ISSN 1808-8546 (ONLINE) 1808-3765 (CD-ROM)

INTEGRATING PHYSICAL AND STATISTICAL APPROACHES FOR SOLAR IRRADIANCE ANALYSIS UNDER VARYING CLOUD CONDITIONS

MARCUS VINÍCIUS CONTES CALÇA[1]; MATHEUS RODRIGUES RANIERO[2] E JOSÉ RAFAEL FRANCO[3]

[1] Departamento de Engenharia Rural e Socioeconomia, Faculdade de Ciências Agronômicas da Universidade Estadual Paulista "Júlio de Mesquita Filho" - UNESP, Avenida Universitária, 3780 - Altos do Paraíso, Botucatu - SP - Brazil, marcus.calca@unesp.br (https://orcid.org/0000-0002-5685-3980).
[2] Departamento de Engenharia Rural e Socioeconomia, Faculdade de Ciências Agronômicas da Universidade Estadual Paulista "Júlio de Mesquita Filho" - UNESP, Avenida Universitária, 3780 - Altos do Paraíso, Botucatu - SP - Brazil, matheus.raniero@unesp.br (https://orcid.org/0000-0001-8338-4887).
[3] Departamento de Engenharia Rural e Socioeconomia, Faculdade de Ciências Agronômicas da Universidade Estadual Paulista "Júlio de Mesquita Filho" - UNESP, Avenida Universitária, 3780 - Altos do Paraíso, Botucatu - SP - Brazil, jose-rafael.franco@unesp.br (https://orcid.org/0000-0002-7129-4304).

1 ABSTRACT

High-quality solar irradiance data are fundamental for precise agrometeorological modeling and renewable energy planning. This study evaluates the integration of physical and statistical approaches for the quality control of 5-minute global (IG), sky diffuse (ID), and direct-beam (IB) solar irradiance measurements on a horizontal surface under varying sky cover conditions. The procedure was applied to a 5-year dataset (2018-2022) from the School of Agricultural Sciences (UNESP), Botucatu - São Paulo, Brazil. The methodology combines International Commission on Illumination (CIE) standards (absolute and consistency checks) with statistical ranges conditioned by cloud cover. The results revealed distinct failure signatures; absolute checks were most effective under overcast sky conditions, capturing cloud enhancement in direct-beam measurements, whereas consistency checks proved rigorous for global solar irradiance under clear skies. The statistical analysis successfully identified outliers, such as shadow band misalignments, that remained within physically valid ranges. Overall, the procedure validated approximately 88% of the dataset, ranging from 83.73% (IG) to 93.94% (ID), ensuring high reliability for agricultural water management and energy systems planning.

Keywords: solar energy, quality control, agrometeorology, data validation, sky cover.

CALÇA, M.V. C.; RANIERO, M. R.; FRANCO, J. R.
INTEGRAÇÃO DE ABORDAGENS FÍSICAS E ESTATÍSTICAS PARA ANÁLISE DA IRRADIÂNCIA SOLAR EM CONDIÇÕES VARIÁVEIS DE NUVENS

2 RESUMO

Medições de irradiância solar de alta qualidade são fundamentais para a modelagem agrometeorológica precisa e o planejamento de energias renováveis. Este estudo avalia a integração de abordagens físicas e estatísticas para o controle de qualidade de dados de irradiância solar global (IG), difusa (ID) e direta (IB) em superfícies horizontais com resolução temporal de 5 minutos, sob diferentes condições de cobertura do céu. O procedimento foi

Recebido em 09/01/2025 e aprovado para publicação em 16/12/2025
DOI: http://dx.doi.org/10.15809/irriga.2025v30p78-93

aplicado a um conjunto de dados de cinco anos (2018-2022) da Faculdade de Ciências Agronômicas (UNESP), Botucatu - São Paulo, Brasil. A metodologia combina padrões da Comissão Internacional de Iluminação - CIE (verificações absolutas e de consistência) com intervalos estatísticos condicionados à cobertura de nuvens. Os resultados revelaram motivos distintas de falhas, as verificações absolutas mostraram-se mais eficazes sob condições de céu nublado, capturando o realce por nuvens nas medições da direta, enquanto as verificações de consistência se mostraram mais rigorosas para a irradiância solar global sob céu aberto. A análise estatística identificou com sucesso valores discrepantes, como desalinhamentos do anel de sombreamento, que permaneciam dentro de limites fisicamente válidos. De forma geral, o procedimento validou aproximadamente 88% do conjunto de dados, variando de 83,73% (IG) a 93,94% (ID), garantindo alta confiabilidade para o manejo hídrico agrícola e o planejamento de sistemas de energia renovável.

Palavras-chave: energia solar, controle de qualidade, agrometeorologia, validação de dados, cobertura do céu.

3 INTRODUCTION

Solar irradiance is an essential parameter for Earth's energy balance, driving fundamental processes such as photosynthesis and the hydrological cycle (Pereira *et al*., 2017). In agrometeorology, accurate solar irradiance data are critical for estimating reference evapotranspiration (ETo), which directly dictates irrigation management and crop yield modeling (Allen *et al*., 1998). Furthermore, the growing need for renewable energy in rural areas requires a more accurate assessment of the solar resources available locally. However, ground-based solar radiometric measurements are prone to systematic and random errors arising from sensor degradation, calibration drifts, tracking misalignments, and environmental soiling (Younes; Claywell; Muneer, 2005). Consequently, the direct use of raw data without rigorous quality control (QC) can generate significant uncertainties in agricultural modeling and renewable energy planning. Therefore, the implementation of robust computational data validation protocols is mandatory to ensure the reliability of the datasets used for decision-making in the sustainable agriculture and renewable energy sectors.

To mitigate measurement uncertainties, several international institutions have established standard quality control procedures to verify solar irradiance reliability. The most widely adopted protocol is the Baseline Surface Radiation Network - BSRN quality control checks (Long; Dutton, 2002), although procedures from the National Renewable Energy Laboratory (1993), International Commission on Illumination - CIE (Tregenza *et al*., 1994), and the Royal Meteorological Institute of Belgium - RMIB (Journeé; Bertrand, 2011) are also extensively applied globally. These conventional methods primarily employ physical limits and cross-component comparisons to detect physically impossible events, assigning a flag to suspected erroneous values (Ohmura *et al*., 1998). While these protocols are excellent tools for identifying gross errors, they are often ineffective at detecting erroneous values that remain masked within the time series, as these anomalies typically fall within valid physical ranges (Urraca *et al*., 2017). Therefore, it is necessary to integrate statistical techniques after physical checks, which are capable of detecting subtle deviations that evade physical thresholds.

Previous research conducted at the Solar Radiometry Laboratory at the School

of Agricultural Sciences (UNESP), in Botucatu, São Paulo, Brazil, successfully implemented a statistical quality control method based on the mean and standard deviation to detect outliers in solar irradiance datasets (Calça *et al*., 2019). This preliminary approach proved highly effective in identifying data points that, while physically possible, were statistically inconsistent with the general data series behavior. However, since solar irradiance is inherently dynamic, its magnitude and variability are heavily modulated by cloud cover. In this sense, applying statistical data filters without considering the sky cover conditions can lead to inaccuracies, as the variance expected under a clear sky differs fundamentally from that of a cloudy sky or other cloud patterns. Consequently, using statistical metrics without this atmospheric context restricts the method's ability to distinguish between natural intermittent shading and actual measurement anomalies, necessitating a refinement of the procedure to include cloudiness patterns as a determining factor.

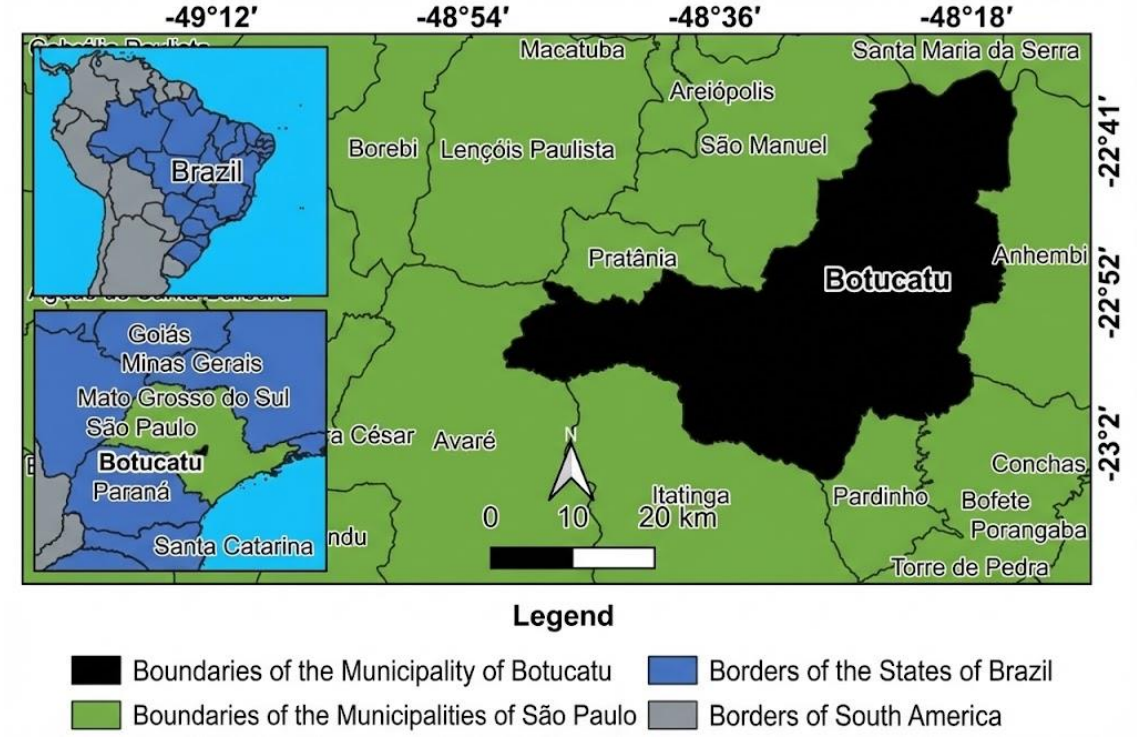
The present study aims to advance the solar irradiance data validation methodology developed by Calça *et al*. (2019) by integrating consistency checks from the International Commission on Illumination (CIE) (Tregenza *et al*., 1994) and a statistical outlier detection technique with a sky cover classification procedure based on the clearness index (KT), according to Escobedo *et al*. (2009). This approach is applied to assess the reliability of 5-minute instantaneous average measurements (Wm^{-2}) of global (IG), sky diffuse (ID), and direct-beam (IB) solar irradiance on a horizontal surface using data from 2018 to 2022 obtained at the Solar Radiometric Station of the School of Agricultural Sciences (UNESP). By establishing dynamic statistical confidence intervals under different atmospheric conditions, ranging from cloudy to clear skies, this study also aims to determine whether stratifying cloudiness patterns increases the sensitivity of the statistical method in detecting subtle errors that would otherwise remain hidden, given that cloud cover is one of the primary factors influencing the magnitude of each measured solar component.

4 MATERIALS AND METHODS

4.1 Location and climate

The proposed physical and statistical quality control procedure was developed on the basis of daytime global (IG), sky diffuse (ID), and direct-beam (IB) solar irradiances measured from 2018 to 2022 at the Solar Radiometry Laboratory (22° 50′ 50″ S, 48° 25′ 54″ W, 786 m) of the School of Agricultural Sciences (São Paulo State University - UNESP) in Botucatu, São Paulo, Brazil. Botucatu (Figure 1) is a town located in the center-south region of São Paulo State and is characterized mainly by intense agricultural activities (sugarcane and eucalyptus) and moderate industrial activities (Codato *et al*., 2008). In accordance with the Köppen climate classification, the region was recently reclassified as Aw (tropical savanna) by Franco *et al*. (2023), featuring a dry winter from June to August and a hot and rainy summer from December to February.

Figure 1. Location of Botucatu Town in São Paulo State, Brazil.



Reflecting this tropical savanna pattern, the 1991–2020 climatological normals indicate a mean annual air temperature of 21.34 °C, ranging from 18.28 °C in the coldest month (August) to 23.80 °C in the warmest month (February). The average relative humidity is 70.00%, reaching a maximum of 75.76% in January and a minimum of 61.87% in August. In terms of precipitation, the mean annual accumulation is approximately 1500 mm and is distributed over approximately 107 rainy days, with a distinct seasonal contrast. January is the wettest month (315 mm), whereas August is the driest month (38 mm), according to Franco *et al*. (2023). This rainfall distribution confirms that the summer season is characterized by significant cloud cover, a critical factor modulating the amount of solar irradiance reaching the Earth's surface.

4.2 Measurement acquisition and instrumentation

Global solar irradiance (IG) was measured on a horizontal surface using an unshaded Eppley Precision Spectral Pyranometer (PSP) (Figure 2a). This instrument type may exhibit a thermal offset, which can be reduced using specific ventilation devices or corrected via pyrheliometer signals (Haeffelin *et al*., 2001). However, in this study, no specific correction was applied to the raw dataset. This approach was adopted to evaluate the performance of the quality control procedure on standard operational data, which are typical of solar radiometric stations where such corrective accessories might not be available. Direct-beam solar irradiance at normal solar incidence (IBN) was measured using an Eppley Normal Incidence Pyrheliometer - NIP (Figure 2b) mounted on an Eppley ST3 single-axis solar tracker. To obtain the direct-beam solar irradiance projected onto a horizontal surface (IB), the normal solar incidence values were

multiplied by the cosine of the solar zenith angle (IBN $\cos\theta_Z$), in degrees (Equation 1), according to Iqbal (1983).

$$\cos\theta_Z = \sin(\delta)\sin(\phi) + \cos(\delta)\cos(\phi)\cos(\omega) \quad (1)$$

where δ is the solar declination (Equation 2), ϕ is the geographical latitude of the measurement acquisition site (-22.85) and ω is the hour angle of the sun (Equation 3), both in degrees.

$$\delta = 0.3964 + 3.631\sin(F) - 22.97\cos(F) + 0.03838\sin(2F) - 0.3885\cos(2F) + 0.07659\sin(3F) - 0.1587\cos(3F) - 0.01021\cos(4F) \quad (2)$$

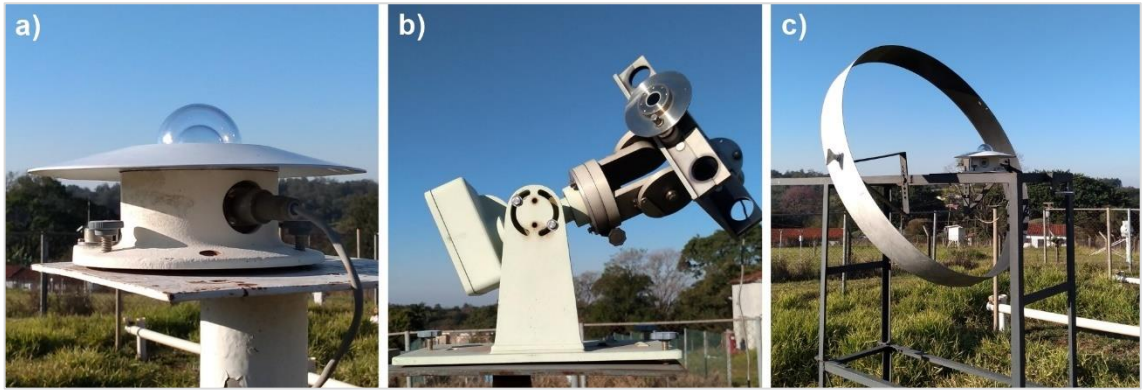
such that F is (360 D)/365, where D is the ordinal day of the year, ranging from 1 to 365.

$$\omega = (12 - Hd)\,15 \quad (3)$$

where Hd is the decimal hour of the day.

The sky-diffuse solar irradiance (ID) was measured using an Eppley Precision Spectral Pyranometer - PSP (Figure 2c) installed on a horizontal surface and equipped with a shadow ring (radius of 0.40 m and width of 0.10 m), developed and geometrically corrected on the basis of the factors proposed by Oliveira, Machado and Escobedo (2002), owing to its cost-effectiveness and ease of operation and maintenance. However, this geometric correction does not account for the anisotropic nature of solar radiation, which leads to an underestimation of the sky-diffuse component, particularly under clear-sky conditions. To mitigate this limitation, Dal Pai *et al*. (2016) corrected the measurements using atmospheric factors, which, on average, yielded improved results. Nevertheless, at the most common data acquisition intervals (1 or 5 minutes), a tendency for data dispersion is observed because of rapid variations in diffuse solar radiation over short time scales. Consequently, depending on the sky cover conditions, both correction approaches may result in either underestimation or overestimation of the actual sky-diffuse solar irradiance.

Figure 2. Pyranometer of global solar irradiance (a), pyrheliometer of direct-beam solar irradiance (b) and pyranometer of sky-diffuse solar irradiance (c) with a shadow.



All solar irradiance devices were installed in a rural area over short grass at a height of at least 1.5 m above ground level. Regular maintenance was performed daily, ensuring optimal solar irradiance monitoring conditions. In addition, auxiliary instruments were used annually to calibrate the routine equipment using the comparative method. Data acquisition was performed using a Campbell Scientific CR3000 Micrologger, which collected, processed, and stored solar irradiance signals at 5-minute intervals from 2018 to 2022, corresponding to averages of 60 measurements obtained at a scan time of 5 seconds during the study period. Table 1 presents the operating specifications of the instrumentation, according to the Eppley Laboratory (1992) and Eppley Laboratory (2019) manuals.

Table 1. Operational specifications of the global, direct-beam and sky diffuse solar irradiance instruments.

Specification	Global Solar Irradiance (IG)	Direct-Beam Solar Irradiance (IB)	Sky-Diffuse Solar Irradiance (ID)
Instrument	Eppley/Precision Spectral Pyranometer (PSP)	Eppley/Normal Incidence Pyrheliometer (NIP)	Eppley/Precision Spectral Pyranometer (PSP)
Classification	Secondary Standard/High Quality	Secondary Standard/High Quality	Secondary Standard/High Quality
Spectral Range	295 - 2800 nm	250 - 3000 nm	295 - 2800 nm
Sensitivity	Approx. 7.45 μV/Wm^{-2}	Approx. 7.59 μV/Wm^{-2}	Approx. 7.47 μV/Wm^{-2}
Output	0 - 12 mV	0 - 10 mV	0 - 12 mV
95% Response Time	5 seconds	5 seconds	5 seconds
Cosine Effect	Approx. 1% ($0° < \theta_Z < 70°$) Approx. 3% ($70° < \theta_Z < 80°$)	-	Approx. 1% ($0° < \theta_Z < 70°$) Approx. 3% ($70° < \theta_Z < 80°$)

4.3 Physical and Statistical Analysis Procedure

Postacquisition data analysis is a prerequisite for investigating solar resources, particularly in subhourly databases where high-frequency sampling increases the volume of potential measurement errors. To mitigate uncertainties inherent to the measurement process, this study applied a sequential quality control (QC) procedure structured

into four distinct stages: (i) physical limit checks, which are based on deterministic limits to detect impossible events; (ii) sky cover classification, which is used to stratify the dataset according to atmospheric conditions; (iii) statistical checks, which are based on probabilistic reasoning conditioned by sky cover to identify questionable events (outliers); and (iv) diagnostics and flagging, which assign reliability labels to each data point for final decision-making.

Measurements of solar irradiance are inherently sensitive to instrumental limitations, particularly at low solar elevations where the directional response (cosine effect) of pyranometers and the signal-to-noise ratio deteriorate. To mitigate these uncertainties and prevent the generation of false positives during the quality control process, this study strictly adhered to the automatic testing protocols established by the International Commission on Illumination - CIE (Tregenza *et al*., 1994). Consequently, consistency checks (CCs) were performed only when the solar elevation angle ($90° - \theta_Z$) exceeded 4° and the horizontal global solar irradiance was greater than 20 Wm^{-2}. This thresholding ensures that discrepancies caused by the physical limits of the sensors during sunrise and sunset are not misclassified as measurement errors, thereby preserving the integrity of the statistical analysis.

4.3.1 Stage 1 - Physical Limit Checks

The first filter applied the physical reasoning rules established by the International Commission on Illumination - CIE (Tregenza *et al*., 1994). These rules include absolute checks, which define the upper and lower physical boundaries of the solar irradiance, and consistency checks, which evaluate the coherence among the three measured components: global (IG), sky diffuse (ID), and direct-beam (IB) solar irradiance. Accordingly, the subscripts G, D, and B indicate the solar irradiance component to which each criterion is applied, yielding the checks ACG (Absolute Checks of Global), ACD (Absolute Checks of Sky-Diffuse), ACB (Absolute Checks of Direct-Beam), CCG (Consistency Checks of Global), CCD (Consistency Checks of Sky-Diffuse), and CCB (Consistency Checks of Direct-Beam). As presented in Table 2, these criteria ensure the removal of gross errors, such as unrealistic peaks or physically impossible solar irradiance values.

Table 2. Solar measurement analysis criteria of absolute checks (AC) and consistency checks (CC).

Label Check	Analysis Criteria	Solar Irradiance
ACG	$0 < IG \leq (1.2\ IE)$	IG
ACD	$0 < ID \leq (0.8\ IE)$	ID
ACB	$0 \leq IB \leq IE$	IB
CCG	$0.75\ (ID + IB) \leq IG \leq (ID + IB)\ 1.25$	IG
CCD	$ID < (1.10\ IG)$	ID
CCB	$IB \leq (1.05\ IG)$	IB

4.3.2 Stage 2 - Sky Cover Classification

Following the removal of gross errors, the dataset was stratified according to atmospheric conditions. This step is crucial for the subsequent statistical analysis, as the variability in solar irradiance is intrinsically linked to cloudiness. The sky cover classification (dimensionless) was determined by the clearness index (KT), defined as the ratio between global (IG) and extraterrestrial (IE) solar irradiances (KT = IG/IE). The classification adopts the four intervals proposed by Escobedo *et al*. (2009), as shown in Table 3.

Table 3. Clearness index interval and sky coverage condition.

Clearness Index Interval	Coverage Condition
$0 \leq KT < 0.35$	Overcast Sky
$0.35 \leq KT < 0.55$	Partially Cloudy Sky
$0.55 \leq KT < 0.65$	Partially Clear Sky
$0.65 \leq KT < 1$	Clear Sky

The extraterrestrial solar irradiance (IE), used to obtain sky cover classification and in the measurement analysis criteria, can be obtained in Wm^{-2} (Equation 4), according to Iqbal (1983).

$$IE = 1361\ EC \cos(\theta_Z) \quad (4)$$

where 1361 is the solar constant (Wm^{-2}), EC (Equation 5) is the dimensionless orbital eccentricity (Earth–Sun distance) and θ_Z is the zenith angle in degrees.

$$EC = 1 - 0.0009467 \sin(F) - 0.01671 \cos(F) - 0.0001489\ (2\ F) - 0.00002917 \sin(3\ F) - 0.0003438 \cos(4\ F) \quad (5)$$

such that F is (360 D)/365, where D is the ordinal day of the year, ranging from 1 to 365.

4.3.3 Stage 3 - Statistical Checks Conditioned by Sky Cover

The core innovation of this quality control procedure lies in the statistical range (SR) analysis applied after physical filtering. Unlike static thresholds, this method establishes dynamic confidence intervals for global (SRG), sky-diffuse (SRD), and direct-beam (SRB) solar irradiances, which are calculated specifically for each sky cover class defined in Stage 2. The allowable variation ranges were determined on the basis of the mean (μ) and standard deviation (σ) of the measurements for each sky cover condition (Table 4). A confidence level of 99.50% was adopted, and a normal probability distribution was assumed. This corresponds to a multiplier k = 2.57, effectively filtering out data points that fall in the extreme 0.5% tails of the distribution, which are flagged as statistical outliers.

Table 4. Statistical ranges (SRs) of the solar measurement analysis criteria.

Label Check	Analysis Criteria	Solar Irradiance
SRG	$\mu_{IG} - (k\,\sigma_{IG}) \leq IG \leq \mu_{IG} + (k\,\sigma_{IG})$	IG
SRD	$\mu_{ID} - (k\,\sigma_{ID}) \leq ID \leq \mu_{ID} + (k\,\sigma_{ID})$	ID
SRB	$\mu_{IB} - (k\,\sigma_{IB}) \leq IB \leq \mu_{IB} + (k\,\sigma_{IB})$	IB

4.3.4 Stage 4 - Diagnostics and Flagging

The quality control procedure assigns a diagnostic flag to each data point. A measurement is classified as Questionable (Q) if it fails any criterion (Physical or Statistical). A measurement is classified as Valid (V) only if it passes all sequential checks (Tregenza *et al*., 1994). Crucially, the automated quality control procedure does not delete questionable data; it flags them for final inspection by the data scientist, preserving the integrity of the raw database while highlighting potential anomalies for decision-making.

If the exclusion of flagged solar irradiance records is considered, employing graphical analysis as a confirmatory step is recommended. This visual inspection allows for a safer assessment; however, the final decision to maintain the removal flag ultimately relies on the data scientist domain expertise regarding the local solar irradiance data distribution, ensuring that actual measurement errors are correctly distinguished from valid extreme climatological events.

5 RESULTS AND DISCUSSION

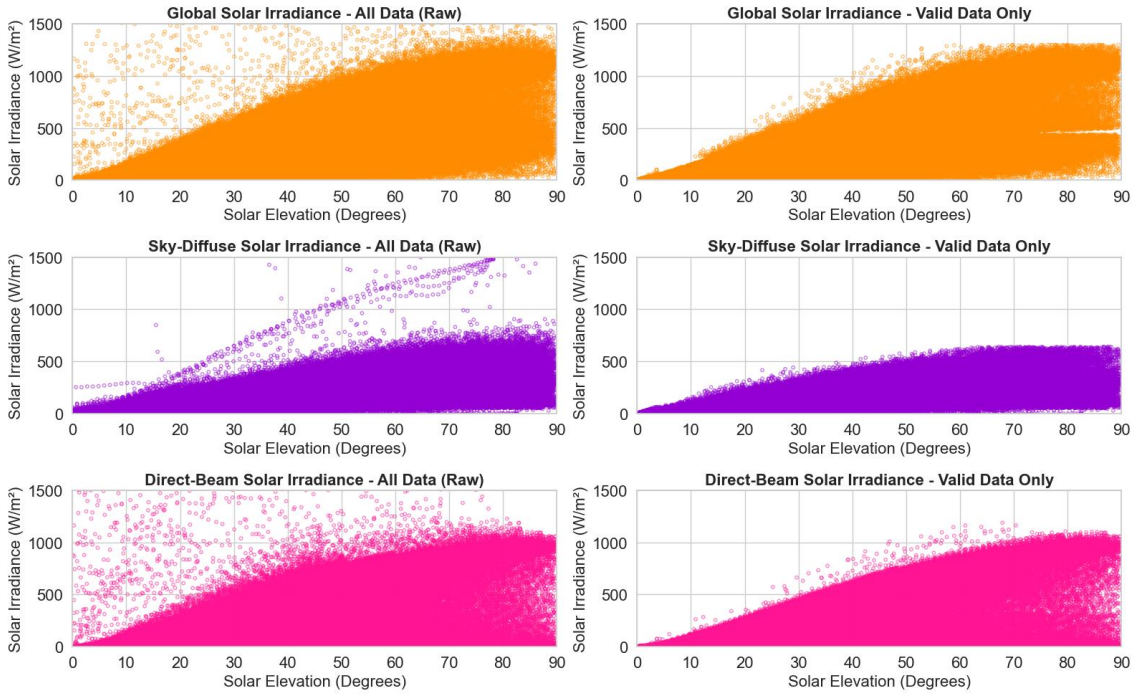
The dataset selected for integrated physical and statistical analysis, covering the period from 2018 to 2022, comprised 237244 data points for global (IG), 231793 for sky diffuse (ID), and 231596 for direct-beam (IB) solar irradiance components. This volume represents high data availability, approximately 90% of the theoretical number of daytime slots, ensuring statistical robustness. Notably, high availability of the direct-beam component was achieved through the physical gap-filling procedure applied during the preprocessing stage (IB = IG - ID). The distribution of raw measurement data points across different sky cover conditions is detailed in Table 5. These figures indicate a higher frequency of occurrence for clear and overcast sky conditions, reflecting a bimodal behavior typical of the local subtropical climate, in contrast to the lower incidence of transitional stages and partially cloudy and partially clear conditions.

Table 5. Numerical results of the solar irradiance data quality control procedure.

Sky Cover Condition	Solar Irradiance	Number of Raw Data Points	Analysis Criteria			Total Amount of Questionable Measures
			Number of Questionable Measurements			
			Absolute Check	Consistency Check	Statistical Range	
Overcast	IG	78452	514	4732	641	5868
	ID	77323	515	5887	1559	7367
	IB	75324	15746	963	2325	17996
Partially Cloudy	IG	44402	30	8247	0	8277
	ID	44152	253	913	372	1407
	IB	43354	987	1355	1502	3336
Partially Clear	IG	30244	12	7005	0	7017
	ID	30055	288	83	1268	1482
	IB	30045	448	1966	450	2577
Clear	IG	81245	1120	13282	935	14533
	ID	80254	1154	243	3201	3777
	IB	80488	1645	2984	554	4302
Total Data Points	IG	237244	4577	33266	1576	38596
	ID	231793	2219	7126	6400	14042
	IB	231596	19720	7268	4831	29105

Notably, quality control procedures, specifically consistency checks, were applied to the sky diffuse solar irradiance data after the application of geometric and atmospheric corrections for the shadow band effect. The use of raw sky-diffuse data typically results in a systematic underestimation of the solar irradiance flux, leading to artificial failures in the energy balance check (IG = ID + IB). By using the corrected dataset, the consistency analysis yielded more robust results, accurately identifying genuine instrumental discrepancies rather than artifacts caused by the shading mechanism. The effectiveness of the solar irradiance data filtering process is fully and visually demonstrated in Figure 3, which contrasts the raw measurement data points against the physically and statistically validated dataset as a function of solar elevation.

Figure 3. Comparison between the raw (left) and validated (right) datasets for global, sky diffuse, and direct-beam solar irradiance as a function of solar elevation.



The global solar irradiance had a greater number of measurements flagged as questionable in the consistency check criterion, with a greater occurrence in clear sky conditions (13282 data points of 81245 in this cloud condition). Sky-diffuse solar irradiance also had a greater number of measures flagged as questionable in the consistency check criterion, with a higher occurrence in the overcast sky (5887 data points of 77323 in this cloud condition). Unlike the solar irradiance already mentioned (global and sky diffuse), the direct-beam solar irradiance measurements were predominantly flagged by the absolute check criterion, with the highest occurrence in the overcast sky (15746 data points of 75324 in this cloud condition). This high incidence suggests the influence of physical phenomena such as the cloud enhancement effect, rather than instrumental failures identified by statistical range, as previously observed.

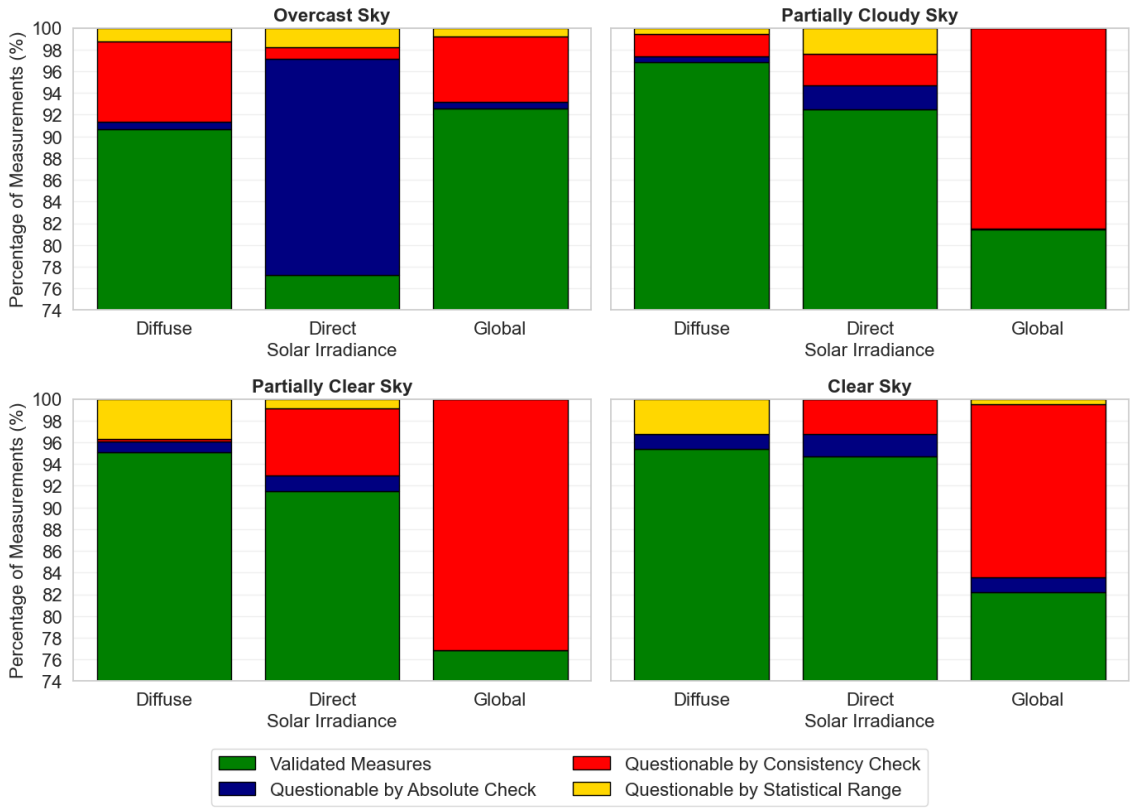
The analysis of daily irregularities revealed distinct patterns of failure. For global and direct-beam solar irradiance, the highest rate of questionable measurements occurred simultaneously on November 11, 2022, when 100% of the data points (156 measurements) were flagged. The primary global trigger was the absolute check criterion (124 flags), indicating a severe instrumental error or data logger malfunction that resulted in physically impossible values. Crucially, the simultaneous failure of the direct beam (156 flags by absolute check) was not an independent event but a propagation of this error. Since the gap-filling procedure for direct-beam solar irradiance relies on the difference method using global solar irradiance, the corrupted global data inevitably generate invalid direct-beam values, which are subsequently flagged by quality control algorithms.

In contrast, the sky-diffuse solar irradiance exhibited the worst performance on a different date, December 22, 2019, with 94.41% of the measurements (152 of 161 data points) flagged as questionable. Unlike the other solar irradiance components, the consistency check criterion (141 flags

registered) was the dominant cause here. This isolation indicates that while the global solar irradiance measurements were within physical limits, before the specific failure mentioned above, the sky-diffuse solar irradiance sensor violated the geometric relationship. This specific failure pattern reinforces the importance of the consistency test in quality control procedures, as it successfully identified a local discrepancy in the method of measuring sky-diffuse solar irradiance, likely caused by shadow band misalignment, that would have gone undetected using only range checks.

Under overcast sky conditions (Figure 4), the direct-beam solar irradiance obtained the highest total amount of questionable measures (22.80%), primarily because of absolute check failures, whereas the global solar irradiance presented the lowest irregularity rate (7.44%). For partially cloudy and partially clear sky conditions, the analysis revealed a shift in stability, where the global solar irradiance indicated the most questionable measures (18.59% and 23.16%, respectively), driven by strict consistency checks. Conversely, the sky-diffuse solar irradiance demonstrated greater stability under these intermediate-sky conditions, with significantly lower questionable rates (3.16% and 4.89%). Under clear-sky conditions, the global solar irradiance maintained the highest frequency of flagged measures (17.78%), whereas the sky-diffuse solar irradiance showed the least questionable data (4.61%).

Figure 4. Distribution of valid and questionable solar irradiance measurements by sky condition.



The comprehensive assessment of quality control outcomes reveals unique sensitivity signatures for each solar irradiance component. Direct-beam solar irradiance is predominantly physically sensitive, with failure rates peaking during overcast sky conditions because of absolute limit checks. In contrast, global solar

irradiance is methodologically sensitive, with the highest frequency of questionable data recorded under clear skies, where a rigorous energy balance consistency check is most unforgiving minor instrumental deviations. Notably, the sky-diffuse solar irradiance proved to be the most resilient variable, maintaining consistently low and stable irregularity rates across the entire spectrum of sky conditions. This robustness strongly validates the efficacy of the postprocessing algorithms applied to correct the shadow band geometry and atmospheric scattering effects.

The statistical analysis complemented the physical checks by identifying outliers on the basis of natural variability, specifically under extreme sky conditions. For the overcast sky, the critical scenario on December 28, 2022, flagged 107 of 126 direct-beam measurements, with a high standard deviation (222.92 Wm^{-2}), indicating rapid fluctuations likely driven by cloud enhancement effects. Conversely, under clear skies (October 22, 2022), the sky-diffuse measurements presented 92 flagged data points with an anomalous average of 739.37 Wm^{-2}. This physically impossible value for sky-diffuse solar irradiance from the sky on a clear day successfully identified a shadow band misalignment event where the sensor captured a direct beam. Thus, the statistical method proved essential for detecting both natural atmospheric instability and gross instrumental failures.

The high incidence of questionable measurements observed under overcast sky conditions aligns with the findings of Lappalainen and Kleissl (2020), who reported that cloud enhancement events frequently trigger absolute range checks by causing solar irradiance spikes that exceed clear-sky limits. Similarly, the rigorous consistency criterion failures observed in the clear-sky analysis corroborate the challenges reported by El Alani *et al*. (2021), where minor geometric deviations or tracking errors result in high rejection rates for otherwise valid data. Consequently, applying statistical range checks serves as a crucial complementary filter. As supported by Roesch *et al*. (2011) regarding baseline surface radiation network (BSRN) protocols, this multilayered approach is essential for accurately distinguishing between natural physical anomalies and real instrumental failures.

Finally, the rigorous application of these quality control procedures has direct implications for agricultural water management and renewable energy planning. Solar irradiance is the primary energy source driving the evapotranspiration process and accounts for the greatest share of the energy balance in the Penman–Monteith method (Allen *et al*., 1998). Any systematic error or instrumental failure in global solar irradiance data propagates nonlinearly into reference evapotranspiration estimates, potentially leading to incorrect irrigation sizing and water waste. By ensuring a validated dataset with known uncertainty limits, this study provides a reliable methodology for processing the solar radiation data used in climatological water requirement models.

6 CONCLUSIONS

In this study, solar irradiance validation was advanced by integrating physical standards with statistical detection stratified by sky cover conditions. The results confirmed that stratification enhances sensitivity, revealing distinct failure signatures. Absolute checks proved most effective under overcast sky conditions for direct-beam measurements, whereas consistency checks were most rigorous for global solar irradiance under clear skies. Statistical analysis served as a vital filter for outliers, such as shadow band misalignments, that remained within physical limits. Notably, sky-diffuse solar

irradiance was the most robust component. Overall, the procedure successfully flagged 83.73% of global, 93.94% of sky diffuse, and 87.43% of direct-beam solar irradiance measurements, confirming the high availability and reliability of the data for renewable energy planning.

The rigorous application of these quality control procedures extends beyond data validation, directly impacting agricultural water management reliability, since solar irradiance is the primary driver of the Penman–Monteith evapotranspiration process. Systematic errors or instrumental failures can propagate into irrigation estimates, leading to water waste or crop stress. By establishing dynamic confidence intervals and filtering out specific anomalies such as tracker failures or sensor gain errors, this approach minimizes the uncertainty in environmental inputs. Consequently, the validated dataset offers a robust basis for hydrological studies in subtropical climates. This study provides a reliable methodology for processing solar radiation data used in climatological water requirement models, ensuring that the final calculated crop water needs reflect the true atmospheric demand.

7 ACKNOWLEDGMENTS

The authors acknowledge the financial support from the São Paulo Research Foundation (FAPESP) and from the Coordination of Superior Level Staff Improvement (CAPES) during the development of this research.

8 REFERENCES

ALLEN, R. G.; PEREIRA, L. S.; RAES, D.; SMITH, M. **Crop evapotranspiration:** Guidelines for computing crop water requirements. Rome: Food and Agriculture Organization of the United Nations, 1998. **(**FAO Irrigation and drainage paper, 56).

CALÇA, M. V. C.; RANIERO, M. R.; FERNANDO, D. M. Z.; RODRIGUES, S. A.; DAL PAI, A. Outliers detection in a quality control procedure for measurements of solar radiation. **IEEE Latin America Transactions**, Piscataway, v. 17, n. 11, p. 1815-1822, 2019. Available at: https://latamt.ieeer9.org/index.php/transactions/article/view/1490. Accessed on: 9 Jan. 2026.

CODATO, G.; OLIVEIRA, A. P.; SOARES, J.; ESCOBEDO, J. F.; GOMES, E. N.; DAL PAI, A. Global and diffuse solar irradiances in urban and rural areas in southeast Brazil. **Theoretical and Applied Climatology**, Wien, v. 93, n. 1-2, p. 57-73, 2008. DOI: https://doi.org/10.1007/s00704-007-0326-0. Available at: https://link.springer.com/article/10.1007/s00704-007-0326-0. Accessed on: 9 Jan. 2026.

DAL PAI, A.; ESCOBEDO, J. F.; DAL PAI, E.; OLIVEIRA, A. P.; SOARES, J. R.; CODATO, G. MEO shadowring method for measuring diffuse solar irradiance: corrections based on sky cover. **Renewable Energy**, Oxford, v. 99, p. 754-763, 2016. DOI: https://doi.org/10.1016/j.renene.2016.07.026. Available at: https://www.sciencedirect.com/science/article/abs/pii/S096014811630622X. Accessed on: 9 Jan. 2026.

EL ALANI, O.; GHENNIOUI, H.; GHENNIOUI, A.; SAINT-DRENAN, Y. M.; BLANC, P. A Visual Support of Standard Procedures for Solar Radiation Quality Control. **International Journal of Renewable Energy Development**, Semarang, v. 10, n. 3, p. 601-613, 2021. DOI: https://doi.org/10.14710/ijred.2021.34806. Available at:

https://ijred.cbiore.id/index.php/ijred/article/view/34806. Accessed on: 9 Jan. 2026.

EPPLEY LABORATORY. **Normal Incidence Pyrheliometer**. Newport: The Eppley Laboratory, 2019. Available at: http://www.eppleylab.com/instrument-list/normal-incidence-pyrheliometer/. Accessed on: 9 Jan. 2026.

EPPLEY LABORATORY. **Precision Spectral Pyranometer**. Newport: The Eppley Laboratory, 1992. Available at: https://s.campbellsci.com/documents/au/manuals/psp.pdf Accessed on: 9 Jan. 2026.

ESCOBEDO, J. F.; GOMES, E. N.; OLIVEIRA, A. P.; SOARES, J. Modeling hourly and daily fractions of UV, PAR and NIR to global solar radiation under various sky conditions at Botucatu, Brazil. **Applied Energy**, Oxford, v. 86, n. 3, p. 299-309, 2009. DOI: https://doi.org/10.1016/j.apenergy.2008.04.013. Available at: http://sciencedirect.com/science/article/abs/pii/S0306261908001086. Accessed on: 9 Jan. 2026.

FRANCO, J. R.; DAL PAI, E.; CALÇA, M. V. C.; RANIERO, M. R.; DAL PAI, A.; SARNIGHAUSEN, V. C. R.; SÁNCHEZ-ROMÁN, R. M. Atualização da normal climatológica e classificação climática de Köppen para o município de Botucatu-SP. **Irriga -** Brazilian Irrigation and Drainage Journal, Botucatu, v. 28, n. 1, p. 77-92, 2023. DOI: https://doi.org/10.15809/irriga.2023v28n1p77-92. Available at: https://revistas.fca.unesp.br/index.php/irriga/article/view/4615. Accessed on: 9 Jan. 2026.

HAEFFELIN, M.; KATO, S.; SMITH, A. M.; RUTLEDGE, C. K.; CHARLOCK, T. P.; MAHAN, R. Determination of the thermal offset of the Eppley Precision Spectral Pyranometer. **Applied Optics**, Washington, DC, v. 40, n. 3, p. 472-484, 2001. DOI: https://doi.org/10.1364/AO.40.000472. Available at: https://opg.optica.org/ao/abstract.cfm?uri=ao-40-4-472. Accessed on: 9 Jan. 2026.

IQBAL, M. **An introduction to solar radiation**. Toronto: Academic Press Canada, 1983.

JOURNÉE, M.; BERTRAND, C. Quality control of solar radiation data within the RMIB solar measurements network. **Solar Energy**, Amsterdam, v. 85, n. 1, p. 72-86, 2011. DOI: https://doi.org/10.1016/j.solener.2010.10.021. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0038092X10003270. Accessed on: 9 jan 2026.

LAPPALAINEN, K.; KLEISSL, J. Analysis of the cloud enhancement phenomenon and its effects on photovoltaic generators based on measured 1-min irradiance data. **Journal of Renewable and Sustainable Energy**, Melville, v. 12, n. 4, article 043502, 2020. DOI: https://doi.org/10.1063/5.0007550. Available at: https://pubs.aip.org/aip/jrse/article-abstract/12/4/043502/284961/Analysis-of-the-cloud-enhancement-phenomenon-and. Accessed on: 9 Jan. 2026.

LONG, C. N.; DUTTON, E. G. **BSRN Global Network recommended quality control tests, V2.0**. Bremerhaven: Alfred Wegener Institute, 2002. (BSRN Technical Report). Available at: https://epic.awi.de/id/eprint/30083/1/BSRN_recommended_QC_tests_V2.pdf. Accessed on: 9 Jan. 2026.

NATIONAL RENEWABLE ENERGY LABORATORY. **Users manual for SERI**

QC software: assessing the quality of solar radiation data. Golden: NREL, 1993. (NREL/TP-463-5608). Available at: https://www.nrel.gov/docs/legosti/old/5608.pdf. Accessed on: 9 Jan. 2026.

OHMURA, A.; DUTTON, E. G.; FORGAN, B.; FRÖHLICH, C.; GILGEN, H.; HEGNER, H.; HEIMO, A.; KÖNIG-LANGLO, G.; MCARTHUR, B.; MÜLLER, G.; PHILIPONA, R.; PINKER, R.; WHITLOCK, C. H.; DEHNE, K.; WILD, M. Baseline Surface Radiation Network (BSRN/WCRP): new precision radiometry for climate research. **Bulletin of the American Meteorological Society**, Boston, v. 79, n. 10, p. 2115-2136, 1998. DOI: https://doi.org/10.1175/1520-0477(1998)079<2115:BSRNBW>2.0.CO;2. Available at: https://journals.ametsoc.org/view/journals/bams/79/10/1520-0477_1998_079_2115_bsrnbw_2_0_co_2.xml. Accessed on: 9 Jan. 2026.

OLIVEIRA, A. P.; MACHADO, A. J.; ESCOBEDO, J. F. A new shadow-ring device for measuring diffuse solar radiation at the surface. **Journal of Atmospheric and Oceanic Technology**, Boston, v. 19, n. 5, p. 698-708, 2002. DOI: https://doi.org/10.1175/1520-0426(2002)019<0698:ANSRDF>2.0.CO;2. Available at: https://journals.ametsoc.org/view/journals/atot/19/5/1520-0426_2002_019_0698_ansrdf_2_0_co_2.xml. Accessed on: 9 jan 2026.

PEREIRA, E. B.; MARTINS, F. R.; GONÇALVES, A. R.; COSTA, R. S.; LIMA, F. L.; RÜTHER, R.; ABREU, S. L.; TIEPOLO, G. M.; PEREIRA, S. V.; SOUZA, J. G. **Atlas brasileiro de energia solar**. 2. ed. São José dos Campos: INPE, 2017. DOI: https://doi.org/10.34024/978851700089. Available at: https://repositorio.unifesp.br/handle/11600/58353. Accessed on: 9 Jan. 2026.

ROESCH, A.; WILD, M.; OHMURA, A.; DUTTON, E. G.; LONG, C. N.; ZHANG, T. Assessment of BSRN radiation records for the computation of monthly means. **Atmospheric Measurement Techniques**, Goettingen, v. 4, n. 2, p. 339-354, 2011. DOI: https://doi.org/10.5194/amt-4-339-2011. Available at: https://amt.copernicus.org/articles/4/339/2011/. Accessed on: 9 Jan. 2026.

TREGENZA, P. R.; PEREZ, R.; MICHALSKY, J.; SEALS, R.; MOLINEAUX, B.; INEICHEN, B. **Guide to recommended practice of daylight measurement**. Vienna: Commission Internationale de l'Eclairage, 1994.

URRACA, R.; GRACIA-AMILLO, A. M.; HULD, T.; MARTINEZ-DE-PISON, F. J.; TRENTMANN, J.; LINDFORS, A. V.; RIIHELÄ, A.; SANZ-GARCIA, A. Quality control of global solar radiation data with satellite-based products. **Solar Energy**, Oxford, v. 158, p. 49-62, 2017. DOI: https://doi.org/10.1016/j.solener.2017.09.026. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0038092X17307983?via%3Dihub. Accessed on: 9 Jan. 2026.

YOUNES, S.; CLAYWELL, R.; MUNEER, T. Quality control of solar radiation data: present status and proposed new approaches. **Energy**, v. 30, n. 9, p. 1533-1549, 2005. DOI: https://doi.org/10.1016/j.energy.2004.04.031. Available at: http://sciencedirect.com/science/article/abs/pii/S0360544204002233. Accessed on: 9 Jan. 2026.