Benchmarking of modeled solar irradiation data in Uruguay at a daily time scale

Iñaki Sarazola, Agustín Laguarda, Juan C. Ceballos, and Rodrigo Alonso-Suárez, *Senior Member, IEEE*

***Abstract*—Accurate solar radiation data are required for solar energy development. In absence of long-term ground measurements, practitioners rely on modeled data, which typically is of unknown uncertainty. This article benchmarks solar irradiation estimation models over Uruguay, analyzing uncertainty behavior and providing recommendations. The performance of six models is evaluated using controlled-quality ground measurements for an extended period, being one of the few benchmark studies in Latin America. The LCIM, a model specially adapted for the region and based on GOES-East satellite images, exhibits the highest accuracy and spatial consistency with a remarkably low root mean squared deviation of 6% and mean bias of less than 1%. The NSRDB and GL1.2 estimates have also a competitive performance and are well-suited alternatives. The MERRA2 database presents high deviations and should not be an option for solar resource assessment in the region without post-processing. This research is a first step towards a South American benchmark and provides information on which estimation models are suitable for large-scale solar energy projects in the region.**

***Index Terms*—Solar resource assessment, satellite estimation, radiation models, GHI.**

I. Introduction

PRECISE on-site solar resource assessment is required for solar energy project sizing, simulation, and financial evaluation. The uncertainty of the solar data sets used for projects' development is the main contributor to the economic risk assessment of solar photovoltaic (PV) large-scale power plants endeavors [1, 2]. This assessment requires controlled-quality solar irradiation data with a span of more than a decade [3]. This allows for predicting the long-term yield of a solar project and its inter-annual and inter-monthly variability, which directly affects revenue and cash flow estimations. As ground measurements with these features are costly in equipment, maintenance, and human resources, it is unlikely that such measured information will be available for arbitrary project locations. To solve this issue, industry practitioners rely on modeled data from different sources, namely, satellite-based estimates or numerical weather models (reanalysis data). Depending on the data source, information without local verification or site adaptation [4] is prone to significant biases and uncertainties, that can undermine the viability of solar energy projects.

Satellite models are based on high time-rate, moderate-resolution, geostationary satellite images. These images provide the cloud's knowledge over large territories with an intra-hour refresh rate (10 or 15 minutes) and 500-1000 m space resolution. This resolution is enough for large-scale solar energy projects, as PV plants typically comprise 1 to 4 image pixels. As satellite models use as input the actual cloud cover and atmospheric information, i.e. it is not simulated by resolving the complex atmospheric dynamic, they are the preferred tool for solar resource assessment [5–8]. Their utilization is recommended, provided the models are locally validated or adjusted to controlled-quality ground measurements.

Several satellite-based models estimate solar irradiation [9–15]. Their performance is uneven around the globe, as their accuracy depends on input data availability, climatic particularities of each site, and satellite view angle, among others, which means that local studies are necessary. Many scientific works address the local evaluation and/or site adaptation of such models in different parts of the world, typically for isolated models or a reduced set of them. These works are difficult to compare, as different climates, metrics, periods, and quality procedures are utilized. Benchmarking initiatives are rare or recent [7, 16, 17]. An effort by Yang & Bright [7] evaluates 8 data sources (six satellite-derived and two reanalysis databases) for 57 BSRN[1] sites worldwide at an hourly scale and with the maximum available data span at each site. Under the popular root mean square metric, a commercial satellite product (SolCast estimates) showed the best overall performance, and the other free-available satellite-derived data were significantly better than reanalysis products. This study only used four Brazilian sites that are currently active in the BSRN. Due to its rather low participation in the BSRN, Latin America is always misrepresented in these kinds of international studies and the region has not yet undergone its own benchmarking initiative with continental coverage.

In this work we make a first step in this direction, addressing the southern part of the South American Pampa Húmeda area (Uruguay's territory). The main contributions of this article are to provide a ranking and typical uncertainty values for modeled solar radiation estimates in a subpart of Southeastern South America (SESA [18]) that can guide solar technology researchers and industry data users. Six modeled data sources are evaluated at a daily scale with a 4-years data span in 7 controlled-quality measuring sites distributed across this region. The estimates are validated against the same data set and metrics. The objective is to quantify the data end-user uncertainty, evaluating the modeled data as obtained from the different web services (data providers). Further analyses regarding the models' seasonal and cloudiness behavior are provided. One reanalysis data source is included to show the significant performance downgrade in comparison to the satellite-based alternatives. The work comprises global horizontal irradiation (GHI) and addresses models with different

[1]Baseline Solar Radiation Network, https://bsrn.awi.de/.

satellite inputs, including GOES-East, Meteosat, and low-orbit Terra and Aqua satellites.

The manuscript is organized as follows. Section II describes the ground measuring stations and the main characteristics of the region under study. Section III briefly introduces the considered models. Performance metrics and quality procedures, applied both to the measured and modeled data, are presented in Section IV. Results and analysis are provided in Section V, and finally, Section IV summarizes our main conclusions.

II. Ground Stations

The measuring sites are listed in Table I, including their station code and precise location. Their geographical distribution can be observed in Fig. 1. These sites are part of the Solar Irradiance Continous Measurement Network (RMCIS, http://les.edu.uy/rmcis/) of the Solar Energy Lab in Uruguay (Universidad de la República). The stations are equipped with class A or B pyranometers (according to the ISO 9060:2018 standard) for GHI measurements. All the sensors are calibrated every two years with the local standard, with traceability to the world's primary radiometric reference in the World Radiation Center (PMOD/WRC[2]). The RMCIS registry and sampling protocol follows the BSRN recommendations [19]. Thus, solar irradiance data is registered at a 1-minute time rate as the average of six 10-second instantaneous samples. This represents a good trade-off between capturing the high temporal solar radiation variability, keeping the data set size operationally manageable, and accounting for the thermopile pyranometers' time constant (which is < 10 seconds for Class A equipment). A time span of four years for all sites is considered for this study (between 01/2018 ad 12/2021). Quality control procedures are applied to the 1-minute irradiance data sets. Then, daily values are computed for the model's validation as explained in Subsection IV-A.

TABLE I
Location of the Measuring Sites

Site	Code	Latitude (°)	Longitude (°)	Altitude (m)
Artigas	AR	-30.398	-56.512	136
Salto	LE	-31.283	-57.918	56
Tacuarembó	TA	-31.739	-55.979	142
Colonia	ZU	-34.338	-57.690	70
Canelones	LB	-34.672	-56.340	38
Montevideo	AZ	-34.918	-56.167	58
Rocha	RC	-34.489	-54.320	20

The area under study is the southeast part of the Pampa Húmeda region of South America. These sites can be considered to be representative of the broader Pampean territory, especially LE, AR, and TA, which do not have coastal influence. Most sites are classified under the updated Köppen-Geiger climate classification [20] as Cfa (temperate with hot summers and no dry seasons), except for the coastal RO site, influenced by the Atlantic Ocean, which is classified as Cfb (same as Cfa but with warm summers). The solar resource short-term variability of the area is intermediate [21], meaning challenging sky conditions for satellite models, as clear-sky, partly cloudy, and overcast conditions alternate.

[2]Physikalisch-Meteorologisches Observatorium Davos, www.pmodwrc.ch.

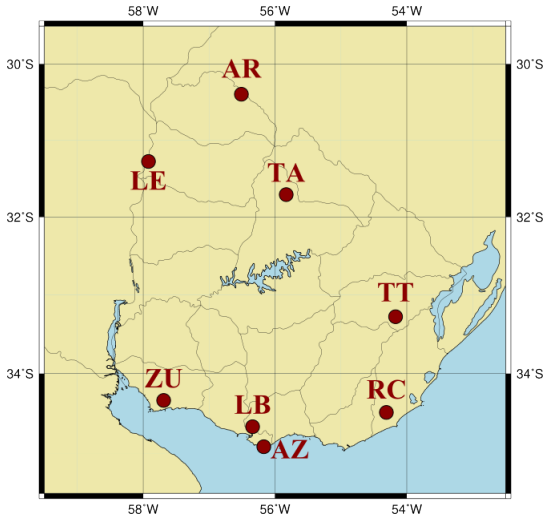


Fig. 1. Geographical distribution of the measurement stations.

III. Models

This section presents the models to be assessed. It includes five satellite-based models whose estimates can be freely downloaded from web portals. It also includes the MERRA2 reanalysis data set which is publicly accessible.

A. LES – LCIM

This model is based on the simple formulation that [9, 15],

$$G_h = G_h^{csk} \times F(\eta) \quad \text{with} \quad \eta = \frac{\rho - \rho_{\min}}{\rho_{\max} - \rho_{\min}} \tag{1}$$

in which the estimates of a clear-sky model (G_h^{csk}) are modulated by a cloud attenuation factor (F). This factor is, in turn, a function of a satellite-derived cloud index, η, defined in Eq. (1) in a similar way than the original index proposal [22]. This family of models is known as Cloud Index Models (CIM), and the acronym LCIM stands for the LES lab implementation [15] with local coefficients and parameters.

The LCIM operational implementation adopts the ESRA model [23] for the GHI clear-sky estimates. This model represents the cloudless atmosphere using Linke Turbidity [24] factors (T_L) for an air mass of 2. For the LCIM, a seasonal T_L cycle derived from the GHI measured data is used [25]. This is a single mean cycle for the whole region, averaged over 5 years of data and 10 measuring stations. This simplification can be done without losing much precision but implies the model's inability to represent small day-by-day clear-sky changes due to aerosols or water vapor content variability. Finally, the LCIM proposes a linear relationship between the attenuation factor ($F(\eta)$ and the cloud index:

$$F(\eta) = a(1 - \eta) + b. \tag{2}$$

The motivation of this linearity is explained in [9, 15, 26]. An example of the relationship between F and $(1 - \eta)$ is showed in Figure 2 [26].

The parameters a and b have low spatial variability within this region [15]. This allows using a single set for the entire area. LCIM operation version uses empirical parameters

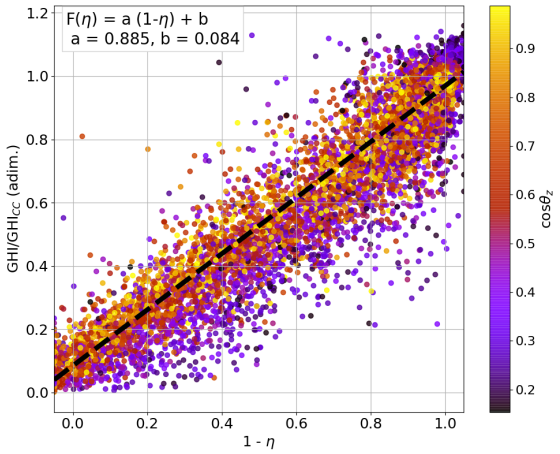


Fig. 2. Example of the linearity between F and $(1-\eta)$ for hourly data (LE site) [26]. The color of the dots denotes the cosine of the solar zenith angle.

spatially averaged over ten Pampean sites. This model runs with GOES-East satellite images (GOES16) with a 10-minutes time rate. The LCIM estimates can be freely downloaded from http://les.edu.uy/online/ghisat/.

B. CAMS – Heliosat-4

Heliosat-4 [14] is a full physical model that uses a fast but still accurate, approximation to the libRadRran radiative transfer model (RTM) [27]. The model uses various satellite inputs and the Copernicus Atmosphere Monitoring Service (CAMS) reanalysis database. Cloud properties are derived from Meteosat Second Generation (MSG) satellite images at their 15-minute time rate using an adapted APOLLO (AVHRR[3] Processing Scheme Over Clouds, Land, and Ocean) scheme [28]. Heliosat-4 operational version takes the form of abacuses (lookup tables), enabling its fast computation. This model has already been validated for the region, but not benchmarked, at hourly [15] and daily [29] time scales. Data can be freely downloaded from the CAMS website on different time bases via interpolation or integration of the original 15-minute time rate. It shall be noted that the all-sky data is only available for the MSG satellite coverage, with increasing uncertainty with the satellite view angle [15].

C. NASA – CERES

The NASA-POWER platform provides a wide range of information concerning renewable energy, including solar radiation. It has different data sets depending on the time range. For the period 2018-2021, the platform supplies data from the CERES SYN1deg[4] product (Edition 4.1) [30, 31]. Solar irradiance (downward shortwave flux in this context) is computed using an RTM with a $1^\circ \times 1^\circ$ space resolution. The aerosol data required for clear-sky modeling are obtained from a chemical transport model that assimilates aerosol optical depth retrievals from the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) radiances. Cloud properties are derived from MODIS and GOES satellite data. Although MODIS data is available only two times at daylight and 3-hourly GOES images are used, the RTM is run with an hourly time step, which is the lowest time resolution for this product. The data are freely accessible in https://power.larc.nasa.gov/, being this a popular product among Latin American solar industry practitioners.

D. INPE/CPTEC – GL1.2

GL1.2 model [11] is a physical solar radiation model developed and operated by the INPE/CPTEC satellite division (http://satelite.cptec.inpe.br/), Brazil. The model poses a simple solution to the shortwave radiative balance in the Earth-Atmosphere system. The solar broadband spectrum is separately considered in its three main intervals; visible, ultraviolet, and infrared. Different modeling hypotheses are applied to each sub-band. Then, the broadband GHI is obtained by the addition of the three components. GL's current version is adapted to work with GOES16 visible channel images. The model considers the clear-sky attenuation by Rayleigh dispersion, water vapor, and stratospheric ozone. These latter two are considered constant across regions in South America. Clouds modeling is done by using the cloud index expression of Eq. (1), with $\rho_{\min} = 0.090$ and $\rho_{\max} = 0.465$. This maximum value is set to represent the albedo transition between stratiform and cumuliform clouds, as seen by the satellite, and was determined by cloud classification of satellite images across Brazil [32, 33]. Calculated in this manner, and following how the cloud information is introduced into the model, this cloud index can be interpreted as cloudy coverage (the fraction of the pixel area covered by clouds). GL 1.2 data for this work were calculated ad-hoc by the CPTEC's team for the stations listed in Table I and were provided with a daily time rate.

E. NREL – NSRDB

The National Solar Radiation Data Base (NSRDB) currently provides the PSM v3 (Physical Solar Model) [34]. This is an RTM that uses various sources of remotely sensed information, including GOES satellite images and the AVHRR Pathfinder Atmospheres-Extended (PATMOS-x) system for cloud properties, and MODIS and MERRA[5]-2 aerosol optical depth products. The available information has a 4 km space resolution and a 30-minutes time rate. To meet this output with serially complete and consistent records, re-gridding, temporal interpolation, time shifting, and gap filling are implemented, including the downscaling of the MERRA2 data set to the 4 km space resolution. Considering the wide range of information that is ingested, especially the large volume of GOES satellite images, the computation of the model requires a high-performance computational method. For this, the FARMS[6] algorithm is utilized [13]. The data are publicly available in https://nsrdb.nrel.gov/.

[3]Advanced Very High-Resolution Radiometer.
[4]Clouds & Earth Radiant Energy System (CERES) Synoptic 1° (SYN1deg).
[5]The Modern-Era Retrospective analysis for Research and Applications.
[6]Fast All-sky Radiation Model for Solar applications.

F. NASA – MERRA2

MERRA2 [35] is NASA's updated global reanalysis data set. It provides physically consistent and complete atmosphere information for a large set of meteorological variables, including ground-level solar irradiance. The model assimilates various data sources; radiances and retrievals from a wide range of satellites, ground, and airborne measurements, and pibal and dropsondes observations, among others. The resulting data set covers the 1980-today period with hourly time step and $0.5^\circ \times 0.625^\circ$ latitude-longitude space resolution. MERRA2 improves aerosol modeling and related input information that affects the clear-sky solar irradiance output. It includes other physics-related improvements that involve water vapor content and clouds' modeling [36], affecting all-sky solar irradiance estimates. Due to these upgrades, the data set is expected to have a better solar irradiance performance in comparison with its predecessor (MERRA versión 1 [37]). The data are freely available at Giovanni's portal: https://giovanni.gsfc.nasa.gov.

IV. Methodology

Satellite estimates are validated here on a daily time basis. This requires the daily data integration of ground measurements and model estimates originally obtained with intra-day frequency (LCIM, NSRDB, and MERRA2). Prerequisites for ground data daily integration are quality inspection on an intra-day time basis and completeness. This latter also applies to the models' estimates. This section describes data processing and performance metrics used for the assessment.

A. Ground data processing

A baseline quality procedure was considered at a 1-minute time rate, including visual inspection and the BSRN filter [19] application for GHI data with local coefficients. This procedure excludes only anomalous, extremely rare, or highly improbable data. As can be seen in Table II, the discarded samples are between 4.2% and 0.3% of the daylight valid samples (first row) depending on the site, which results in an overall discard of 1.6%.

The daily integration procedure requires complete daylight records, e.g. there should not be daylight gaps or missing values to perform the integration of a given day. Although the discard rate is low, gaps with short lengths may affect a large number of days under this prerequisite. Solar assessments are importantly influenced by the samples' distribution regarding clear-sky, partly cloudy, and overcast conditions. This distribution is affected by quality control procedures and the number of years used for evaluation. Relying on a single data year or biased decimated daily data can lead to non-representative evaluations. Therefore, daily completeness to perform the integration and longer evaluation periods should be favored, particularly when only a few intra-day minutes are required to recover a complete day. In most cases, accepting little interpolation, say, some minutes gap length, can recover several daily values. An interesting analysis regarding the daily data recovery capacity of a few minutes gap-filling is presented in Fig. 3. It shows the available complete days (y axis) in each station as a function of the admitted daily interpolated samples (x axis). It is observed that the interpolation of small gaps quickly recovers days, meaning that non-complete days with small gaps are more common. Also, it shows that the interpolation of a large number of missing samples has little value, recovering a few days with the additional cost of adding higher uncertainty. A trade-off maximum interpolation of 30 1-minute samples a day (black dotted line) was set in this work, which allows having at least 1200 days in all sites. Most of these cases are not sequential samples, having an uneven distribution during the day. The resulting interpolation does not exceed 0.47% of the filtered data in any station, being 0.23% in overall terms.

TABLE II
Quality control and daily integration summary

	AR	LE	TA	ZU	LB	AZ	RC
Raw (min)	1012550	1047507	1045943	1047792	1048631	1035811	1043687
Disc. (min)	42254	24091	7152	3661	2869	22061	16273
Disc. (%)	4.17	2.30	0.68	0.35	0.27	2.13	1.56
Valid. (min)	970296	1023416	1038791	1044131	1045762	1013750	1027414
Interp. (min)	3843	4850	1277	1576	753	3451	596
Interp. (%)	0.40%	0.47%	0.12%	0.15%	0.07%	0.34%	0.06%
Valid days	**1240**	**1300**	**1432**	**1447**	**1454**	**1333**	**1433**

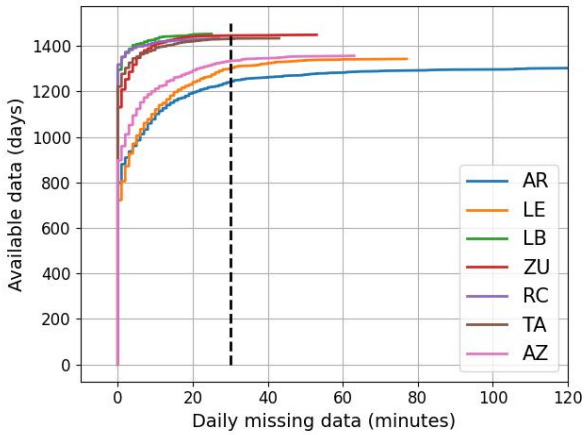


Fig. 3. Visual Analysis of interpolation limit for a day validity. The vertical line is traced to show how the discriminant factor will affect the data set.

Interpolation can be done in a smart way using the clearness index, k_t, defined as the ratio between the GHI and the corresponding top of the atmosphere irradiance [38]. This allows the interpolation of a quantity that does not have an intra-day geometrical trend associated with the Sun's apparent movement. Daylight linear interpolation of k_t was applied, from which the GHI values can be calculated using the extraterrestrial irradiance. After interpolation, only complete days are considered for daily integration. Integration was performed with the common Riemann sum. It shall be noted that the procedure described above, based on 1-minute data and a sparse 1-minute gap-filling, is a more accurate daily data retrieval approach than simply using hourly data (measured or averaged) or daily means from the available samples, which are rather common procedures.

B. Estimates processing

Models' estimates are provided in different time scales: hourly, 10-minute, or daily. Those that come on a daily scale are ready to use (considering quality flags, if any). Those that have a lower (intra-day) time scale, need to be daily integrated. Most of the estimates are based on satellite models, and satellite information is not always available due to technical issues. This produces gaps in some models' data sets. To avoid discarding almost complete days, the same criterion of 30-minute gap interpolation tolerance was used.

C. Performance metrics

An extended set of performance indicators are considered, covering the most common metrics utilized in the solar resource assessment field [39]; mean bias deviation (MBD), mean absolute deviation (MAD), root mean squared deviation (RMSD), Kolmogorov-Smirnov integral (KSI), and Pearson's correlation. As researchers do not always select the same metrics, the choice here is for a large set, so baseline values are given for all of them and performance comparisons are facilitated. The first three metrics are defined as follows,

$$\text{MBD} = \frac{1}{N}\sum_{i=1}^{i=N} d_i, \qquad \text{MAD} = \frac{1}{N}\sum_{i=1}^{i=N} |d_i|, \qquad \text{RMSD} = \sqrt{\frac{1}{N}\sum_{i=1}^{i=N} d_i^2}, \tag{3}$$

where $d_i = \widehat{y_i} - y_i$ is the difference between the model estimates ($\widehat{y_i}$) and the reference measured data (y_i), and N is the number of samples in the daily data series. The MBD measures the systematic bias that a model can introduce into the long-term assessment, while the MAD and RMSD measure the error dispersion with the absolute and quadratic norms, respectively. Due to its increased sensitivity to outliers, the RMSD is frequently employed in this field. Here, both are reported for completeness. These metrics account for data in individual terms (i.e. sample to sample) and are expressed here in their relative form (rMBD, rMAD, and rRMSD) as a percentage of the measurement's mean value. The KSI, on the other hand, measures statistical similarity. It is defined as [40],

$$\text{KSI} = \int |\hat{F}(y) - F(y)|\ dy, \tag{4}$$

thus integrating the absolute difference between the cumulative probability function of the estimates ($\hat{F}$) and the measurements (F) in the whole range of the target variable y (which in this case is the GHI). This provides a negatively-oriented metric (the lower, the better) that quantifies the statistical difference between both data sets. Usually, a low RMSD also leads to a low KSI, although they formally measure different characteristics of the estimated performance. Finally, Pearson's correlation coefficient (Corr.) is also included, quantifying the intensity of the linear relationship between estimates and measurements.

V. Results and Discussion

The assessed metrics are presented in Table III for each model and site, including also each metric site-average (last column). The validation is performed over the period 01/2018–12/2021. Due to availability reasons, the periods of the GL1.2 and NSRDB estimates are shorter than the rest (04/2019-12/2021 and 1/2019-12/2020, respectively), and the last two rows of Table III provide further information on this issue. This is, the data set size and measurement's average when considering the different spans.

GL1.2 estimates show the smallest rMBD through sites, in the range of -2.0% and $+0.5\%$ with a site mean of -0.2%. MERRA2 estimates, on the contrary, present the highest bias by large, being positive and consistent across all sites, from $\simeq +5\%$ to $+11\%$. The other models present small bias in overall terms, with two groups: LCIM and NSRDB having values in the range of $[-1.5, +2.6]\%$ and a site average of $+1.3\%$, and CERES and Heliosat-4 with values within $\simeq \pm 5\%$ and slightly higher site averages (in absolute terms). Regarding the rRMSD metric, three groups can be distinguished. The LCIM presents the lowest values across sites and an impressive site average of 6%. Then, NSRDB, GL1.2, CERES, and Heliosat-4 are in the second group, with average values between 7.3% and 8.3%. Some rRMSD exceptions can be observed within models and sites, for instance, a low 4.9% value for the NSRDB in LE and some values above 9.5%, as for GL1.2 in ZU, Heliosat-4 in AZ, and CERES in LB, AZ, and RC. Finally, MERRA2 estimates are in the last group, showing important inaccuracies with rRMSD values surpassing 20% in almost all sites. The rest of the metrics provide similar insights, in particular, the rMAD follows closely the rRMSD. It shall be noticed that the NSRDB estimates have the lowest KSI, being similar to the one obtained by the LCIM, and that the LCIM estimates present remarkably high correlations. The previous analysis suggests that the LCIM, GL1.2, and NSRDB are the best-performing models for daily scale GHI estimates in the region, with an important rRMSD gain in favor of the LCIM. It also discards the MERRA2 GHI estimates utilization, as there are several much better-performing options freely available. The CERES model provides an interesting performance considering it only uses 3h GOES-East satellite images. The estimates' behavior can be observed graphically in Fig. 4. This figure provides scatter plots (estimates vs. reference) for each model in the LE site, and enables to arise similar conclusions but in a qualitative manner. The color code is used to show the four seasons of the year.

A. Seasonal performance

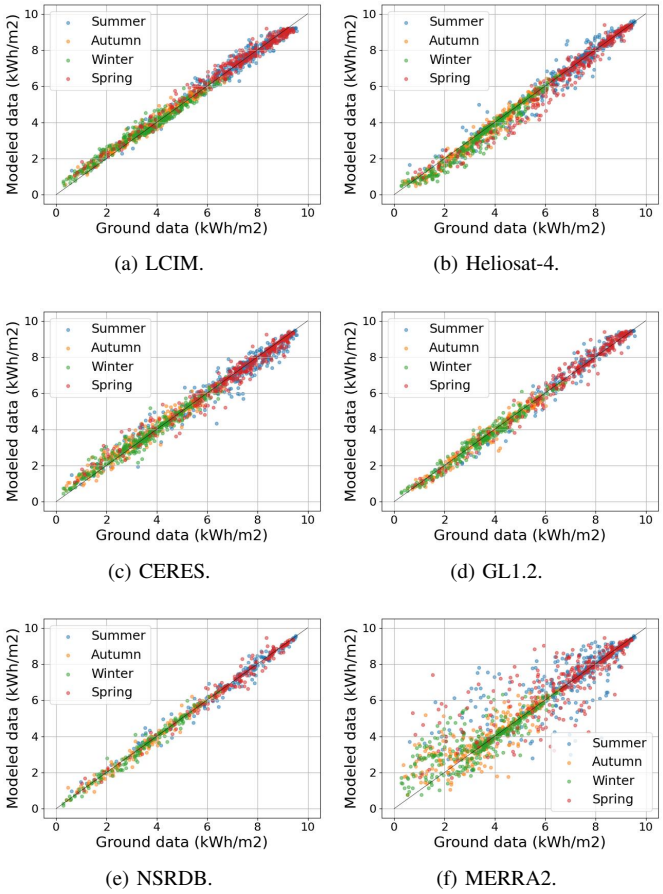
Seasonal performance is generally of interest in meteorological analysis, even more for solar radiation, as for the GHI there is a well-known seasonal geometrical behavior. In the previous analysis, the rMBD and rRMSD indicators were mostly discussed, as they are generally preferred in the solar resource assessment field. Also, observed performances are fairly uniform across sites. Following this and due to ease of presentation, only these two metrics are discussed from now on and as site averages.

TABLE III
MODELS' PERFORMANCE IN ESTIMATING GHI AT A DAILY TIME SCALE.

Model	metric	AR	LE	TA	ZU	LB	AZ	RC	Average
LCIM	rMBD (%)	2.1	1.6	2.1	-0.2	1.2	0.5	1.7	**1.3**
	rRMSD (%)	5.9	5.5	7.0	5.7	5.6	6.0	6.2	**6.0**
	rMAD (%)	4.6	4.3	5.3	4.7	4.4	4.8	4.9	**4.7**
	Corr.	0.994	0.994	0.992	0.994	0.995	0.994	0.994	**0.994**
	KSI (kWh/m^2)	0.122	0.108	0.114	0.104	0.091	0.103	0.092	**0.105**
NSRDB*	rMBD (%)	2.6	0.1	2.7	0.3	2.3	-1.5	2.6	**1.3**
	rRMSD (%)	7.9	4.9	9.1	6.3	7.4	6.9	8.8	**7.3**
	rMAD (%)	4.5	3.3	5.7	4.4	5.3	4.6	6.1	**4.8**
	Corr.	0.989	0.994	0.986	0.993	0.992	0.991	0.989	**0.990**
	KSI (kWh/m^2)	0.131	0.052	0.134	0.076	0.123	0.082	0.129	**0.104**
GL1.2*	rMBD (%)	0.5	0.1	-0.1	0.3	-0.9	0.3	-1.8	**−0.2**
	rRMSD (%)	5.8	5.9	6.9	13.5	6.9	7.1	8.6	**7.8**
	rMAD (%)	4.3	4.3	5.1	9.0	5.0	5.3	6.2	**5.6**
	Corr.	0.994	0.993	0.992	0.969	0.992	0.991	0.989	**0.988**
	KSI (kWh/m^2)	0.081	0.082	0.128	0.158	0.133	0.090	0.145	**0.117**
CERES	rMBD (%)	0.4	0.3	1.1	1.3	5.3	4.1	3.6	**2.3**
	rRMSD (%)	6.5	7.0	7.5	7.3	9.8	11.6	9.8	**8.5**
	rMAD (%)	4.8	4.8	5.5	5.2	7.1	8.0	7.0	**6.1**
	Corr.	0.992	0.989	0.990	0.990	0.988	0.976	0.985	**0.987**
	KSI (kWh/m^2)	0.101	0.111	0.112	0.089	0.251	0.209	0.164	**0.148**
Heliosat-4	rMBD (%)	-0.8	-2.3	0.2	-2.5	-1.6	-5.0	-1.3	**−1.9**
	rRMSD (%)	8.3	7.3	7.6	7.9	8.6	10.1	8.6	**8.3**
	rMAD (%)	5.9	4.9	5.6	5.6	6.0	7.0	6.1	**5.9**
	Corr.	0.987	0.990	0.991	0.990	0.988	0.985	0.988	**0.988**
	KSI (kWh/m^2)	0.137	0.148	0.124	0.152	0.133	0.251	0.116	**0.152**
MERRA2	rMBD (%)	6.3	4.7	10.3	6.3	8.6	6.1	11.0	**7.6**
	rRMSD (%)	20.5	19.6	24.6	20.5	21.9	19.7	23.9	**21.5**
	rMAD (%)	13.3	12.0	16.3	13.3	14.6	12.8	16.4	**14.1**
	Corr.	0.917	0.907	0.903	0.924	0.925	0.927	0.921	**0.918**
	KSI (kWh/m^2)	0.319	0.248	0.496	0.308	0.407	0.303	0.497	**0.368**
# Data		1220	1279	1409	1422	1430	1311	1410	**1354**
Ground measurements mean (kWh/m^2)		5.0	5.3	4.8	4.9	4.7	4.9	4.5	**4.9**
# Data* GL1.2 | NSRDB		800|520	870|644	959|721	971|724	976|728	898|660	958|706	**919|672**
Mean* GL1.2 | NSRDB (kWh/m^2)		5.0|5.0	5.2|5.2	4.8|4.8	4.8|4.9	4.7|4.7	4.8|5.0	4.5|4.5	**4.8|4.9**

Fig. 5 displays the average rMBD and rRMSD discriminated by seasons. The measurements' mean values for each season are provided in the caption as a reference. Fig. 5a shows that the relative bias is generally more significant (in magnitude) in autumn and winter for most models. The only exception to this is the NSRDB, which remains low in both seasons. GL1.2 confirms its lower bias across seasons and CERES shows a high downgrade in the autumn-winter period. The rRMSD behavior (Fig. 5b) is consistent with worse metrics in autumn-winter, as a noticeable increase of $\simeq$ 2-4% can be observed between spring-summer and autumn-winter for all models. LCIM presents the lowest rRMSD across almost all seasons, with the only exception of winter, in which the NSRDB outperforms. GL1.2 estimates, although not having the lowest rRMSD in any of the seasons, remain competitive in all of them. The seasonal analysis, in overall terms, is consistent with the analysis using the whole data set span; LCIM and NSRDB stand out for their low rRMSD, closely followed by GL1.2, which also stands out for its low bias. MERRA2 has no season with competitive performance.

The higher relative metrics in autumn-winter may be associated with lower average values. However, when absolute metrics are analyzed, there are only slight differences from the previous analysis. For the MBD the results are, in fact, analogous to those of the rMBD. The RMSD shows some different behavior, it remains nearly constant throughout the year for the LCIM and CAMS, but increases a little in spring-summer in comparison to autumn-winter for other models.

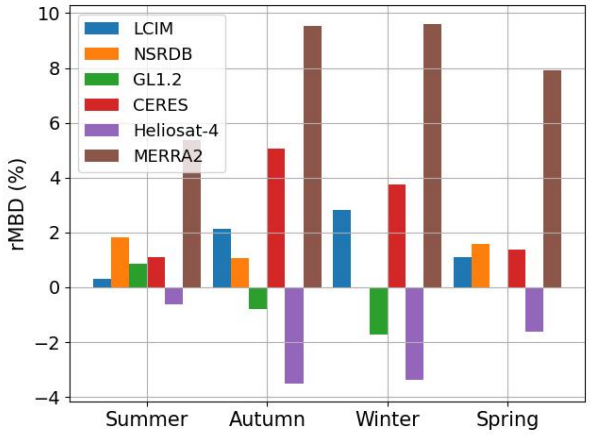


(a) LCIM. (b) Heliosat-4.

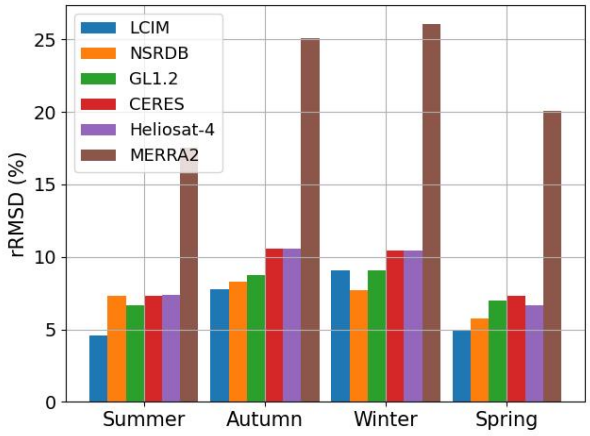
(c) CERES. (d) GL1.2.

(e) NSRDB. (f) MERRA2.

Fig. 4. Model estimates vs. measurements for each model in the LE site. Seasons are shown in different colors.



(a) rMBD metric.



(b) rRMSD metric.

Fig. 5. Performance metrics in each season. The measurement's mean in kWh/m^2 are: 6.7 (summer), 3.3 (autumn), 3.2 (winter) and 6.5 (spring).

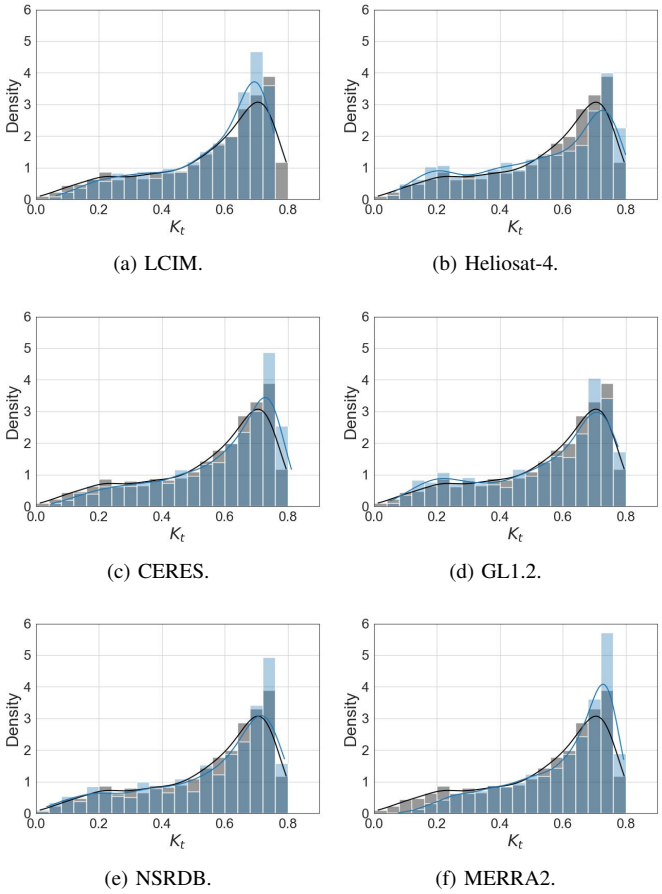
B. Performance dependence on cloudiness

Cloudiness discrimination is addressed by using the daily clearness index (K_t) as a proxy. This index is calculated by normalizing the daily GHI measured data by its corresponding extraterrestrial daily value [38]. K_t is therefore a dimensionless quantity that removes the seasonal geometrical behavior from the daily GHI and accounts for variations due to cloudiness. In overall terms, K_t values above 0.70 and below 0.20 represent days with mostly clear sky and overcast conditions, respectively, while the middle values are associated with mixed conditions. As a first approach, Fig. 6 shows the K_t probability density function of each model's estimates (in sky blue) in comparison to the measurements (in gray) for the LB site, although are similar throughout sites. This is a typical diagram for sites experiencing mixed sky conditions, in which days with clear sky, partly cloudy, and overcast conditions alternate. Almost all models misrepresent the clear sky distribution (last 4 bars in Fig. 6), either by not being able to account for the higher values (LCIM), having unbalanced frequencies in each clear sky bin in comparison to the ground data (NSRDB, Heliosat-4, GL1.2, and LCIM) or estimating a larger number of clear days than observed (MERRA2 and CERES). MERRA2 also estimates fewer overcast days than observed. For mid K_t values (between 0.2 and 0.6), LCIM, NSRDB, GL1.2, and CERES show a good agreement between density curves.

A final analysis is presented in Fig. 7, showing the rMBD and rRMSD dependence with K_t. The MERRA2 is excluded from the plot as it results in over-ranged bars. A clear correlation between cloudiness and uncertainty is observed; the models amplify their relative deviations with increasing cloudiness. GL1.2 and Heliosat-4 are the only models that change the sign of the bias with the sky conditions, being negative for mid K_t values. The LCIM is the model with the lowest rRMSD in most K_t intervals but is one of the worse for the lowest K_t interval (along with Heliosat-4 and CERES). GL1.2 and NSRDB estimates are the most consistent models across all conditions. NSRDB and Heliosat-4 are the best at modeling clear sky conditions (highest K_t interval). The anti-correlation found for all models between cloudiness and accuracy is not strange, since clouds are usually the more difficult phenomena to be precisely characterized by solar radiation models.

VI. Conclusions

A benchmark study of daily solar irradiation data sources was presented for Uruguay, southeast of South America.

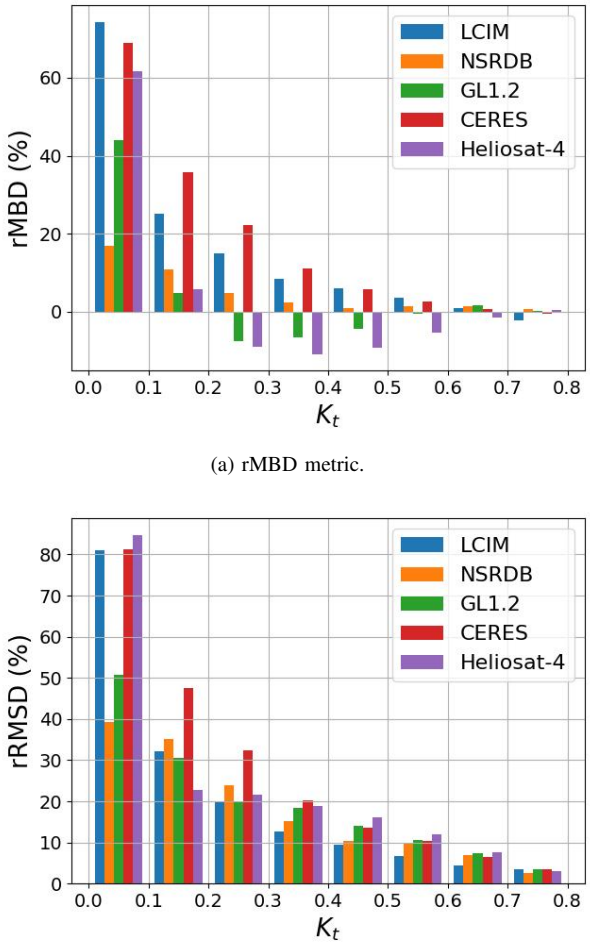


(a) LCIM. (b) Heliosat-4.

(c) CERES. (d) GL1.2.

(e) NSRDB. (f) MERRA2.

Fig. 6. Daily clearness index distribution in superposition with each model's estimates for the LB site. Blue: modeled data. Gray: ground data.



(a) rMBD metric.

(b) rRMSD metric.

Fig. 7. Performance metrics on cloudiness bands (spatially averaged).

Six modeled data sources were analyzed, including different satellite-based schemes and a reanalysis data source. These data sets uncertainty was evaluated and analyzed using four years of controlled-quality ground measurements from seven sites. The regionally-adjusted semi-empirical LCIM model based on GOES-East satellite images resulted as the best option, showing better overall metrics with spatial consistency. However, the NSRDB and GL1.2 estimates are well-positioned in the overall metrics, with even better consistency across sky conditions. Their performance is only slightly worse than the LCIM. In particular, GL1.2 exhibit exceptionally low bias, a remarkable feature of this model. A third group of models is conformed by CERES and Heliosat-4, with higher rMBD and rRMSD than the previous ones, but still accurate. From this, it can be inferred that models using high time-rate (10 or 15 minutes) GOES-East images are the preferred option for this region. MERRA2 reanalysis data show a very high uncertainty (observed in all metrics), so it should not be used for solar resource assessment in the region, at least not without an important post-processing procedure based on quality ground measurements. As a general takeaway, except for MERRA2, all models obtain a reasonable performance and, depending on the use, can be considered for solar resource assessment provided the estimates' uncertainty is enough for the intended application. These uncertainties are provided in this work, so industry practitioners can evaluate their utilization. For instance, from the models analyzed here, the LCIM, NSRDB, and GL1.2, are recommended in this region for applications in which the solar input uncertainty is critical.

ACKNOWLEDGMENT

R. Alonso-Suárez, A. Laguarda and I. Sarazola acknowledge financial support from the CSIC Group's Program, Universidad de la República, and PEDECIBA, Uruguay.

REFERENCES

[1] M. Schnitzer, C. Thuman, and P. Johnson. The impact of solar uncertainty on project financeability: mitigating energy risk through on-site monitoring. In *Proceedings of the American Solar Energy Society (ASES)*, pages 1–5, Denver, Colorado, United States., 13-17 May 2012.

[2] A. McMahan, C. Grover, and F. Vignola. Evaluation of resource risk in solar-project financing. In J. Kleissl, editor, *Solar Energy Forecasting and Resource Assessment*, pages 81–95. Academic Press, Boston, 2013.

[3] F. Vignola, C. Grover, N. Lemon, and A. McMahan. Building a bankable solar radiation dataset. *Solar Energy*, 86(8):2218–2229, 2012.

[4] J. Polo et al. Benchmarking on improvement and site-adaptation techniques for modeled solar radiation datasets. *Solar Energy*, 201:469–479, 2020.

[5] G. Huang, Z. Li, X. Li, S. Liang, K. Yang, D. Wang, and Y. Zhang. Estimating surface solar irradiance from satellites: past, present, and future perspectives. *Remote Sensing of Environment*, 233:111371, 2019.

[6] G. Salazar, C. Gueymard, J. Galdino, O. de Castro Vilela, and N. Fraidenraich. Solar irradiance time series derived from high-quality measurements, satellite-based models, and reanalyses at a near-equatorial site in Brazil. *Renew. and Sust. Energy Reviews*, 117:109478, 2020.

[7] D. Yang and J.M. Bright. Worldwide validation of 8 satellite-derived and reanalysis solar radiation products: A preliminary evaluation and overall metrics for hourly data over 27 years. *Solar Energy*, 210:3–19, 2020.

[8] C. Gueymard. Solar radiation resource: measurement, modeling, and methods. In T.M. Letcher, editor, *Comprehensive Renewable Energy*, pages 176–212. Elsevier, Oxford, second edition, 2022.

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

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

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

[12] D. Kratz, P. Stackhouse, S. Gupta, A. Wilber, P. Sawaengphokhai, and G. McGarragh. The fast longwave and shortwave flux (FLASHFlux) data product: Single-scanner footprint fluxes. *Journal of Applied Meteorology and Climatology*, 53(4):1059 – 1079, 2014.

[13] Y. Xie, M. Sengupta, and J. Dudhia. A fast all-sky radiation model for solar applications (FARMS): Algorithm and performance evaluation. *Solar Energy*, 135:435–445, 2016.

[14] Z. Qu et al. Fast radiative transfer parameterisation for assessing the surface solar irradiance: the Heliosat-4 method. *Meteorologische Zeitschrift*, 26(1):33–57, 02 2017.

[15] A. Laguarda, G. Giacosa, R. Alonso-Suárez, and G. Abal. 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, 2020.

[16] P. Ineichen. Long term satellite global, beam and diffuse irradiance validation. In *Elsevier Energy Procedia*, volume 48, pages 1586–1596, 2014. Proceedings of the 2nd International Conference on Solar Heating and Cooling for Buildings and Industry (SHC 2013).

[17] A. Forstinger et al. Expert quality control of solar radiation ground data sets. In *ISES Solar World Congress*, Oct 2021.

[18] M. L. Bettolli, S. A. Solman, R. P. da Rocha, M. Llopart, J. M. Gutierrez, J. Fernández, M. E. Olmo, A. Lavin-Gullon, S. C. Chou, D. Carneiro Rodrigues, E. Coppola, R. Balmaceda Huarte, M. Barreiro, J. Blázquez, M. Doyle, M. Feijoó, R. Huth, L. Machado, and S. Vianna Cuadra. The CORDEX Flagship Pilot Study in southeastern South America: a comparative study of statistical and dynamical downscaling models in simulating daily extreme precipitation events. *Climate Dynamics*, 56:1589–1608, 2021.

[19] L.J.B. McArthur. Baseline Surface Radiation Network (BSRN) Operations Manual. Td-no. 1274, wrcp/wmo, World Meteorological Organization (WMO), 2005. www.wmo.org.

[20] M. Peel, B. Finlayson, and T. Mcmahon. Updated world map of the köppen-geiger climate classification. *Hydrology and Earth System Sciences Discussions*, 11:1633–1644, 2007.

[21] F. Marchesoni-Acland and R. Alonso-Suárez. Intra-day solar irradiation forecast using rls filters and satellite images. *Renewable Energy*, 161:1140–1154, 2020.

[22] D. Cano, J.M. Monget, M. Albuisson, H. Guillard, N. Regas, and L. Wald. A method for the determination of the global solar radiation from meteorological satellite data. *Solar Energy*, 37:31–39, 1986.

[23] C. Rigollier, O. Bauer, and L. Wald. On the clear sky model of the ESRA—European Solar Radiation Atlas—with respect to the Heliosat method. *Solar energy*, 68(1):33–48, 2000.

[24] F. Linke. Transmissions-koeffizient und trubungsfaktor. *Meteoreorological Magazine Beitraege zur Physik der AtmosphaereBeitr*, 10:91–103, 1922.

[25] A. Laguarda and G. Abal. Clear-Sky broadband irradiance: first model assessment in Uruguay. In *ISES Conf. Proceedings, Solar World Congress*, pages 1–12, 2017.

[26] Agustiín Laguarda. *Modelado de la irradiancia solar sobre la superficie terrestre: Modelos físicos e híbridos utilizando información satelital sobre la Pampa Húmeda*. PhD thesis, Facultad de Ingeniería, Universidad de la República, 2 2021. PhD Thesis in Energy Engineering.

[27] B. Mayer and A. Kylling. Technical note: The libRadtran software package for radiative transfer calculations – description and examples of use. *Atmospheric Chemistry and Physics*, 5:1855–1877, 2005.

[28] K. Kriebel, G. Gesell, M. Kästner, and H. Mannstein. The cloud analysis tool APOLLO: Improvements and validations. *International Journal of Remote Sensing*, 24(12):2389–2408, 2003.

[29] J. Gonzalez, V. Teixeira-Branco, and R. Alonso-Suárez. Evaluation of the Heliosat-4 and FLASHFlux models for solar global daily irradiation estimate in Uruguay. In *ISES Conf. Proceedings, Solar World Congress*, 2019.

[30] D. Rutan et al. CERES synoptic product: Methodology and validation of surface radiant flux. *Journal of Atmospheric and Oceanic Technology*, 32(6):1121 – 1143, 2015.

[31] D. Fillmore, D. Rutan, S. Kato, F. Rose, and T. Caldwell. Evaluation of aerosol optical depths and clear-sky radiative fluxes of the CERES Edition 4.1 SYN1deg data prod-

uct. *Atmospheric Chemistry and Physics*, 22(15):10115–10137, 2022.

[32] A.C.S Porfirio and J.C Ceballos. A method for estimating direct normal irradiation from goes geostationary satellite imagery: Validation and application over northeast brazil. *Solar Energy*, 155(1):178–190, 2017.

[33] Classification of GOES imagery (Imager and ABI) at DISSM/CGCT/INPE: Related centroids (in portuguese). Relatorio Técnico GSTAR RT-002-2022. http://pururuca.cptec.inpe.br/gstar/GSTAR_dados.html. Accessed: March 2023.

[34] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin, and J. Shelby. The National Solar Radiation Data Base (NSRDB). *Renewable and sustainable energy reviews*, pages 51–60, 2018.

[35] R. Gelaro et al. The Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2). *Journal of Climate*, 30:5419–5454, 2017.

[36] F. Feng and K. Wang. Does the Modern-Era Retrospective analysis for Research and Applications-2 aerosol reanalysis introduce an improvement in the simulation of surface solar radiation over China? *International Journal of Climatology*, 39(3):1305–1318, 2019.

[37] M. Rienecker et al. MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*, 24(14):3624 – 3648, 2011.

[38] J.A. Duffie and W.A. Beckman. *Solar Engineering of Thermal Processes*. Wiley and Sons, Inc., Hoboken, New Jersey, third edition, 2006.

[39] Jie Zhang, Anthony Florita, Bri-Mathias Hodge, Siyuan Lu, Hendrik F Hamann, Venkat Banunarayanan, and Anna M Brockway. A suite of metrics for assessing the performance of solar power forecasting. *Solar Energy*, 111:157–175, 2015.

[40] B. Espinar, L. Ramírez, A. Drews, H. G. Beyer, L. F. Zarzalejo, J. Polo, and L. y Marín. Analysis of different comparison parameters applied to solar radiation data from satellite and german radiometric stations. *Solar Energy*, 83(1):118–125, 2009.

Iñaki Sarazola is an advanced mechanical engineering student at the Faculty of Engineering (FING), University of the Republic (Udelar), Uruguay. He holds a teaching and research assistant position at FING's Mechanical Engineering and Industry Production Institute (IIMPI) and an intern position at Udelar's Solar Energy Laboratory. His passion for sustainable solutions and renewable energies has led him to contribute to this research.

Agustín Laguarda holds a tenured Assistant Professor position at the Physics Institute, Faculty of Engineering, Udelar. He has a physics degree and a Ph.D. in Energy Engineering. He is also a researcher at the Solar Energy Laboratory (LES), specializing in satellite-based solar irradiance modeling, clear sky and spectral modeling, and solar resource assessment. He leads the solar irradiance modeling research area at LES.

Juan C. Ceballos is a physicist from the National University of Tucumán, Argentina; Dr. in Sciences (Meteorology) from the University of São Paulo, Brazil. Mainly interested in solar and atmospheric radiation modeling and climatology. He is with the National Institute of Space Research (INPE), Brazil, as a research collaborator of the G-STAR (Group of Solar, Terrestrial and Atmospheric Research), Meteorological Satellites and Sensors Division.

Rodrigo Alonso-Suárez holds a tenured Associate Professor position at the Physics Institute, Faculty of Engineering, Udelar, and is Director of the Solar Energy Laboratory, Udelar. He has a degree and a Ph.D. in electrical engineering and is currently an IEEE Senior Member. As a researcher, he is specialized in solar resource assessment and forecasting, and satellite-based solar irradiance modeling.