Enhanced site-adaptation for reliable PV assessment: Comparative study for Brazil and Germany

Gustavo X. A. Pinto[1,3], Helena F. Naspolini[1], Oliver Pfeiffer[2], Marília Braga[1], Amanda M. F. Gomes[1,3], Rafael A. Campos[1], Ulf Blieske[2], Ingo Stadler[2] and Ricardo Rüther[1]

1 Solar Energy Research Laboratory/Universidade Federal de Santa Catarina, Florianópolis (Brazil)

2 Cologne Institute for Renewable Energy/TH Köln-University of Applied Sciences, Cologne (Germany)

3 Visiting Scholar at TH Köln-University of Applied Sciences, Cologne (Germany)

Abstract

While satellite-based irradiance methods typically offer high confidence levels, they may still show considerable discrepancies when compared with ground-based measurements. To mitigate this and improve the reliability of localized generation estimates, satellite data correction techniques (site adaptation) that integrate short-term ground measurements are employed. This paper proposes a method for enhancing the linear regression site adaptation method by incorporating solar irradiance band separation and clear-sky classification index (Kc) analysis. The methodology begins with data acquisition and control. Subsequently, a site adaptation method is applied, supplemented by a Kc analysis, data segmentation based on solar irradiance bands and error metrics are calculated for each case. A comparison analysis between a location in Brazil and in Germany is carried out. In all cases, the proposed methodology was able to significantly improve satellite bias on solar resource assessment. Additionally, results show that using two years of measured data offers significant improvements over a one-year period and could enhance the reliability of solar generation long term forecasts.

Keywords: Site adaptation, Solar radiation, Linear regression, Solar energy

1. Introduction

The technical and financial feasibility of photovoltaic (PV) solar modules is directly related to the assessment of the available solar resource. Accurate data estimation in the target region is critical for producing reliable energy generation forecasts. Accurate projections are crucial in energy auctions, as they bolster investor and financier confidence by minimizing the risks associated with under or overestimation of energy generation.

In Brazil, the Energy Research Company (EPE) establishes technical standards and requirements that solar power plants must fulfill to achieve certification. Measurement efforts are focused on specific locations in preparation for energy auction projects and require a minimum data collection period of one year (EPE, 2025). In Germany, PV yield predictions are legally binding, if they are part of the contract between a solar project developer and a costumer (Higher Regional Court, Munich, of 1/21/2014). Therefore, precise forecasts are of great interest to PV project developers. However, a measurement period of one-year may not adequately reflect the typical meteorological conditions at a given location. In the planning phase of solar PV power plants, it is essential to account for at least ten years of solar resource data at the site (Mieslinger et al., 2014).

Given this context, the objective of this study is to validate the site adaptation method presented by Polo et al. (2015) and Aguiar et al. (2019) to Brazil and Germany. Additionally, the clear-sky index analysis proposed by Aguiar et al. (2019) will be applied, alongside the method proposed by Pinto et al. (2023; 2024), for dividing the analyses by irradiance ranges is proposed and tested.

2. Data acquisition

Data acquisition for this study was carried out using both satellite data and ground-based measurements. The analysis period spans a total of six years, from January 1, 2018, to December 31, 2023. This continuity in both

10.18086/swc.2025.04.18 Available at http://proceedings.ises.org

the measured and satellite data sets, along with the considered temporal resolution (30 min), facilitates the analysis and comprehension of trends throughout the considered period. The satellite data used in this study for both locations were obtained from the Solcast® solar irradiance historical data provider (*https://solcast.com/*). The methodology and validation of the satellite-derived irradiance dataset from Solcast® are detailed by Bright (2019).

2.1. Brazil

The measured GHI data were obtained from a Kipp & Zonen pyranometer, model SMP22, installed at the solar radiation measurement station of the UFSC Solar Energy Research Laboratory (27.43°S, 48.44°W), located in Florianópolis, Brazil, as shown in Figure 1. For a more detailed and comprehensive description, additional information is available in Mantelli et al. (2019), which outlines the technical and operational aspects of the meteorological station's implementation.

(a)

(b)

Figure 1: UFSC Solar Energy Research Laboratory (a) and its meteorological and solar radiation measurement station (b).

2.2. Germany

The German irradiance data were obtained from the German Weather Service (DWD), which runs a network of weather stations throughout Germany. The weather station used for this study is located in southern Germany (station ID 2638, Klippeneck), and the data can be accessed via the CDC portal (Deutscher Wetterdienst 2025a). The irradiance is measured using a scanning pyranometer/pyrheliometer (ScaPP), which was installed at a height of 986 meters above sea level in 2016. A SCaPP differs from a conventional pyranometer in that it utilizes a small solar cell instead of a thermal sensor (Deutscher Wetterdienst 2025b). The standardized dataset is described in detail in (Deutscher Wetterdienst 2021) and the quality and uncertainty of the values are evaluated in (Kaspar et al. 2013).

2.3. Data availability

Figure 2 illustrates the timeline of satellite-derived and ground-based data availability at a 30-minute temporal resolution.

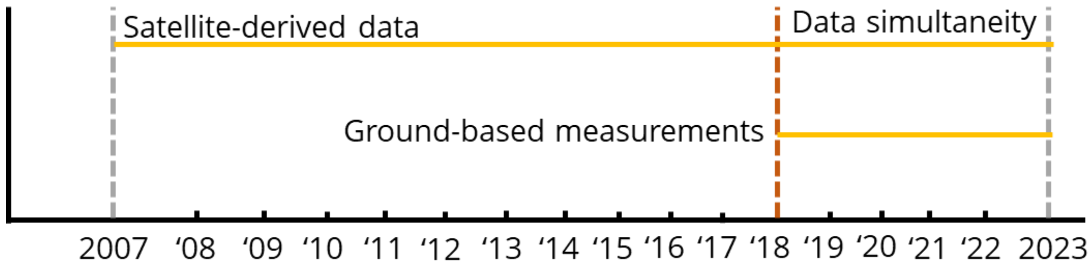


Figure 2: Data availability timeline.

3. Methodology

In this study, the linear regression method is employed to estimate correction factors that mitigate the bias between the two data sets, derived from a scatter plot of ground-based measurements and satellite data. The points of the linear fit exhibit deviations from the ideal linear regression. The linear regression derived from the scatter plot between satellite and ground-based data is estimated using the least squares method. These regression coefficients are applied to achieve an optimal adjustment of the satellite data, as outlined in Equation (1). Once the corrected satellite data are evaluated over the period of the measured data acquisition, it is essential to extend the correction to the entire historical data set. This is achieved by generating a scatter plot between the corrected satellite data and the original satellite data. The updated linear regression coefficients derived from this relationship are applied according to Equation (2).

$$\boldsymbol{Y_{satellite}^{corr} = Y_{satellite} - \left[(a-1) * X_{ground\_measured} + b\right]} \quad \text{(eq. 1)}$$

$$\boldsymbol{Y_{satellite}^{new} = C * Y_{satellite} + D} \quad \text{(eq. 2)}$$

For this analysis, the ground-based measurement dataset is divided into different sky classification categories. The aim is to achieve greater accuracy for clear sky days, as these are the days when the PV system generates the most energy. Clear sky days were selected based on the Clear Sky Index (Kc), presented by Equation (3). To calculate the clear sky index, ground-based measured data were used, along with clear sky irradiance implemented using *PVlib* library (Jensen et al., 2023). Each sky condition undergoes an independent linear regression, and the historical dataset is then created by combining these datasets. The Kc classification ranges (Babar et al., 2019; Smith et al., 2017; Engerer and Mills, 2014) are presented in Table 1. If Kc > 0.8, it's a clear-sky; if 0.4 ≤ Kc ≤ 0.8, its intermediate cloudy-sky; and if Kc ≤ 0.4, it's cloudy-sky.

$$\boldsymbol{Kc = (Ground\ Based\ Measured\ Irradiation\ )/(Clear\ Sky\ Daily\ Irradiation)} \quad \text{(eq. 3)}$$

Table 1: Kc classification.

Sky condition	Kc range
Clear-sky	> 0.8
Intermediate cloudy-sky	0.4 > Kc ≥ 0.8
Cloudy-sky	0 ≥ Kc ≥ 0.4

This work continues to apply an additional division of the dataset, stratified by irradiance bands (Pinto et al. 2023; Pinto et al., 2024). The measured data are segmented into groups, each corresponding to a 200 W/m² irradiance interval, ranging from 0 to 1,200 W/m², resulting in six distinct groups. Following this segmentation, the linear regression method was independently applied to each group, which, when aggregated, constitutes the complete historical dataset.

Additionally, two error evaluation metrics were employed in comparing the ground-measured data with the satellite data adapted for the corresponding analysis period. The Mean Bias Error (MBE) is presented in Equation (4), indicating the direction in which the model deviates and the Root Mean Square Error (RMSE), which provides insights into the data dispersion relative to the best-fit line, is presented in Equation (5). To determine the relative errors, %rMBE and %rRMSE, the absolute errors were divided by the average GHI values measured on the ground.

$$\boldsymbol{MBE = \frac{1}{N}\left[\sum_{i=1}^{N}\left(GHI_{satellite,i} - GHI_{ground\_measured,i}\right)\right]} \qquad \text{(eq. 4)}$$

$$\boldsymbol{RMSE = \sqrt{\frac{1}{N}\left[\sum_{i=1}^{N}\left(GHI_{satellite,i} - GHI_{ground\_measured,i}\right)^{2}\right]}} \qquad \text{(eq. 5)}$$

Figure 3 displays the overview of the methodology.

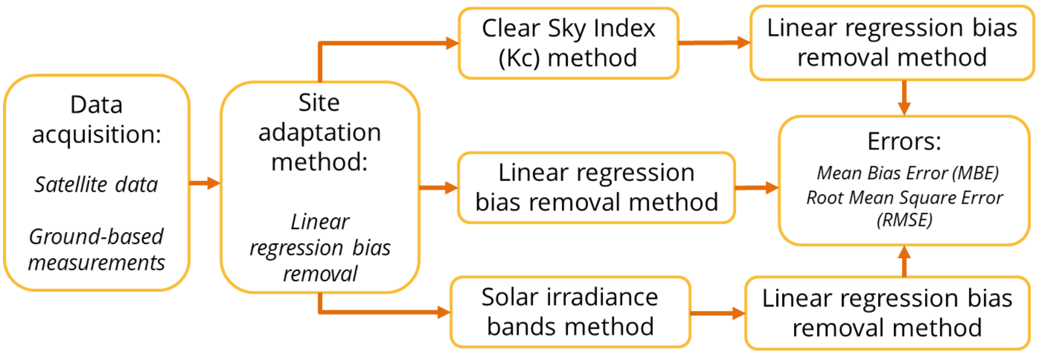


Figure 3: Overview of the applied method.

4. Results

In order to evaluate the Kc profile for each site, Figures 4 and 5 illustrate the distribution of days as classified by the Kc for the measurement period of 2018 throughout 2023 for the analyzed sites in Brazil and Germany, respectively. The cumulative data for Germany appear relatively constant for Kc values between 0.2 and 0.9. In contrast, in Brazil, the number of days tends to increase with each 10% increment in Kc values up to 1.0. At the site in Brazil, clear sky conditions were observed on 48% of the days during the study period, while intermediate cloudy conditions occurred on 34% of the days, and fully cloudy conditions on the remaining 18%. In contrast, the Germany site experienced clear sky days on 36% of the total, with intermediate cloudy days making up 43%, and cloudy days accounting for the remaining 21%.

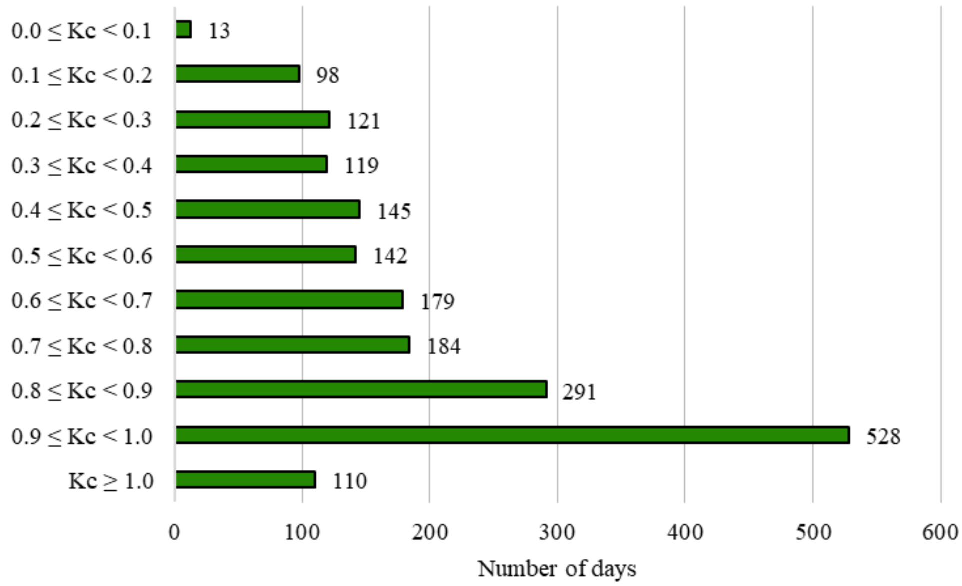


Figure 4: Distribution of days classified by the Clear Sky Index (Kc) in Brazil for 2018-2023.

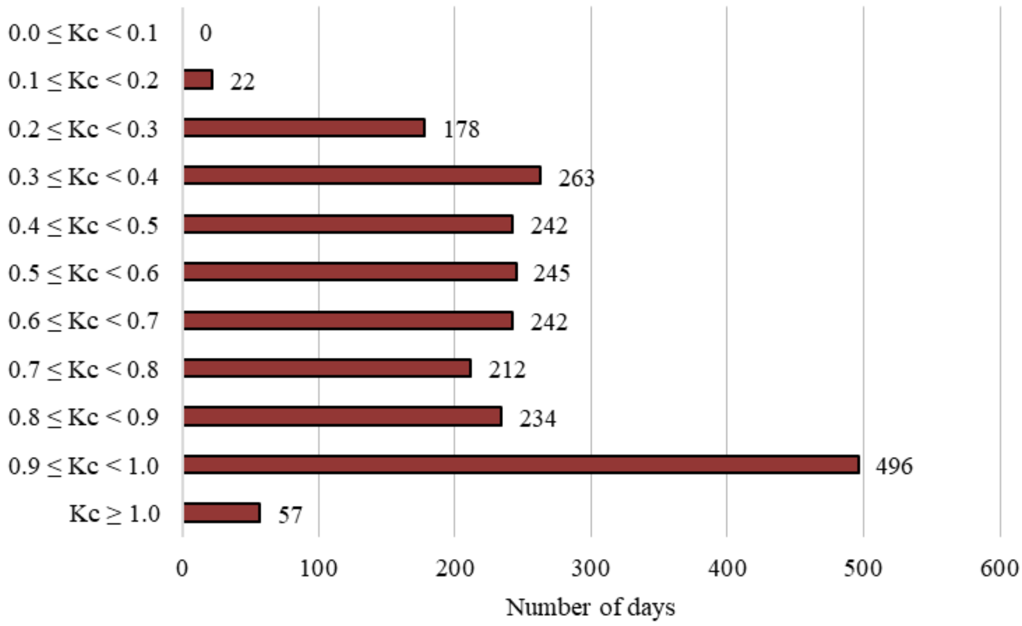
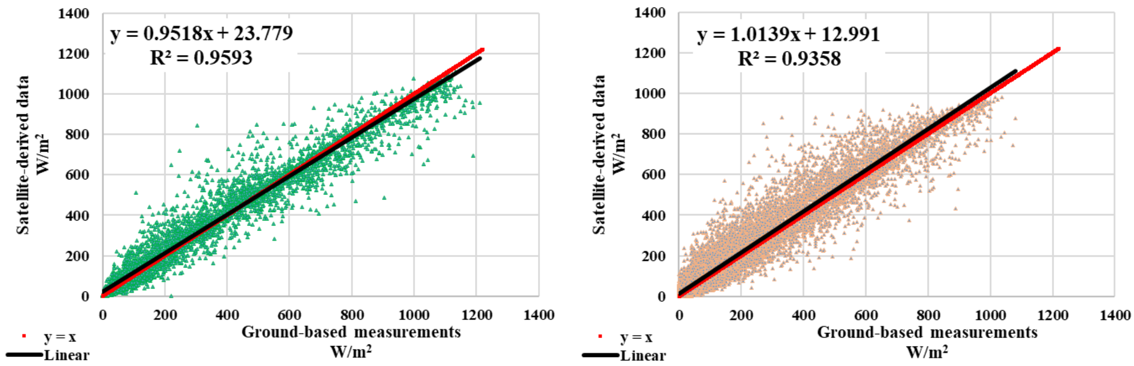


Figure 5: Distribution of days classified by the Clear Sky Index (Kc) in Germany for 2018-2023.

As an exemplification of the method, Figure 6 presents for Brazil (Fig. 6a) and Germany (Fig. 6b) the scatter plots between the satellite-derived data and the ground-based measurements for the year 2023. For the Brazilian site, the chart suggests that satellite data tend to overestimate lower irradiance values and underestimate higher irradiance values. In contrast, for the German site, the satellite data systematically overestimate irradiance values across the entire range.

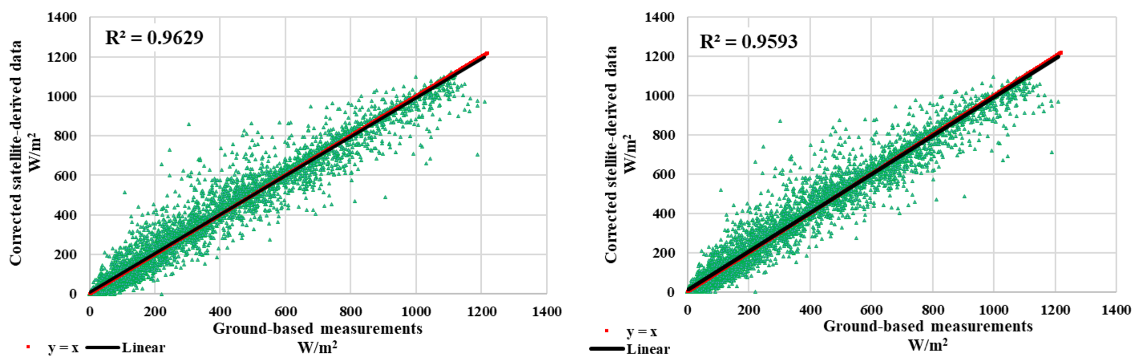


(a) (b)

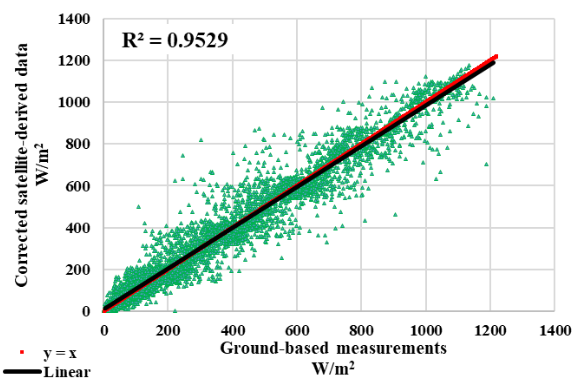
Figure 6. Scatter plots of the available data in 2023 for Brazil (a) and Germany (b).

Figure 7 presents the scatter plots and coefficients of determination (R^2) between the corrected satellite-derived data and ground-based measurements for the methods evaluated at the Brazilian site. The results show that the R^2 values vary by no more than 0.01 across the methods, indicating that all approaches exhibit comparable and satisfactory performance, as R^2 serves as a statistical measure of the goodness of fit between the model and the observed data. Additionally, Figure 7(c) (solar irradiance band method) reveals a distinct “step” pattern occurring at intervals of 200 W/m² along the y-axis.

Figure 8 presents the scatter plots and coefficients of determination (R^2) between the corrected satellite-derived data and ground-based measurements for the methods evaluated at the German site. The results indicate that the R^2 values vary by no more than 0.037 across the methods, which is approximately four times the variation observed for Brazil. Figure 8(c) also exhibits the characteristic “step” pattern. In comparison, the results for Germany show a greater concentration of data in the range up to 600 W/m², with only the solar irradiance band method (Fig. 8c) yielding corrected satellite-derived values exceeding 1,000 W/m².

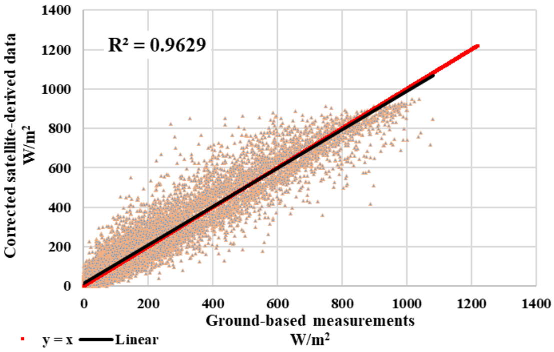


(a) (b)

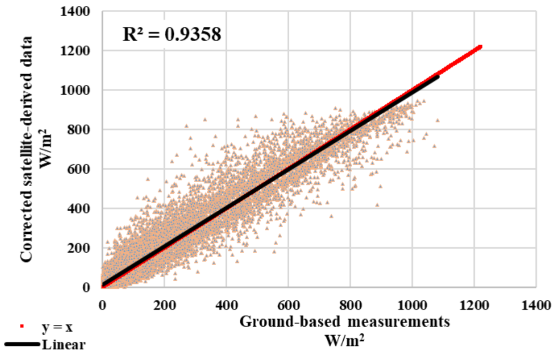


(c)

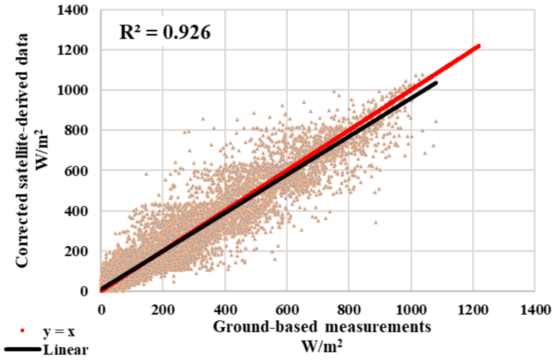
Figure 7. Scatter plots of the corrected satellite-derived data and the ground-based measurement data in 2023 for Brazil for the original method (a), Kc classification (b) and solar irradiance bands (c).



(a)



(b)



(c)

Figure 8. Scatter plots of the corrected satellite-derived data and the ground-based measurement data in 2023 for Germany for the original method (a), Kc classification (b) and solar irradiance bands (c).

The rMBE and rRMSE values for the investigated site adaptation methods in the analyzed sites in Brazil and Germany are shown in Table 2 and 3 (best results highlighted) and takes into account a time frame of 1 to 6 years of ground-based measured data.

Table 2. rMBE(%) metrics for each site adaptation method utilized for the collective period spanning 1 to 6 years (best results highlighted).

Year	Brazil				Germany			
	No data correction	Original method	Kc Classification	Solar irradiance bands	No data correction	Original method	Kc Classification	Solar irradiance bands
Error rMBE (%)								
1 year (23)	1.99	1.04	1.14	0.31	11.26	6.39	6.49	3.22
2 years (23-22)	0.08	0.75	0.86	0.30	7.64	5.63	5.98	2.92
3 years (23-21)	0.62	0.68	0.73	0.31	6.29	5.52	6.03	3.11
4 years (23-20)	0.43	0.67	0.72	0.31	7.00	5.39	6.06	3.25
5 years (23-19)	-0.16	0.57	0.62	0.31	8.07	5.53	6.26	3.43
6 years (23-18)	-0.88	0.45	0.57	0.32	8.06	5.58	6.32	3.49

Table 3. rRMSE(%) metrics for each site adaptation method utilized for the collective period spanning 1 to 6 years (best results highlighted).

Year	Brazil				Germany			
	No data correction	Original method	Kc Classification	Solar irradiance bands	No data correction	Original method	Kc Classification	Solar irradiance bands
Error rRMSE (%)								
1 year (23)	*17.49*	17.48	17.51	18.77	*45.46*	42.96	43.02	45.30
2 years (23-22)	*18.82*	19,07	19.10	20.41	*42.72*	42.01	42.24	44.33
3 years (23-21)	*20.12*	20.50	20.50	21.87	*43.48*	43.34	43.69	45.77
4 years (23-20)	*20.19*	20.57	20.60	21.98	*44.41*	43.89	44.29	46.42
5 years (23-19)	*20.17*	20.58	20.66	21.99	*45.1*	44.09	44.49	46.73
6 years (23-18)	*20.31*	20.75	20.89	22.41	*45.43*	44.45	44.83	47.06

For both Brazil and Germany, the rMBE generally decreased with longer data periods, indicating that extended datasets enhance estimation accuracy (up to three to four years in Germany). In Brazil, the proposed solar irradiance bands method yielded the lowest rMBE values, ranging from 0.30% to 0.32%, and remained stable across all time intervals, showing minimal variation compared to other methods. Similarly, in Germany, this method achieved the lowest rMBE values, ranging from 2.92% to 3.49%, across the periods analyzed. In terms of relative Root Mean Square Error (rRMSE), the original method produced the most favorable individual results for both countries.

5. Conclusion

The objective of this paper was to evaluate the linear regression site adaptation method and to propose a complementary analysis based on solar irradiance bands. Additionally, the impact of the clear sky index was assessed. As a case study, two locations, one in Brazil and one in Germany, were analyzed and compared.

The main findings of the study are presented below:

- The observed differences in data dispersion (RMSE) and bias (MBE) between sites can be attributed to variations in the precision of the irradiance sensors used.
- Extending the data period did not lead to a noticeable reduction in data dispersion (RMSE) in either case. Because the rRMSE is particularly sensitive to outliers, incorporating additional years of measurements may introduce interannual variability that maintains, rather than decreases, the overall error level.
- The proposed method employing solar irradiance bands resulted in lower MBE values when compared to other methods, demonstrating the potential to improve site adaptation accuracy.
- In specific periods, the rMBE of the original data were lower than that after method application. This outcome is due to negative values in $Y_{satellite}^{new}$, which are not physically meaningful for solar applications, being set to zero, resulting in a slight increase in inaccuracy.
- Additionally, for both locations, using two years of measured data would offer improvements over a one-year period and could enhance the reliability of generation forecasts.

This study aimed to advance knowledge on improving the linear regression site adaptation method by incorporating solar irradiance band separation and evaluating the impact of the irradiance data collection period on solar energy projects, with a focus on applications relevant to industry planners and researchers in both governmental and non-governmental organizations.

6. Acknowledgments

The authors wish to acknowledge CTG Brasil for the financial support for this study through project PD-10381-0620/2020, conducted in partnership with SENAI-RN ISI-ER and UNESP-Ilha Solteira as part of the Research, Development, and Innovation (RDI) program of the Brazilian Electricity Regulatory Agency (Agência Nacional de Energia Elétrica - ANEEL). A. M. F. Gomes is also grateful for the support received from the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) doctoral scholarship program. This research was carried out within the scope of the activities of the National Institute for Technological Development of Photovoltaic Solar Energy Applications, INCT-DAESF, CNPq/MCTI/Brazil, Process 408486/2024-4. This work was conducted during a scholarship supported by the International Cooperation Program PROBRAL fostering a collaborative exchange between TH Köln-University of Applied Sciences and the Universidade Federal de Santa Catarina (UFSC). The partnership was financed by Capes – Brazilian Federal Agency for Support and Evaluation of Graduate Education within the Ministry of Education of Brazil and the German Academic Exchange Service (DAAD).

7. References

Aguiar, M.L., Polo, J., Vindel, J.M., Oliver, A., 2019. Analysis of satellite derived solar irradiance in islands with site adaptation techniques for improving the uncertainty. Renew. Energy 135, 98–107. https://doi.org/10.1016/j.renene.2018.11.099

Babar, B., Graversen, R., Boström, T., 2019. Solar radiation estimation at high latitudes: Assessment of the CMSAF databases, ASR and ERA5. Solar Energy 182, 397–411. https://doi.org/10.1016/j.solener.2019.02.058

Bright, J.M., 2019. Solcast: Validation of a satellite-derived solar irradiance dataset. Sol. Energy 189, 435–449. https://doi.org/10.1016/j.solener.2019.07.086

Deutscher Wetterdienst, 2021. 10-minute sum of station observations of global radiation in J/cm² for Germany. https://cdc.dwd.de/sdi/pid/OBS_DEU_PT10M_RAD-G/DESCRIPTION_OBS_DEU_PT10M_RAD-G_en.pdf

Deutscher Wetterdienst, 2025a. CDC – Climate Data Center. https://cdc.dwd.de/portal/202209231028/view1

Deutscher Wetterdienst, 2025b. Measurement methods. https://www.dwd.de/DE/presse/pressebilder/messtechnik_node.html

Engerer, N.A., Mills, F.P., 2014. KPV: A clear-sky index for photovoltaics. Solar Energy 105, 679–693. https://doi.org/10.1016/j.solener.2014.04.019

EPE, n.d. Leilões de energia: instruções para cadastramento nos leilões. https://www.epe.gov.br/pt/leiloes-de-energia/leiloes/instrucoes-para-cadastramento

Higher Regional Court Munich, 2014. Case number 32 O 2248/11, 21 January 2014.

Jensen, A.R., et al., 2023. pvlib iotools—Open-source Python functions for seamless access to solar irradiance data. Sol. Energy 266. https://doi.org/10.1016/j.solener.2023.112092

Kaspar, F., Müller-Westermeier, G., Penda, E., Mächel, H., Zimmermann, K., Kaiser-Weiss, A., Deutschländer, T., 2013. Monitoring of climate change in Germany – data, products and services of Germany's National Climate Data Centre. Adv. Sci. Res. 10 (1), 99–106. https://doi.org/10.5194/asr-10-99-2013

Mantelli, S.L.N., Martins, G.L., Rüther, R., 2020. Uncertainty analysis of solar monitoring station: A case study. In: Proceedings of the ISES Solar World Congress 2019 and IEA SHC International Conference on Solar Heating and Cooling for Buildings and Industry 2019. International Solar Energy Society, pp. 2135–2140. https://doi.org/10.18086/swc.2019.44.03

Mieslinger, T., Ament, F., Chhatbar, K., Meyer, R., 2014. A new method for fusion of measured and model-derived solar radiation time-series. Energy Procedia, 1617–1626. https://doi.org/10.1016/j.egypro.2014.02.182

Pinto, G.X.A., Gomes, A.M.F., Braga, M., Verissimo, J.P.A., Nascimento, L., Rüther, R., 2023. Site adaptation: metodologias e estudo de caso para Florianópolis-SC. Rev. Bras. Energ. Sol. 14 (1), 1–10. https://doi.org/10.59627/rbens.2023v14i1.418

Pinto, G.X.A., Gomes, A.M.F., Braga, M., Sergio, L.A.Z., Naspolini, H.F., Rüther, R., 2024. Avaliação do impacto do período e resolução de dados em método de site adaptation: estudo de caso para Florianópolis-SC. Rev. Bras. Energ. Sol. 15 (1), 17–26. https://doi.org/10.59627/rbens.2024v15i1.449

Polo, J., Martín, L., Vindel, J.M., 2015. Correcting satellite derived DNI with systematic and seasonal deviations: Application to India. Renew. Energy 80, 238–243. https://doi.org/10.1016/j.renene.2015.02.031

Smith, C.J., Bright, J.M., Crook, R., 2017. Cloud cover effect of clear-sky index distributions and differences between human and automatic cloud observations. Solar Energy 144, 10–21. https://doi.org/10.1016/j.solener.2016.12.055