Analysis of the GHI variability index for the city of Natal - RN

Ana Cleide Bezerra Amorim[1], Samira de Azevedo Santos Emiliavaca[1], Bruno do Nascimento e Silva[1], Alan Rodrigues de Sousa[1], João Vítor Fonseca de Mendonça[1], Nicholas Medeiros Lopes[1], Gabriela Rocha Alves da Silva[1], André Luiz de Oliveira Lira[1]

[1]SENAI Innovation Institute for Renewable Energies, Natal (Brazil)

Abstract

Natal, a city located in Northeastern Brazil, is characterized by high levels of solar radiation, making it a strategic site for solar energy generation. However, its coastal setting promotes frequent cloud formation, introducing substantial variability in solar irradiance and influencing the performance of photovoltaic systems. This study investigates cloud cover variability in Natal using the Variability Index (VI) and the Clear-Sky Index (Kt*), based on irradiance data collected by a solarimetric station between 2015 and 2023. Cluster analysis was employed to classify sky conditions. Additionally, an investigation into overirradiance events was conducted for each type of sky. Results indicate that clear-sky conditions are uncommon, with daily VI values typically ranging from 20 to 30. Seasonal patterns reveal a reduction in VI during the rainy season (April–July) and higher variability during the dry months. Four distinct sky condition classes were identified: Clear Sky, Overcast Sky, Variable Part of the Day, and Highly Variable All Day. Overirradiance events occurred across all categories but were most frequent under highly variable conditions. These findings advance the understanding of cloud dynamics in coastal tropical environments and provide valuable insights for optimizing solar energy generation and enhancing grid stability in regions with high solar variability.

Keywords: Solar Energy, Cloudiness, Clear Skies, Northeast Brazil, Overirradiance

1. Introduction

The city of Natal, located in the state of Rio Grande do Norte, is popularly known as the "City of the Sun" due to its high levels of solar radiation throughout the year. Its privileged geographical location, close to the equator and the northeastern coast, results in some of the highest levels of solar radiation in Brazil. However, its proximity to the coast also influences the region's atmospheric conditions due to the frequent formation of clouds, especially during most of the day. This condition contributes to high variability in the amount of solar radiation reaching the ground, which directly affects the performance of photovoltaic systems that depend on consistent solar radiation to generate energy.

The literature addresses various aspects of irradiance variability and related processes in time and space, such as the variability of energy production from photovoltaic systems (Marcos *et al*., 2011; Lave *et al*., 2013a, 2013b; Haaren *et al*., 2014). Factors such as climatic conditions and the position of the sun can cause significant changes in the amount of energy produced. In Brazil, photovoltaic solar energy has grown year after year, and even after the introduction of Law 14.300/2022, which establishes the legal framework for distributed microgeneration and minigeneration, photovoltaic solar energy has continued to expand and now represents 7.75 percent of electricity generation in the country (ANEEL, 2024).

The dependence of photovoltaic (PV) energy production on meteorological parameters affects energy output due to climate-induced variability (Poddar *et al*., 2023). During the day, variations in irradiance lead to intermittency in the energy produced, exacerbating issues of grid stability and resilience in locations with a high concentration of PV systems (Da Rocha *et al*., 2020). In addition, solar resource variability can cause imbalances between generation and load, making it difficult to maintain frequency control and leading to moments of energy surplus or shortage, thus affecting system reliability (Hossain *et al*., 2023).

The variation in cloudiness throughout the day is one of the main challenges for solar energy, as it can affect the efficiency of photovoltaic systems installed in the region, creating difficulties in solar energy generation planning. Moreover, the presence of clouds is an important factor as it can lead to overirradiance events, where peaks of solar radiation exceed typical levels, followed by periods of low irradiance.

The main objective of this work is to classify the variability of cloudiness in the city of Natal. A specific

10.18086/swc.2025.04.03 Available at http://proceedings.ises.org

objective to be investigated is: do overirradiance events occur under all sky conditions, or only under certain patterns of cloud variability? Answering this question will contribute to a better understanding of the dynamics of cloud cover in the city and could help optimize the performance of local photovoltaic systems.

2. Materials and methods

2.1 Study area

The state of Rio Grande do Norte (RN) has a long coastline to the north and east, bordered by the waters of the Atlantic Ocean. Most of the population (785,363 inhabitants) lives in the capital, Natal, which is located in the eastern part of the state and covers an area of 167.26 km², characterized by the Atlantic Forest biome (IBGE, 2025) (Figure 1).

RN is, after Pernambuco, the state with the largest semi-arid (BSh) area in the Northeast Region of Brazil (NEB), with approximately 61.2% of its territory under this classification (Figure 1). However, the city of Natal is part of the tropical dry-summer climate region ("As"; s – dry summer) with a coastal character (Alvarez et al., 2013) and a constant formation of sea breezes, which directly influences local meteorological conditions. Even on sunny days, cloud formation is common in the region because the land surface heats up faster than the nearby ocean, creating a low-pressure area over the coast that generates rising convective currents over the continent, favoring cloud formation throughout the day (Varejão-Silva, 2006).

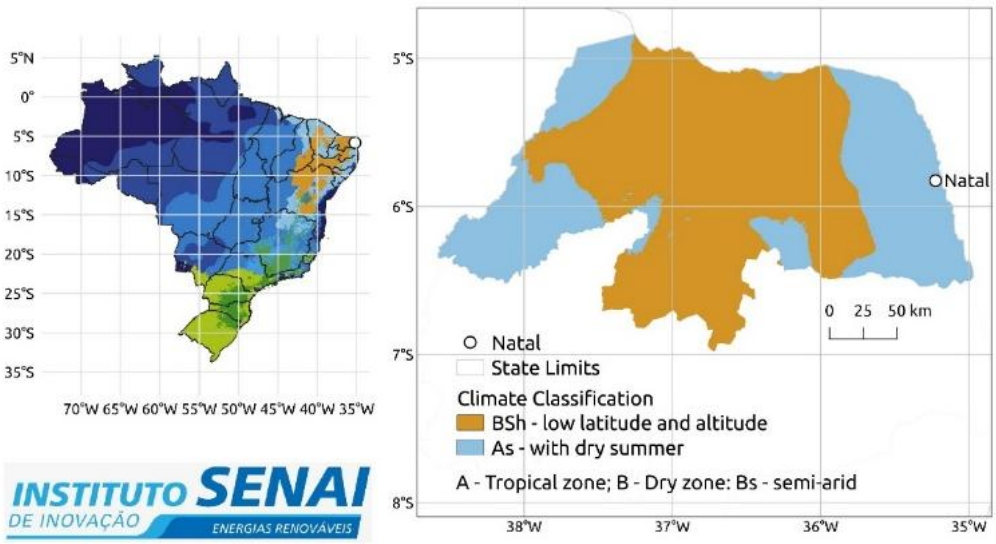


Figure 1: Köppen climate classification. Adapted from Alvarez *et al*. (2013).

2.2 Data

Data from the solarimetric station located at the SENAI Institute for Innovation in Renewable Energy (ISI-ER) were used from April 2015 to December 2023. The station is equipped with sensors that measure temperature, relative humidity, wind speed and global horizontal irradiance (GHI). Data is collected every second and integrated every minute by a CR3000 data logger, which can be accessed via the internet. Global horizontal irradiance data were measured using a Kipp & Zone CMP22 pyranometer and adjusted according to the Baseline Surface Radiation Network (BSRN) algorithm (Long and Shi, 2008).

Clear sky global irradiance on the horizontal plane, with 1-minute resolution, is used from the Copernicus Atmosphere Monitoring Service (CAMS) (Gschwind *et al*., 2019).

2.3 Methodology

2.3.1 Variability Index (VI)

The Variability Index (VI) was proposed by Stein *et al*. (2012) and is used to quantify the variability of the solar resource, as shown in Equation (1). Conceptually, the VI is defined as the ratio between the length of the measured irradiance curve and the length of the corresponding curve under clear-sky conditions. High values of VI indicate high irradiance variability throughout the day, while low values of VI can occur on both clear and very cloudy days (Stein *et al*., 2012). It is a simple measure of irradiance variability over time, calculated here on a daily basis using 1-minute resolution data. Equation (1) shows the calculation of the VI:

$$VI = \frac{\sum_{k=2}^{n}\sqrt{(GHI_k - GHI_{k-1})^2 + \Delta t^2}}{\sum_{k=2}^{n}\sqrt{\left(GHI_{CS_k} - GHI_{CS_{k-1}}\right)^2 + \Delta t^2}} \quad \text{(eq. 1)}$$

Where GHI is the Global Horizontal Irradiance and GHI_{CS} is the Clear-Sky Horizontal Irradiance.

2.3.2 Clear-Sky Index (Kt*)

The clear-sky index (K_t*) represents a normalization of GHI to a theoretical clear-sky GHI (GHI_{CS}) (i.e., irradiance at a particular location assuming a cloud-free atmosphere) (Lohmann, 2018). Clear-sky irradiance depends on atmospheric parameters and is therefore specific to each model. Many different clear-sky models have been proposed (Badescu *et al*., 2012; Gueymard, 2012; Reno *et al*., 2012; Lefèvre *et al*., 2013), and it should be emphasized that the use of any of them introduces uncertainties. In this study, the CAMS clear-sky GHI is used at 1-minute resolution (Gschwind *et al*., 2019).

$$K_\tau{}^{*} = \frac{GHI}{GHI_{CS}} \quad \text{(eq. 2)}$$

2.3.3 Extraterrestrial Irradiance (I_o)

To identify overirradiance episodes (Gueymard, 2017a; Nascimento *et al*., 2019), the measured GHI values exceeding the extraterrestrial horizontal irradiance (I_o) were analyzed. The I_o values were estimated using the following expression:

$$I_0 = I_{sc}.\, \varepsilon_0 \cos(\Theta_z) \quad \text{(eq. 3)}$$

Where I_{sc} is the solar constant, equal to 1367 W/m², ε_0 is the eccentricity of the Earth's orbit, and $\cos(\Theta_z)$ is the cosine of the zenith angle.
GHI values above 1000 W/m² were considered for the analysis of overirradiance events, as values significantly lower than the Standard Test Conditions (STC) irradiance (1000 W/m^2) do not pose relevant or potential issues that affect the operational performance of photovoltaic plants.

2.3.4 Cluster Analysis

Cluster analysis (CA) is a multivariate technique aimed at separating groups based on a measure of dissimilarity between them (Wilks, 2006). This technique was used to identify characteristics of sky conditions considered homogeneous during 2015–2023. It was applied to daily VI and K_t^* data calculated from 1-minute measurements at the Natal solarimetric station.

This study employs the Euclidean distance, a widely used measure of dissimilarity in CA applications on meteorological data (Santos *et al*., 2015; Amorim *et al.*, 2020, 2024), and expressed in Equation (4):

$$d_{i,j} = \left[\sum_{k=1}^{k} W_k \left(X_{i,k} - X_{j,k}\right)^2\right]^{\frac{1}{2}} \quad \text{(eq. 4)}$$

where $X_{i,k}$ and $X_{j,k}$ are the time series for each grid point compared within group k; W_k =1, ..., k is a weighting factor, and the indices i and j represent the grid points of the group.

In addition to the dissimilarity measure, we applied Ward's minimum variance hierarchical clustering method (Mingoti, 2005), defined according to Equation (5). Ward's method seeks to minimize the within-group sum of squares, which is equivalent to finding the smallest standard deviation among the data in each group.

$$W = \sum_{g=1}^{Gr} \sum_{i=1}^{n_g} \left\| X_i - \overline{X}_g \right\|^2 = \sum_{g=1}^{Gr} \sum_{i=1}^{n_g} \sum_{K=1}^{K} \left(X_{i,K} - \overline{X}_{g,K} \right)^2 \quad \text{(eq. 5)}$$

where Gr and g represent the groups of grid points; Xi is each vector of grid points of variable K; Xg is the average of all the grid points of group g of variable K.

The Silhouette graph (Rousseeuw, 1987) was then used to evaluate the quality of each grouping (Santos *et al.*, 2015; Amorim *et al.*, 2020, 2024). The Silhouette value represents the average contour value of each observation. Properly grouped observations have values close to 1, whereas poorly grouped observations have values close to -1. The Silhouette value of object Xi (each observation) is calculated using Equation (6):

$$S(X_i) = \frac{b(X_i) - a(X_i)}{max\{a(X_i), b(X_i)\}} \quad \text{(eq. 6)}$$

N is the total number of data.

3. Results and Discussion

3.1. VI behavior

Figure 2 shows the daily VI values for Natal between 2015 and 2023. Figure 2a is a histogram of these values, showing that most days have relatively high VI values (between 20 and 30). This indicates that clear skies are uncommon in the city. Cloudiness favours the generation curve of a solar farm, which has numerous peaks and valleys throughout the day. According to the literature, this can be smoothed out by balancing with hydraulic turbines or using a battery bank (Alencar *et al.*, 2018).

The graph in Figure 2b shows that the monthly average VI (external variability) fluctuates, with an overall mean of around 25. There is a decrease in VI values between April and July, reaching a minimum average of 19 in June (red dots). This behavior is associated with the higher concentration of clouds during the rainy season, which occurs between the austral autumn and winter. It is worth noting that the main rainy season in the coastal region of Northeast Brazil (LNEB) coincides precisely with the period from April to July (Moura *et al.*, 2009; Amorim *et al.*, 2014, 2020), which corresponds to the lowest VI values. The period from September to February is characterized as the driest interval for the city of Natal, during which the highest VI values are observed. The months of October, November, and December stand out, with average VI values above 30, peaking at 35 in November.

The internal variability, on the other hand, shows large fluctuations in VI values throughout the months, indicating that different types of classifications are present in each month, such as clear, cloudy, and partly cloudy days. This classification will be discussed in more detail in the following results.

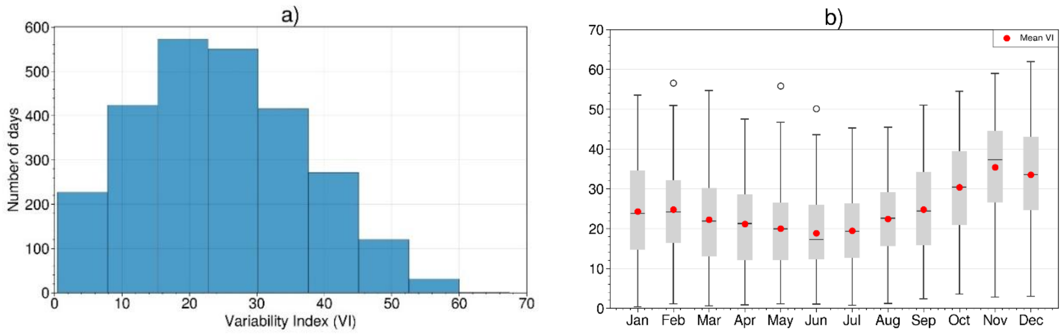


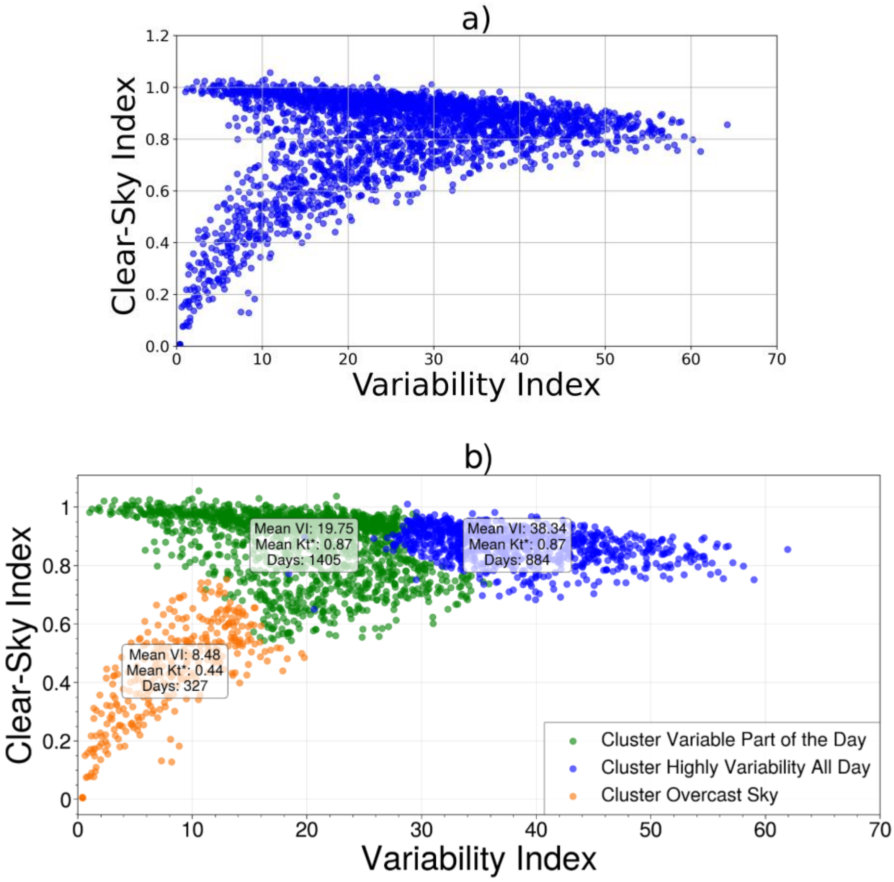
Figure 2: Köppen climate classification. Adapted from Alvarez *et al*. (2013).

3.2 Variability Rating

Stein *et al*. (2012) point out that the VI has a limitation when it comes to classifying cloudy conditions, suggesting a relationship with the clear-sky index (*Kt**). Figure 3 illustrates the relationship proposed in the literature, where an arrow-shaped scatter plot can be identified (Figure 3a), with days showing high variability in GHI behaviour represented by the tip of the arrow.

The results of the cluster analysis are shown in Figures 3b and 3c, suggesting the presence of three to four different types of cloud conditions for the city of Natal. Based on the average silhouette coefficient (S) from Equation 7, the analysis indicates three groups, with a value of S = 0.44 (Figure 4b). However, when four groups are considered (S = 0.38), clear-sky days are more distinctly defined in the results (Figure 4c). The cluster analysis with four groups identified 1,101 days classified as Variable Part of the Day, 884 as Highly Variable All, 327 as Overcast Sky, and 304 as Clear Sky - the lowest among the categories.

The limitation of the VI classification between clear-sky and overcast-sky conditions can be observed in the results, which show average daily values close to 10 and 8, respectively. This distinction is made possible by the Kt* values, which have mean values of 0.95 for clear-sky days and 0.44 for overcast-sky days, as suggested by Stein *et al*. (2012). On the other hand, the groups with greater variability, such as Variable Part of the Day and Highly Variable All Day, presented higher average VI values—22.40 and 38.34, respectively—although their Kt* values were similar, with means of 0.85 and 0.87.



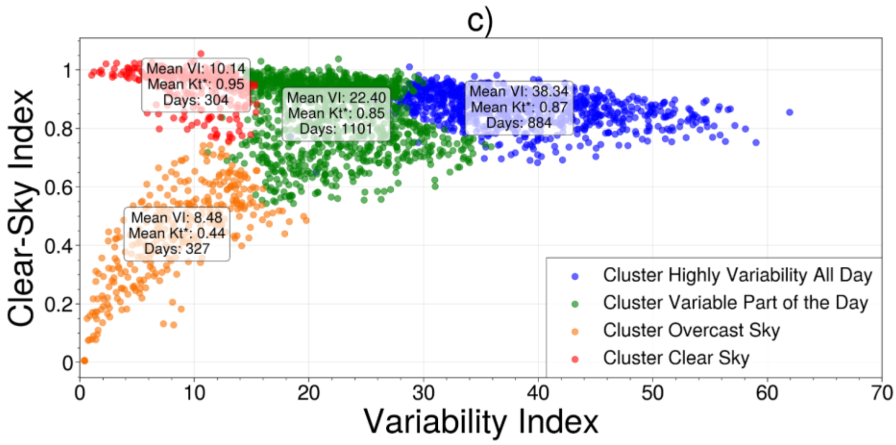


Figure 3: Scatter plot (arrowhead) of daily clear-sky index (*Kt) and VI for each day from 2015 to 2023: a) without clustering technique; b) with 3 groups; and c) with 4 groups.**

Figure 4 presents a representative selection of four days of GHI daily behaviour for each cluster, with their respective VI and Kt* values shown in the legend. The VI values for overcast-sky conditions are slightly lower (Rocha *et al*., 2020) because, on overcast days, cloudiness predominates throughout the day, leading to reduced radiation and small variability. Accordingly, Kt* values are lower than 0.5. On clear-sky days (Figure 4b), especially in coastal cities such as Natal, cloud formation is common during the day, resulting in greater variability than under overcast conditions, which is reflected in the slightly higher daily VI value. In contrast, on days with strong intermittency—such as Variable Part of the Day and Highly Variable All Day—the daily VI values were 22.37 and 38.1, respectively.

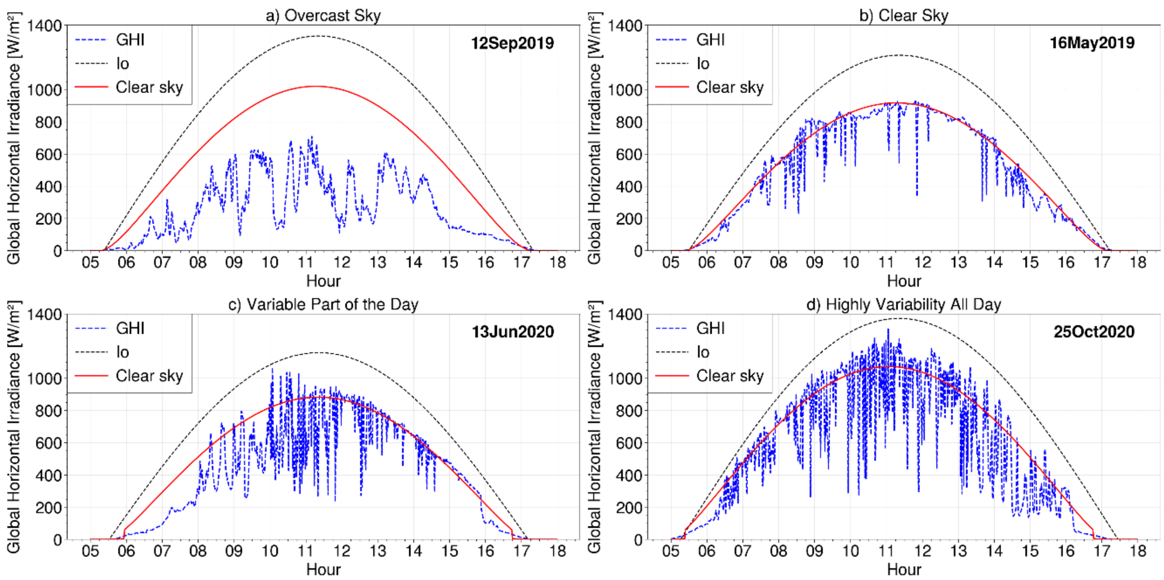


Figure 4: Days representing the average behaviour of each group: a) Overcast Sky [VI = 8.89; Kt* = 0.45]; b) Clear Sky [VI = 10.16; Kt* = 0.95]; c) Variable Part of the Day [VI = 22.37; Kt* = 0.85]; and d) Highly Variability All Day [VI = 38.10; Kt* = 0.87].

3.3 Variability Seasonal

Figure 5 shows the seasonality of each cluster for the city of Natal. Days with high variability in GHI are prominent across all seasons, with the Highly Variable All-Day group prevailing over the others in spring and austral summer, while the Variable Part of the Day group outnumbers Highly Variable All-Day during autumn and winter. Regarding the Clear Sky and Overcast Sky groups, Clear Sky is the least frequent condition in summer and the most frequent in winter, whereas Overcast Sky shows the opposite pattern - being more frequent in autumn and less frequent in spring.

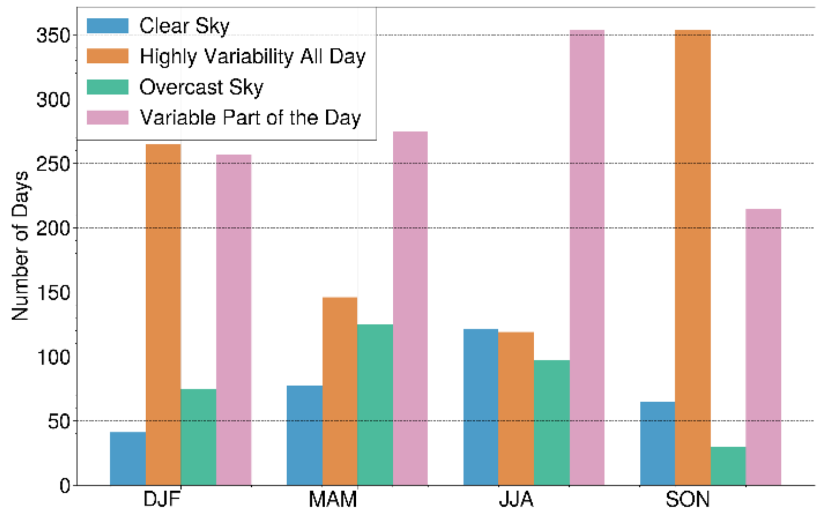


Figure 5: Seasonal frequency of the number of days in each group (Clear Sky, Highly Variability All Day, Overcast Sky and Variable Part of the Day).

3.4 Overirradiance Events

Table 1 presents the records of days on which overirradiance events occurred for each group. Each day with G greater than I_0 for one minute or longer was counted as an event day. Although the exploratory data analysis revealed days with more than two events occurring at different times, the analysis was limited to categorizing each day simply as having or not having an overirradiance event.

The results show that events were recorded in all groups, with the Variable Part of the Day and Highly Variable All Day groups exhibiting high GHI variability throughout the day. The Clear Sky group recorded 15 event days, 14 of which occurred during the region's rainy season, between March and July.

Table 1: Numbers of overirradiance events

SKY	Record days
Overcast Sky	76
Clear Sky	15
Variable Part of the Day	277
Highly Variability All Day	214

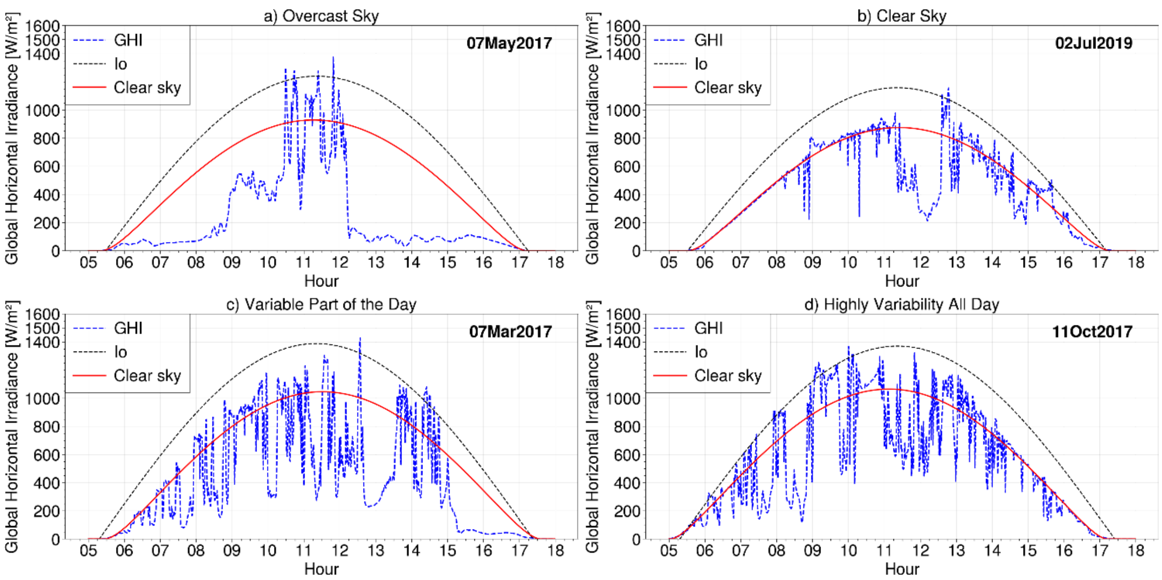


Figure 6: Overirradiance events in the different types of sky conditions in the city of Natal: a) Overcast sky, b) Clear sky, c) Variable Part of the Day and d) Highly Variability All Day.

4. Conclusions

This study analysed the variability of cloudiness and the occurrence of overirradiance events in Natal, a coastal city popularly known as the 'City of the Sun'. Despite its reputation for high solar availability, Natal's proximity to the coast leads to frequent cloud formation, which directly affects the variability of solar irradiance and consequently the performance of photovoltaic systems.

Based on 1-second GHI data from the solarimetric station at the SENAI Institute for Innovation in Renewable Energies (ISI-ER), aggregated into one-minute intervals, the results revealed relatively high Variability Index (VI) values (between 20 and 30) on a daily basis, confirming that clear-sky conditions are rare in the region. VI values followed local climatological trends, with lower values during the rainy season (April to July) and higher values in the drier months, reflecting the influence of seasonal cloud dynamics.

Cluster analysis combined with the silhouette coefficient revealed that GHI variability in Natal can be effectively categorised into four cloudiness groups: Clear Sky, Overcast Sky, Variable Part of the Day, and Highly Variable All Day. Although the silhouette metric initially suggested three clusters, using four groups provided a clearer separation of clear-sky days. The results showed that 1,101 days were classified as 'Variable part of the day', 884 as 'Highly variable all day', 327 as 'Overcast sky', and 304 as 'Clear sky'.

The limitation of the VI classification between 'Clear Sky' and 'Overcast Sky' was overcome by using the clear-sky index (Kt^*), which averaged 0.95 and 0.44 respectively, consistent with findings reported in the literature. Meanwhile, the 'Variable part of the day' and 'Highly variable all day' groups exhibited high VI values (22.40 and 38.34, respectively) and similar Kt^* values (0.85 and 0.87), confirming strong intra-day variability.

The seasonal analysis revealed that variable conditions during the day and highly variable conditions throughout the day dominate across all seasons, reflecting the persistent influence of local meteorological dynamics. Overirradiance events were observed in all groups, occurring most frequently in the variable-sky clusters. The Clear Sky group recorded only 15 event days, 14 of which occurred during the rainy season.

Overall, the results highlight that, even in a region with high solar potential such as Natal, the stability of solar resources is significantly affected by the intraday and seasonal variability of cloud cover. These findings are relevant for improving short-term solar forecasting, optimising photovoltaic system design and planning energy generation in tropical coastal environments.

5. Acknowledgments

This research was carried out as part of the activities of the National Institute for Technological Development of Photovoltaic Solar Energy Applications (INCT-DAESF), funded by the Brazilian agencies CNPq and MCTI under process number 408486/2024-4. The authors would like to thank the SENAI Institute for Innovation in Renewable Energy for its support and infrastructure; CTG Brazil, for its financial contribution to the ANEEL PD-10381-0620/2020 project; and TotalEnergies, for its current financial contributions to the ANP 23980-6 and ANP 24118-2 projects. The authors would like to thank ISI-ER for their support in the development of this work and TotalEnergies for the current financial support under the ANP PD-24118-2 project.

6. References

Alvarez, C. A., Stape, J. L., Sentelhas, C., Gonçalves, J. L. M., Sparovek, G., 2013. Köppen's climate classification map for Brazil. Meteorologische Zeitschrift, 22(6), 711–728.

ANEEL, 2024. Sistema de Informações de Geração da ANEEL - SIGA. https://app.powerbi.com/view?r=eyJrIjoiNGE3Nj

Badescu, V., *et al.*, 2012. Computing global and diffuse solar hourly irradiation on clear sky: Review and testing of 54 models. Renewable and Sustainable Energy Reviews, 16, 1636–1656.

Da Rocha, V. R., *et al.*, 2020. Avaliação da variabilidade do recurso solar em território brasileiro. In: Congresso

Brasileiro de Energia Solar - CBENS.
Gueymard, C. A., 2012. Clear-sky irradiance predictions for solar resource mapping and large-scale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models. Solar Energy, 86, 2145–2169.
Hossain, M. B., *et al*., 2023. Dynamic electrical circuit modeling of a proton exchange membrane electrolyzer for frequency stability, resiliency, and sensitivity analysis in a power grid. IEEE Transactions on Industry Applications, 59(6), 7271–7281.
IBGE, 2025. Instituto Brasileiro de Geografia e Estatística. https://www.ibge.gov.br/
Lave, M., Kleissl, J., 2013. Cloud speed impact on solar variability scaling - Application to the wavelet variability model. Solar Energy, 91, 11–21.
Lave, M., Kleissl, J., Stein, J. S., 2013. A wavelet-based variability model (WVM) for solar PV power plants. IEEE Transactions on Sustainable Energy, 4, 501–509.
Lefèvre, M., *et al*., 2013. McClear: A new model estimating downwelling solar radiation at ground level in clear-sky conditions. Atmospheric Measurement Techniques, 6, 2403–2418.
Lohmann, G. M., 2018. Irradiance variability quantification and small-scale averaging in space and time: A short review. Energies, 11(7), 1836.
Long, C. N., Shi, Y., 2008. An Automated Quality Assessment and Control Algorithm for Surface Radiation Measurements. The Open Atmospheric Science Journal, 2, 23–37.
Marcos, J., *et al*., 2011. From irradiance to output power fluctuations: The PV plant as a low pass filter. Progress in Photovoltaics: Research and Applications, 19, 505–510.
Mingoti, S. A., 2005. Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada. Belo Horizonte: Editora UFMG.
Nascimento, L. R. do, Viana, T. de S., Campos, R. A., Rüther, R., 2019. Extreme solar overirradiance events: Occurrence and impacts on utility-scale photovoltaic power plants in Brazil. Solar Energy, 188, 612–620.
Poddar, S., *et al*., 2023. Changes in solar resource intermittency and reliability under Australia's future warmer climate. Solar Energy, 266, 112039.
Reno, M. J., Hansen, C. W., Stein, J. S., 2012. Global horizontal irradiance clear sky models: Implementation and analysis. Sandia National Laboratories.
Rousseeuw, P. J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Computational and Applied Mathematics, 20, 53–65.
Stein, J. S., Hansen, C. W., 2017. The Variability Index: A New and Novel Metric for Quantifying Irradiance and PV Output Variability. Sandia National Laboratories.
Strunk Jr., W., White, E.B., 1979. The Elements of Style, third ed. Macmillan, New York.
Van Haaren, R., Morjaria, M., Fthenakis, V., 2014. Empirical assessment of short-term variability from utility-scale solar PV plants. Progress in Photovoltaics: Research and Applications, 22, 548–559.
Varejão-Silva, M. A., 2006. Meteorologia e Climatologia. Versão digital 2. Recife, PE.