Solar irradiation regionalization in Uruguay: understanding the interannual variability and its relation to El Niño climatic phenomena

A. Laguarda[a,*], R. Alonso-Suárez[a], R. Terra[b]

[a]*Laboratorio de Energía Solar, Instituto de Física, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay.*
[b]*Instituto de Mecánica de los Fluídos e Ingeniería Ambiental, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay.*

Abstract

A regionalization of different characteristics of the solar resource is performed for Uruguay and surrounding areas (Southeastern South America). The input information consists of daily satellite estimates of Global Horizontal Irradiation (GHI) generated in a regular grid using a low uncertainty empirical satellite-based model which was specifically adapted for the region. Clusters are derived from the climatological annual cycle of monthly irradiation and clearness index and, separately, from time-series of monthly variability. The solar irradiation variability in each cluster is compared with El Niño South Oscillation (ENSO) signal. A high negative correlation is observed between ENSO and solar irradiance, most predominantly over February to May and November to December, particularly for the latter. This means that in a strong El Niño/La Niña year, solar irradiation values for the November-December period in Uruguay will be smaller/higher than the climatological

*Corresponding author: A. Laguarda, laguarda@fing.edu.uy

average. These results are in agreement with the ones obtained for rainfall in other studies.
Keywords: Solar Radiation, Clustering, ENSO, k-means.

1. Introduction

The classification of an area into regions with similar climatic characteristics can guide the design of ground measurement networks suited for long-term renewable energy studies. The location of the measuring sites, if not adequate, may affect climatological studies related to the renewable energies technology potential and long-term resource assessment. This is the case of solar irradiation and wind intensity, that highly vary in both space and time. These measurements, along with satellite images and/or atmospheric models, are the main tools to assess renewable energies resources. Associated uncertainties are directly translated to the financial risk of renewable energy projects. The regionalization also allows to better understand the impact of different large-scale climatic phenomena, such as the El Niño South Oscillation (ENSO) on the regional renewable resource availability. These effects are better characterized over regions than over specific sites. This work focuses on the monthly mean of global solar irradiation in a horizontal plane at ground level (GHI) and its monthly variability. The input information consists of daily satellite estimates of GHI obtained from a low uncertainty empirical satellite-based model, which was specifically adapted for the regional characteristics (Alonso-Suárez et al., 2012, 2014). The original hourly estimates, which are computed for this work for a 15 years time span on a regular grid of 31 points covering the region, are monthly averaged to analyze its annual cycle and interannual variability.

Clustering techniques identify patterns in data without any a priori information, namely expert knowledge of predefined patterns. These techniques have been used for regionalization around the globe (Diabate et al., 2004; Zagouras et al., 2013, 2014, 2015; Polo et al., 2015; Watanabe et al., 2016). In Diabate et al. (2004) a set of hierarchical clustering methods are applied to annual cycles of clearness index's monthly averages at 62 sites in Africa to obtain different climatic zones across the continent. Other authors prefer to use non-hierarchical clustering techniques, such as the k-means method. Polo et al. (2015) use this methodology with daily sunshine hourly records to obtain a set of well-defined climatic regions for Vietnam. The clustering is based on a variability index for each site that is calculated from the Cumulative Distribution Function (CDF) of daily solar irradiation obtained from the sunshine hours. Zagouras et al. (2014, 2015) also used the k-means algorithm, but applied to daily gridded irradiation records estimated with the satellite SUNY model (Perez et al., 2002). The clustering is applied after reducing the dimensionality with Principal Component Analysis (PCA). The output was a set of regions for solar energy large-scale electricity applications in the state of California. Zagouras et al. (2013) used the same algorithm with satellite cloud information for Greece to guide the design of a solar radiation monitoring network. Watanabe et al. (2016) classified 47 ground measurement stations in Japan according to different metrics of solar radiation series for engineering applications. All these studies highlight the importance of identifying regions with similar solar variability, as it is associated with solar power fluctuations that can be a major challenge for their grid integration. Clustering techniques have proven to be an efficient approach for regionalization purposes.

In this article we provide a first solar irradiation regionalization of Uruguay and surrounding areas. Different characteristics of solar irradiation are analyzed, including resource monthly averages and anomalies. One of the goals of regionalization is to improve the description and understanding of the impact of large-scale climatic phenomena on the solar resource. In particular, the impact of the El Niño South Oscillation (ENSO) on this area is well known (Aceituno, 1988; Grimm et al., 2000), being the main climatic forcing in the region. Its impacts on local climate has been well established for some meteorological variables, in particular, on precipitation (Pisciottano et al., 1994) and streamflow (Mechoso and Pérez-Iribarren, 1992). However, the impact of ENSO on the regional solar irradiation has not been documented yet, and is a major contribution of this work.

This article is organized as follows. Section 2 and Section 3 describe the data and methodology, respectively. The results of the regionalization are presented in Section 4; Subsection 4.1 is based on the monthly annual cycles of mean irradiation and clearness index, while Subsection 4.2 focuses the variability of the monthly clearness index (monthly anomalies). Section 5 describes the relationship between the ENSO and the clearness index variability through correlation (Subsection 5.1), conditional stratification (Subsection 5.2) and seasonal averages (Subsection 5.3). Finally, in Section 6 and Section 7, we present the final discussion and conclusions, respectively.

2. Data

The region under study is the country of Uruguay and close areas, a part of the broader Pampa Humeda region of the Southeastern part of South America. It is located between 30°S and 35°S (latitude) and 53°W and 58°W (longitude). The region is mainly rural and of low altitude, barely exceeding 500 meters in few places. Figure 1 illustrates the area location with a topography map for further reference.

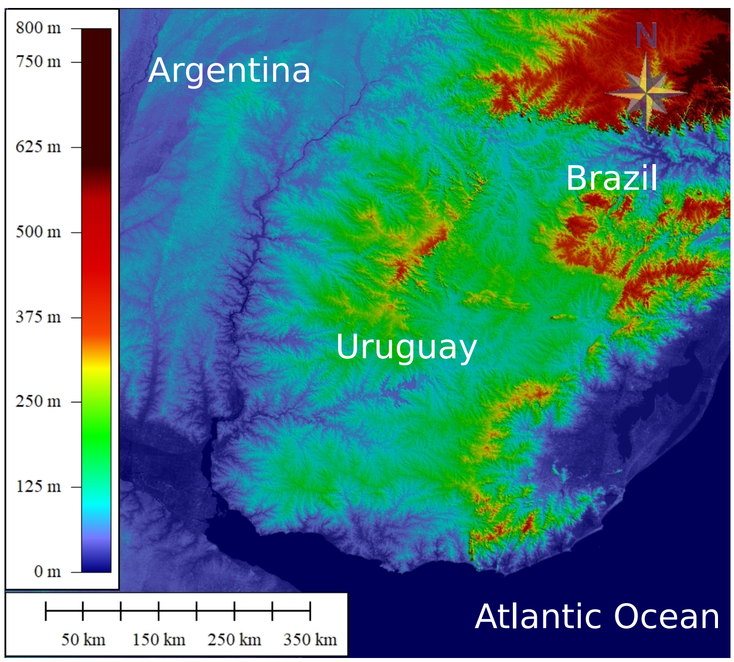


Figure 1: Topography map of the region with altitude data retrieved by the Terra-Aster NASA satellite information. Most of the area show heights under 500 m above the sea level.

The study requires long-term solar radiation data with adequate spatial coverage over the region. Solar radiation networks were first established in the country circa 2010, which implies that the ground measurements time span is inadequate to define a climatology or appropriate to describe inter-annual variability. Further, ground measurements are not able to provide a high space resolution as the sites are located in sparse points in a larger territory. We, therefore, turn to solar irradiation estimates based on satellite imagery which can provide two decades data with a spatial resolution of few kilometers. Subsection 2.1 describes the methodology to generate hourly and daily GHI estimates, while Subsection 2.2 introduces the climate index associated to ENSO.

2.1. Solar irradiation estimates

Solar irradiation at the surface can be accurately estimated using models that use as input geostationary satellite images (Perez et al., 2002; Rigollier et al., 2004; Ceballos et al., 2004; Cebecauer et al., 2010; Alonso-Suárez et al., 2012; Qu et al., 2017; Laguarda et al., 2020). As the images are regularly available at a rate of more than two per hour and a nominal resolution of 1 km or less, they can be used as input to model a highly fluctuating phenomena like solar radiation. Satellite based models for snow free areas can use solely the visible channel images to quantify the cloudiness, which is the first factor affecting solar radiation availability at ground level. In colder sites, where snow might occur, the infrared images can be used to differentiate snow from clouds, as they both present a high albedo in the visible spectrum but differ in their temperature. More complex models (Qu et al., 2017) use multi-spectral satellite images to quantify cloud properties and then infer the ground solar irradiation. Satellite based models for solar irradiation assessment can be classified into physical (Noia et al., 1993b) or empirical (Noia et al., 1993a), depending on their formulation. Physical models attempt to model the radiative transfer thorough the atmosphere, so they require a detailed knowledge of the atmospheric state, including cloud type and cloud phase, water vapour content, aerosol optical depth, among others. The uncertainty of these models highly depend on the availability and quality of their input information. On the other hand, empirical models rely on simple statistical relationships between the ground solar irradiation and different variables that can be modelled, measured or calculated, including satellite information. These models are potentially accurate if high quality measurements are available in the target area to adjust their coefficients. The main disadvantage of statistical models is that the local adjustment can not be extrapolated globally. Some of the state-of-the-art models that are used worldwide for solar resource assessment do not fit perfectly in these two categories, like the SUNY model (Perez et al., 2002) or the former Heliosat versions (Rigollier et al., 2004; Beyer et al., 1996), and they are referred as hybrid. These hybrid models have a physically-based formulation but some of their parameters are adjusted using ground measurements.

In this work we use an empirical satellite-based model that was specifically adapted for the region to estimate ground level solar global irradiation at a horizontal plane (Alonso-Suárez et al., 2012, 2014). It uses GOES-East visible channel images of South America to estimate hourly irradiation via a statistical regression. This model was originally proposed by Tarpley (1979) and Justus and Paris (1986), and afterward was modified by Alonso-Suárez et al. (2012), resulting in a significant improvement on its performance. The model coefficients were adjusted using ground data from the Continuous Solar Irradiance Measurement Network (RMCIS) administrated by Universidad de la República's Solar Energy Laboratory (LES, http://les.edu.uy). The stations of this network are located at semi-rural environments and record data at one minute intervals as the average of 15 seconds samples using spectrally flat Class A pyranometers (according to the ISO 9060:2018 standard). These pyranometers are calibrated every two years as recommended by the World Meteorological Organization (WMO). The calibration is done in accordance with the ISO 9847:1992 standard using a Secondary Standard pyranometer that the LES maintain with traceability to the World Radiometric Reference (WRR, World Radiation Center, Davos, Switzerland). This simple regression model can provide hourly GHI estimates for the region with a negligible mean bias and a mean uncertainty of 14% (root mean square deviation expressed as a percentage of the mean measurements value). Daily values are obtained from the accumulation of the intra-day hourly values. The daily estimates (H) also show a negligible bias and have a reduced uncertainty of 7%. This performance is satisfactory when compared with commercial models commonly used in Europe and the USA, such as the Heliosat series (Rigollier et al., 2004; Qu et al., 2017) or the SUNY model (Perez et al., 2002). For this work we generate the satellite-derived daily irradiation data (H) in a regular grid of $1^\circ \times 1^\circ$ latitude-longitude distance (see Figure 2) for the period between January 2000 to December 2016 (17 years). In Figure 2 the sites referred with 'U' lie within the Uruguayan territory, while A and B sites correspond to Argentina and Brazil respectively. The monthly averages ($\overline{H}$) are obtained by averaging the daily values over each month. Months containing less than 15 valid days are discarded, leaving 189 out of the 204 months of the period considered.

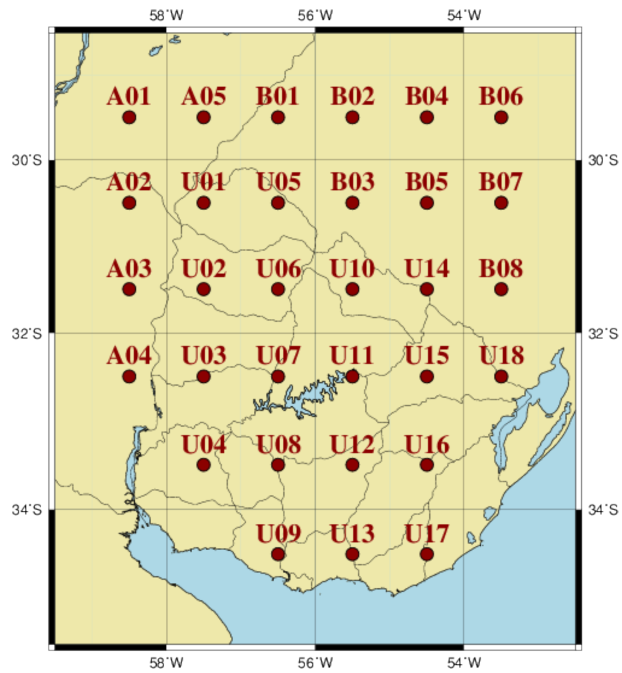


Figure 2: Locations used for generating monthly means of irradiation data (18 points in Uruguay, 8 points in Brazil and 5 in Argentina).

To remove the daily and seasonal geometrical behavior of the solar irradiation due to the sun's apparent movement, it is common practice to use the clearness index, K_t, which isolates the variability due to changes in the atmosphere (cloudiness, water content, aerosols, etc.) from the deterministic geometrical trend. This variable is defined at a monthly basis as,

$$K_t = \frac{\overline{H}}{\overline{H}_0} \tag{1}$$

where $\overline{H}$ is the monthly average of the daily GHI measured values and $\overline{H}_0$ is the monthly average of the daily irradiation on an horizontal plane at the top of the atmosphere, H_0. This daily variable can be calculated analytically for each site knowing the day of the year and latitude (Iqbal, 1983).

2.2. El Niño

There are several indices to quantify ENSO phenomenon. However, the most widely used index in research, monitoring and reporting is the Niño 3.4 index (N3.4) which represents the average sea surface temperature anomaly over a region in the central equatorial Pacific Ocean, between 5°N–5°S and 170°W–120°W (Trenberth, 1997). The N3.4 information is publicly available at www.cpc.ncep.noaa.gov, from where we retrieved the data. Each monthly value corresponds to the average of the trimester centered in the corresponding month.

3. Methodology

In this work three solar radiation regionalizations are obtained through well known clustering techniques, based on yearly cycles of $\overline{H}$ and K_t, and on K_t monthly anomaly time series. Using the latter, the interannual variability of the solar resource and its seasonal varying association with ENSO climatic phenomena is studied in detail.

3.1. Regionalization

A first set of two regionalizations is performed over the annual cycles of K_t and $\overline{H}$, i.e. 12 dimensions vectors, one for each of the 31 sites showed in Figure 2. We start with all the 31 sites and use the Ward hierarchical agglomerative algorithm (Wilks, 2011) to sequentially group the individuals, minimizing intra-group variance based on the Euclidian distance on the 12-dimensional space, up to a final number of clusters set to three. The centroids of these clusters constitute the seeds of the non-hierarchical k-means methods that regroups all individuals to the nearest centroid (using the same distance as before), recalculates the centroids and proceeds iteratively until individuals no longer change of cluster. The combined algorithms do not include any a priori information, such as seeds, being the final number of clusters the only subjective parameter, which will prove a reasonable choice in view of results in Section 4. Further details of both clustering methods can be found in Duda et al. (2001); Wilks (2011).

Next, a different regionalization is made using the K_t interannual anomaly time series as the attribute at each site. First, we perform a principal component analysis (PCA) of the initial data to reduce the dimensionality without loosing significant information. The initial data here is a set of 31 vectors - one per grid point- of 202 monthly anomaly values of K_t. The PCA converts an initial set of correlated variables into a sequence of uncorrelated linear combinations of said variables, the principal components, each contributing a decreasing portion of the total variance. It is, then, possible to rank the principal components and select a reduced set that explains most of the variance of the original data set. More details of the method can be found in Jolliffe (2002); Bishop (2006). As it is shown in Subsection 4.2, five principal components are retained, which explain most of the associated variance. Then, the same combined Ward plus k-means methodology is used with the only difference that the distance d selected in this case, for both clustering methods, is based on Pearson correlation instead of the Euclidean distance:

$$d(\mathbf{x_i}, \mathbf{x_j}) = 1 - corr(\mathbf{x_i}, \mathbf{x_j}), \qquad (2)$$

where $\mathbf{x_i}$ and $\mathbf{x_j}$ are monthly K_t anomaly series. This distance is appropriate in this case because we aim to find clusters whose mean solar irradiation is well correlated with ENSO.

3.2. Impact of ENSO in the regional solar radiation

The second part of this work focuses on the study of the influence of ENSO on the monthly anomalies of K_t behavior. The work considers several stages, from exploratory to deeper analysis. Firstly, the Pearson correlation is computed between each cluster's monthly K_t anomalies and the N3.4 index time-series. Each cluster's centroid is considered as it represents the average value within the cluster. An exploration of how these K_t anomalies stratify with the N3.4 index is also done. For stratification, two situations on N3.4 index are distinguished: greater than $+0.5$ °C (named positive) or less than -0.5 °C (named negative). By using the Wilcoxon-Mann-Whitney (WMW) classical non parametrical test, stratified according to the sign of simultaneous N3.4 values, we tested if the two stratified data sets came or not from the same probability distribution. Unfortunately, this test hypothesis could not be rejected in this work to a 95% confidence level using the monthly data time-series. Instead, we grouped the data considering that the influence of ENSO on the region ranges from October to next year July (Pisciottano et al., 1994), so we use the N3.4 index averaged during the peak season of extreme events (November to January, N3.4-NDJ) to stratify the monthly K_t anomalies from the preceding September to the following August. In other words, we tag a complete year (from September to August) according to N3.4-NDJ's sign only, and then we apply the WMW test for each cluster. This is applied to the yearly monthly time-series and does not provide information of the seasonality of the effect.

To obtain more information from the stratification, in particular regarding seasonality, the test is applied also to stratified data sets for each month and for each cluster separately. In order to have enough data, the test is performed over twelve 3-months moving-window periods. For instance, for March, the test is performed with data corresponding to February, March and April stratified with N3.4-NDJ$>$ $+0.5$ °C and N3.4-NDJ$<$ -0.5 °C. Using this procedure, we identify seasons in which the stratified behavior of K_t anomalies are categorically different. After identifying the season with noticeable ENSO effect, the correlation is also computed for the periods identified. To correlate these seasonal K_t anomalies with N3.4-NDJ index, the K_t anomalies are averaged over each season and year, and Pearson's correlation is computed between each clusters seasonal averages and N3.4-NDJ. Finally, this calculation is repeated for the same identified periods, but for each of the 31 sites individually, without considering the clusters. This allow us to obtain a correlation map in where the spatial distribution of the ENSO effect can be seen.

4. Clustering results

Clustering results are shown in this section, first applied to the annual cycles of $\overline{H}$ and K_t and then to the time series of monthly anomalies of K_t.

4.1. Clustering by climatological annual cycles

To distinguish regions according to the mean solar irradiation, the annual cycle of $\overline{H}$ is used as input for the k-means/Ward algorithm. Figure 3 shows the clustering result when the irradiation values are used. The procedure is repeated using the monthly average of the K_t, to exclude the geometrical effect of latitude and the resulting clusters are shown in Figure 4. As can be seen in Figure 3, solar resource increases mainly from South to North (closer to Equator) with a slight east/west gradient, resulting in a south-east to north-west increasing trend. This result is in accordance with local solar maps (Abal et al., 2010; Alonso-Suárez et al., 2014). The arrangement of the clusters in Figure 4 (due mainly to cloudiness) has a stronger tilt in the spatial trend than in the case of the irradiation clustering, more aligned with the North-South line. Lowest K_t values are observed along the Atlantic Ocean coast (coast eastern to 55°W longitude) while highest K_t values are observed North-West inland. Both spatial clustering provide an interesting interpretation of the solar resource space distribution in the region: the irradiation trend is a result of combining a cloudiness trend (K_t), more aligned to increase for West to East and more affected by the Atlantic Ocean proximity, with the latitude effect, that is associated with a South to North increasing trend of the extraterrestrial irradiation (H_o).

Figures 5 and 6 show, respectively for irradiation and K_t, the annual cycle for the entire domain (left) and the anomalous annual cycle for each cluster's centroid (right). Irradiation annual cycle over the entire region has a minimum peak in June-July and a maximum peak in December-January, as expected. K_t annual cycle shows that, compared to the annual average, slightly more clouds are observed in local winter (June) and slightly more clear skies are observed in local summer (December-January). The center cluster has a similar behavior to the entire domain in both cases. The other two clusters exhibit an above and below trend in relation with the average. Irradiation largest anomalies are observed during intermediate seasons, and

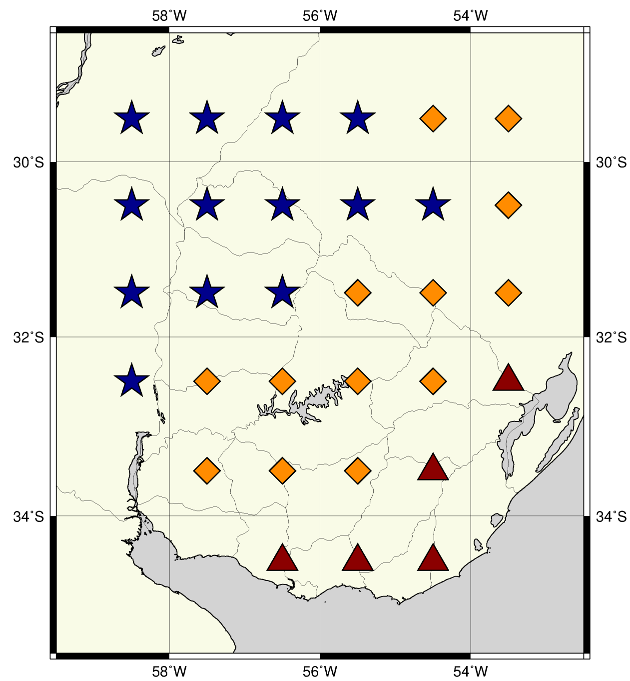


Figure 3: The three clusters found (see shapes and/or colors) according to annual cycles of monthly irradiation.

in August in particular. Cloudiness largest anomalies are observed also during August-September, but also December-January. The minimum difference between clusters is observed in local winter (June).

4.2. Clustering by seasonal and interannual variability

To identify regions with coherent seasonal and interannual variability of the solar resource, the monthly anomalies of K_t in the period 2000-2016 are used as the grid-point attributes. To simplify the time-series (31 vectors -one per grid point- of 202 monthly anomaly values of K_t), leaving only the more dominants effects, a PCA is applied. PCA returns a smaller dimensional subspace of the data that retains most of the original data information, reducing

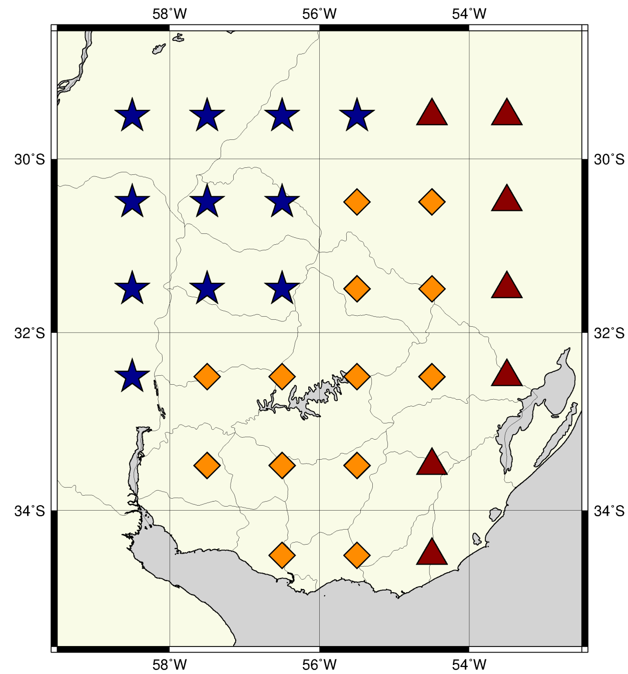


Figure 4: The three clusters found (see shapes and/or colors) according to annual cycles of monthly clearness index.

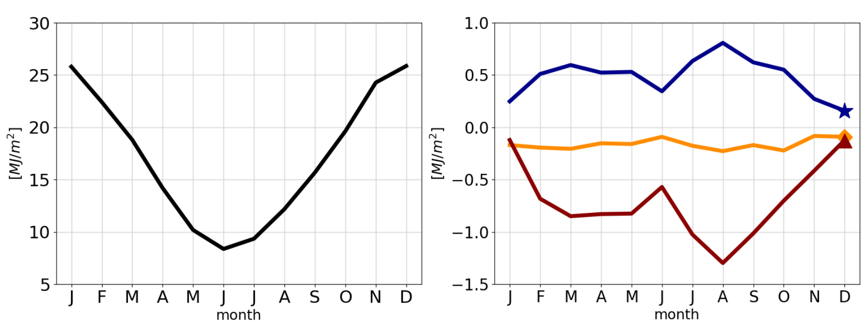


Figure 5: Mean annual cycle of monthly irradiation over the entire domain (left). Anomalous annual cycle for each cluster centroid (right). The triangle, diamond and star shapes correspond to Fig. 3.

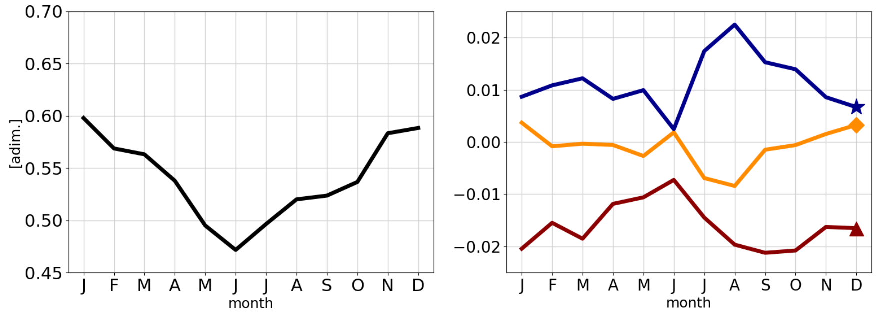


Figure 6: Mean annual cycle of monthly clearness index over the entire domain (left). Anomalous annual cycle for each cluster centroid (right). The triangle, diamond and star shapes correspond to Fig. 4

the complexity of the signal and filtering non typical effects. In this case, the first five principal components represent 92.6% of the total variance, so the clustering analysis is performed in the reconstructed time series based on these 5 main modes. The clusters obtained (see Figure 7) show a dominant latitudinal arrangement with a tilt perpendicular to the coast, contrary to the clusters performed based on the irradiation annual cycle.

Considering that ENSO is the main source of interannual variability in the region and its effects are known to vary both spatially and seasonally (Aceituno, 1988; Cazes-Boezio et al., 2003), these clusters will be used in the next section to quantify the ENSO effect. The rationale behind this is that these clusters represent better the interannual variability of solar irradiation in the region. These clusters were derived by using the correlation as distance (Eq. (2)), thus they are better suited for the correlation analysis in the following section.

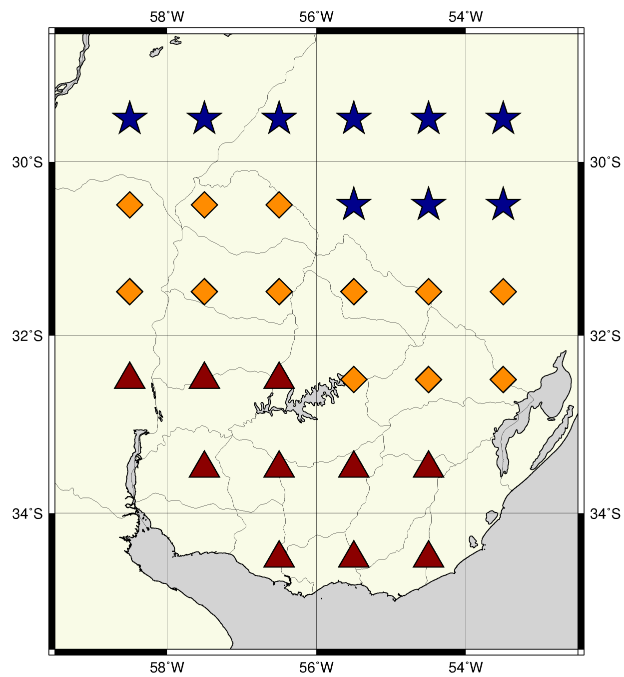


Figure 7: The three clusters found (see shapes and/or colors) according to monthly clearness index anomaly time series.

5. Influence of ENSO

In this section we explore the interannual variability of the solar resource and its seasonal variations in association with ENSO climatic phenomena. For this analysis the K_t anomalies are used, filtering the seasonal cycle and climatological trend, and isolating the results from the seasonal influence.

5.1. Correlation

Table 1 shows the Pearson correlation between the monthly N3.4 Index and the clearness index anomaly for each cluster average (Figure 7), together with p-value of the Student test. As mentioned in Section 3, correlations computed in this way do not account for the seasonality of ENSO signal in the local climate, which is known to be significant (Cazes-Boezio et al., 2003). The values obtained are still negative in all cases, although rather small. Note that high N3.4 signal is associated with wetter and cloudier weather in the region, therefore with lower clearness indices. Notably, the northern cluster, where a strongest ENSO signal is expected (Pisciottano et al., 1994), shows a correlation that is significant to a 99% confidence level ($1 - p \approx 99\%$), which is a very high value.

	Correlation	p-value
Southern Cluster (triangles)	-0.105	0.1386
Central Cluster (diamonds)	-0.137	0.0512
Northern Cluster (stars)	-0.183	**0.0091**

Table 1: Correlation between the cluster's centroids in terms of K_t anomaly and the N3.4 index. The third column indicates the p-value of the t-Student test.

5.2. Stratification

In this subsection, we explore how the clearness index stratifies with ENSO. As mentioned in Section 3, the WMW test results are not clear to quantify if data, stratified according to simultaneous N3.4 ($> +0.5$ °C and < -0.5 °C), come from the same statistical distribution or not. According to the local period of influence of N3.4 signal (Pisciottano et al., 1994), the test is performed over the sets of data stratified according to N3.4-NDJ for each cluster. In this scenario, the test rejects the null hypothesis with a confidence of 95% for all the clusters. This means that the data behaves differently according to the sign of N3.4-NDJ (p-values of 0.0021 for star-shaped sites in

Figure 7, 0.0029 for diamond-shaped and 0.0426 for triangle-shaped). Figure 8 shows the empirical Cumulative Distribution Functions (CDF) of K_t anomaly for each cluster for all data discriminated according to N3.4-NDJ $> +0.5$ °C (red) and N3.4-NDJ < -0.5 °C (blue), jointly with the all data CDF (black). The WMW test for each month separately was repeated and the three clusters showed different distributions (i.e. discard null hypothesis) mainly for March, April, November and December. In Table 2 the p-Values of the test performed over each month over the clusters are shown. We further analyze the influence of N3.4-NDJ within the year. Figure 9 showing the annual cycles of K_t anomaly conditioned to N3.4-NDJ.

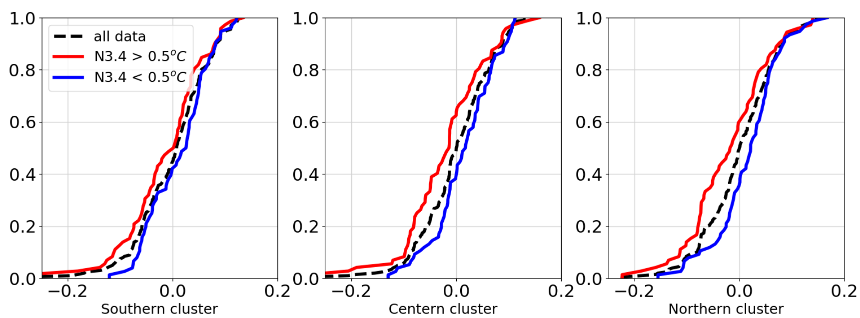


Figure 8: Empirical Cumulative Distribution Function (ECDF) of the monthly K_t anomaly for each cluster discriminating according to ENSO-NDJ. The black dashed line denotes the ECDF without ENSO discrimination.

5.3. *Seasonability*

In light of the previous subsection's results, we further analyze the influence of ENSO signal on the K_t anomaly for specific seasons: November-December and February to May, which largely coincide with the main seasons

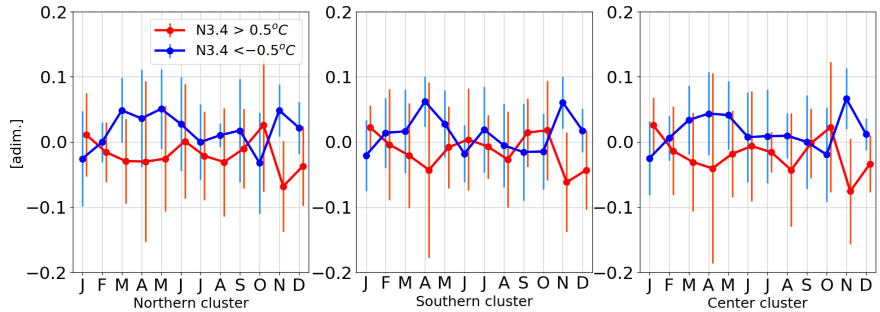


Figure 9: Each chart shows the annual cycles of K_t anomaly for each cluster discriminating according to N3.4-NDJ. The x-axis corresponds to the month. The dots represent the mean value for each month, while the bars represent the standard deviation.

of influence of ENSO on the local climate (Pisciottano et al., 1994). In order to correlate the K_t anomalies with N3.4-NDJ index, the K_t anomalies are averaged over each season. The p-values for these seasons are shown in Table 3, which shows that the K_t anomaly stratified according to the sign of N3.4-NDJ belong to different probability distribution with a high confidence level ($1 - p \geq 99\%$) for November-December season, but not for February to May. Results, which capture the interannual variability of the seasonally averaged influence of ENSO on the clearness index, are shown in Table 4.

Lastly, we analyze the spatial variability of ENSO influence on K_t anomalies for these selected seasons independently of the clusters. A correlation map between anomaly K_t series over the Nov-Dec and Feb-May periods with the N3.4-NDJ index is performed and results are shown in Figure 10. For the case of the Nov-Dec season, all sites are correlated with a 95% of confidence. The same holds for the Feb-May case with the exception of the three coastal points. Figure 11 shows the scatter plot of the clearness index, spatially

	Northern cluster (star)	Center cluster (diamond)	Southern cluster (triangle)
Jan	0.339	0.773	0.554
Feb	0.335	0.693	0.987
Mar	**0.026**	**0.035**	**0.064**
Apr	**0.006**	**0.012**	**0.030**
May	0.056	0.056	0.169
Jun	0.103	0.117	0.537
Jul	0.235	0.179	0.837
Aug	0.176	0.133	0.939
Set	0.988	0.826	0.511
Oct	0.266	0.231	0.455
Nov	0.088	**0.041**	**0.020**
Dec	**0.017**	**0.042**	**0.022**

Table 2: p-values of the WMW test for twelve trimesters discriminated according to N3.4-NDJ $> 0.5^oC$ and N3.4-NDJ $< -0.5^oC$ for each cluster.

	November-December series	February to May series
Southern cluster (triangle)	**0.0006**	0.8235
Center cluster (diamond)	**0.0008**	0.5034
Northern cluster (star)	**0.0009**	0.1388

Table 3: p-values of the WMW test for November-December and February to May of clearness index anomaly discriminated according to ENSO $> +0.5°$C and ENSO $< -0.5°$C for each cluster.

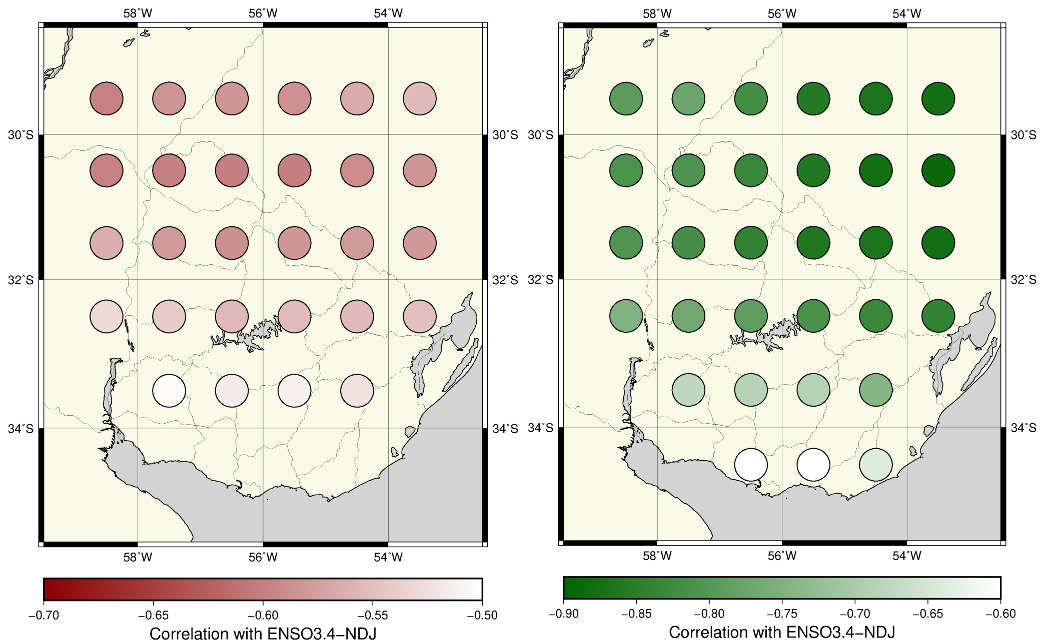
	November-December		February to May	
	correlation	p-value	correlation	p-value
Southern cluster (triangles)	-0.704	**0.0016**	-0.519	**0.0395**
Central cluster (diamonds)	-0.852	<**0.0001**	-0.586	**0.0170**
Northern cluster (stars)	-0.868	<**0.0001**	-0.594	**0.0152**

Table 4: Correlation and p-value for between N3.4-NDJ and seasonally averaged clearness index for each cluster. All correlations are statistically significant at 95% of confidence.

averaged over all sites that exhibit significant correlations (those depicted in Figure 10), and the N3.4-NDJ index, evidencing the strong influence of ENSO on the interannual variability of the solar resource in the selected seasons. To emphasize this behavior, the linear regression is added to the plot in order to visualize the trend.

6. Discussion

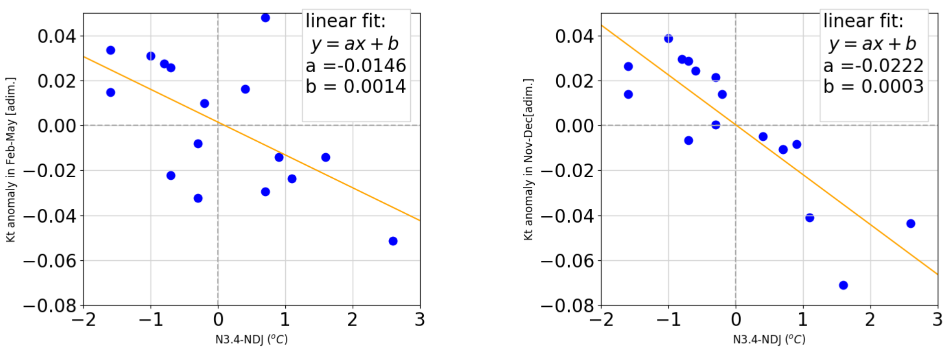
We characterized the spatio-temporal variability of the solar resource over a region of Southeastern South America -containing Uruguay- based on 17-years-long monthly series of solar irradiation. The 31-points gridded data set was derived from hourly estimates based on satellite imagery with a spatial resolution of few kilometers, validated against high accuracy radiometric land stations. A cluster analysis was performed based on the mean annual cycle (12 monthly values) of both solar irradiation and clearness index, rendering three regions in each case. Solar irradiation grows from southeast to northwest, consistent with existent solar maps, showing minimum spatial



(a) Feb-May season (b) Nov-Dec season

Figure 10: Correlation map between each site and the N3.4-NDJ for February-May (left) and November-December (right) seasons. Only sites with correlations significant to a 95% confidence level are shown.

differences between clusters during -austral- summer. This is due to the latitudinal variation of daily irradiation at the top of the atmosphere combined with the spatial distribution of the clearness index (K_t). The latter shows clusters that arrange roughly from east to west, with K_t increasing inland throughout the year, a reflection of the larger humidity and cloudiness closer to the Atlantic coastline. Largest differences between clusters occur from late winter through early fall, when the clearness index in the most continental cluster is approximately between 6% and 9% grater that of the more coastal region.



(a) Feb-May season

(b) Nov-Dec season

Figure 11: Linear regression of spatial average of the K_t February-May (left) and November-December (right) anomaly of the sites with a 95% confidence correlation with N3.4-NDJ. The linear regression added to the plot is not intended to model the behavior of the K_t anomaly but to illustrate its trend.

Next, we focused on interannual variability, which may be of interest for the management of an interconnected electric system with a large participation of solar energy. A cluster analysis was again performed, this time using the entire K_t anomalies time series as the attribute for each grid point and a correlation-based distance. The resulting clusters arrange from south-southwest to north northeast, with a different orientation of those based on the climatological annual cycle. Only the time series of the northernmost cluster shows marginally significant anti-correlation with ENSO, the largest source of interannual variability in the region. However, it is well known that ENSO signal on the regional climate is seasonal dependent, a property that is not captured in this cluster analysis. Therefore, an ENSO based stratification of the clearness index time series associated with each cluster was performed on moving trimesters. The results confirm previous studies of ENSO influence on the local climate (mainly on precipitation) that have identified late spring and early summer as the main season of influence and fall as a second -less significant- one. For November-December (February-May), the correlation of N3.4-NDJ with the centroid of every cluster is statistically significant to a 99% (95%) level, with increasing correlation to the north.

Motivated by the previous results, correlation between N3.4-NDJ and seasonal mean K_t for every grid point were computed for each season (November-December and February-May). The resulting correlation maps and associated scatter plots quite clearly show the strength and spatial distribution of ENSO signal in those seasons. This is consistent with previous studies for other variables, although the strength of the correlation during February-May seems higher than what has been reported for precipitation.

7. Conclusion

A first comprehensive regionalization of Uruguay's solar resource was performed with three different attributes of this variable, in all cases obtaining three spatial clusters. In the first case, based on GHI mean annual cycle, the results show the expected Southeast to Northwest gradient in total irradiation, consistent with previous solar resource climatological maps (Abal et al., 2010; Alonso-Suárez et al., 2014). In the second case, based on the clearness index mean annual cycle, the arrangement of the clusters shows an increase in K_t from East to West, associated with decreasing cloudiness inland. The third case, based on the interannual variability of the clearness index, results in clusters arranged from the Southwest to the Northeast. This last regionalization depicts clusters with similar clearness index anomalies (indicating temporal variability) at the monthly time-scale and can thus contribute to the management of the solar resource in the region. It is well established that ENSO has a significant seasonal impact on the climate of the region during extreme phases of this cuasi-periodic oscillation of the climate system, known as El Niño and La Niña. ENSO signal in the local hydro-climate is of such relevance, that the associated N3.4 index is already included in the electric system simulator used for planning and dispatch (Maciel et al., 2015), in particular due to its impact in the hydroelectric component of the system. It is thus natural to explore the relation between the solar resources and N3.4 index. In this sense, the most important result is the determination of a high negative correlation between ENSO (as represented by N3.4 index) and solar irradiation variations (as represented by K_t anomalies). This anti-correlation is more significant over the February-May and November-December periods, and allows to conclude with high confidence level ($\geq 99\%$ for the former period and $\geq 96\%$ for the latter) that during El Niño years lower solar resource will be available in the region, specially over the center and Northern region. The opposite occurs during La Niña years.

These results are very promising and potentially useful for the design of solar irradiation measurements networks and for the management of integrated electrical systems with increasing contributions of solar energy. A second stage of this study, in order to obtain further and more detailed information, would be to reassess the results with longer time series and higher resolution grids, which may allow to include sub-seasonal variability in the ENSO analysis.

Acknowledgments

The authors thank Mr. Martin Benenati that kindly provided the topography map of Figure 1 based on the Terra-Aster NASA satellite data.

References

Abal, G., D'Angelo, M., Cataldo, J., and Gutiérrez, A. (2010). *Mapa Solar del Uruguay, versión 1.0.* UCUR, Montevideo, Uruguay.

Aceituno, P. (1988). On the functioning of the southern oscillation in the south american sector. part 1: surface climate. *Monthly Weather Review*, 116:505–524.

Alonso-Suárez, R., Abal, G., Musé, P., and Siri, R. (2014). Satellite-derived solar irradiation map for Uruguay. In *Elsevier Energy Procedia*, volume 57, pages 1237–1246.

Alonso-Suárez, R., Abal, G., Siri, R., and Musé, P. (2012). Brightness-dependent Tarpley model for global solar radiation estimation using GOES satellite images: application to Uruguay. *Solar Energy*, 86(11):3205–3215.

Beyer, G. B., Costanzo, C., and Heinemann, D. (1996). Modifications of the Heliosat procedure for irradiance estimates from satellite images. *Solar Energy*, 56:207–212.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, Berlin, Heidelberg.

Cazes-Boezio, G., Robertson, A. W., and Mechoso, C. R. (2003). Seasonal dependence of ENSO teleconnections over South America and relationships with precipitation in Uruguay. *Journal of climate*, 16(8):1159–1176.

Ceballos, J. C., Bottino, M., and de Souza, J. (2004). A simplified physical model for assessing solar radiation over Brazil using GOES 8 visible imagery. *Journal of Geophysical Research: Atmospheres*, 109(D2).

Cebecauer, T., Šúri, M., and Perez, R. (2010). High performance MSG satellite model for operational solar energy applications. In *Proceedings of the American Solar Energy Society (ASES)*, pages 1–5, Phoenix, Arizona, United States.

Diabate, L., Blanc, P., and Wald, L. (2004). Solar radiation climate in africa. *Solar Energy*, 76.

Duda, R. O., Hart, P. E., and Stork, D. G. (2001). *Pattern Classification*. Wiley, New York, 2 edition.

Grimm, A. M., Barras, V. R., and Doyle, M. E. (2000). Climate variability in southern South America associated with El Niño and La Niña events. *Journal of Climate*, 13:35–58.

Iqbal, M. (1983). *An introduction to solar radiation*. Academic Press.

Jolliffe, I. T. (2002). *Principal Component Analysis, Second Edition*. Springer.

Justus, C. and Paris, M. V. (1986). Satellite-measured insolation in the united states, mexico, and south america. *Remote Sensing of Environment*, 20:57–83.

Laguarda, A., Giacosa, G., Alonso-Suárez, R., and Abal, G. (2020). Performance of the site-adapted CAMS database and locally adjusted cloud index models for estimating global solar horizontal irradiation over the Pampa Húmeda. *Solar Energy*, 199:295–307.

Maciel, F., Terra, R., and Chaer, R. (2015). Economic impact of considering el Niño-Southern Oscillation on the representation of stream flow in an electric system simulator. *International Journal of Climatology*, 35(14):4094–4102.

Mechoso, C. R. and Pérez-Iribarren, G. (1992). Streamflow in southeastern South America and the Southern Oscillation. *Journal of Climate*, 5:1535–1539.

Noia, M., Ratto, C., and Festa, R. (1993a). Solar irradiance estimation from geostationary satellite data: I. statistical models. *Solar Energy*, 51(6):449 – 456.

Noia, M., Ratto, C., and Festa, R. (1993b). Solar irradiance estimation from geostationary satellite data: Ii. physical models. *Solar Energy*, 51(6):457 – 465.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., F., and Vignola (2002). A new operational model for satellite-derived irradiances: description and validation. *Solar Energy*, 73:307–317.

Pisciottano, G., Díaz, A., Cazes, G., and C.R., M. (1994). El Niño-Southern Oscillation impact on rainfall in Uruguay. *Journal of Climate*, 7:1286 – 1302.

Polo, J., Gastón, M., Vindel, J., and Pagola, I. (2015). Spatial variability and clustering of global solar irradiation in vietnam from sunshine duration measurements. *Renewable and Sustainable Energy Reviews*, 42:1326–1334.

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

Rigollier, C., Lefevre, M., and Wald, L. (2004). The method Heliosat-2 for deriving shortwave solar radiation from satellite images. *Solar Energy*, 77(2):159–169.

Tarpley, J. (1979). Estimating incident solar radiation at the surface from geostationary satellite data. *Journal of Applied Meteorology*, 18(9):1172–1181.

Trenberth, K. E. (1997). The definition of El Niño. *Bulletin of the American Meteorological Society*, 78(12):2771–2778.

Watanabe, T., Takamatsu, T., and Nakajima, T. Y. (2016). Evaluation of variation in surface solar irradiance and clustering of observation stations in japan. *Journal of Applied Meteorology and Climatology*, 55(10):2165–2180.

Wilks, D. (2011). *Statistical Methods in the Atmospheric Sciences, 3rd edition.* International geophysics series (v100). Academic Press.

Zagouras, A., Inman, R., and Coimbra, C. (2014). On the determination of coherent solar microclimates for utility planning and operations. *Solar Energy*, 102:173–188.

Zagouras, A., Kazantzidis, A., Nikitidou, E., and Argiriou, A. (2013). Determination of measuring sites for solar irradiance, based on cluster analysis of satellite-derived cloud estimations. *Solar Energy*, 97:1–11.

Zagouras, A., Kolovos, A., and Coimbra, C. (2015). Objective framework for optimal distribution of solar irradiance monitoring networks. *Renewable Energy*, 80:153–165.