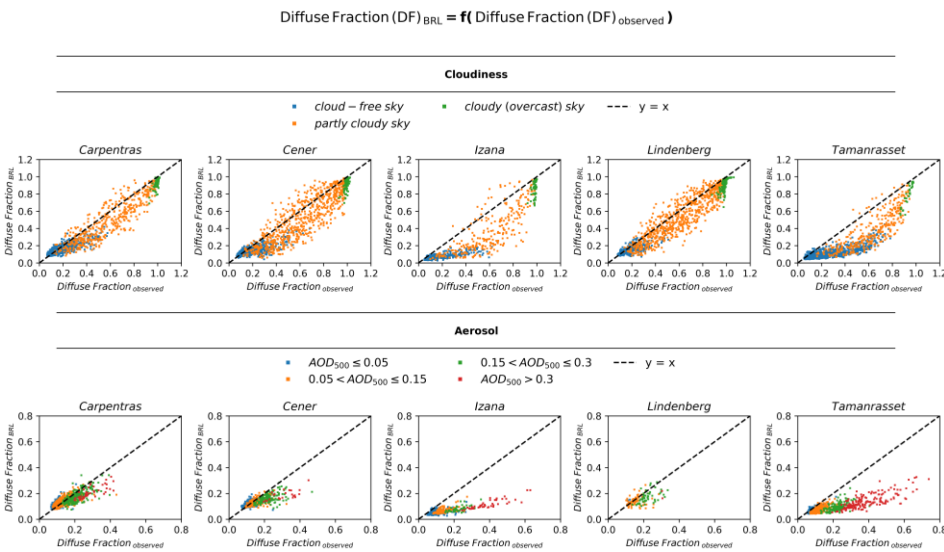

Figure 3. Comparison of the diffuse fraction estimated using BRL with ~~the actual one calculated directly from~~that estimated by the ground-based measurements under diverse atmospheric conditions for SZA < 60°: (top) classification with respect to cloudiness and (bottom) classification with respect to aerosol optical depth

From Figure 3, a distinct dependency of BRL's reliability on the atmospheric conditions can be observed. Under all-sky conditions, the presence of clouds has a notable impact on the ~~model~~model's performance. Partly cloudy conditions result in greater dispersion of the values from the identity line respectively, likely due to the complexity of such sky scenes. Under overcast conditions, where the sky can be considered homogeneous and isotropic, the model ~~performs slightly better.~~ in most cases performs slightly better. However, the limitations of the DNI-based classification methodology, related to the complexity of the cloud scenes, the spatiotemporal variability during the hourly periods, and the 3D variability of cloud properties, would require additional observational tools for a more detailed investigation. More specifically, the vast majority of overcast cases where the BRL diffuse fraction is below 0.8 while the observed is close to 1 correspond to periods involving rapid transitions between partly cloudy and overcast skies, occurring either during the hour itself or immediately before or after it. Furthermore, a limited number of cases identified during intense dust events at Tamanrasset and Izana, where the reduction of DNI was so pronounced that the applied DNI-based criterion classified these conditions as overcast. However, these cases are not further investigated, as the energy production levels during such periods are very low.

Under cloud-free skies, BRL tends to underestimate, and this bias becomes more pronounced as aerosol load increases. Aiming to highlight this dependency, Figure 4 shows the difference between the estimated and the observed diffuse fraction as function of AOD_{500}, emphasizing also the extent to which it is related to the aerosol type by providing FMF_{500}. ~~A negative trend~~A decrease for increasing AOD_{500} is evident across all cases. In Tamanrasset and Izaña, associated with the influence of Saharan dust, the coarse mode dominates, and a more distinct and well-defined curve is depicted~~.~~ compared to other sites.

It is important to clarify that for assessing the impact of aerosols we have assumed entirely cloud-free conditions. However, the criterion applied based on DNI does not fully guarantee the absence of small, scattered clouds within the sky dome. Such clouds could induce slight enhancements in DHI. A more rigorous assessment of the impact associated exclusively with aerosols could be achieved by integrating images from ground-based co-located all-sky cameras. On the other hand, the presence of aerosols even under cloudy scenes, introduces an additional uncertainty which is difficult to investigate accurately.

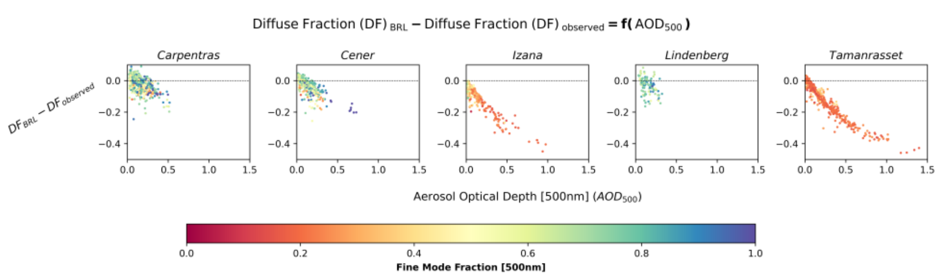

Figure 4. Difference between the estimated using BRL and ~~actual diffuse~~thediffuse fraction estimated by the ground-based measurements as function of AOD_{500} and FMF_{500}

3.2 Sensitivity analysis of the BRL performance under cloud-free sky conditions from RT simulations

The uncertainties in estimating diffuse fraction under cloud-free sky conditions, as discussed in section 3.1, are further investigated. We performed RT simulations using libRadtran to calculate GHI and DHI under various aerosol scenarios. The resulting GHI values were then used as input to BRL to estimate the diffuse fraction, which was subsequently compared to the diffuse fraction derived directly from the ratio of DHI to GHI computed by libRadtran.

To ensure a comprehensive analysis, we considered three representative latitudes (25°, 35° and 45°). Since BRL requires an hourly time-series of GHI as input, the analysis was conducted for the summer solstice. On this day, a sufficient number of hourly values are available, corresponding to a wide range of SZA values, allowing for a robust assessment of the methodology. The sensitivity analysis was performed for surface albedo values of 0.2 and 0.8 as well as for altitudes of 0~~,~~.1 and 2 km. For aerosol parameterization, we examined completely clear-sky conditions as a reference, alongside scenarios with AOD_{500} values of 0.2, 0.6, and 1, while varying the SSA and AE. Specifically, the scenarios included SSA values of 0.7, 0.9 and 1, combined with AE values of 0, 1 and 2. The results of this sensitivity analysis for an albedo of 0.2 are provided in Figure 5, while the results for an albedo of 0.8 are included in the supplement (Figure S1).

The results confirm that BRL performs well under clear sky conditions and for SZA below 60°, while the incorporation of aerosols in the sky scene introduces larger uncertainties. In all scenarios, we observe that lower values of AE correspond to higher uncertainties. Moreover~~, regarding SSA~~, when SSA is 0.9 or 1 BRL gradually tends to underestimate the diffuse fraction as aerosol load increases. Instead, when SSA is 0.7, BRL exhibits a different behavior, shifting toward an overestimation of the diffuse fraction at high aerosol loads.

The findings of this sensitivity analysis are consistent with the evaluated BRL performance from ground-based measurements presented in section 3.1, especially at SZA smaller than 60° - 70°, and underscore the role of aerosol in the accuracy of diffuse fraction estimations. Differences between the results shown in Figures 2 and 5 at SZA between 60° - 80° can be due to a number of site-related reasons. For example, enhancement of the diffuse component due to scattering by underlying atmospheric layers and clouds in the case of Izaña may compensate the observed overestimation of the diffuse fraction by BRL. Concerning the impact related to AE and SSA, we confirm that the higher underestimations observed for Tamanrasset and Izaña are associated with the optical properties of desert dust aerosol particles. While AE and SSA alone are not sufficient to fully characterize the aerosol type, they serve as strong indicators, aligning with the classification framework of Dubovik et al. (2002). The same comparison for albedo 0.8 (Figure S1 in the supplement) reveals a significant broadening of the discrepancies. Moreover, we observe the presence of a systematic error, even under clear sky conditions.

The resulting differences were practically identical across the three selected latitudes, indicating that the BRL model is largely independent of latitude and can therefore be considered as a reliable solution over a wide range of latitudes. Furthermore, the effect of altitude was found to be small. Finally, the outcomes of this analysis highlight potential inconsistencies arising from aerosols with different optical properties. Although the updated parameters of the BRL's model (as implemented in the GSEE model) reported by Lauret et al. (2013) were derived using data from nine worldwide locations, encompassing a broad range of sky conditions that capture a fully representative set of optical properties remain challenging.

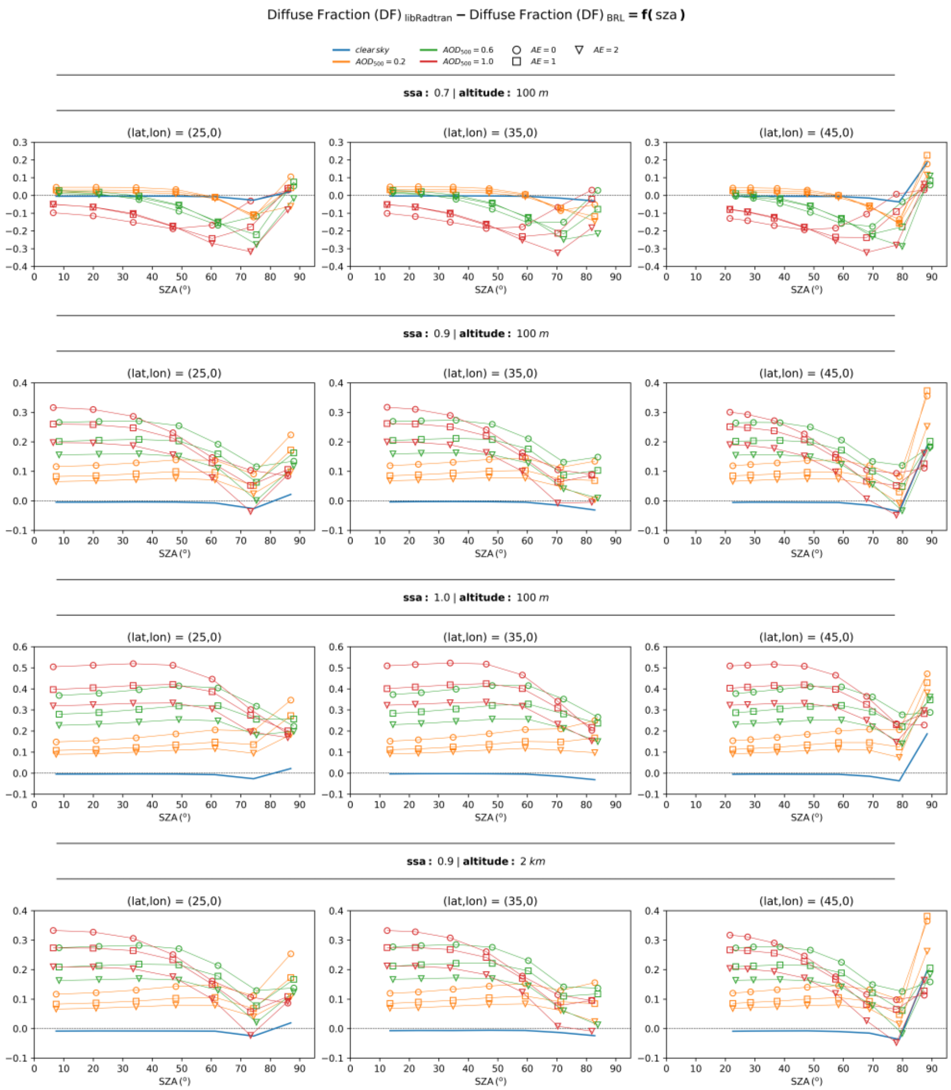

Figure 5. Difference between the diffuse fraction derived directly from the computations of DHI and GHI using libRadtran and the one estimated by applying BRL to the libRadtran-computed GHI

3.3 Analysis of the differences in energy production using hourly integrals within the modelling of PV plants

Uncertainties in estimating the diffuse fraction influence the calculation of the total irradiance received by an inclined panel's surface, thereby affecting the accuracy of the PV power simulations. In this section, we employ the main submodule of GSEE, used for modelling the electric output from a PV panel, aiming to assess the extent to which these uncertainties propagate to the estimation of the hourly power production. We analyze discrepancies arising from using only GHI from BSRN as input radiation data to the model, instead of both DHI and GHI. More specifically, we compare the total energy produced per hour per unit, expressed in watt-hours (Wh), per unit of nominal power (kWp). The energy production is evaluated for both fixed panels and 2-axis tracking systems.

The results of this comparison for c-Si based technology PV panels for different atmospheric conditions are presented in Figure 6, illustrating the impact of cloudiness, and in Figure 7, demonstrating the effect of aerosols. The corresponding results for CdTe technology are provided in the supplement (Figures S2 and S3 respectively). In the modelling of 2-axis solar tracking systems, where the panel is continuously adjusted to maintain a perpendicular orientation to incoming solar radiation, the system becomes more sensitive to uncertainties in the estimation of the diffuse fraction, leading to more significant differences in energy production. Specifically, the contribution of the direct irradiance is maximized in such systems, as the panel exploits the entirety of the available direct irradiance. On the other hand, in the simulation of static panels, the contributions of direct and diffuse components are more evenly distributed, making the impact of diffuse fraction uncertainties less pronounced in energy production.

Regarding the uncertainties related to the atmospheric conditions, from Figure 6 we confirm that the highest dispersion occurs in partly cloudy conditions, while from Figure 7, where we examine cloud-free conditions, we note that further improvement achieved as aerosol load decreases. Under totally overcast skies the energy production is extremely low, rendering errors practically negligible. Moreover, accuracy is influenced by aerosols, where a gradual decline in accuracy is detected as aerosol load increases. However, assessing the extent of aerosol loading impact is complex, depending on the interaction of solar radiation with particles of varying optical properties, as extensively analyzed in the previous sections. This effect becomes particularly evident in cases of high aerosol loading, where a noticeable offset is observed, while under certain conditions, the associated uncertainty is comparable to that found in partly cloudy conditions.

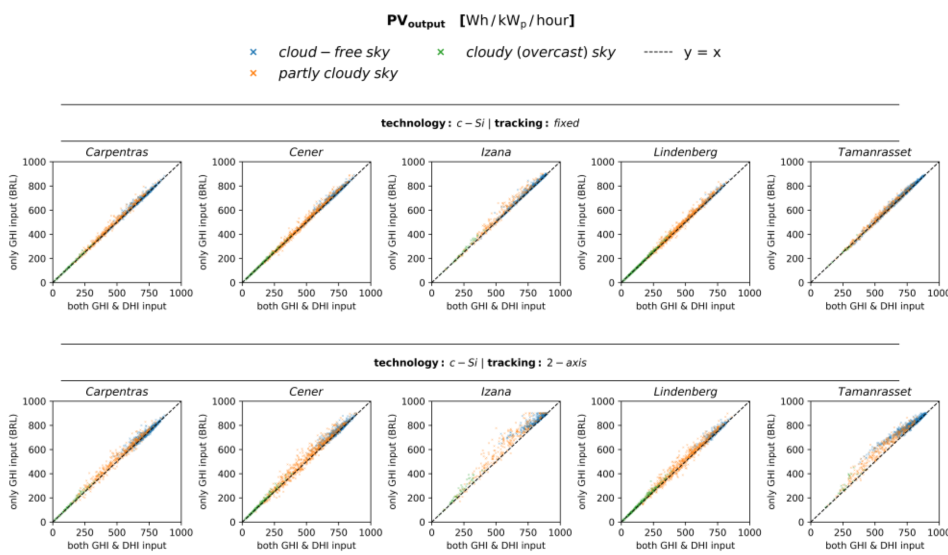

Figure 6. Comparison of the estimated hourly PV power generation between simulations performed using GSEE with input data consisting of either only GHI or both GHI and DHI under varying cloudiness conditions: (top) fixed panels (bottom) 2-axis tracking systems

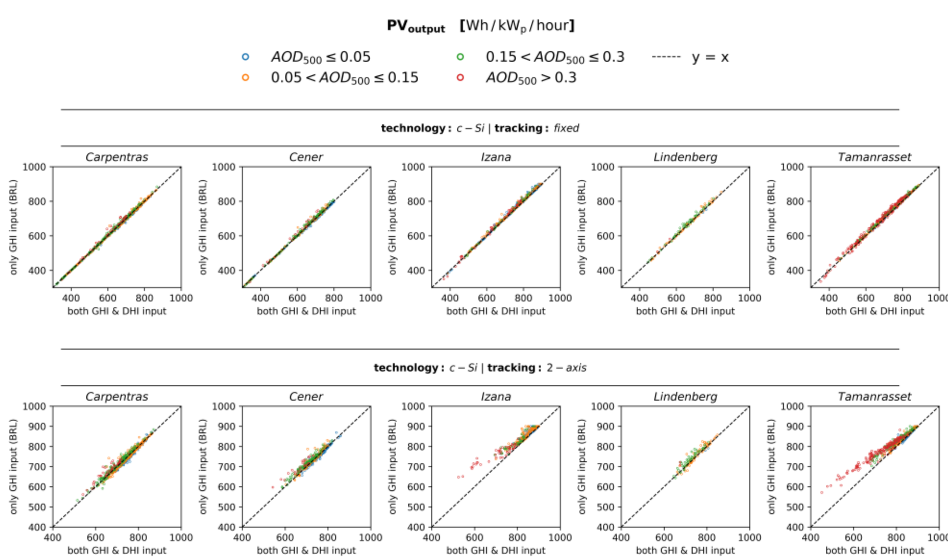

Figure 7. Comparison of the estimated hourly PV power generation between simulations performed using GSEE with input data consisting of either only GHI or both GHI and DHI under varying aerosol conditions: (top) fixed panels (bottom) 2-axis tracking systems

The PV systems considered in this study have a nominal capacity of 1 kWp. The PV model applies a default system loss factor of 10%. This effectively limits the maximum achievable power output to approximately 90% of the nominal capacity (i.e., around 900 W/kWp). This effect becomes apparent at the Izaña site due to its low latitude combined with its specific geographical and atmospheric conditions, which lead to high irradiance levels. As a result, the simulated PV output in some cases appears capped around 900 Wh/kWp per hour when only GHI is used.

Additionally, Tables ~~S1-S5 in the supplement,~~4 and 5 present the validation results~~, including computed~~ for Carpentras and Tamanrasset, selected as representative locations that encompass a wide variety of sky conditions. Validation results for the remaining stations are available in the supplement (Tables S1-S3). All the evaluation metrics ~~that quantify the errors. All the computations~~ correspond to simulations of PV panels with c-Si technology.

Table 4. Evaluation metrics for GSEE performance within hourly intervals in Carpentras, comparing simulations with diffuse fraction from measurements and from the BRL model

STATION: Carpentras		fixed panels			2-axis tracking		
		RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)	RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)
All-Sky scenes		12.6	6.6	0.8	20.8	12.5	1.2
All-Sky scenes (cloudiness)	cloud-free	9.2	4.6	0.4	14.8	8.7	0.5
	partly cloudy	19.5	12.5	2.3	32.5	23.9	3.8
	cloudy (overcast)	5.8	3.0	2.0	10.5	6.1	4.6
Cloudless-Sky scenes (aerosol load)	low	4.7	3.4	-0.4	9.5	7.5	-0.8
	moderate	4.3	2.2	0.1	7.8	4.7	0.0
	high	6.4	4.0	0.6	11.0	7.8	0.9
	very high	14.9	10.2	1.6	22.7	17.2	2.6

Table 5. Evaluation metrics for GSEE performance within hourly intervals in Tamanrasset, comparing simulations with diffuse fraction from measurements and from the BRL model.

STATION: Tamanrasset	fixed panels			2-axis tracking		
	RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)	RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)

All-Sky scenes		13.6	9.3	1.0	40.4	27.8	3.8
All-Sky scenes (cloudiness)	cloud-free	11.5	8.0	0.8	35.3	23.4	2.9
	partly cloudy	20.1	15.0	2.0	56.1	45.7	8.1
	cloudy (overcast)	8.4	5.2	-0.1	45.3	30.1	11.2
Cloudless-Sky scenes (aerosol load)	low	3.2	2.0	0.2	6.6	4.0	0.3
	moderate	5.4	4.6	0.6	13.0	10.5	1.2
	high	12.5	11.7	1.6	30.1	27.4	3.4
	very high	18.0	16.2	1.9	57.0	49.2	6.8

Based on the calculated statistical indices, the Root Mean Square Error (RMSE) values for fixed panels range from 4.7 Wh/kWp/hour (clear sky) to 19.5 Wh/kWp/hour (partly cloudy) in Carpentras, and from 3.2 to 20.1 Wh/kWp/hour in Tamanrasset. Under very high aerosol loading, RMSE reaches 14.9 and 18.0 Wh/kWp/hour, respectively. For 2-axis tracking systems, RMSE values vary significantly, ranging from 9.5 to 32.5 Wh/kWp/hour in Carpentras and from 6.6 to 56.1 Wh/kWp/hour in Tamanrasset, with peaks of 22.7 and 57.0 Wh/kWp/hour under very high aerosol loading conditions. Similarly, the Mean Absolut Error (MAE) values are generally lower for fixed panels (3.4-12.5 Wh/kWp//hour in Carpentras, 2.0-15.0 in Tamanrasset) and substantially higher for 2-axis tracking (7.5-23.9 and 4.0-45.7 Wh/kWp/hour, respectively). Notably in Tamanrasset, MAE values under very high aerosol loading exceed those observed under partly cloudy conditions, with values increasing from 15.0 to 16.2 Wh/kWp/hour for fixed panels and from 45.7 to 49.2 Wh/kWp/hour for 2-axis tracking systems. Regarding the relative mean bias (rMBE), this remains mostly within ± 4.6% for fixed panels but can reach up to 11.2% for 2-axis tracking, particularly in aerosol-laden conditions.

3.4 Estimating total daily PV power output using the Climate Interface

Validation of the estimated daily energy production using the Climate Interface is achieved by comparing the estimates with the results obtained from the direct summation of the hourly simulations with input both GHI and DHI.

The Climate Interface generates the hourly profile of GHI for each day as a sinusoidal function. Then, the BRL is applied to the hourly time-series, and the hourly power generation is computed. Finally, these values are summed up to provide an estimate of the total daily output power. As shown in ~~Figure~~Fig. 8, which illustrates the differences between the Climate Interface estimates and the sums of the hourly simulations, this approach introduces a variability throughout the year. Furthermore, Figure S6 in the supplement presents the percentage differences between the two approaches, using the latter as the reference.

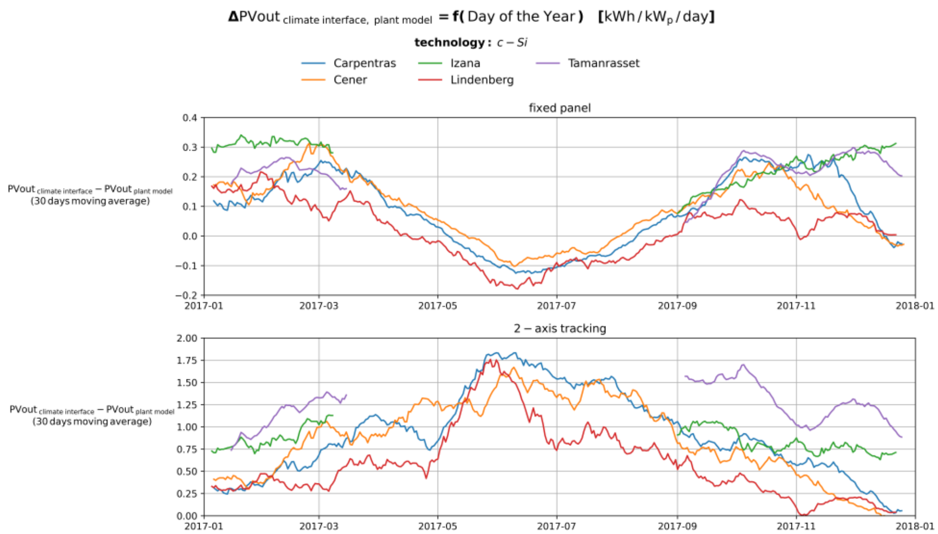

Figure 8. Time-series of the differences between the daily PV output estimated using the climate interface and the corresponding daily sums from hourly simulations.

The time-series represent the centered 30-day moving average. To ensure that the values are representative of the reference period, we have applied all conditions requiring at least 20 days of available data within each 30-days interval. In Tamanrasset and Izaña, especially during the summer months, there are significant data gaps on several days, often occurring around solar noon.

More precisely, from ~~Figure~~Fig. 8, we observe that within the modelling of PV plants with fixed panels, there is a tendency to overestimate in winter, with deviations of approximately 0.3 kWh/kWp/day, and to slightly underestimate in summer, where deviations are around 0.1 kWh/kWp/day. In contrast, for 2-axis solar tracking systems, the resulting deviations are significantly larger, with a general tendency toward overestimation that peaks during summer, reaching approximately 1.75 kWh/kWp/day. The percentage differences span from -10 to 20 % for fixed panels and from -5 to 35 % for 2-axis tracking systems.

The variability in the percentage difference between the daily PV output estimated using the climate interface and the corresponding daily sums is mainly a function of the minimum SZA, while especially in the case of modeling for 2-axits tracking systems, the variation is also influenced by aerosol loading, with differences tending to increase as aerosol load rises (Figures S4 and S5 in the supplement).

Additional validation results are provided in the supplement (Tables ~~S6-S10).~~S4-S8). Indicatively, for Carpentras and Tamanrasset, representative results are discussed below. For fixed panels, RMSE is minimized at 0.18 kWh/kWp/day under very-low aerosol conditions, compared to the overall 0.22 kWh/kWp/day for Carpentras. In Tamanrasset, the lowest RMSE is observed at 0.15 kWh/kWp/day under very low aerosol conditions, while the overall reaches 0.24. In the case of 2-axis tracking, a significant increase is observed from low-aerosol to aerosol-laden conditions, ranging from 0.82 to 1.28 kWh/kWp/day in Carpentras and from 0.66 to 1.37 in Tamanrasset. Similar widening trends are also evident in the MAE values across different aerosol loading conditions. The computed statistical indices confirm that the differences are minimized under sunny and nearly aerosol-free sky conditions. Comparing the performance on low-aerosol days to that on aerosol-laden, we conclude that, particularly in the case of modelling 2-axis tracking systems, errors increase significantly. In Tamanrasset, in particular, the errors are more than double.

3.5 Evaluation of the reliability of using the CAMS solar radiation time-series product in modelling PV power potential

The aim of this section is to inspect the reliability of using the CAMS solar radiation time-series product in modelling the PV power potential adapted to a certain location~~,~~. A review of the existing literature indicates a lack of studies directly examining the accuracy of using CAMS data for assessing PV power potential. This is addressed by comparing the output power obtained from using CAMS solar radiation data with that calculated using ground-based measurements. The analysis focuses on the capability of CAMS to provide accurate estimates of both GHI as well as its individual components.

In this section, we have excluded Izaña, because, due to its high altitude – as indicated through a personal communication with Yves-Marie Saint-Drenan (2025) – comparable results would require adjusting the measurements to the elevation of the stations, which is a complicated process and beyond the scope of this study.

The CAMS-based diffuse fraction, compared to the observed, is presented in Figure 9 under different prevailing conditions. We observe that the calculation of the diffuse component is subject to significant uncertainty. Cloudiness is the primary uncertainty source, particularly under partly cloudy conditions. Additionally, notable discrepancies related to aerosols emerge only in cases of very high aerosol loading.

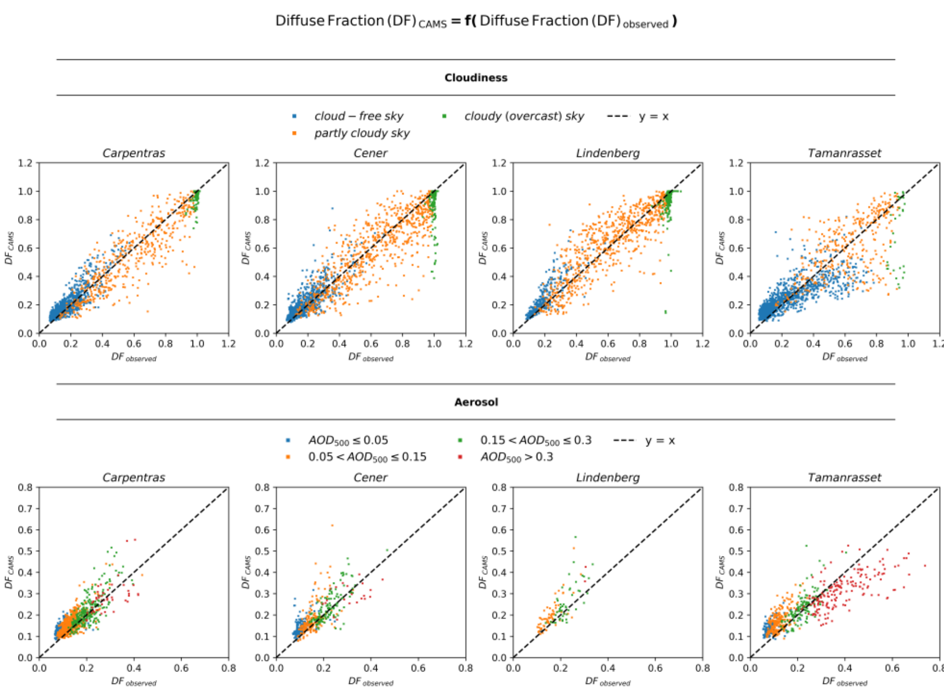

Figure 9. Comparison of the CAMS-based diffuse fraction estimated using BRL with the actual one under diverse atmospheric conditions

In ~~Figure~~Fig. 10 we provide density scatter plots comparing the CAMS-based PV output power with that computed from the ground-based BSRN data, aiming to illustrate how the uncertainty in the diffuse component estimates propagate to the calculation of power generation. Notably, there is a much greater dispersion from the y=x line in the case of simulating PV plants with 2-axis tracking system, compared to that within the modelling of fixed panels. This outcome is attributed to the increased sensitivity of the 2-axis tracking systems to the partitioning of global irradiance into its components. Nevertheless, correlation coefficients are in all cases better than 0.9. ~~Additional evaluation metrics are provided in the supplement (Tables S11-S14).~~

Formatted: Font: Not Bold

Additional evaluation metrics are provided in the supplement (Tables S9-S12). Indicatively, we observe that under cloudless conditions, for fixed panels, RMSE ranges between 25.0 to 42.3 Wh/kWp/hour in Carpentras and 16.6 and 31.0 Wh/kWp/hour in Tamanrasset, with variations linked to aerosol loading. Similarly, MAE ranges from 20.0 to 36.9 Wh/kWp/hour in Carpentras and 11.9 to 22.9 Wh/kWp/hour in Tamanrasset. For 2-axis systems, RMSE and MAE follow similar trend, ranging from 28.8 to 49.9 Wh/kWp/hour and 22.3 to 44.1 Wh/kWp/hour, respectively, in Carpentras, and from 20.8 to 48.0 Wh/kWp/hour and 15.3 to 35.5 Wh/kWp/hour, respectively, in Tamanrasset. Conversely, under cloudy conditions the errors are significantly increasing. In Carpentras, as well as in Cener, and Lindenberg (according to the corresponding tables in the supplement) the errors peak under partly cloudy conditions, with RMSE reaching up to 94.2 Wh/kWp/hour in Carpentras. However, in Tamanrasset, the highest errors occur under overcast conditions, where RMSE and MAE for 2-axis solar tracking systems reach 210.7 and 151.6 Wh/kWp/hour, respectively. This exception can be interpreted through Figure 15, which illustrates that in the rare overcast scenes in Tamanrasset, CAMS occasionally reports low diffuse fraction values instead of values close to 1, suggesting that CAMS did not accurately represent cloudiness in these cases.

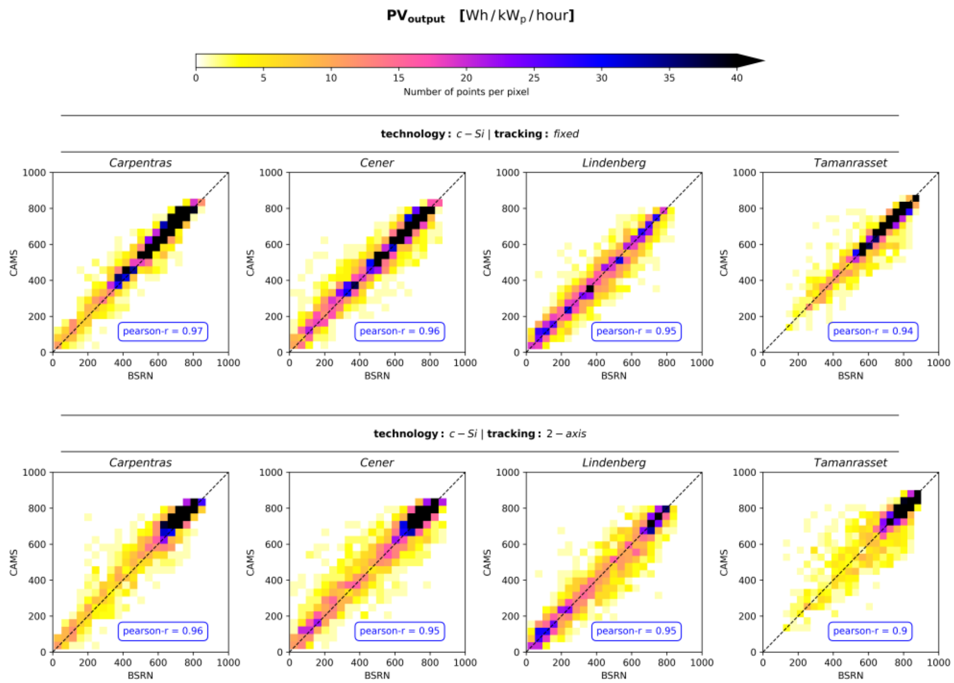

Figure 10. Overview of the reliability of the CAMS-based PV power simulations

4. Conclusions

~~This study evaluated different solar radiation information that is commonly used for PV power modelling, and their implications for PV modelling accuracy.~~ The optimal approach to include solar radiation information to PV power models such as GSEE is to use actual in-situ measurements of global and diffuse solar irradiance. Since measurements of the diffuse component are rarely available, it is common to use measurements of the GHI (if available) and retrieve the diffuse component using a model such as BRL. In the absence of in-situ measurements, other options include the use of datasets such as CAMS or even a radiative transfer model, provided that atmospheric inputs such as clearness index, aerosol optical depth (AOD), and other aerosol properties are available. This study evaluated these options and their implications for PV modelling accuracy.

The results highlighted the importance of having precise information for the distribution of solar irradiance among its components in PV power modelling. The implementation of the BRL diffuse fraction within GSEE serves as a practical, and under certain conditions, reliable solution to the absence of detailed information for each component separately. Moreover, the integrated Climate Data Interface submodule offers valuable prospects for investigating fluctuations in the solar PV power generation across various timescales. In this context, the use of BRL has a key contribution alongside the other computational procedures in processing climate datasets. Previous studies on PV power modelling approaches have not examined their reliability under diverse atmospheric conditions, including the effects associated with cloudiness, aerosol loading, as well as aerosol optical properties.

The evaluation of the BRL's performance revealed a dependency of its reliability on the prevailing sky conditions. ~~As a result, discrepancies arising from inconsistencies in diffuse fraction estimation propagate to PV power generation. Within the modelling of PV plants equipped with 2-axis solar tracking system, the deviations are much more pronounced relative to optimally inclined panels.~~ BRL has excellent accuracy under totally clear sky scenes and still performs well for cloudless scenes with moderate aerosol loading. In general, its accuracy is inversely proportional to the complexity of the cloud scene. However, the model systematically underestimates the diffuse fraction under high-loading conditions, such as during dust events. ~~Under such circumstances, this bias can potentially lead to significant overestimation of power generation by up to 49.2 Wh/kWp/hour~~The discrepancies arising from diffuse fraction estimation propagate to PV power generation and become particularly pronounced in the modelling of 2-axis tracking systems. Indicatively, MAE under cloud-free scenes with moderate aerosol loading, ranges between 2.2 to 6.6 Wh/kWp/hour for fixed panels and 4.7 to 15.0 Wh/kWp/hour for 2-axis tracking systems. Under partly cloudy conditions, where the cloud scene is more complex, the MAE increases substantially, ranging from 12.4 to 25.8 Wh/kWp/hour for fixed panels and from 23.5 to 55.1 Wh/kWp/hour for 2-axis tracking systems. Moreover, during intense dust events, MAE can reach up to 49.2 Wh/kWp/hour in Tamanrasset, which is comparable to that computed under partly cloudy conditions. Overall, the rMBE remains within the $\pm 5\%$, with the exception of a limited cases under overcast conditions. The same analysis applied to CdTe panels yielded similar results, with minor differences.

Aiming to provide an indicative assessment of the financial impacts of the effect of desert dust aerosols, we assume that the statistical indices calculated for Tamanrasset are representative of a large-scale solar farm located in the Sahara region, with 500 MW installed PV capacity and systems equipped with 2-axis solar tracking system. For this hypothetical solar farm, according to the value of the Mean Absolute Error (MAE) on Table 4 for very high aerosol loading, we estimate that the produced energy is $0.0492\ [kWh/kWp/hour] \times 500 \times 10^3\ [kWp] = 24600\ [kWh/hour] \overset{supposing\ 12\ sunlight\ hours\ per\ day}{\Longrightarrow} \sim 295200\ [kWh/day]$ less than the expected from the PV power simulations. According to the global average auction prices for selling produced energy back to the grid in 2021 (IRENA, n.d.), the overestimations are equivalent to a financial loss of $0.039\ [USD/kWh] \times 295200 [kWh/day] \approx 11{,}500\ USD/day$. Therefore, site assessments that do not correctly account for the ~~impact~~distribution of surface solar irradiance in the sky under desert dust ~~aerosols~~aerosol conditions may overestimate financial performance and the annual financial deficit could be accumulated to hundreds of thousands of US dollars per year.

Comparing the range of computed errors, we observe that the errors arising from employing CAMS rather than using ground-based measurements, even when the diffuse fraction is not provided, are higher across the overwhelming majority of the considered sky conditions. More specifically, regarding the overall performance, MAE when using CAMS ranges between 33.7 and 46.1 Wh/kWp/hour, while with ground-based GHI measurements, MAE remains below 10 Wh/kWp/hour within the modelling of systems with fixed panels and can reach up to 27.8 Wh/kWp/hour within the modelling of 2-axis tracking systems. This outcome highlights the value of ground-based measurements.

To sum up, achieving the highest quality PV power simulations necessitates high-quality, concurrent measurements of solar irradiance components. In absence of this, the submodules included in the GSEE package enable reliable simulations under the vast majority of prevailing sky conditions. CAMS serves as a valuable data source for PV power modelling, but it cannot fully replace the precision and reliability of using ground-based measurements. The integration of aerosol correction within the BRL model opens new possibilities for further improvements in the modelling of solar energy systems. A more comprehensive assessment would require measured PV output data; however, acquiring simultaneous direct and diffuse irradiance measurements at the same location as the solar farms remains challenging.

Data availability

The BSRN data are freely available on the BSRN web-page (https://bsrn.awi.de/). The AERONET version 3 products are freely available from the AERONET website (https://aeronet.gsfc.nasa.gov/). The CAMS radiation time-series are available from the Atmosphere Data Store (https://ads.atmosphere.copernicus.eu). The rest of the data used in this paper are available upon request from the authors.

Author Contributions

Conceptualization: NP and IF; Data curation: NP and KP; Formal analysis: NP; Funding acquisition: CZ; Investigation: NP; Methodology: NP, IF, SK, AK and AG; Project administration: CZ; Resources: SP, KP and LD; Software: NP; Supervision: IF; Validation: NP, IF and SP; Visualization: NP; Writing – original draft: NP; Writing – review & editing: all authors

Funding

This work has been supported by the action titled "Support for upgrading the operation of the National Network for Climate Change (CLIMPACT II)", funded by the Public Investment Program of Greece, General Secretary of Research and Technology/Ministry of Development and Investments. Part of this work was also supported by the COST Action Harmonia (CA21119) supported by COST (European Cooperation in Science and Technology). This work was partially funded by the Copernicus Climate Change Service under contracts C3S2 _461-1_GR (Seasonal to decadal predictions for national renewable energy management).

Acknowledgments

We thank the teams of the AERONET for ground measurements and maintenance, and CAMS for the data production and distribution. We would like to thank the five site instrument operators and technical staff of the BSRN network stations who made the ground-based measurements feasible. A. Gkikas, J. Kapsomenakis, and C.S. Zerefos also acknowledge "CAMS2_82 Project: Evaluation and Quality Control (EQC) of global products."

References

Anderson, G., Clough, S., Kneizys, F., Chetwynd, J., & Shettle, E. (1986). *AFGL atmospheric constituent profiles (0-120 km)* (Tech. Rep. AFGL-TR-86-0110). Air Force Geophysics Laboratory, Hanscom Air Force Base.

Anderson, K. S., Hansen, C. W., Holmgren, W. F., Jensen, A. R., Mikofski, M. A., & Driesse, A. (2023). *pvlib python: 2023 project update*. *Journal of Open Source Software, 8*(92), Article 5994. https://doi.org/10.21105/joss.05994

Ångström, A. (1929). On the atmospheric transmission of sun radiation and on dust in the air. *Geografiska Annaler*, *11*(2), 156–166. https://doi.org/10.1080/20014422.1929.11880498

Barreto, Á., García, R. D., Guirado-Fuentes, C., Cuevas, E., Almansa, A. F., Milford, C., Toledano, C., Expósito, F. J., Díaz, J. P., & León-Luis, S. F. (2022). Aerosol characterisation in the subtropical eastern North Atlantic region using long-term AERONET measurements. *Atmospheric Chemistry and Physics*, *22*(17), 11105–11124. https://doi.org/10.5194/acp-22-11105-2022

Blaga, R., Mares, O., Paulescu, E., Boata, R., Sabadus, A., Hategan, S.-M., Calinoiu, D., Stefu, N., & Paulescu, M. (2024). Diffuse fraction as a tool for exploring the sensitivity of parametric clear-sky models to changing aerosol conditions. *Solar Energy (Phoenix, Ariz.)*, *277*(112731), 112731. https://doi.org/10.1016/j.solener.2024.112731

Blanc, P., Remund, J., & Vallance, L. (2017). Short-term solar power forecasting based on satellite images. In *Renewable Energy Forecasting* (pp. 179–198). Elsevier.

Boland, J. W., Scott, L., & Luther, M. (2001). *Modelling the diffuse fraction of global solar radiation on a horizontal surface*. *Environmetrics, 12*(2), 103–116. https://doi.org/10.1002/1099-095X(200103)12:2<103::AID-ENV447>3.0.CO;2-2

Buras, R., Dowling, T., & Emde, C. (2011). New secondary-scattering correction in DISORT with increased efficiency for forward scattering. *Journal of Quantitative Spectroscopy & Radiative Transfer*, *112*(12), 2028–2034. https://doi.org/10.1016/j.jqsrt.2011.03.019

Cañadillas-Ramallo, D., Moutaoikil, A., Shephard, L. E., & Guerrero-Lemus, R. (2022). The influence of extreme dust events in the current and future 100% renewable power scenarios in Tenerife. *Renewable Energy*, *184*, 948–959. https://doi.org/10.1016/j.renene.2021.12.013

Copernicus Atmosphere Monitoring Service. (2020). *CAMS solar radiation time-series*. Copernicus Atmosphere Monitoring Service (CAMS) Atmosphere Data Store. https://doi.org/10.24381/5cab0912

Cuevas, E., Romero-Campos, P. M., Kouremeti, N., Kazadzis, S., Räisänen, P., García, R. D., Barreto, A., Guirado-Fuentes, C., Ramos, R., Toledano, C., Almansa, F., & Gröbner, J. (2019). Aerosol optical depth comparison between GAW-PFR and AERONET-Cimel radiometers from long-term (2005–2015) 1 min synchronous measurements. *Atmospheric Measurement Techniques*, *12*(8), 4309–4337. https://doi.org/10.5194/amt-12-4309-2019

Driemel, A., Augustine, J., Behrens, K., Colle, S., Cox, C., Cuevas-Agulló, E., Denn, F. M., Duprat, T., Fukuda, M., Grobe, H., Haeffelin, M., Hodges, G., Hyett, N., Ijima, O., Kallis, A., Knap, W., Kustov, V., Long, C. N., Longenecker, D., … König-Langlo, G. (2018). Baseline Surface Radiation Network (BSRN): structure and data description (1992–2017). *Earth System Science Data*, *10*(3), 1491–1501. https://doi.org/10.5194/essd-10-1491-2018

Dubey, S., Sarvaiya, J. N., & Seshadri, B. (2013). Temperature dependent photovoltaic (PV) efficiency and its effect on PV production in the world – A review. *Energy Procedia*, *33*, 311–321. https://doi.org/10.1016/j.egypro.2013.05.072

Dubovik, O., Holben, B., Eck, T. F., Smirnov, A., Kaufman, Y. J., King, M. D., Tanré, D., & Slutsker, I. (2002). Variability of absorption and optical properties of key aerosol types observed in worldwide locations. *Journal of the Atmospheric Sciences*, *59*(3), 590–608. https://doi.org/10.1175/1520-0469(2002)059<0590:voaaop>2.0.co;2

Dubovik, O., & King, M. D. (2000). A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements. *Journal of Geophysical Research*, *105*(D16), 20673–20696. https://doi.org/10.1029/2000jd900282

Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Seyboth, K., Kadner, S., Zwickel, T., Eickemeier, P., Hansen, G., Schlomer, S., & Von Stechow, C. (Eds.). (2011). *Renewable energy sources and climate change mitigation: Special report of the intergovernmental panel on climate change*. Cambridge University Press.

Emde, C., Buras-Schnell, R., Kylling, A., Mayer, B., Gasteiger, J., Hamann, U., Kylling, J., Richter, B., Pause, C., Dowling, T., & Bugliaro, L. (2016). The libRadtran software package for radiative transfer calculations (version 2.0.1). *Geoscientific Model Development*, *9*(5), 1647–1672. https://doi.org/10.5194/gmd-9-1647-2016

Faid, A., Smara, Y., Caselles, V., & Khireddine, A. (2012). Evaluation of the Saharan aerosol impact on solar radiation over the Tamanrasset area, Algeria. *International Journal of Advanced Research in Engineering and Technology, 3*(1), 24–32.

Fountoulakis, I., Kosmopoulos, P., Papachristopoulou, K., Raptis, I.-P., Mamouri, R.-E., Nisantzi, A., Gkikas, A., Witthuhn, J., Bley, S., Moustaka, A., Buehl, J., Seifert, P., Hadjimitsis, D. G., Kontoes, C., & Kazadzis, S. (2021). Effects of aerosols and clouds on the levels of surface solar radiation and solar energy in Cyprus. *Remote Sensing*, *13*(12), 2319. https://doi.org/10.3390/rs13122319

Fountoulakis, I., Papachristopoulou, K., Proestakis, E., Amiridis, V., Kontoes, C., & Kazadzis, S. (2022). Effect of aerosol vertical distribution on the modeling of solar radiation. *Remote Sensing*, *14*(5), 1143. https://doi.org/10.3390/rs14051143

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., Eck, T. F., Holben, B. N., Lewis, J. R., Campbell, J. R., Welton, E. J., Korkin, S. V., & Lyapustin, A. I. (2019). Advancements in the Aerosol Robotic Network (AERONET) Version 3 database – automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. *Atmospheric Measurement Techniques*, *12*(1), 169–209. https://doi.org/10.5194/amt-12-169-2019

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., & Smirnov, A. (1998). AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sensing of Environment*, *66*(1), 1–16. https://doi.org/10.1016/s0034-4257(98)00031-5

Hou, X., Wild, M., Folini, D., Kazadzis, S., & Wohland, J. (2021). Climate change impacts on solar power generation and its spatial variability in Europe based on CMIP6. *Earth System Dynamics*, *12*(4), 1099–1113. https://doi.org/10.5194/esd-12-1099-2021

Huld, T., Gottschalg, R., Beyer, H. G., & Topič, M. (2010). Mapping the performance of PV modules, effects of module type and data averaging. *Solar Energy (Phoenix, Ariz.)*, *84*(2), 324–338. https://doi.org/10.1016/j.solener.2009.12.002

Intergovernmental Panel on Climate Change (IPCC). (2023). *Climate change 2022 – impacts, adaptation and vulnerability: Working group II contribution to the sixth assessment report of the* *intergovernmental panel on climate change*. Cambridge University Press. https://doi.org/10.1017/9781009325844

Jacovides, C. P., Tymvios, F. S., Assimakopoulos, V. D., & Kaltsounides, N. A. (2006). Comparative study of various correlations in estimating hourly diffuse fraction of global solar radiation. *Renewable Energy*, *31*(15), 2492–2504. https://doi.org/10.1016/j.renene.2005.11.009

Kakran, S., Rathore, J. S., Sidhu, A., & Kumar, A. (2024). Solar energy advances and CO2 emissions: A comparative review of leading nations' path to sustainable future. *Journal of Cleaner Production*, *475*(143598), 143598. https://doi.org/10.1016/j.jclepro.2024.143598

Kato, S., Ackerman, T. P., Mather, J. H., & Clothiaux, E. E. (1999). The k-distribution method and correlated-k approximation for a shortwave radiative transfer model. *Journal of Quantitative Spectroscopy & Radiative Transfer*, *62*(1), 109–121. https://doi.org/10.1016/s0022-4073(98)00075-2

Kazantzidis, A., Tzoumanikas, P., Blanc, P., Massip, P., Wilbert, S., & Ramirez-Santigosa, L. (2017). Short-term forecasting based on all-sky cameras. In *Renewable Energy Forecasting* (pp. 153–178). Elsevier.

Kosmopoulos, P., Kazadzis, S., El-Askary, H., Taylor, M., Gkikas, A., Proestakis, E., Kontoes, C., & El-Khayat, M. (2018). Earth-Observation-based estimation and forecasting of particulate matter impact on solar energy in Egypt. *Remote Sensing*, *10*(12), 1870. https://doi.org/10.3390/rs10121870

Kouklaki, D., Kazadzis, S., Raptis, I.-P., Papachristopoulou, K., Fountoulakis, I., & Eleftheratos, K. (2023). Photovoltaic spectral responsivity and efficiency under different aerosol conditions. *Energies*, *16*(18), 6644. https://doi.org/10.3390/en16186644

Lauret, P., Boland, J., & Ridley, B. (2013). Bayesian statistical analysis applied to solar radiation modelling. *Renewable Energy*, *49*, 124–127. https://doi.org/10.1016/j.renene.2012.01.049

Liu, B. Y. H., & Jordan, R. C. (1960). The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. *Solar Energy (Phoenix, Ariz.)*, *4*(3), 1–19. https://doi.org/10.1016/0038-092x(60)90062-1

Logothetis, S.-A., Salamalikis, V., & Kazantzidis, A. (2020). Aerosol classification in Europe, Middle East, North Africa and Arabian Peninsula based on AERONET Version 3. *Atmospheric Research*, *239*(104893), 104893. https://doi.org/10.1016/j.atmosres.2020.104893

Long, C., & Dutton, E. (2010). *BSRN Global Network recommended QC tests, V2.x*. https://epic.awi.de/id/eprint/30083/1/BSRN_recommended_QC_tests_V2.pdf

Mayer, B., & Kylling, A. (2005). Technical note: The libRadtran software package for radiative transfer calculations - description and examples of use. *Atmospheric Chemistry and Physics*, *5*(7), 1855–1877. https://doi.org/10.5194/acp-5-1855-2005

McMahan, A. C., Grover, C. N., & Vignola, F. E. (2013). Evaluation of resource risk in solar-project financing. In *Solar Energy Forecasting and Resource Assessment* (pp. 81–95). Elsevier.

Owusu, P. A., & Asumadu-Sarkodie, S. (2016). A review of renewable energy sources, sustainability issues and climate change mitigation. *Cogent Engineering*, *3*(1), 1167990. https://doi.org/10.1080/23311916.2016.1167990

Papachristopoulou, K., Fountoulakis, I., Bais, A. F., Psiloglou, B. E., Papadimitriou, N., Raptis, I.-P., Kazantzidis, A., Kontoes, C., Hatzaki, M., & Kazadzis, S. (2024). Effects of clouds and aerosols on downwelling surface solar irradiance nowcasting and short-term forecasting. *Atmospheric Measurement Techniques*, *17*(7), 1851–1877. https://doi.org/10.5194/amt-17-1851-2024

Papachristopoulou, K., Fountoulakis, I., Gkikas, A., Kosmopoulos, P. G., Nastos, P. T., Hatzaki, M., & Kazadzis, S. (2022). 15-year analysis of direct effects of total and dust aerosols in solar radiation/energy over the Mediterranean Basin. *Remote Sensing*, *14*(7), 1535. https://doi.org/10.3390/rs14071535

Paulescu, E., & Blaga, R. (2019). A simple and reliable empirical model with two predictors for estimating 1-minute diffuse fraction. *Solar Energy (Phoenix, Ariz.)*, *180*, 75–84. https://doi.org/10.1016/j.solener.2019.01.029

Pedro, H. T. C., Inman, R. H., & Coimbra, C. F. M. (2017). Mathematical methods for optimized solar forecasting. In *Renewable Energy Forecasting* (pp. 111–152). Elsevier.

Pfenninger, S., & Staffell, I. (2016). Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy (Oxford, England)*, *114*, 1251–1265. https://doi.org/10.1016/j.energy.2016.08.060

Raptis, I.-P., Kazadzis, S., Fountoulakis, I., Papachristopoulou, K., Kouklaki, D., Psiloglou, B. E., Kazantzidis, A., Benetatos, C., Papadimitriou, N., & Eleftheratos, K. (2023). Evaluation of the solar energy nowcasting system (SENSE) during a 12-months intensive measurement campaign in Athens, Greece. *Energies*, *16*(14), 5361. https://doi.org/10.3390/en16145361

Qu, Z., Oumbe, A., Blanc, P., Espinar, B., Gesell, G., Gschwind, B., Klüser, L., Lefèvre, M., Saboret, L., Schroedter-Homscheidt, M., & Wald, L. (2017). Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method. *Meteorologische Zeitschrift*, *26*(1), 33–57. https://doi.org/10.1127/metz/2016/0781

Renewables Ninja. (n.d.). *Climate data interface*. GSEE Documentation. https://gsee.readthedocs.io/en/latest/climatedata-interface/

Ridley, B., Boland, J., & Lauret, P. (2010). Modelling of diffuse solar fraction with multiple predictors. *Renewable Energy*, *35*(2), 478–483. https://doi.org/10.1016/j.renene.2009.07.018

Schroedter-Homscheidt, M., Azam, F., Betcke, J., Hanrieder, N., Lefèvre, M., Saboret, L., & Saint-Drenan, Y. -M. (2022). Surface solar irradiation retrieval from MSG/SEVIRI based on APOLLO Next Generation and HELIOSAT-4 methods. *Meteorologische Zeitschrift*, *31*(6), 455–476. https://doi.org/10.1127/metz/2022/1132

Shettle, E. (1989). Models of aerosols, clouds, and precipitation for atmospheric propagation studies. In *Atmospheric propagation in the UV, visible, IR and mm-region and related system aspects* (AGARD Conference Proceedings No. 454). NATO Advisory Group for Aerospace Research and Development.

Stoffel, T. (2013). Terms and Definitions. In *Solar Energy Forecasting and Resource Assessment* (pp. 1–19). Elsevier.

Toledano, C., González, R., Fuertes, D., Cuevas, E., Eck, T. F., Kazadzis, S., Kouremeti, N., Gröbner, J., Goloub, P., Blarel, L., Román, R., Barreto, Á., Berjón, A., Holben, B. N., & Cachorro, V. E. (2018). Assessment of Sun photometer Langley calibration at the high-elevation sites Mauna Loa and Izaña. *Atmospheric Chemistry and Physics*, *18*(19), 14555–14567. https://doi.org/10.5194/acp-18-14555-2018

WMO. (2021). *Guide to instruments and methods of observation* (WMO-No. 8). https://library.wmo.int/doc_num.php?explnum_id=57838

Yang, D. (2019). SolarData package update v1.1: R functions for easy access of Baseline Surface Radiation Network (BSRN). *Solar Energy (Phoenix, Ariz.)*, *188*, 970–975. https://doi.org/10.1016/j.solener.2019.05.068

Supplement

Nomenclature

Acronym	**Definition**
~~AEMET~~	~~Agencia Estatal de Meteorología: Meteorological State Agency of Spain~~
AERONET	AErosol RObotic NETwork
AE	Angström Exponent
AOD	Aerosol Optical Depth
BRL	Boland-Ridley-Lauret diffuse fraction model
BSRN	Baseline Surface Radiation Network
CAMS	Copernicus Atmosphere Monitoring Service: ECMWF tool for atmospheric composition knowledge
~~CENER~~	~~Centro Nacional de Energías Renovables, National Renewable Energy Centre of Spain~~
DHI	Diffuse Horizontal Irradiance
DNI	Direct Normal Irradiance
DWD	Deutscher Wetterdienst: German Meteorological Service
ECMWF	European Centre for Medium-Range Weather Forecasts
FMF	Fine Mode Fraction
gg	Assymetry factor
GHI	Global Horizontal Irradiance
GSEE	Global Solar Energy Estimator
~~IARC~~	~~Izaña Atmospheric Research Center: AEMET Observatory of Izaña, Tenerife, Spain~~
MAE	Man Absolute Error
MBE	Mean Bias Error
~~Md~~	~~Median~~
MOL-RAO	Meteorologisches Observatorium Lindenberg, Richard-Aßmann-Observatorium: DWD Observatory Lindenberg, Lindenberg (Tauche), Germany
NEO	Navarino Environmental Observatory, Messinia, Greece

~~OMN~~	~~Office national de la météorologie, "Météo Algérie": National Meteorological Office of Algeria~~
PDFs	Probability Density Functions
PMOD/WRC	Physikalisch-Meteorologisches Observatorium Davos / World Radiation Center, Davos, Switzerland
PV	Photovoltaic
QC	Quality Check
r	Weighted correlation coefficient
rMBE	Relative Mean Bias Error
RMSE	Root Mean Square Error
RTM	Radiative Transfer Model
SDA	Spectral Deconvolution Algorithm
SSA	Single Scattering Albedo
SV	Solar Visibility
SZA	Solar Zenith Angle
TOC	Total Ozone Column (in DU)
TPM	Faculty of Technology, Policy, and Management, Delft, the Netherlands
UTC	Coordinated Universal Time
WMO	World Meteorological Organisation

Evaluation Metrics

The formulas for the evaluation metrics used are the following:

1. Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum(x_{mod} - x_{obs})^2}$$

2. Mean Absolute Error (MAE)

$$MAE = \frac{1}{N}\sum|x_{mod} - x_{obs}|$$

3. relative Mean Bias Error (rMBE)

$$rMBE = \frac{1}{N}\sum\left(\frac{x_{mod} - x_{obs}}{x_{obs}}\right) \times 100\%$$

…

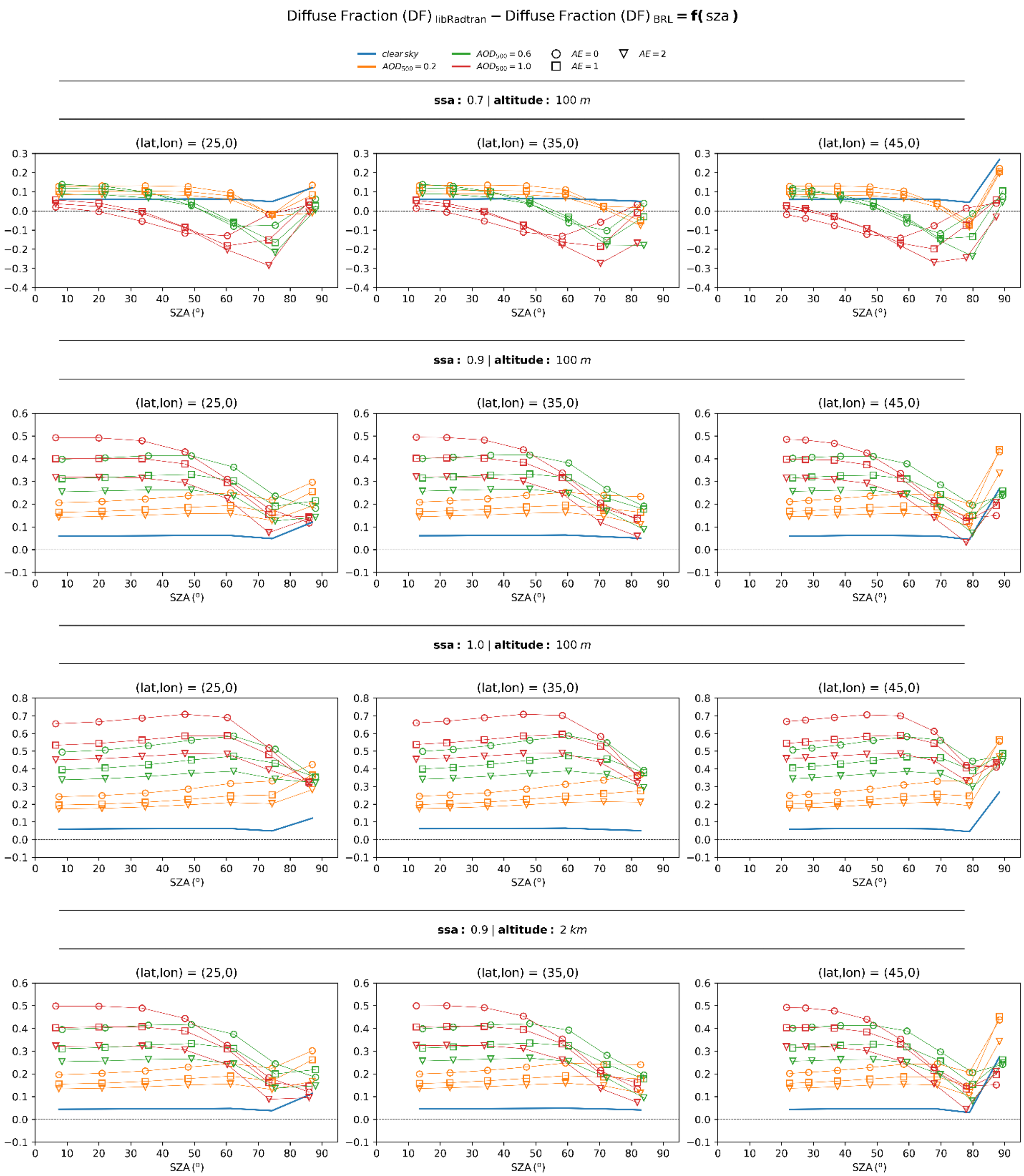

Figure S1. Difference between the diffuse fraction derived directly from the computations of DHI and GHI using libRadtran and the one estimated by applying BRL to the libRadtran-computed GHI for surface albedo 0.8

...

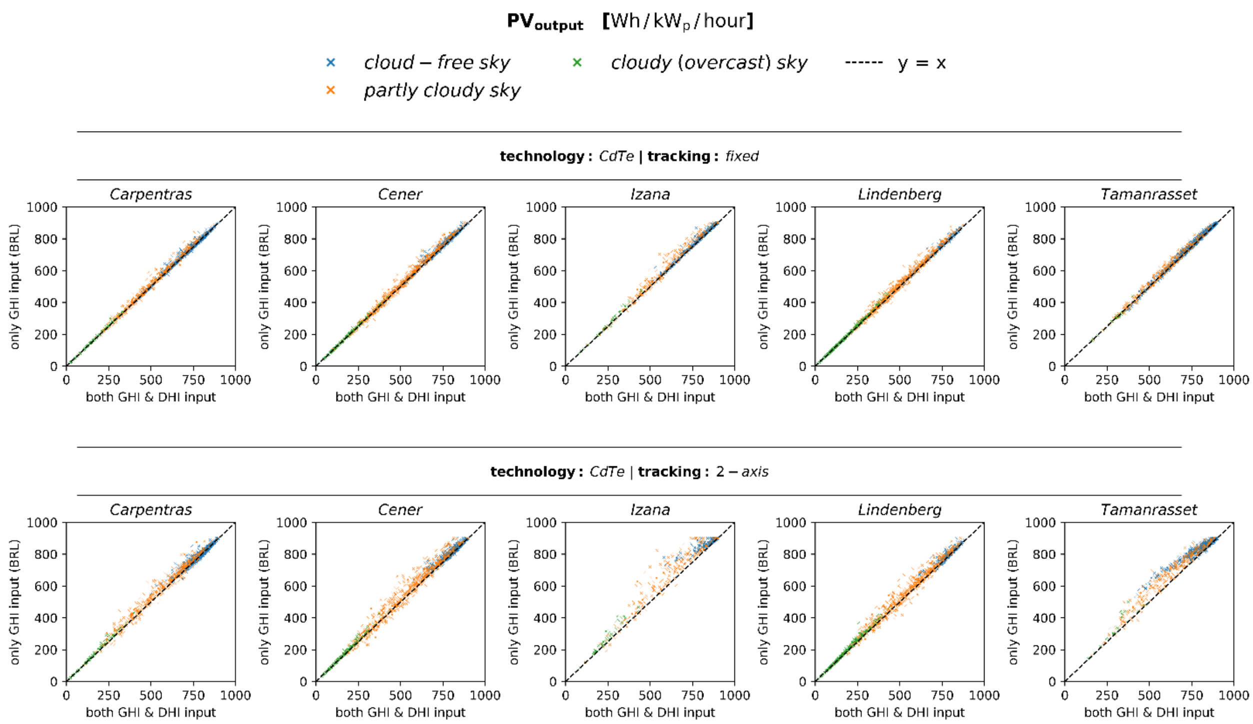

Figure S2. Comparison of the estimated hourly PV power generation between simulations performed using GSEE with input data consisting of either only GHI or both GHI and DHI under varying cloudiness conditions for panels with CdTe technology

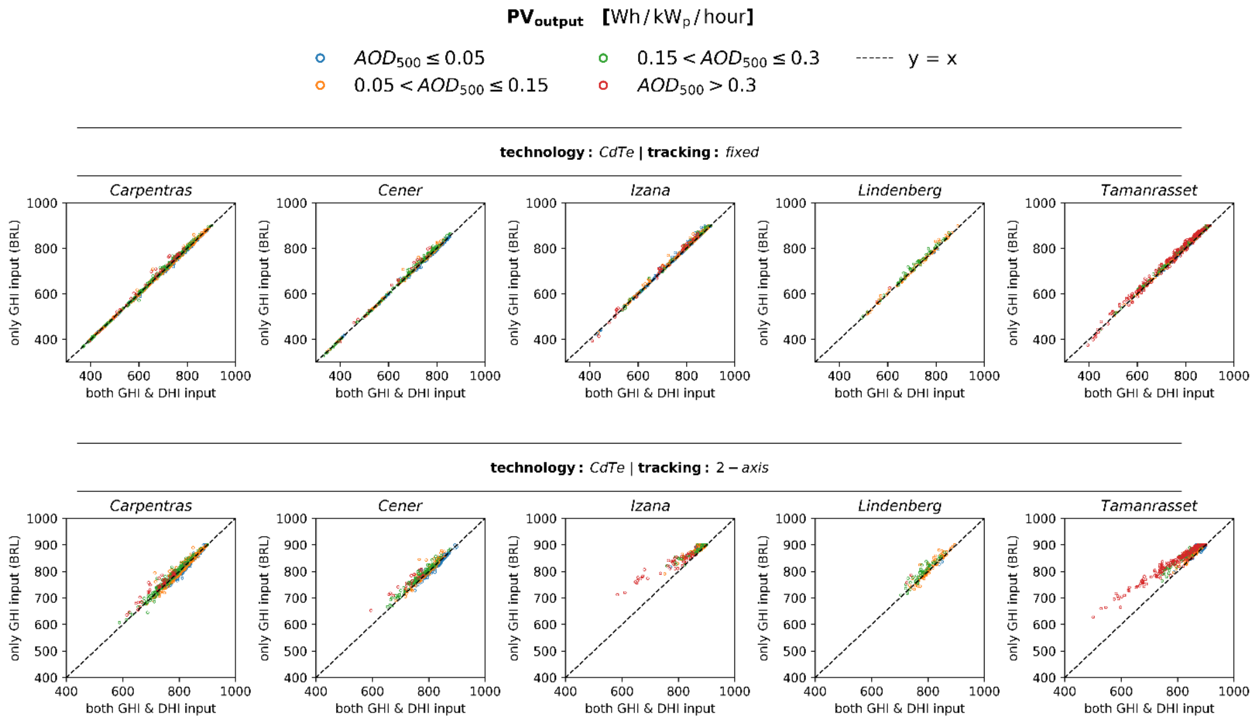

Figure S3. Comparison of the estimated hourly PV power generation between simulations performed using GSEE with input data consisting of either only GHI or both GHI and DHI under varying aerosol conditions for panels with CdTe technology

~~**Table S1.** Evaluation metrics for GSEE performance within hourly intervals in Carpentras, comparing simulations with diffuse fraction from measurements and from the BRL model~~

~~STATION: Carpentras~~		~~fixed panels~~			~~2-axis tracking~~		
		~~RMSE (Wh/kWp/hour)~~	~~MAE (Wh/kWp/hour)~~	~~rMBE (%)~~	~~RMSE (Wh/kWp/hour)~~	~~MAE (Wh/kWp/hour)~~	~~rMBE (%)~~
~~All-Sky scenes~~		~~12.6~~	~~6.6~~	~~0.8~~	~~20.8~~	~~12.5~~	~~1.2~~
~~All-Sky scenes (cloudiness)~~	~~cloud-free~~	~~9.2~~	~~4.6~~	~~0.4~~	~~14.8~~	~~8.7~~	~~0.5~~
	~~partly cloudy~~	~~19.5~~	~~12.5~~	~~2.3~~	~~32.5~~	~~23.9~~	~~3.8~~
	~~cloudy (overcast)~~	~~5.8~~	~~3.0~~	~~2.0~~	~~10.5~~	~~6.1~~	~~4.6~~
~~Cloudless-Sky scenes (aerosol load)~~	~~low~~	~~4.7~~	~~3.4~~	~~-0.4~~	~~9.5~~	~~7.5~~	~~-0.8~~
	~~moderate~~	~~4.3~~	~~2.2~~	~~0.1~~	~~7.8~~	~~4.7~~	~~0.0~~
	~~high~~	~~6.4~~	~~4.0~~	~~0.6~~	~~11.0~~	~~7.8~~	~~0.9~~
	~~very high~~	~~14.9~~	~~10.2~~	~~1.6~~	~~22.7~~	~~17.2~~	~~2.6~~

~~**Table S2.** Evaluation metrics for GSEE performance within hourly intervals in Tamanrasset, comparing simulations with diffuse fraction from measurements and from the BRL model.~~

~~STATION: Tamanrasset~~		~~fixed panels~~			~~2-axis tracking~~		
		~~RMSE (Wh/kWp/hour)~~	~~MAE (Wh/kWp/hour)~~	~~rMBE (%)~~	~~RMSE (Wh/kWp/hour)~~	~~MAE (Wh/kWp/hour)~~	~~rMBE (%)~~
~~All-Sky scenes~~		~~13.6~~	~~9.3~~	~~1.0~~	~~40.4~~	~~27.8~~	~~3.8~~
~~All-Sky scenes (cloudiness)~~	~~cloud-free~~	~~11.5~~	~~8.0~~	~~0.8~~	~~35.3~~	~~23.4~~	~~2.9~~
	~~partly cloudy~~	~~20.1~~	~~15.0~~	~~2.0~~	~~56.1~~	~~45.7~~	~~8.1~~
	~~cloudy (overcast)~~	~~8.4~~	~~5.2~~	~~-0.1~~	~~45.3~~	~~30.1~~	~~11.2~~
~~Cloudless-Sky scenes (aerosol load)~~	~~low~~	~~3.2~~	~~2.0~~	~~0.2~~	~~6.6~~	~~4.0~~	~~0.3~~
	~~moderate~~	~~5.4~~	~~4.6~~	~~0.6~~	~~13.0~~	~~10.5~~	~~1.2~~

Table ~~S3~~S1. Evaluation metrics for GSEE performance within hourly intervals in Cener, comparing simulations with diffuse fraction from measurements and from the BRL model.

STATION: Cener		fixed panels			2-axis tracking		
		RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)	RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)
All-Sky scenes		14.5	8.2	1.2	27.1	16.7	2.3
All-Sky scenes (cloudiness)	cloud-free	11.7	6.4	0.8	19.5	11.9	1.3
	partly cloudy	19.3	12.4	2.0	37.5	26.4	4.1
	cloudy (overcast)	4.7	2.7	1.6	11.2	6.3	4.6
Cloudless-Sky scenes (aerosol load)	clear sky / low	4.0	2.5	-0.2	7.9	5.5	-0.4
	moderate	6.9	3.1	0.4	11.4	6.2	0.6
	high	8.7	6.2	1.0	15.4	12.8	1.8
	very high	NaN	NaN	NaN	NaN	NaN	NaN

Table ~~S4~~S2. Evaluation metrics for GSEE performance within hourly intervals in Lindenberg, comparing simulations with diffuse fraction from measurements and from the BRL model.

STATION: Lindenberg		fixed panels			2-axis tracking		
		RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)	RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)
All-Sky scenes		16.0	9.7	1.8	26.1	17.0	2.7
All-Sky scenes (cloudiness)	cloud-free	11.8	6.7	0.9	20.7	13.5	1.4
	partly cloudy	20.4	13.9	2.3	32.4	23.5	3.5
	cloudy (overcast)	6.7	3.6	2.4	11.7	6.5	4.6

Cloudless-Sky scenes (aerosol load)	clear sky / low	NaN	NaN	NaN	NaN	NaN	NaN
	moderate	8.9	5.5	0.5	14.9	9.7	0.6
	high	12.6	10.3	1.4	19.5	15.9	2.0
	very high	NaN	NaN	NaN	NaN	NaN	NaN

Table ~~S5~~S3. Evaluation metrics for GSEE performance within hourly intervals in Izana, comparing simulations with diffuse fraction from measurements and from the BRL model.

STATION: Izana		fixed panels			2-axis tracking		
		RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)	RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)
All-Sky scenes		20.0	11.3	1.5	41.5	22.3	2.8
All-Sky scenes (cloudiness)	cloud-free	12.3	7.2	0.9	26.6	12.7	1.5
	partly cloudy	36.1	25.8	4.3	73.4	55.1	9.3
	cloudy (overcast)	16.8	11.8	4.6	35.5	26.0	11.8
Cloudless-Sky scenes (aerosol load)	clear sky / low	6.8	4.8	0.6	7.8	3.7	0.4
	moderate	9.3	6.5	0.9	20.8	15.0	1.8
	high	11.2	8.6	1.1	31.4	26.2	3.3
	very high	14.1	11.8	1.4	64.5	52.1	7.3

....

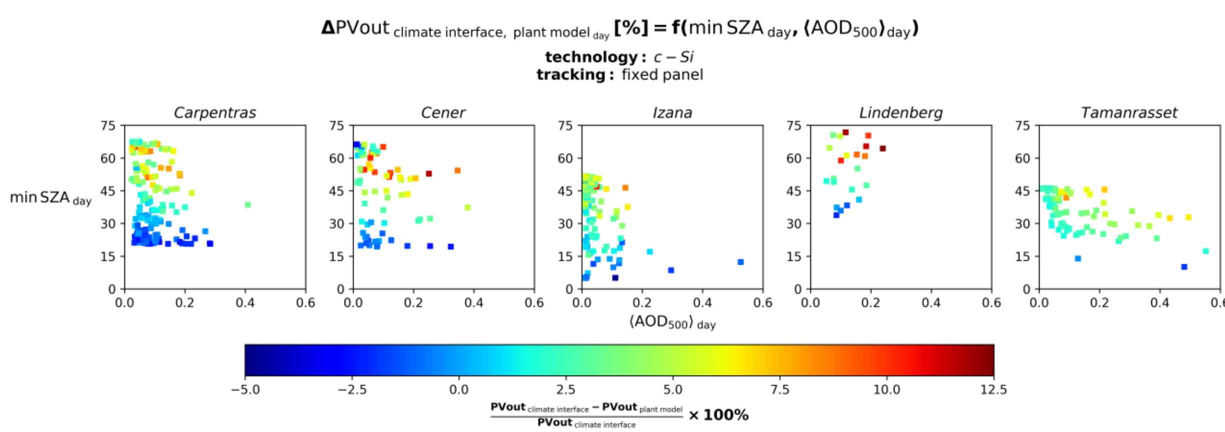

Figure S4. Percentage differences between the daily PV output estimated using the climate interface and the corresponding daily sums from hourly simulations as function of minimum daily SZA and mean daily aerosol load for fixed panels

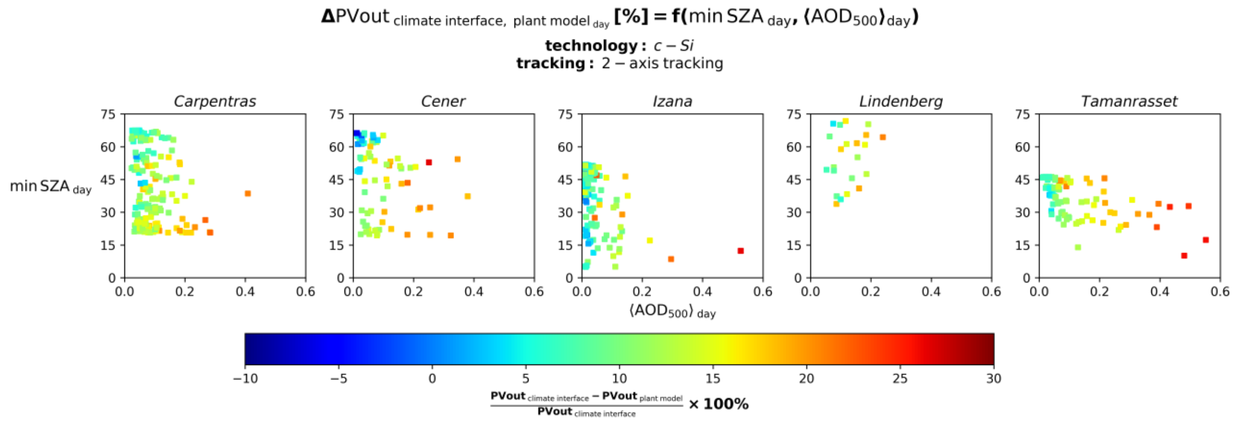

Figure S5. Percentage differences between the daily PV output estimated using the climate interface and the corresponding daily sums from hourly simulations as function of minimum daily SZA and mean daily aerosol load for panels with 2-axis tracking

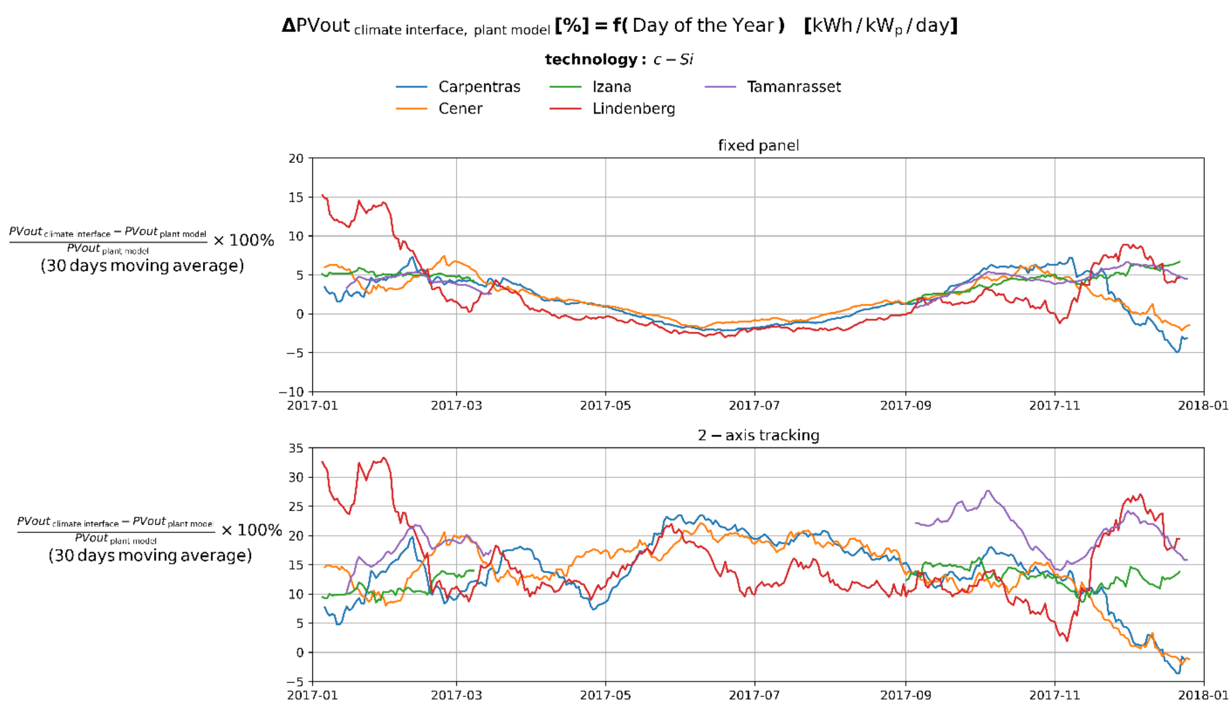

Figure S6. Time-series of the percentage differences between the daily PV output estimated using the climate interface and the corresponding daily sums from hourly simulations

...

Table ~~S6~~S4. Evaluation metrics assessing the reliability of GSEE Climate Data Interface in estimating total daily PV output power in Carpentras

STATION: Carpentras		fixed panels			2-axis tracking		
		RMSE (kWh/kWp/day)	MAE (kWh/kWp/day)	rMBE (%)	RMSE (kWh/kWp/day)	MAE (kWh/kWp/day)	rMBE (%)
All Days		0.22	0.17	1.7	1.24	1.08	15.5
Sunny (cloudless) Days		0.19	0.15	1.8	1.19	1.08	14.1
Sunny Days: average aerosol load	very-low aerosol	0.18	0.15	2.1	0.82	0.72	10.0
	aerosol-laden	0.19	0.16	1.7	1.28	1.20	15.3

Table ~~S7~~S5. Evaluation metrics assessing the reliability of GSEE Climate Data Interface in estimating total daily PV output power in Tamanrasset

STATION: Tamanrasset		fixed panels			2-axis tracking		
		RMSE (kWh/kWp/day)	MAE (kWh/kWp/day)	rMBE (%)	RMSE (kWh/kWp/day)	MAE (kWh/kWp/day)	rMBE (%)
All Days		0.24	0.22	3.4	1.45	1.34	19.0
Sunny (cloudless) Days		0.22	0.20	3.4	1.20	1.11	14.4
Sunny Days: average aerosol load	very-low aerosol	0.15	0.14	2.5	0.66	0.62	7.7
	aerosol-laden	0.24	0.22	3.7	1.37	1.31	17.3

Table ~~S8~~S6. Evaluation metrics assessing the reliability of GSEE Climate Data Interface in estimating total daily PV output power in Cener

STATION: Cener		fixed panels			2-axis tracking		
		RMSE (kWh/kWp/day)	MAE (kWh/kWp/day)	rMBE (%)	RMSE (kWh/kWp/day)	MAE (kWh/kWp/day)	rMBE (%)
All Days		0.24	0.18	2.5	1.28	1.08	16.8
Sunny (cloudless) Days		0.26	0.21	3.3	1.15	1.00	13.6
Sunny Days (average aerosol load)	aerosol-free	0.18	0.15	2.4	0.73	0.60	7.6
	aerosol-laden	0.30	0.24	4.0	1.37	1.28	17.8

Table ~~S9~~S7. Evaluation metrics assessing the reliability of GSEE Climate Data Interface in estimating total daily PV output power in Lindenberg

STATION: Lindenberg		fixed panels			2-axis tracking		
		RMSE (kWh/kWp/day)	MAE (kWh/kWp/day)	rMBE (%)	RMSE (kWh/kWp/day)	MAE (kWh/kWp/day)	rMBE (%)
All Days		0.24	0.18	2.1	0.99	0.81	15.2
Sunny (cloudless) Days		0.29	0.23	4.4	1.04	0.96	14.6
Sunny Days (average aerosol load)	aerosol-free	NaN	NaN	NaN	NaN	NaN	NaN
	aerosol-laden	NaN	NaN	NaN	NaN	NaN	NaN

Table ~~S10~~S8. Evaluation metrics assessing the reliability of GSEE Climate Data Interface in estimating total daily PV output power in Izana

STATION: Izana		fixed panels			2-axis tracking		
		RMSE (kWh/kWp/day)	MAE (kWh/kWp/day)	rMBE (%)	RMSE (kWh/kWp/day)	MAE (kWh/kWp/day)	rMBE (%)
All Days		0.28	0.23	3.4	1.12	0.94	11.2
Sunny (cloudless) Days		0.25	0.21	3.1	0.92	0.80	8.9
Sunny Days (average aerosol load)	aerosol-free	0.25	0.22	3.3	0.75	0.65	7.2
	aerosol-laden	0.24	0.19	2.3	1.38	1.31	15.3

…

Table ~~S11~~S9. Evaluations metrics accessing the reliability of using CAMS solar radiation time-series for modelling PV output power in Carpentras

STATION: Carpentras		fixed panels			2-axis tracking		
		RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)	RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)
All-Sky scenes		49.7	36.6	3.9	60.7	43.0	3.8
All-Sky scenes (cloudiness)	cloud-free	35.9	28.4	3.2	41.7	32.5	2.7
	partly cloudy	74.1	56.7	5.7	94.2	70.3	7.1
	cloudy (overcast)	46.7	37.4	11.4	49.2	38.3	13.9
Cloudless-Sky scenes (aerosol load)	low	25.0	20.0	2.1	28.8	22.3	1.5
	moderate	32.5	25.9	3.4	36.7	29.1	2.9
	high	41.8	36.0	5.0	49.0	41.3	4.7
	very high	42.3	36.9	6.1	49.9	44.1	6.3

Table ~~S12~~S10. Evaluations metrics accessing the reliability of using CAMS solar radiation time-series for modelling PV output power in Tamanrasset

STATION: Tamanrasset		fixed panels			2-axis tracking		
		RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)	RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)
All-Sky scenes		55.9	33.7	-1.0	75.2	46.1	-0.6
All-Sky scenes (cloudiness)	cloud-free	32.4	20.8	-0.7	45.2	29.8	-0.7
	partly cloudy	87.3	67.5	-4.4	111.8	86.9	-3.0
	cloudy (overcast)	124.9	89.2	23.3	210.7	151.6	48.1
Cloudless-Sky scenes (aerosol load)	low	16.6	11.9	0.3	20.8	15.3	-0.6
	moderate	19.7	14.9	-0.4	28.2	20.9	-1.2
	high	29.8	18.9	-0.9	34.9	23.2	-1.0
	very high	31.0	22.9	0.7	48.0	35.5	2.4

Table ~~S13~~S11. Evaluations metrics accessing the reliability of using CAMS solar radiation time-series for modelling PV output power in Cener

STATION: Cener		fixed panels			2-axis tracking		
		RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)	RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)
All-Sky scenes		63.1	43.0	2.0	75.3	51.2	2.1
All-Sky scenes (cloudiness)	cloud-free	35.6	27.0	0.9	45.6	33.5	0.4
	partly cloudy	82.7	61.1	2.1	99.6	73.4	3.4
	cloudy (overcast)	77.3	50.8	20,7	83.5	53.4	24.8
Cloudless-Sky scenes (aerosol load)	clear sky / low	28.0	21.5	1.7	32.1	25.0	1.1
	moderate	38.5	30.4	2.8	48.3	38.8	2.3
	high	34.3	28.4	2.9	41.7	35.1	2.3
	very high	NaN	NaN	NaN	NaN	NaN	NaN

Table ~~S14~~S12. Evaluations metrics accessing the reliability of using CAMS solar radiation time-series for modelling PV output power in Lindenberg

STATION: Lindenberg		fixed panels			2-axis tracking		
		RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)	RMSE (Wh/kWp/hour)	MAE (Wh/kWp/hour)	rMBE (%)
All-Sky scenes		66.1	46.7	-1.2	76.2	53.8	-1.4
	cloud-free	37.9	24.9	-1.7	50.7	33.4	-2.5

All-Sky scenes (cloudiness)	partly cloudy	76.4	57.5	-2.6	88.9	67.0	-2.3
	cloudy (overcast)	60.9	42.3	8.0	63.9	43.2	9.5
Cloudless-Sky scenes (aerosol load)	clear sky / low	NaN	NaN	NaN	NaN	NaN	NaN
	moderate	42.3	28.7	-2.5	53.4	24.7	-3.0
	high	40.9	26.4	-2.4	52.6	32.0	-2.8
	very high	NaN	NaN	NaN	NaN	NaN	NaN