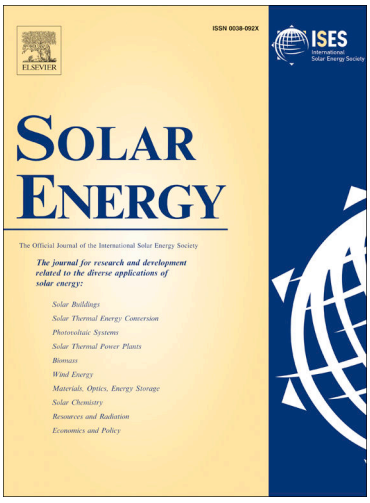
Journal Pre-proofs




# Benchmarking of meteorological indices for sky cloudiness classification

Andrés Suárez García, Montserrat Díez Mediavilla, Diego Granados López, David González Peña, Cristina Alonso Tristán

PII: S0038-092X(19)31160-0
DOI: https://doi.org/10.1016/j.solener.2019.11.060
Reference: SE 8486

To appear in: *Solar Energy*

Received Date: 18 December 2018
Revised Date: 8 November 2019
Accepted Date: 16 November 2019

Please cite this article as: Suárez García, A., Díez Mediavilla, M., Granados López, D., González Peña, D., Alonso Tristán, C., Benchmarking of meteorological indices for sky cloudiness classification, *Solar Energy* (2019), doi: https://doi.org/10.1016/j.solener.2019.11.060

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier Ltd on behalf of International Solar Energy Society.

Journal Pre-proofs

# Benchmarking of meteorological indices for sky cloudiness classification

Andrés Suárez García[1,2], Montserrat Díez Mediavilla[2], Diego Granados López[2], David González Peña[2], Cristina Alonso Tristán[2,*]

[1] Centro Universitario de la Defensa. Marín, Pontevedra, Spain. [2] Research Group Solar and Wind Feasibility Technologies (SWIFT). Electromechanical Engineering Department, Universidad de Burgos, Spain.

* Corresponding author: catristan@ubu.es; cristinaalonso.tristan@gmail.com

## ABSTRACT

Sky classification is a complex problem, due in part to such abstract conceptual definitions as clear, intermediate, and overcast, as well as other intermediate ranges. The CIE (Commission Internationale de L'Éclairage) Standard classification offers a solution to this problem, although its application requires data on the luminance distribution of the whole sky that are less commonly available. A benchmarking and classification system of ten meteorological indices is introduced in this study to classify the sky types from overcast to clear. The indices can be calculated from measurements of global, diffuse, and direct irradiance that are widely available from meteorological ground stations. The classification system uses confusion matrices, a machine-learning tool that generates a visual display of the results of supervised-learning algorithms. The CIE Standard skies classification, applied to half hourly sky-scanner measurements in Burgos (Spain), over the period June 2016 - May 2017, is used in this study as a baseline reference for a comparative review of the results from the meteorological indices and their results. They are classified by four performance ratings: Accuracy, Jaccard, Cohen, and Matthews, which feature both classification similarity and the randomness of any agreement**.** All meteorological indices yielded a high average degree of accuracy - close to 80% - in a detailed review of their classification. Neverthless, the results suggested that Perez's Clearness Index based on global, diffuse and direct radiation measurements offered the most precise classification of the skies, followed closely by the Klucher Clearness Index and the Perraudeau Nebulosity Index.

Journal Pre-proofs

***Keywords**:* clearness index, sky classification, CIE standard, cloudiness classification, meteorological indices.

## NOMENCLATURE

| | |
|---|---|
| **Greek symbols:** | $D_V$ Diffuse luminance ($cd/m^2$) |
| $\chi$ Scattering angle | $D(O)$ Diffuse horizontal irradiance (W/m²) |
| $\varphi(Z)$ Gradation function | $F$ Clearness Function |
| $\gamma$,$\gamma_s$ Angle of elevation, solar elevation. | $\boldsymbol{FK}$ Klucher Clearness Index |
| $\alpha$, $\alpha_s$ Azimuth angle, solar azimuth. | $\boldsymbol{FP}$ Perraudeau Nebulosity Index, |
| $f(\chi)$ Indicatrix function | $G_{st}$ Standard global irradiance |
| $\in$ Perez's clearness index | $G_{clear}(0)$ Clear-sky irradiance (W/m²) |
| $\varepsilon_0$ Average value of the orbital eccentricity of the Earth. | $G(O)$ Global horizontal irradiance (W/m²) |
| Δ Perez's brightness index | $\boldsymbol{k_b}$ Horizontal direct fraction |
| $\kappa$ Cohen Kappa | $k_d$ Horizontal diffuse fraction |
| $Z$, $Z_s$ Angle from zenith, angle between sky zenith and sun. | $k_{d0}$ Horizontal diffuse fraction for a clear sky. |
| a, b, c, d, e Coefficients of the gradation and indicatrix function for CIE standard skies classification | $\boldsymbol{k_k}$ Batlles Clearness index |
| | $\boldsymbol{k_t}$ Clearness Index |
| | $k_{tt}$ Auxiliary parameter for Batlles Clearness Index calculation |
| $b_p$ Reference number of each band. | $L_p$ Luminance of a sky patch measured by the Sky-scanner ($cd/m^2$) |
| $B_{sc}$ Extraterrestrial irradiance constant (1361.1 W/m²). | $L_z$ Zenith luminance ($cd/m^2$) |
| $B(n)$ Beam irradiance (W/m²) | MI Meteorological index |
| $B_{ext}(n)$ Extraterrestrial direct solar irradiance (W/m²) | $m$ Optical air mass |
| | $n_p$ Number of patches in band b. |
| $Ces$ Standard cloudiness fraction | NR Luminance normalization ratio ($cd/m^2$) |
| $G_{ext}(0)$ Extraterrestrial global solar irradiance on a horizontal plane | p Reference number of a scanned sky patch |
| $Cle$ Cloudless Index | |

| | |
|---|---|
| $d_n$ Day number of year, 1 on 1st January and 365 on 31st December. February has 28 days. | $\boldsymbol{Si}$ Igawa Index |

## 1. INTRODUCTION

A key aspect in the modelling of solar radiation and daylighting is sky classification. Many models for the calculation of global, direct, and diffuse irradiation and illumination (i.e. luminous efficacy) are defined for different sky types, based on the values of different climatic parameters. Searching for parameters that can quantify abstract concepts such as clear, partially cloudy, and overcast skies, as well as all possible intermediate classes, is a complex problem that researchers have addressed using different strategies. Sky conditions of the same category should have similar solar radiation and sky luminance distributions and the corresponding climatic parameters should be within certain ranges (Li and Lam, 2001).

In 2003, 15 standard sky types were defined in the CIE categorization (Uetani et al., 2003). The classification included five types of clear sky, five intermediate types, and five types of cloudy skies. Sky types of the same category have the same well-defined sky luminance pattern. Once the sky types are identified, the basic solar irradiance and daylight illuminance on the surfaces of interest can be obtained through simple mathematical expressions (Li et al., 2013). The luminance distribution for each standard sky type can help arrive at accurate determinations of daylight illuminance (Kittler et al., 1997). Several works have reported that the CIE standard sky classification provides a good overall framework for representing the actual sky conditions and covers the whole probable spectrum of skies found in nature (Alshaibani, 2011; Li and Cheung, 2006; Li et al., 2008; Li et al., 2007; Markou et al., 2005; Markou et al., 2004; Torres et al., 2010a, b; Tregenza, 2004; Wong et al., 2012). Each CIE

General Standard Sky is well defined by the straightforward approach for sky classification: the sky luminance pattern. The standard instrument for measuring sky luminance distribution is a sky scanner and basic sky luminance data are available at many locations across the world.

When interpreting sky conditions, meteorological data are initially used as weighting factors to show the degree of sky clearness, but different researchers have each adopted different kinds of Meteorological Indices (MIs), each with a different range (Li et al., 2007; Lou et al., 2017; Umemiya and Kanou, 2008). Their selection depended on the availability of meteorological variables. Some previous attempts have been made to use specific MIs to classify sky conditions. The diffuse fraction, $k_d$, and the clearness index, $k_t$, were used in (Brunger and Hooper, 1993) for the classification of sky conditions. Igawa (Igawa et al., 2004) defined a clear-sky index, *Si*, that can be used as an index with no dependency on the solar altitude for the classification of sky conditions. (Baharuddin et al., 2010) and (Rahim et al., 2004) classified daylight data into three sky conditions – clear, intermediate, and overcast – using two methods: sunshine duration and cloud ratio methods. The sunshine duration method estimates the frequencies of the occurrence of clear, intermediate, and overcast sky from the monthly mean value of the relative sunshine duration. The cloud ratio method classifies the sky condition into three sky classes, based on the observation of the pattern of diffuse fraction graphs. (Kong and Kim, 2013) evaluated the ability of the diffuse fraction, Perez's clearness index, and the clearness index to classify the skies by means of a frequency distribution. (Umemiya and Kanou, 2008) introduced a sky classification method, by using different combinations of nine "insolation" or irradiation indices. In (Lou et al., 2017), the CIE Standard Skies identified by the luminance scan were correlated with fourteen meteorological parameters (solar altitude angle, clearness index, diffuse fraction,

turbidity, air temperature, relative humidity, wet bulb temperature, and direct normal solar irradiance, among others) using the Classification Tree algorithm. (Li and Lam, 2001) investigated the prevailing sky conditions in Hong-Kong in terms of different climatic parameters (cloud cover distribution, hours of sunshine, $k_t$ and $k_d$) and highlighted the merits of each one. In terms of modelling solar radiation and outdoor illuminance components, the authors concluded that $k_t$ was the best parameter from among the four indices under study. (Li et al., 2014) analysed different alternative approaches to perform the CIE Standard skies classification, the clearness index, and the turbidity index, among others.

No previous studies have presented analyses of each MI and its sky classification capability. The frequency of occurrence of clear, intermediate, and overcast skies is determined by combining various MIs through mathematical algorithms. Likewise, no simultaneous comparison between the different approaches to sky classification has been conducted. The comparison is usually done in terms of frequency of appearance. Very few of the works just reviewed included comparisons of their results with the CIE Standard sky classification. Nevertheless, the CIE Standard skies are internationally considered as sufficiently comprehensive to simulate skylight luminance and radiance distributions, crucial factors in passive energy efficient building designs and in active solar energy applications. Even a simple sky classification with only three categories (clear, partial, and cloudy) would allow a major improvement in lighting control systems (Li, 2010; Li et al., 2008), a precise selection of the models to estimate solar irradiance and for its prediction (Djafer et al., 2017; Ruiz-Arias and Gueymard, 2018), more accurate determination of the spectral components of solar radiation (Escobedo et al., 2009; Jacovides et al., 2007), and better prediction of photovoltaic production, to improve grid integration (Perveen et al., 2018).

In this study, a review of the different MIs used for sky classification and their benchmarking has led to the selection of those calculated with three standard variables recorded at many meteorological stations: global, beam, and diffuse irradiance. Sky conditions and characterization of CIE Standard sky types were reported in a previous publication with a full year of data recorded at Burgos, northwestern Spain, (Suárez-García et al., 2018). The sky classifications using these MIs were compared to the CIE cloudiness categories (cloudy, partial, and clear), adapting the original intervals to match the number of categories.

The main objective of this study is to assess the selected MIs presented as alternatives to the CIE classification, so as to perform sky classifications of acceptable accuracy. For sky classification, the appropriate MIs would be expected to require less expensive and more common equipment at meteorological ground stations than the sky scanner that is used to measure sky luminance and its distribution. Calculation of simple MIs is fast and allows an efficient implementation of control systems of daylighting and active solar energy applications. The use of confusion matrices as a benchmaking tool is a novelty in the study of sky conditions. In this way, both the MI classification and the CIE Standard, may be compared against the same timestamp, rather than through a frequency distribution that can only be done at the end of the data collection period.

This study is structured as follows: the CIE Standard classification will be described in Section 2. In Section 3, the MI derived from global, beam, and diffuse irradiation will be used to define the sky types. Section 4 will examine the main characteristics of the confusion matrices classification and the indicators used for the MIs benchmarking process. The experimental facility and data used in this work, as well as the results of the study, will be

described in Section 5 and in Section 6, respectively. Finally, the principal observations will be presented in Section 7 together with the main contributions of the study.

## 2. CIE STANDARD CLASSIFICATION

The CIE standard sky classification describes the luminance ratio of any given sky patch, $L_p$ $(kcd/m^2)$, normalized by the sky's zenith luminance, $L_z$ $(kcd/m^2)$, as the product of the relative gradation function, $\varphi(Z)/\varphi(0)$, and the relative scattering indicatrix function, $f(\chi)/f(0)$, as shown in Eq. 1:

$$\frac{L_p}{L_z} = \frac{f(\chi)\cdot\varphi(Z)}{f(0)\cdot\varphi(0)} \quad \text{Eq. 1.}$$

The gradation function (Eq. 2) gives the luminance/radiance variation from horizon to zenith and the indicatrix function (Eq. 3) expresses the decrease of luminance from the solar disc to sky patches far away from the sun:

$$\varphi(Z)/\varphi(0) = \frac{1 + a\cdot\exp\left(b/\cos Z\right)}{1 + a\cdot\exp b} \quad \text{Eq. 2.}$$

$$f(\chi)\Big/f(0) = \frac{1 + c\cdot\left[\exp\left(d\cdot\chi\right) - \exp\left(d\cdot\frac{\pi}{2}\right)\right] + e\cdot cos^2\chi}{1 + c\cdot\left[\exp\left(d\cdot Z_s\right) - \exp\left(d\cdot\frac{\pi}{2}\right)\right] + e\cdot cos^2 Z_s} \quad \text{Eq. 3.}$$

Z is the angle between sky zenith and the sky patch under scrutiny, $(rad)$. $Z_s$ is the angle between sky zenith and the Sun. Coefficients *a, b, c, d,* and *e* can be adapted to depict the 15 CIE sky conditions: five overcast, five partly cloudy, and five clear sky types, as shown in Table 1. $\chi$ is the scattering angle $(rad)$ that represents the shortest distance from the sky patch to the solar disc and is calculated from Equation 4:

$$\chi = \arccos\left(\cos Z_s \cdot \cos Z + \sin Z_s \cdot \sin Z \cdot \cos\left|\alpha - \alpha_s\right|\right) \quad \text{Eq. 4,}$$

where, $\alpha$ is the azimuth angle of the sky patch and $\alpha_s$ is the solar azimuth.

Figure 1 shows a sky image of three different sky conditions classified by the CIE Standard as: a) Clear (IV.4 type); b) partial (III.2); and, c) cloudy (II.2).

Table 1. Parameters of CIE standard Sky types (Uetani et al., 2003).

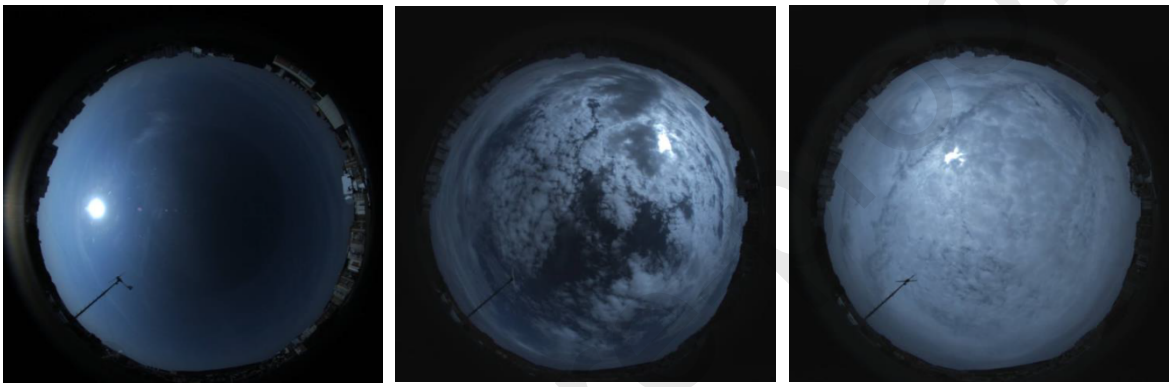
| | Type | a | b | c | d | e | Description |
|---|---|---|---|---|---|---|---|
| CLOUDY | I.1 | 4.0 | -0.70 | 0 | -1.0 | 0.00 | Overcast with a steep gradation & azimuthal uniformity |
| | I.2 | 4.0 | -0.70 | 2 | -1.5 | 0.15 | Overcast with a steep gradation & slight brightening toward the Sun |
| | II.1 | 1.1 | -0.80 | 0 | -1.0 | 0.00 | Overcast with a moderate gradation & azimuthal uniformity |
| | II.2 | 1.1 | -0.80 | 2 | -1.5 | 0.15 | Overcast with a moderate gradation & slight brightening toward the Sun |
| | III.1 | 0.0 | -1.00 | 0 | -1.0 | 0.00 | Overcast, foggy or cloudy, with overall uniformity |
| PARTIAL | III.2 | 0.0 | -1.00 | 2 | -1.5 | 0.15 | Partly cloudy with a uniform gradation & slight brightening toward the Sun |
| | III.3 | 0.0 | -1.00 | 5 | -2.5 | 0.30 | Partly cloudy with a uniform gradation & a brighter circumsolar effect |
| | III.4 | 0.0 | -1.00 | 10 | -3.0 | 0.45 | Partly cloudy, rather uniform with a clear solar corona |
| | IV.2 | -1.0 | -0.55 | 2 | -1.5 | 0.15 | Partly cloudy with a shaded sun position |
| | IV.3 | -1.0 | -0.55 | 5 | -2.5 | 0.30 | Partly cloudy with brighter circumsolar effect |
| CLEAR | IV.4 | -1.0 | -0.55 | 10 | -3.0 | 0.45 | White-blue sky with a clear solar corona |
| | V.4 | -1.0 | -0.32 | 10 | -3.0 | 0.45 | Very clear / unturbid with a clear solar corona |
| | V.5 | -1.0 | -0.32 | 16 | -3.0 | 0.30 | Cloudless polluted with a broader solar corona |
| | VI.5 | -1.0 | -0.15 | 16 | -3.0 | 0.30 | Cloudless turbid with a broader solar corona |
| | VI.6 | -1.0 | -0.15 | 24 | -2.8 | 0.15 | White-blue turbid sky with a wide solar corona effect |

The luminance distributions of individual standard skies were modelled and compared with the scanned sky luminance readings. The standard sky that was assigned had the lowest Mean-Square Error (RMSE), (Tregenza, 2004). The original criterion to define the sky type, known as the Standard Sky Luminance Distribution method (SSLD) (Kittler et al., 1997), uses a theoretical assemblage of curves that represent the relation between the zenith luminance/diffuse luminance ($L_z/D_V$) ratio and the solar elevation angle. These curves converge at solar elevation values higher than 35º, making it difficult to apply this method in



Journal Pre-proofs

certain areas and at times when the solar elevation angle is higher than 35º (Li et al., 2013), which is the case at the location under study, especially in summer. Various procedures to circumvent this issue are proposed, using various methods of normalization (Li et al., 2014). In a previous paper (Suárez-García et al., 2018), the Normalization Ratio (NR) introduced by Littlefair (Littlefair, Paul J., 1994; Littlefair, P. J., 1994) was used to obtain the CIE Standard sky types in Burgos, Spain.



a) Clear sky b) Partly cloudy sky c) Overcast sky

Figure 1: Sky Image of diferent sky conditions classified by CIE Standard as a) Clear (IV.4 type), b) partial (III.2) and c) cloudy (II.2), taken by a SONA201D All-Sky Camera-Day in Burgos, Spain, on 6/22/17 at 10:30 UTC, 3/11/2017 at 16:45 UTC and 8/23/2017 at 13:00 UTC, respectively.

## 3. METEOROLOGICAL INDICES FOR SKY CLASSIFICATION.

The sky classification by luminance distribution is reliable, but it has several restrictions (Li et al., 2014). The most significant problem is that luminance measurements are only available at a few sites in the world and only over short measurement periods. Alternatively, sky conditions can be evaluated using MIs calculated from meteorological variables readily accessible from most weather stations. In this study, different MIs traditionally used to classify the sky conditions were used to correlate the CIE Standard Skies determined by the

luminance scan in terms of confusion matrices. The correlation can be easily interpreted and used to analyse the long-term sky conditions at sites that share similar climatic variables with the location under study. In the following paragraphs, the different MIs contemplated here are briefly described. For comparative purposes, a common classification is used in all cases, i.e., the CIE cloudiness categorization (cloudy, partial and clear, as shown in Table 1). As highlighted in Section 1, the definition of clear, partial, and cloudy sky conditions may be ambiguous, and the limits between these words are fuzzy. However, *cloudy sky* can be defined at one end of a continuous scale with *clear sky* at the other end.

The problem addressed in this work is the homogenization of the sky categories in the same number of classes without changing the original limits established by their authors. This follows a similar approach to that in (Gueymard et al., 2019) where the authors compared different MI models in an attempt to classify clear skies. Different intervals were tested, in order to mitigate the semantic effects and to maximize MI performance, keeping the class limits that the respective authors set in their original works. Figure 2 shows the possible combinations of the limits of the intervals for the different MIs studied, keeping the ordinal relationship between cloudy, partial, and clear-sky conditions. For some of the MIs, no adaptation was needed. When more than three intervals were defined in the original work, all possible combinations for merging the intervals were tested. The adapted interval is the one that maximized the classification metrics for each MI. The original and the adapted values of the intervals selected by each index are summarized in Table 2. The MIs are calculated using global horizontal irradiance, $G(0)$; diffuse horizontal irradiance, $D(0)$; and beam irradiance, $B(n)$.

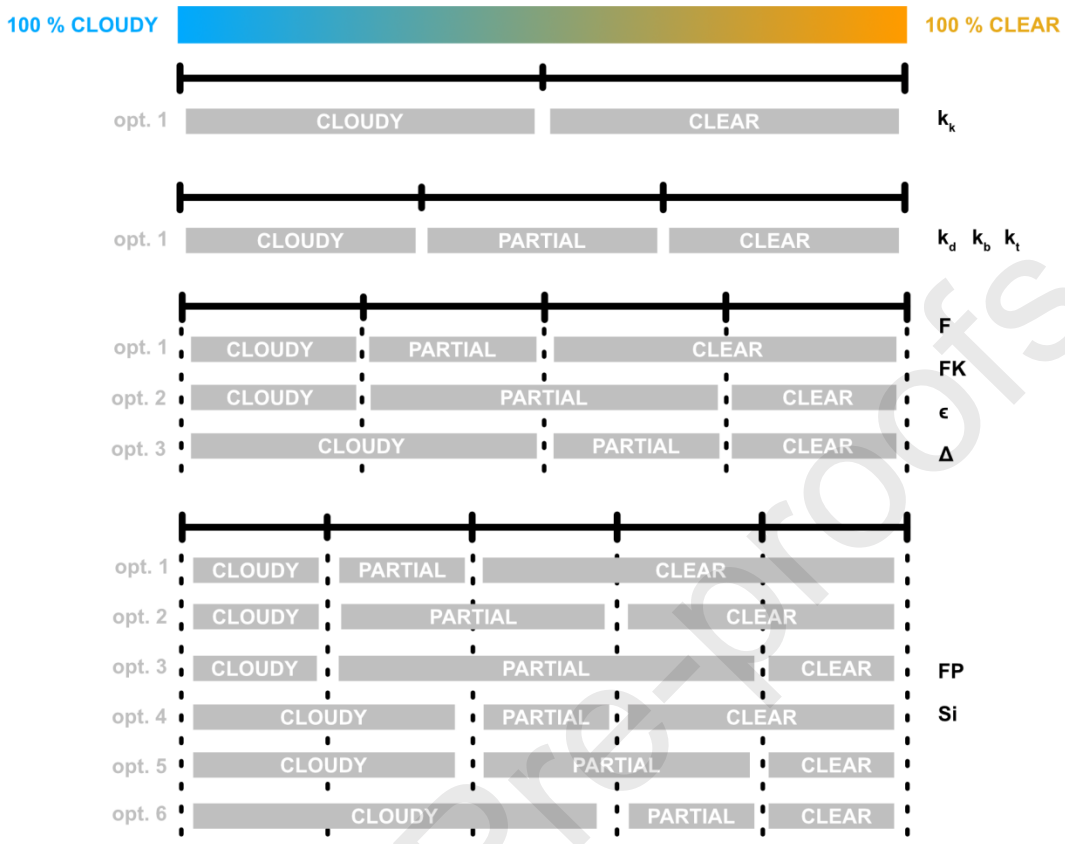



Figure 2: Interval adaptation procedure of the original MIs.

### 3.1. Horizontal Diffuse Fraction, $k_d$.

The horizontal diffuse fraction (or cloud ratio or cloudiness index or diffuse ratio), $k_d$ (Eq. 5), is defined as the ratio between the diffuse horizontal irradiance, $D(0)$, and the global horizontal irradiance, $G(0)$ (Erbs et al., 1982). $k_d$ refers to the cloudiness of the sky and/or the turbidity of the atmosphere: the higher the proportion of the diffuse radiation in the global one, the higher the $k_d$ (Kambezidis, 2018).

$$k_d = \frac{D(0)}{G(0)} \qquad \text{Eq. 5.}$$

Conversely, low values mean that the global radiation mainly consists of the direct component that predominates under clear skies. The cloud ratio method classifies the sky condition into three skies based on the observation of the pattern of $k_d$ graphs defining a sky as either clear or overcast when the value of $k_d$ remains close to either 0 or 1, respectively, and as intermediate when $k_d$ changes frequently and rapidly (Baharuddin et al., 2010). Different works have used the diffuse fraction for the classification of radiation data into three sky conditions. Hence, its value can be divided into three intervals from 0 (clear sky) to 1 (overcast sky). (Baharuddin et al., 2010; Kong and Kim, 2013; Li and Lam, 2001; Rahim et al., 2004). In this work, three equal intervals were applied assigning a cloudiness type to each one, which unifies the intervals used in the previously mentioned studies. Both, the original and the adapted intervals are shown in Table 2.

### 3.2. Horizontal Direct Fraction $k_b$.

As an alternative to $k_d$, the horizontal direct fraction, $k_b$, is defined as the ratio between the direct horizontal irradiance, $B(0)$, and the global horizontal irradiance, $G(0)$ as shown in Equation 6. Hence, the direct horizontal irradiance can be calculated through the direct irradiance on a plane facing the Sun, $B(n)$, and the solar altitude angle, $\gamma_s$. As with $k_b$, a number of intervals can be defined between its minimum (0, overcast sky) and its maximum (1, clear sky) value. Three equal intervals were taken assigning a cloudiness type to each one of them as can be seen in Table 2. The pattern is the inverse of the one used for $k_d$, due to the complementarity of both indices ($k_b = 1 - k_d$).

$$k_b = \frac{B(0)}{G(0)} = \frac{B(n)\cos Z_s}{G(0)} = \frac{B(n)\sin \gamma_s}{G(0)} \quad \text{Eq. 6.}$$

### 3.3. Clearness index, $k_t$.

The clearness index, $k_t$, (Iqbal, 1983) (Eq. 7) is the ratio between the global horizontal irradiance, $G(0)$, and the extraterrestrial global horizontal irradiance, $G_{ext}(0)$.

$$k_t = \frac{G(0)}{G_{ext}(0)} = \frac{G(0)}{B_{sc}\,\varepsilon_0 \cos Z_s} \quad \text{Eq. 7.}$$

$B_{sc}$ is the extraterrestrial irradiance constant (1361.1 W/m$^2$ (Gueymard, 2018; Gueymard and Ruiz-Arias, 2016)), $\varepsilon_0$ is the average value of the orbital eccentricity of the Earth, calculated from Eq. 8, and $Z_s$ is the angle between sky zenith and sun.

$$\varepsilon_0 = 1 + 0.033 \cdot cos\left[2 \cdot \pi \cdot d_n/365\right] \quad \text{Eq. 8,}$$

where, $d_n$ is the day the year. $k_t$ has often been adopted to indicate the relative clearness of the atmosphere for sky categorization (Djafer et al., 2017; Escobedo et al., 2009; Kong and Kim, 2013; Wang et al., 2013) and it indicates the percentage of solar irradiance that radiates through the atmosphere. (Mellit et al., 2008) suggested that the choice of $k_t$ interval values would also differ from one site to another. In general, when the atmosphere is clear, a small fraction of the solar radiation is scattered, resulting in a predominance of direct sunlight yielding a high $k_t$ reading. Under overcast skies, a large portion of the solar radiation is dispersed, so the main component is diffuse with a small $k_t$ value. Several empirical relationships between $k_d$ and $k_t$ have been developed to calculate diffuse irradiation on horizontal and tilted surfaces from the global irradiation using a different time basis (Iqbal, 1983) that has been the object of numerous reviews (Khorasanizadeh and Mohammadi, 2016; Khorasanizadeh et al., 2016; Muneer et al., 2007; Tapakis et al., 2016; Torres et al.,

2010a). As many categories as required may be generated for the indices $k_d$ and $k_b$. The selected intervals are depicted in Table 2.

### 3.4. Clearness Function $F$

The anisotropic sky-diffuse models use the horizontal diffuse fraction, $k_d$, and the clearness index, $k_t$, to describe the prevailing sky conditions. Under non-overcast conditions, the constituent components of sky-diffuse irradiance are a circumsolar (Sun's aureole) part and background diffuse irradiance. The sky clarity indices are used to relate the above-mentioned components (Muneer, 2007), the most common of which is the clearness function, $F$, defined by Eq. 9:

$$F = \frac{G(0) - B(0)}{G_{ext}(0)} = \frac{G(0) - B(n)\cos Z_s}{B_{sc}\,\varepsilon_0 \cos Z_s} \quad \text{Eq. 9.}$$

Muneer (Muneer, 2007) established four types of sky in function of the $F$ value that were reduced to three, which are used in this work, as shown in Table 2.

### 3.5. Batlles Clearness Index $k_k$

Two new indices, $k_{tt}$ and $k_k$, have been introduced by (Batlles et al., 2000), based on solar altitude, and are defined in Eqs. 10 and 11 as:

$$k_{tt} = -0.3262 - 0.0032\,\gamma_s + 0.6843\log(\gamma_s) \quad \text{Eq. 10,}$$

$$k_k = 1.0827 - 0.3893\log(\gamma_s) \quad \text{Eq. 11.}$$

$k_k$ index is restricted to clear skies, simultaneously defined by $k_t > k_{tt}$ and $k_d < k_k$ (Muneer et al., 2004). This index therefore only distinguishes two categories of skies: clear and overcast. The original criteria were maintained for this study.

### 3.6. Klucher Clearness Index $FK$

Klucher (Klucher, 1979) proposed a model for estimating irradiance on a tilted surface and developed the function $FK$, depending on $D(0)$, and $G(0)$, as defined by Equation 12. The original and the adapted classification of cloudiness using this index, shown in Table 2, is equal to the one established in the clearness function $F$.

$$FK = 1 - \left(\frac{D(0)}{G(0)}\right)^2 = 1 - k_d^2 \qquad \text{Eq. 12.}$$

### 3.7. Perez's Clearness Indices $\epsilon$ and $\Delta$

Equations 13 and 14 introduce the sky clearness index, $\epsilon$, and the brightness factor, $\Delta$, respectively, both defined by Perez in 1987 (Perez et al., 1987) and revised in 1990 (Perez et al., 1990). $\epsilon$ predicts cloud conditions using the ratio between the diffuse horizontal irradiance and the direct one on the same plane. $\Delta$ quantifies cloud thickness or aerosol loading .

$$\epsilon = \frac{\frac{D(0) + B(n)}{D(0)} + k\, Z_s^3}{1 + k\, Z_s^3} \qquad \text{Eq. 13,}$$

$$\Delta = \frac{m\, D(0)}{B_{ext}(n)} \qquad \text{Eq. 14.}$$

$Z_s$ is the angle between sky zenith and sun $(\mathrm{rad})$ and $k = 1.04$ (or $k = 5.53 \cdot 10^{-6}$ if $Z_s$ is expressed in degrees). $B_{ext}(n)$ is the extraterrestrial direct irradiance ($B_{ext}(n) = B_{sc}\, \varepsilon_0$) and $m$ is the optical air mass calculated using the Kasten model (Kasten, 1993). Classification of the cloudiness sky types using $\epsilon$ and $\Delta$ are shown in Table 2

### 3.8. Perraudeau Nebulosity Index, $FP$

Derived from the original work of Perraudeau (Perraudeau and Chauvel, 1986), the nebulosity index, $FP$, is defined by Eq. 15 (Kambezidis et al., 1998). This index models the degree of the sky covered by clouds using the diffuse horizontal fraction, $k_d$, and the diffuse horizontal fraction for a clear sky, $k_{d0}$, given by Eq. 16.

$$FP = \frac{1 - k_d}{1 - k_{d0}} \quad \text{Eq. 15,}$$

$$k_{d0} = \frac{G_{clear}(0)}{G_{clear}(0) + B(n)} \quad \text{Eq. 16.}$$

Here $G_{clear}(0)$ , given by eq. 17, is the clear-sky irradiance and $B(n)$ is the beam irradiance.

$$G_{clear}(0) = (0.5528 + 0.8785 \cdot \gamma - 0.01322 \cdot \gamma^2 + 0.0003434 \cdot \gamma^3) \cdot$$
$$(6.9731 + 0.042496 \cdot \gamma - 8.5275 \cdot 10^{-4} \cdot \gamma^2 - 8.6088 \cdot 10^{-5} \cdot \gamma^3 + 1.984 \cdot 10^{-6} \cdot \gamma^4 - 1.6222 \cdot 10^{-8} \cdot \gamma^5 + 4.7823$$

Eq. 17

Depending on the $FP$ value, five sky types can be identified that are shown in Table 2.

### 3.9. Igawa Index $Si$

The Igawa Index (Igawa et al., 2004), $Si$, defined by Eq. 18, uses the standard global irradiance, $G_{st}$, the cloudless index, $Cle$, and the standard cloudiness fraction, $Ces$, calculated from equations 18-21.

$$Si = \frac{G(0)}{G_{st}} + \sqrt{Cle} \quad \text{Eq. 18,}$$

$$G_{st} = 0.84 \frac{B_{sc}}{m} e^{-0.0657\, m} \quad \text{Eq. 19,}$$

$$Cle = \frac{1 - k_d}{1 - Ces} \quad \text{Eq. 20,}$$

$$Ces = 0.01299 + 0.07698\, m - 0.003857\, m^2 + 0.0001054\, m^3 - 0.000001031\, m^4 \quad \text{Eq. 21.}$$

$G_{st}$ is calculated considering the Linke turbidity factor as 2.5 in the Kasten model (Kasten, 1993) of global irradiance for a clear sky; $Ces$ is given by a polynomial fit adjustment using the optical air mass as the independent variable and $B_{sc}$ is the extraterrestrial irradiance constant (1361.1 W/m$^2$ ). The original criterion used by Igawa and the adapted one stablished following the procedure described by Figure 2 (Section 3) are shown in Table 2.

Table 2. Summary of the MIs used to classify the skies, the original intervals used to define the clear, partial and cloudy sky conditions and the adapted ones used in this work.

| Symbol | MI | Ref. | ORIGINAL | ADAPTED |
|---|---|---|---|---|
| $k_d$ | Diffuse fraction | (Kong and Kim, 2013) | (0.00, 0.33] clear<br>(0.33, 0.8) partial<br>[0.8, 1) cloudy | (0.00, 0.33] clear<br>(0.33, 0.8) partial<br>[0.8, 1) cloudy |
| $k_b$ | Direct fraction | | | [0.66, 1) clear (0.33, 0.66) partial<br>(0, 0.33] cloudy |
| $k_t$ | Clearness Index | (Wang et al., 2013) | [0.65, 1) clear<br>(0.35, 0.65) partial<br>(0, 0.35] cloudy | [0.65, 1) clear<br>(0.65, 0.35) partial<br>(0, 0.35] cloudy |
| $F$ | Clearness Function | (Muneer, 2007) | [0.61, 1.00) completely clear<br>[0.51, 0.61) clear<br>[0.18, 0.51) partial<br>[0.00, 0,18) completely cloudy | [0.51, 1.00) clear<br>[0.18, 0.51) partial<br>(0.00, 0.18) cloudy |
| $k_k$ | Batlles Clearness Index | (Batlles et al., 1998) | $k_t > k_{tt}$<br>$k_d < k_k$ clear | $k_t > k_{tt}$<br>$k_d < k_k$ clear |
| $FK$ | Klucher Clearness Index | (Klucher, 1979) | [0.61, 1.00) completely clear<br>[0.51, 0.61) clear | [0.51, 1.00) clear<br>[0.18, 0.51) partial<br>(0.00, 0.18) cloudy |



Journal Pre-proofs

| | | | | |
|---|---|---|---|---|
| | | | [0.18, 0.51) partial<br>[0.00, 0,18)<br>completely cloudy | |
| $\epsilon$ | Perez's clearness index | (Perez et al., 1990; Perez et al., 1987) | [6.20, ∞)<br>completely clear<br>[2.80, 6.20) clear<br>[1.50, 2.80) partial<br>[1.065, 1.50) cloudy<br>[1.00,1.065)<br>completely cloudy | [2.4, ∞) clear<br>[1.50, 2.4) partial<br>[1.00, 1.50) cloudy |
| $\Delta$ | Sky brightness | (Perez et al., 1990; Perez et al., 1987) | [0.48, ∞) very bright<br>[0.30, 0.48) bright<br>[0.10, 0.30) partial<br>[0.00, 0.10) very dark | [0.30, ∞) clear<br>[0.10, 0.30) partial<br>[0.00, 0.10) cloudly |
| $FP$ | Nebulosity Index | (Kambezidis et al., 1998) | [0.90, 1.00] blue sky<br>[0.70, 0.90) partial<br>[0.20, 0.70) partially blue<br>[0.05, 0.20) partially cloudy<br>[0.00, 0.05) totally cloudy | [0.70, 1.00] clear<br>[0.05, 0.70) partial<br>[0.00, 0.05) cloudy |
| $S_i$ | Igawa Index | (Igawa et al., 2004) | [1.70, ∞) clear<br>[1.50, 1.70) almost clear<br>[0.60, 1.50) partially clear<br>[0.30, 0.60) partially cloudy<br>[-∞, 0.30] totally cloudy | [1.70, ∞) clear<br>(0.30, 1.70) partial<br>[-∞, 0.30] cloudy |

## 4. CONFUSION MATRICES

In the field of machine learning, a confusion matrix is used for measuring the performance of classification algorithms. Classification is known as a supervised learning approach, because the machine is trained with selected examples of the data from which it is said to learn. After the learning process, the model that has been tuned by the data inputs will classify a new observation or sample into a given number of classes. The classification can be binary (e.g. a dichotomous classification of email as either spam or not spam) or multi-



Journal Pre-proofs

class (e. g. recognition of handwritten numbers). When the algorithm attempts to classify a sample, it processes an individual measurable property or feature. For example, in binary classification the algorithm attempts to detect the presence of a relevant feature. If it is detected, the sample is labelled as “positive” and, if otherwise, “negative”, (Figure 3 ). Comparing the prediction with reality, there are four possible scenarios: the predicted positive will agree with the actual one (True Positive or TP), the predicted positive will not agree with the actual one (False Positive or FP), the predicted negative will agree with the actual one (True Negative or TN) and the predicted negative will not agree with the real one (False Positive or FP). Each row of the matrix represents the instances of the predicted class, while each column represents instances of the reference class (Powers, 2011). The matrix was so-named, because it easily visualizes whether the algorithm is confusing or mislabelling two classes.

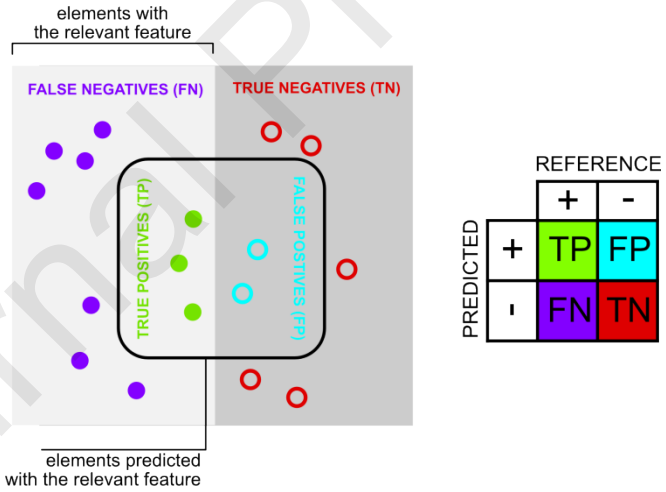



Figure 3 : Confusion matrix: possible scenarios in the comparison of the prediction to the actual data.

In this present work, the reference labels are established by the CIE methodology for sky classification and the predicted labels by the MIs. Hence, the latter are the algorithms or

predictive models under analysis. As mentioned above, the CIE defines fifteen types of sky that can be grouped into three types of cloudiness: clear, partial, or cloudy. The MIs distinguish between the features in an attempt to define these three classes. This task can be analyzed as a multiclass problem that can be decomposed as a multiple dichotomous classification where each cloudiness categorization is predicted against the other remaining ones (Figure 4). At the end of the process, the dichotomous classification of the MIs is evaluated using several performance ratings explained in the following sections.

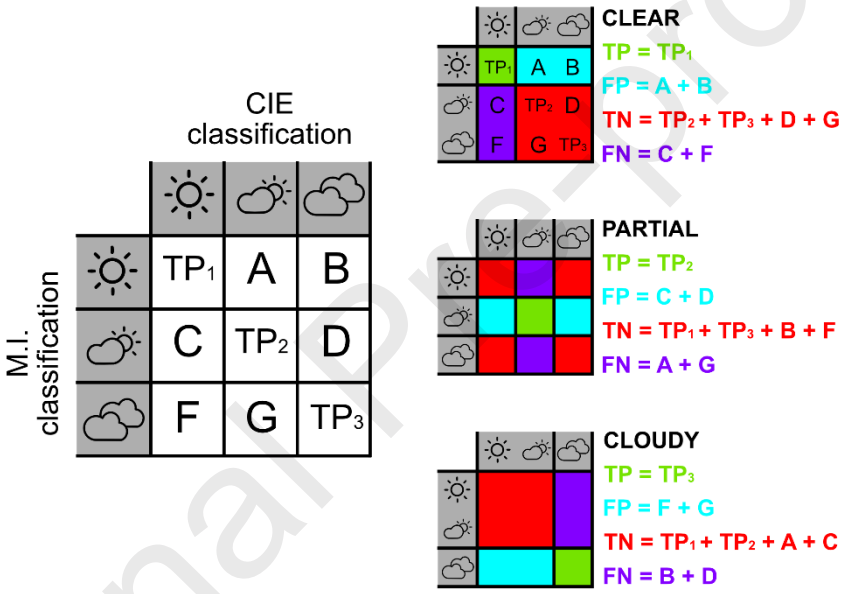



Figure 4 : Multiclass to dichotomous transformation.

### 4.1. Accuracy Index

The Accuracy Index, also known as the Simple Matching Coefficient, represents the ratio of correct predictions, positive or negative, amongst all the cases evaluated by the algorithm expressed by Eq. 22:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{Eq. 22.}$$

As can be appreciated, it represents the overall portion of agreement and it is usually the starting point for analyzing the quality of a predictive model. The main weakness of the so-called Accuracy Paradox is that models of a given accuracy may have greater predictive power than others of higher accuracy (Kundel and Polansky, 2003). An illustration of this fact would be a comparison of two different algorithms designed to detect insurance frauds for the same set of 100 samples, represented in Figure 5. Given a confusion matrix (with notation [TP FP; FN TN]) for the first one of $M_1 = [1\ 2;2\ 95]$ and for the second one of $M_2 = [0\ 2;0\ 98]$, then the accuracy of the first and the second algorithm would be 0.95 and 0.98, respectively. In comparison with the first model, the second model would show fewer incorrect predictions and improved accuracy, globally; however, its fraud detection power would be weaker. This paradox arises in imbalanced data where there is a substantial difference in the size of the categories. The use of other metrics may therefore be advisable.

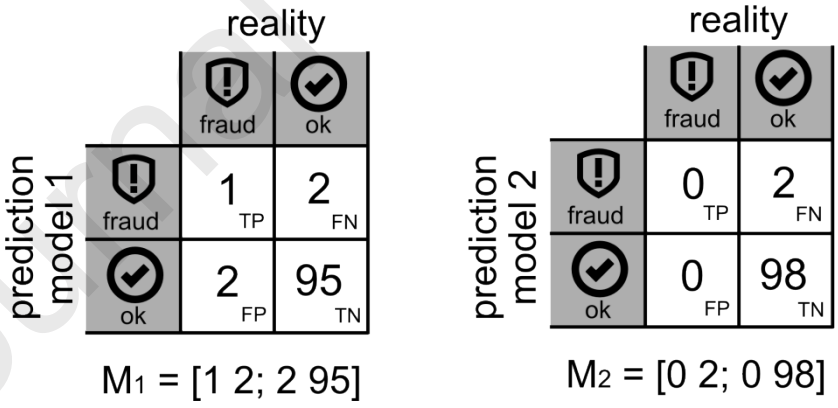



Figure 5: Confusion matrices for a fraud detection example.

In multiclass classification performance, there are two types of averages: micro-averages and macro-averages. The micro-average sums up all the TP, TN, FP, and FN cases of all the classes, aggregating their contributions before calculating the metrics. The result would be dominated by the performance of the common categories and it would be similar to weighting it against the population size of each category. A bad performance labelling one of the classes could be masked. The macro-average - the average type used in this work- is computed by averaging the metrics after their calculation. Therefore, equal weight was given to each category, regardless of its frequency. Using the latter, the classification performance must be good in all of the (cloudy, partial and clear) classes, if it is to be among the highest positions of the metric ranking.

### 4.2. Jaccard Index

The Jaccard Index (Jaccard, 1912) expressed by Eq. 23, measures the similarity between the number of true positives in all the cases under evaluation. The main difference to the Accuracy Index is the omission of $TN$, meaning that the Jaccard Index only takes into account successful detection of the relevant attribute. In the above insurance fraud example, the Jaccard Index is 0.20 for $M_1$ and 0 for $M_2$, showing a better reflection of the predictive power of the first algorithm.

$$Jacard\ Index = \frac{TP}{TP + FP + FN} \quad \text{Eq. 23.}$$

The Jaccard Index is focused on the detection of the relevant feature or true positives. It makes no distinction between FP and TN. There are some scenarios where such a distinction is crucial. For example, in medical diagnostics, a model with a confusion matrix [90 10; 0 0] and another characterized by [90 0; 10 0] would have the same Jaccard Index

of 0.9. However, the first model would have 10 FP cases and the second one 10 FN. The second model would have harmful consequences, so the use of the first one would be recommendable. Depending on the field of application, the use of another metric or full information would be recommendable before taking the final decision.

### 4.3. Cohen's Kappa

Cohen's Kappa, $\kappa$, is a measure of true agreement rather than a true prediction that reflects the possibility that the algorithm agrees with the reference by chance. It indicates the proportion of agreement beyond that expected by chance, as shown in Eq. 24:

$$\kappa = \frac{observed\ agreement - chance\ agreement}{1 - chance\ agreement} = \frac{p_o - p_e}{1 - p_e} \quad \text{Eq. 24,}$$

where $p_o$ is the overall agreement given by Eq. 25; $p_e$ is the expected agreement by chance as calculated by Eq. 26 , $1 - p_e$ is the fraction of cases on which agreement is not expected to occur by chance, and $N$ (Eq. 27) is the total number of cases analyzed by the algorithm.

$$p_o = \frac{TP + TN}{N} \quad \text{Eq. 25,}$$

$$p_e = \left(\frac{TP + FP}{N}\right)\left(\frac{TP + FN}{N}\right) + \left(\frac{TN + FP}{N}\right)\left(\frac{TN + FN}{N}\right) \quad \text{Eq. 26,}$$

$$N = TP + TN + FP + FN \quad \text{Eq. 27.}$$

Table 3 shows the strength of agreement for various ranges of $\kappa$ suggested by (Landis and Koch, 1977). The choice of intervals is arbitrary, but is now in wide use (Kundel and

Polansky, 2003). In the previous example, Cohen's Kappa scored 0.313 (fair strength) for $M_1$ and 0 (poor strength) for $M_2$, avoiding the Accuracy Paradox.

Table 3 Strength of agreement indicated with $\kappa$ values

| $\kappa$ | **strength** |
|---|---|
| $(-\infty, 0]$ | poor |
| [0, 0.20) | slight |
| [0.20, 0.40) | fair |
| [0.40, 0.60) | moderate |
| [0.60, 0.80) | substantial |
| [0.80, 1.00] | almost perfect |

Cohen's Kappa takes into account the four classes of the confusion matrix and is more informative than other confusion-matrix measures. It can produce drastically different results to Accuracy and the Jaccard Index in scenarios where the positives cases are the predominant class. For example, a classification algorithm that produces the confusion matrix $M_3 = [90\ 5;4\ 1]$ would produce Accuracy 0.91, Jaccard Index 0.91, and $\kappa$ 0.13. Nevertheless, the Jaccard Index and Accuracy will provide indicative assessments of almost perfect predictive power and $\kappa$ will predict a performance near a random classifier with null predictive power, because of its incapacity to detect true negatives (Sim and Wright, 2005).

### 4.4. Matthews Correlation Coefficient

The Matthews Correlation Coefficient (Matthews, 1975) given by Eq. 28 is a measure of the quality of the binary classification:

$$Matthews\ Correlation = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FN)(TP + FP)(TN + FP)(TN + FN)}} \quad \text{Eq. 28.}$$



Journal Pre-proofs

Essentially, it is a correlation coefficient between the reference and predicted classification returning a value between +1 (perfect prediction) and -1 (total disagreement). The 0 value is representative of a random prediction. Returning again to the Accuracy Paradox example, the Matthews Correlation is 0.31 for $M_1$ and 0 for $M_2$. As with Cohen's Kappa, the Matthews Correlation takes into account the four classes of the confusion matrix with the same implications as those seen in the Section 4.3 (Chicco, 2017).

### 4.5. Combination of the confusion matrix metrics

In the preceding sections, various metrics for a classification algorithm have been explained. All of them are attempts to sum up the confusion matrix associated with the algorithm using only one number. Inevitably, the process is associated with a loss of information, because a four dimensional matrix is collapsed into one number. Each dimension attempts to highlight one aspect of interest. However, they are correlated in some way because the starting data are all the same. In fact, the ranking offered by each one is very similar.

The different rankings produced by all the metrics are combined to arrive at the best MI for classifying sky cloudiness. At the end of the process, four rankings are thus obtained, one per confusion matrix. The one designated as best model holds the best positions in all of them. So, by using all of them simultaneously, the deficiencies of the metric and its biases are avoided.

Journal Pre-proofs

# 5. EXPERIMENTAL SECTION

## 5.1. The meteorological facility

The experimental data for this study were gathered at a meteorological weather station located on the roof of the Higher Polytechnic School building at Burgos University (42°21′04″N; 3°41′20″O; 856 m above mean sea level). This five-storey building is in an area with no other buildings of comparable height, free from any external obstructions. The experimental equipment is shown in Figure 6 . The following meteorological data were measured: temperature, wind velocity and direction, atmospheric pressure, humidity, and rainfall. Global, beam, and diffuse horizontal irradiation ($G(0)$, $B(n)$, $D(0)$) were all recorded using first-class Hukseflux SR11 pyranometers and a Hukseflux DR01 pyrheliometer. The facility includes a SONA201D All-Sky Camera-Day, from Sieltec Canarias and an MS-321LR sky scanner, from EKO. The experimental variables were recorded with a CAMPBELL CR3000 datalogger. Tables 4, 5 and 6, respectively, show the technical specifications of the sky scanner, the pyranometers, and the pyrheliometer used for this study.

Journal Pre-proofs

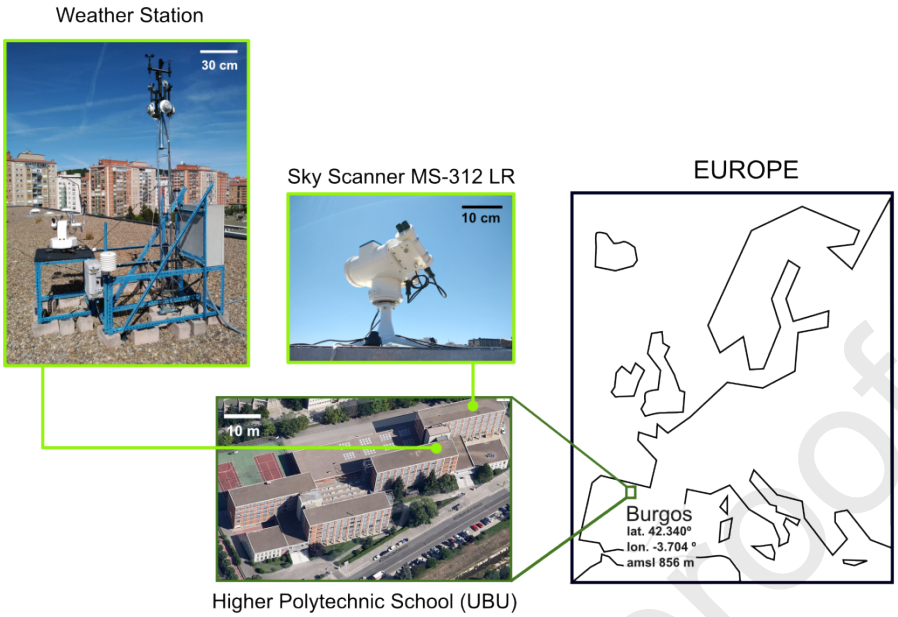



Figure 6 : Location of the meteorological station on the roof of the Higher Polytechnic School building at the University of Burgos, Spain.

Table 4. Sky Scanner specifications

| | |
|---|---|
| Model | MS-321LR EKO Instruments |
| Dimensions (W x D x H) | 430 mm x 380 mm x 440 mm |
| Mass | 12.5 kg |
| Aperture | 11 º |
| Luminance | 0 to 50 kcd/m$^2$ |
| Radiance | 0 to 300 W/m$^2$/sr |
| A/D Convertor | 16 bit |
| Calibration Error | 2% |

Table 5. Pyranometer specifications

| | |
|---|---|
| Model | SR11 |
| ISO classification | first class |
| Spectral range | 300 to 2800 nm |
| Irradiance range | 0 to 2000 W/m$^2$ |
| Sensitivity | 15x10$^{-6}$ V/(Wm$^{-2}$) |
| Calibration uncertainty | < 1.8% |

Table 6. Pyrheliometer specifications

| | |
|---|---|
| Model | DR01 |
| ISO classification | first class |
| Spectral range | 200 to 4000 nm |
| Irradiance range | 0 to 2000 W/m$^2$ |
| Sensitivity | 7-15x10$^{-6}$ V/(Wm$^{-2}$) |
| Calibration uncertainty | < 0.3% |

### 5.2. Data Processing

Global, beam, and diffuse irradiance data are measured from June 2016 to May 2017, for the classification and benchmarking of the different MIs selected for this study. Both the pyranometers and the pyrheliometer are classified as “first class” in the ISO classification (ISO 9060:1990) with a WMO performance level that is of “good quality” (WMO, 2010). The calibration and the management of the meteorological facility is done following ISO (1992) and the WMO Guide to Meteorological Instruments and Methods of Observation (WMO, 2008. (Updated 2010)). Global, diffuse, and beam irradiation data are recorded every ten minutes (averaging recorded scans of thirty seconds). The sky-scanner completes a full scan in four minutes and starts a new scan every fifteen minutes. So as to match simultaneous records of global, diffuse, and beam irradiation, only half-hourly and hourly sky-scanner measurements are used in this study. The irradiance data are analysed and filtered using traditional quality criteria (Gueymard and Ruiz-Arias, 2016). If irradiance data (global, diffuse, and beam irradiance data) fail to pass the quality criteria, then the three simultaneous data sets are rejected.

The number of indices, once the data from both sources had passed the quality criteria, was above 300 per month, as shown in Figure 7. In the summer months there were fewer estimated indices than expected, because the weather station and Sky Scanner

experienced several shutdowns due to servicing works. Overall, over 3600 indices were correlated with their respective sky cloudiness.

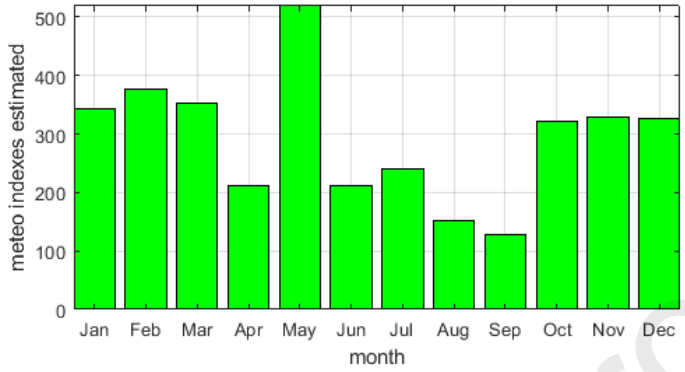



Figure 7 . Number of sets of the MIs estimated per month

### 5.3. CIE standard classification of skies in Burgos, Spain

The sky scanner divides the sky into 145 patches or sectors (*p*) that cover the whole dome. The sectors are grouped into eight bands, named $b_p$, and by their solar altitude, $\gamma = (\frac{\pi}{2} - Z)$, where *Z* is the zenith angle. Figure 8 shows the location of the sectors in the whole dome. A luminance measurement (kcd/m$^2$) of each patch is taken four times per hour. Half-hourly and hourly measurements were recorded between June 2016 and May 2017. Continuous scanning yielded luminance data corresponding to the 145 patches (see Figure 8) recommended for the CIE in the Guide to Daylight Measurements (Tregenza et al., 1994), which were measured and registered. Likewise, the luminance corresponding to each of the commonly considered 15 standard sky types presented in Table 1 was calculated at the same time and for the same 145 patches. The standard sky type ascribed to each record expressed the lowest RMSD (Root Mean Square Deviation) between the 145 normalized luminance values that were measured and calculated.

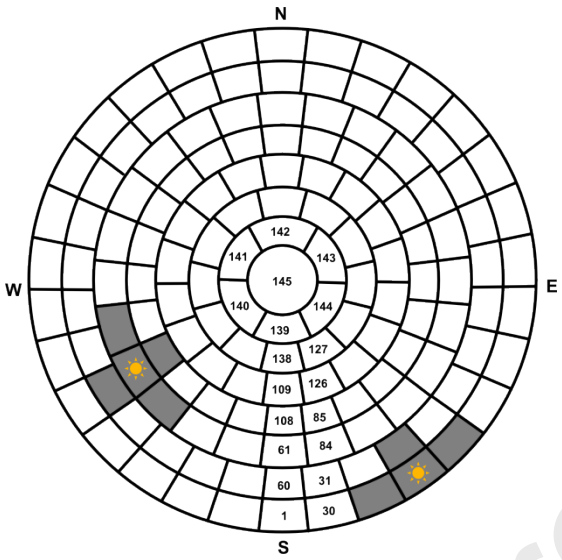



Figure 8 : Sky divided into 145 sectors ($p$) grouped into 8 bands ($b_p$). The number of patches per band ($n_b$) is shown in the figure. Figure shows two ficticious cases illustrating how, depending on sun position, the patches adjacent to the position of the sun are excluded for the luminance calculation.

The sky scanner is adjusted each month to measure daylight hours from sunrise to sunset. The first and last measurement of the day (solar elevation angle equal or lower than 5º) were discarded, as were measurements higher than 50 kcd/m$^2$ or lower than 0.1 kcd/m$^2$, following the specification of the equipment. Figure 9 gives the frequency of occurrence (FOC) of the sky classification results by the method. One previous work (Suárez-García et al., 2018) described the details of the sky luminance measurement and classification methodology. All sky types of the CIE classification, shown in Table 1, can be found in Burgos, from overcast to very clear. The lowest frequency is for type I.2, corresponding to Overcast with the steep gradation and slight brightening toward the sun (1.7 %), and the highest frequency is for type V.5. (18.3 %) (Cloudless polluted with a broader solar corona). Cloudiness labelling was done with the CIE sky types: I.1 to III.1 classified as cloudy, III.2 to IV.3 as partially cloudy, and IV.4 to VI.6 as clear skies. These three categories reflect the characteristically

clear skies that are predominant in Burgos (56.3 %), while cloudy skies are presented in 23.7% and partially cloudy in 20 % of cases, as shown in Figure 10.

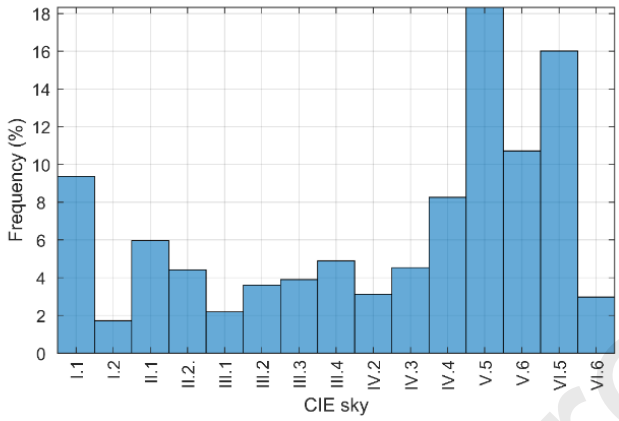



Figure 9: Frequency of occurrence (FOC) from June 2006 to May 2017 Burgos CIE Standard Skies

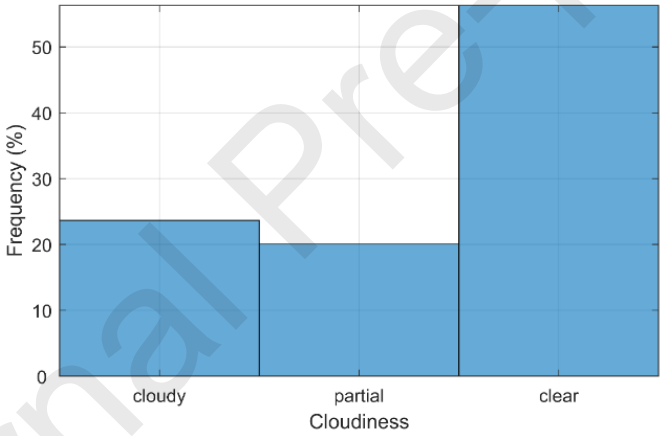



Figure 10: Comparative study of cloudiness classifications in Burgos.

## 6. RESULTS AND DISCUSSION

### 6.1. Results of cloudiness classification from the meteorological indices

In this section, the results of the calculation of each MI used for sky classification are presented. The calculated MIs have values inside the intervals defined by their authors



Journal Pre-proofs

(Table 2). The box plots, presented in Figure 11, show the data organized by quartiles. The high variability of their interquartile ranges (IQR) can be observed: $k_t$, $k_b$ , $k_d$ , $FK$, $\epsilon$ and $Si$. It might be thought that a wider dispersion of the values would mean a higher number of different cloudiness classifications. As will be seen later (Table 8), $\epsilon$, $FK$, and $k_t$ are ranked first, third, and fourth in the cloudiness classification. However, the second best single index ($FP$) had an IQR of low variance. The explanation could be a better adaption of its intervals originally established by the CIE types of cloudiness.

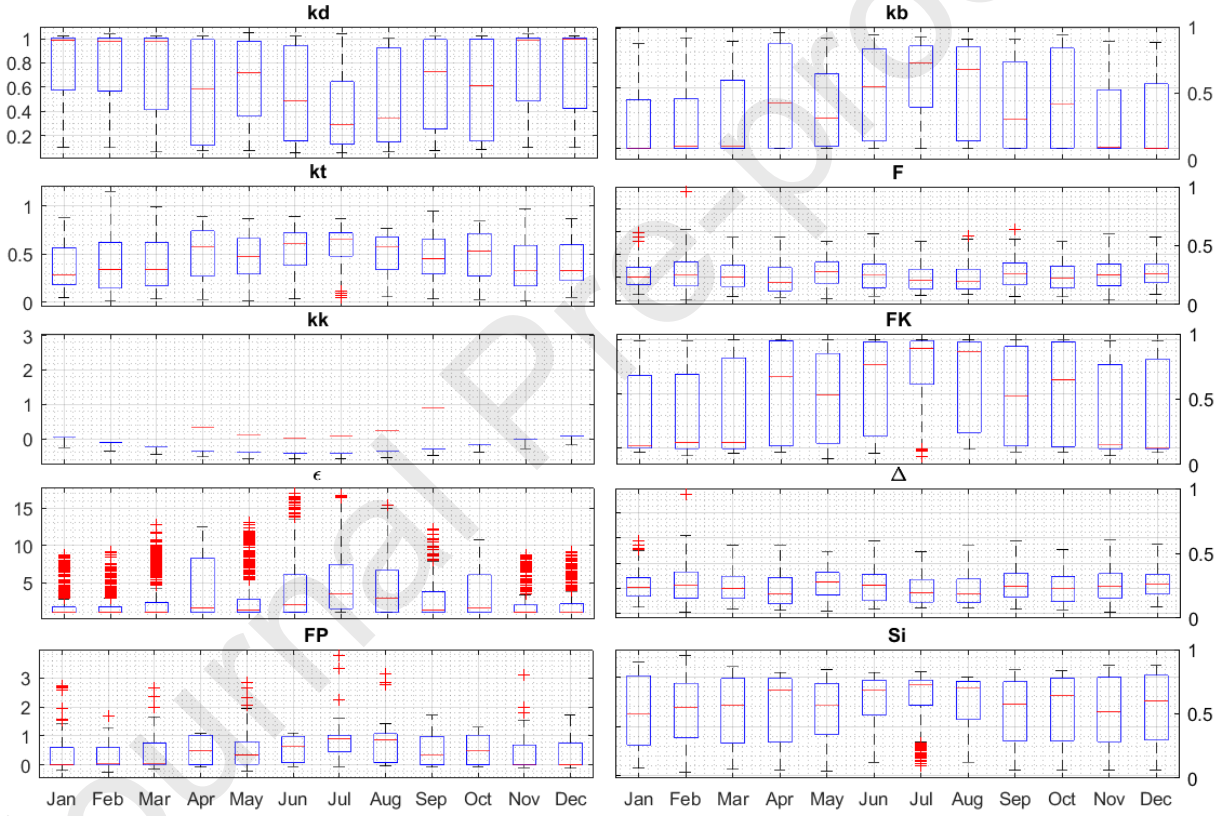



Figure 11: MIs calculated from the experimental data of global, beam, and diffuse irradiance measured in Burgos from June 2016 to May 2017.

The results of each MIs for the cloudiness classification are summarized in Figure 12 and presented by month. Figure 13 shows the monthly cloudiness classification obtained from the standard CIE.

Journal Pre-proofs

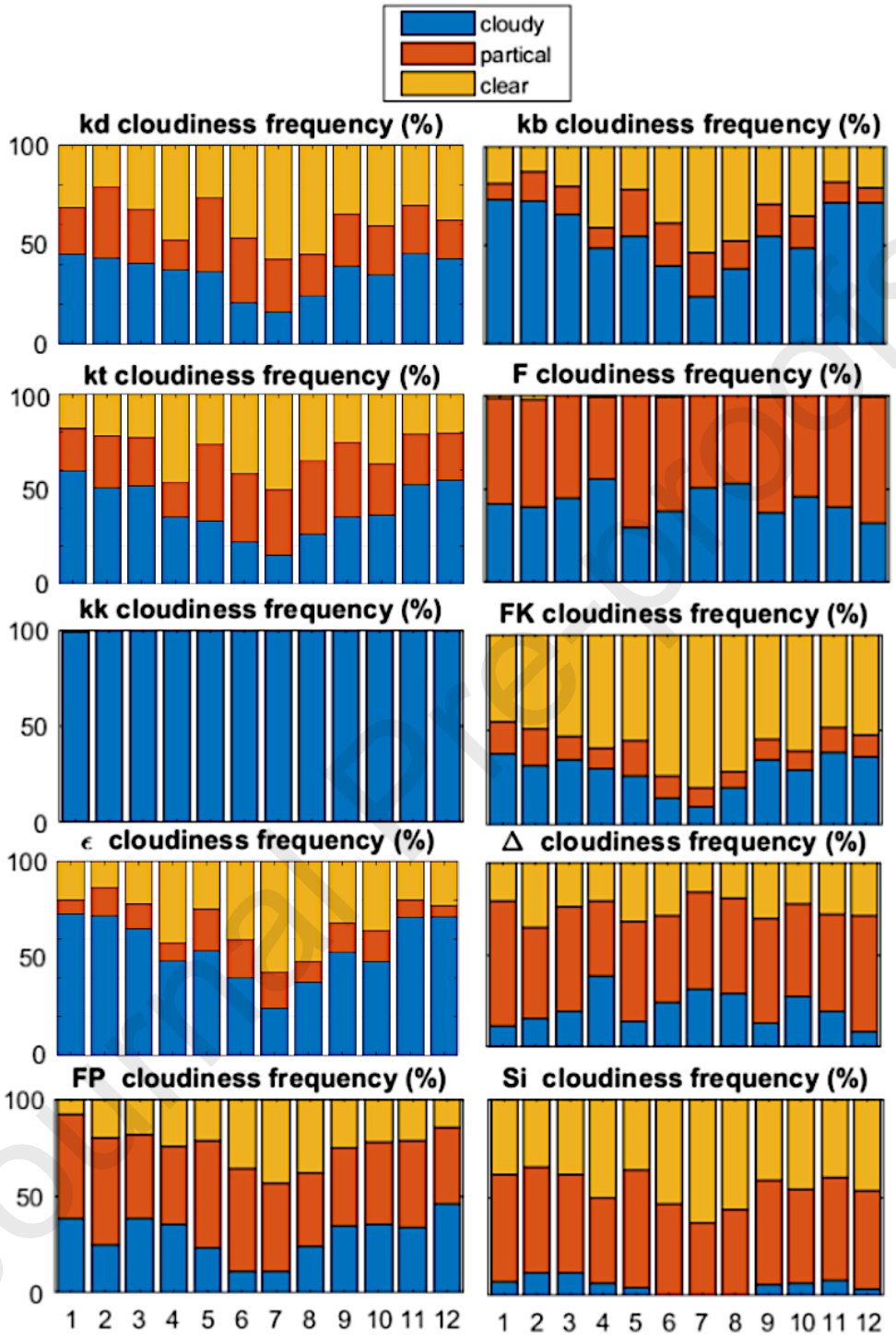



Figure 12 : Monthly cloudiness classification obtained from each meteorological index calculated from the experimental data of global, beam, and diffuse irradiance measured in Burgos from June 2016 to May 2017.

Individual analysis of sky classification with the MIs shows that practically no index is able to identify the high percentage of clear skies indicated by the CIE classification. The Klucher clearness index ($FK$) is the one that identifies a higher percentage of clear skies, and at the opposite extreme, the Batlles clearness index identifies practically all cases as overcast. $\Delta$, $Si$, $FP$, and $F$, identify most of the cases as partial cloudiness. The stacked bar graph shows an imperfect match between the different approaches. A frequency graph would be insufficient for a complete comparison: the energy contribution of the Sun depends on its elevation above the horizon, which varies according to the time of day and day of year. It is therefore necessary to compare the classification at each timestamp.

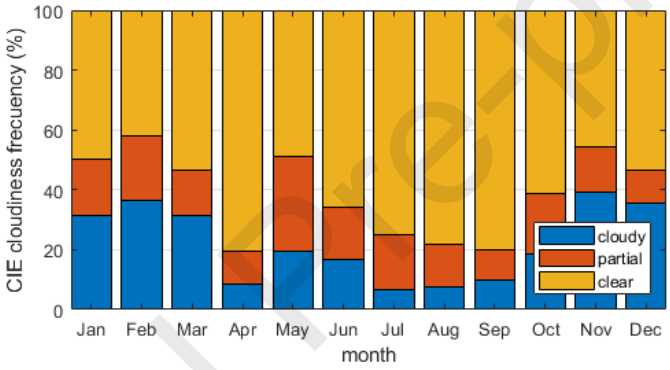



Figure 13: CIE monthly cloudiness classification

### 6.2. Confusion Matrices

Finally, the comparison between classifications at each instant is obtained and presented in terms of the confusion matrices and the rating variables: Accuracy, $\kappa$, Jaccard, and Matthews. The overall power prediction of the MIs is shown using the data for all the months. Then, their estimations are compared with the CIE classification for all the samples that were registered. The maximum percentage of true positives in each sky type is limited by the CIE classification shown in Figure 10. For this reason, an ideal algorithm with no errors in its classification would obtain the following classification matrices, expressed as percentages:

$M_{clear} = [57\ 0;0\ 43]$, $M_{partial} = [24\ 0;0\ 76]$ and $M_{cloudy} = [19\ 0;0\ 81]$. In the matrices, the values are the maximum true positives and true negatives that the algorithms can obtain. In Figure 14 the confusion matrices for each MI used as a sky cloudiness classifier are presented. Percentages of true positives (tp), true negatives (tn), false positives (fp) and false negatives (fn) are depicted. The lowercase abbreviation is used to distinguish the percentage of the raw recount where uppercase is preferred. To give as much information as possible, the number of samples analyzed, N, is detailed in the title of each figure. For example, the confusion matrix for the $k_d$ detecting clear skies is $M_{clear,\ kd} = [15\ 24;43\ 18]$ in percentage terms or $M_{clear,\ kd} = [330\ 527;945\ 395]$ in samples counting.

A predominant behavior can be observed in each one of the sky types. The false negatives are predominant in the clear skies (Figure 14). There is a considerable number of cases where the MIs predict a non-clear sky (i.e. partial or cloudy) when the CIE model predicts a clear sky. Better than all of the other MIs, $FK$ stands out from the rest with a true positive ratio near 45%, ten percentage points higher than the rest.

The behavior of the MIs slightly changes in the partly cloudy skies classification shown in Figure 14, where the false positives and true negatives are predominant. Clear and cloudy skies are clearly distinguished. However, MIs can hardly classify a partial cloudiness sky correctly. Finally, the true negatives predominate in the cloudy scenarios, with the rest of cases evenly spread out between the true positives, false positives, and false negatives. In both cloudiness scenarios, partial and cloudy skies, the true positive ratio is below 20 %.

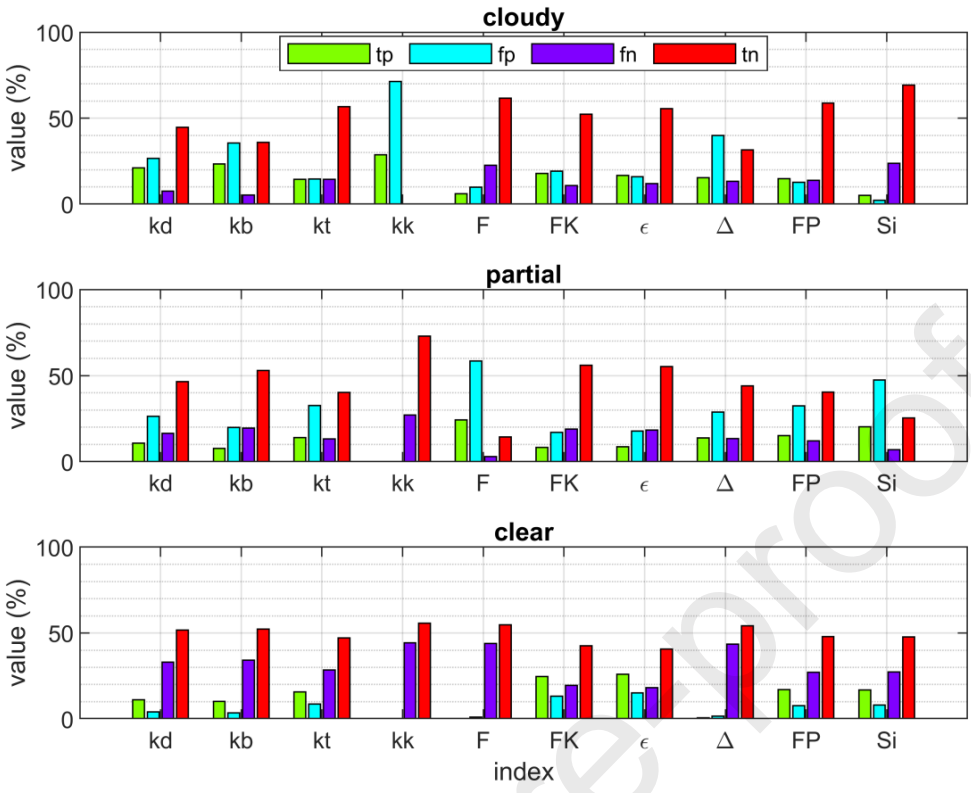



Figure 14: Results of the Confusion matrices for cloudy skies (N = 2197), partly cloudy skies (N = 2197), and clear skies (N = 2197).

The confusion matrix metrics, shown in Table 7, summarize the performance of the MI classifications. The results are discussed in two groups: the indices that characterize similarity (Accuracy and Jaccard) and the indices that analyze randomness in the agreement (Cohen Kappa and Matthews Correlation). The results obtained by Accuracy, in some cases, show values, above 80 %, such as the classification of cloudy skies by the $FP$ and the $Si$ indices, and the average value was above 60% in five indices ($FK$, $FP$, $Si$, $k_t$, and $k_b$). Nonetheless, these high values are reached, because the true negative ratios value are included in the accuracy computation. In the present study, where there are three classes (clear, partial, and cloudy), a true negative obtained in one class implies that the sample

could be in any of the other two. The Jaccard index shows an average of the meteorological classifiers below 35%. This considerable reduction is because the Jaccard Index only measures the similarity of the true positives, omitting the true negatives. The Jaccard index emphasizes the scoring capability of the algorithm to detect the desired value of the attribute under analysis. The Jaccard Index is nevertheless 45% when the Klucher Clearness classifies clear skies. Cohen's Kappa and Matthews Correlation Coefficient both show a fair strength of agreement for most of the classifiers, as shown in Table 7. Poor performance under partly cloudy skies is also very visible from the values, below 10% obtained for both classifiers. The results obtained in the cloudy skies are the best ones, followed by the clear ones. In other words, the extreme situations are those that are detected with best accuracy.

Table 7: Results of the confusion matrix indicators obtained by each of the MIs studied for the different cloudiness categories and on average by all cloudiness classification methods.

| | MI | $k_d$ | $k_b$ | $k_t$ | $F$ | $kk$ | $FK$ | $\epsilon$ | Δ | $FP$ | $Si$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **ACCURACY** | **Av.** | **0.6190** | **0.6077** | **0.6267** | **0.5375** | **0.5242** | **0.6779** | **0.6758** | **0.5397** | **0.6474** | **0.6147** |
| | cloudy | 0.6579 | 0.5926 | 0.7105 | 0.6763 | 0.2863 | 0.7000 | 0.7211 | 0.7205 | 0.7358 | 0.7421 |
| | partial | 0.5711 | 0.6063 | 0.5421 | 0.3858 | 0.7284 | 0.6658 | 0.6389 | 0.4511 | 0.5553 | 0.4558 |
| | clear | 0.6279 | 0.6242 | 0.6274 | 0.5505 | 0.5579 | 0.6679 | 0.6674 | 0.4474 | 0.6511 | 0.6463 |
| **JACCARD** | **Av.** | **0.2701** | **0.2465** | **0.2871** | **0.1494** | **0.0954** | **0.3268** | **0.3357** | **0.1596** | **0.3143** | **0.2525** |
| | cloudy | 0.3810 | 0.3651 | 0.3309 | 0.1575 | 0.2863 | 0.3722 | 0.3735 | 0.0635 | 0.3597 | 0.1624 |
| | partial | 0.1994 | 0.1624 | 0.2342 | 0.2827 | 0.0000 | 0.1511 | 0.1948 | 0.1813 | 0.2542 | 0.2713 |
| | clear | 0.2298 | 0.2119 | 0.2962 | 0.0081 | 0.0000 | 0.4570 | 0.4387 | 0.2341 | 0.3289 | 0.3239 |
| **$\kappa$** | **Av.** | **0.1734** | **0.1436** | **0.1856** | **0.0423** | **0.0000** | **0.2366** | **0.2482** | **-0.0431** | **0.2318** | **0.1664** |
| | cloudy | 0.3020 | 0.2431 | 0.2940 | 0.0857 | 0.0000 | 0.3247 | 0.3442 | 0.0671 | 0.3456 | 0.1862 |
| | partial | 0.0275 | 0.0086 | 0.0551 | 0.0541 | 0.0000 | 0.0573 | 0.0796 | -0.0777 | 0.0909 | 0.0638 |
| | clear | 0.1906 | 0.1792 | 0.2078 | -0.0128 | 0.0000 | 0.3278 | 0.3207 | -0.1186 | 0.2590 | 0.2493 |
| **MATTHEWS** | **Av.** | **0.2006** | **0.1813** | **0.1944** | **0.0812** | **0.0000** | **0.2392** | **0.2488** | **0.1120** | **0.2432** | **0.2051** |
| | cloudy | 0.3278 | 0.2943 | 0.2940 | 0.0919 | 0.0000 | 0.3305 | 0.3456 | 0.1283 | 0.3457 | 0.2532 |
| | partial | 0.0283 | 0.0086 | 0.0602 | 0.1043 | 0.0000 | 0.0593 | 0.0796 | 0.0890 | 0.1000 | 0.0892 |
| | clear | 0.2456 | 0.2411 | 0.2291 | 0.0473 | 0.0000 | 0.3279 | 0.3213 | 0.1186 | 0.2840 | 0.2728 |

Table 8 summarizes the average ranking obtained by the MIs. $\epsilon$ was in first place in the Jaccard Index, $\kappa$ and Matthews evaluations. It was only surpassed by the $FK$ index in the Accuracy ranking. Bearing in mind the similarity of the numerical values obtained by $\epsilon$, $FK$ and $FP$ in the evaluation indices, these three MIs may be identified as the best classifiers of sky cloudiness. When studying only the capacity of the MIs to identify clear skies, $FK$ obtained first position in all the indicators. This result was already shown in Figure 12. None of the MIs under analysis have simultaneously shown good results in the rating variables, for the identification of either partial or cloudy skies,.

Table 8. Ranking average summary

| | $k_d$ | $k_b$ | $k_t$ | $F$ | $kk$ | $FK$ | $\epsilon$ | Δ | $FP$ | $Si$ |
|---|---|---|---|---|---|---|---|---|---|---|
| **ACCURACY** | 5 | 7 | 4 | 9 | 10 | 1 | 2 | 8 | 3 | 6 |
| **JACCARD** | 5 | 7 | 4 | 9 | 10 | 2 | 1 | 8 | 3 | 6 |
| $\kappa$ | 5 | 7 | 4 | 8 | 9 | 2 | 1 | 10 | 3 | 6 |
| **MATTHEWS** | 5 | 7 | 6 | 9 | 10 | 3 | 1 | 8 | 2 | 4 |
| **GLOBAL** | **5** | **7** | **4** | **9** | **10** | 2 | 1 | **8** | 3 | **6** |

## 7. CONCLUSIONS

A key aspect in the modelling of solar radiation and daylighting is sky classification. The use of MIs as an alternative to the CIE standard classification has been investigated in this work. Various MIs for sky classification have been reviewed. Ten classical MIs obtained from global, diffuse, and/or beam irradiance measurements, recorded at weather stations have been used. The baseline reference in the one-year experimental measurement campaign was the standard CIE classification for homogenous skies. Only three cloudiness sky categories have been established: clear, partial and cloudy. This homogenization has required scaling down the number of intervals of some MIs, keeping the class limits stated in the original works. An exhaustive analysis of all the possibilities has been made to merge the intervals, always respecting the original limits and selecting the option that optimized the

result of each MI. Confusion matrices have been used as a tool for benchmarking the MIs under study and four different quality metrics have been calculated, to determine the performance of each MI as a sky classifier.

Although the classification of the skies was limited to three categories, none of the MIs under analysis distinguished the large global number of clear skies that the CIE determined in the city of Burgos; the $FK$ index was the closest to this classification, in terms of frequency of occurrence of each of the sky categories. The rest of the indices, in general, indicated partially or fully covered skies as the predominant ones in the area under study. The monthly distribution over the year-long analysis of the types of sky calculated by the different indices also differed from that offered by the CIE standard, which predicted predominantly clear skies in summer.

The metrics for the cloudiness classification have shown that the performance of the MIs may, at best, be considered as "fair". Classification accuracy reached values of over 70% for cloudy skies. However, this metric can lead to misunderstandings, as it accounts for the correct classification of true positives and true negatives (see the Accuracy Paradox in section 4.1). As Table 8 shows, the best MI for the sky classification was $\epsilon$ followed closely by $FK$ and $FP$ indices. $k_t$ and $F$ showed the worst results for the classification of cloudiness of the skies into three categories. No correlation was observed between the mathematical complexity of the MIs or the number of input parameters and the classification result of the MIs. As can be observed, the original definition of the intervals is one of the main factors that justifies the difference in the performance of the MIs.

The experimental data of the present study were taken from a different location other than that used in the original studies for the definition of each MI. However, if subtle language



Journal Pre-proofs

differences are neglected, the definition of "clear" or "cloudy" sky should be independent of the location. Attending to the metrics, regardless of the accuracy, the results were quite similar. The average accuracy almost reached 70 % in $FK$, and was above 50 % in all the MIs. However, this result is overoptimistic because of the accuracy paradox. The average value was below 33% for the Jaccard Index and below 25% for Cohen's Kappa and Matthews correlation coefficient. In short, the capability of the MIs at classifying sky cloudiness can be considered "fair". As demonstrated, their isolated use would not be advisable for the classification of sky cloudiness.

Several possible lines may be followed to improve on the results that have been presented here. Future work must address different combinations of these indices linked with other variables related to climatic and atmospheric conditions, which were not considered in the calculation of the indices. Another option would be to modify the original interval limits of each sky type to obtain the best classification results. The use of machine learning techniques should also be considered, to obtain models that will help to classify day cloudiness in real time. The final objective must be to reuse the big data available from meteorological stations, so as to classify cloudiness in the most reliable manner without the use of expensive sky scanner devices.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support provided by the Regional Government of Castilla y León under the "Support Program for Recognized Research Groups of Public Universities of Castilla y León" (ORDEN EDU/667/2019) and the Spanish Ministry of Science, Innovation & Universities under the I+D+i state programme "Challenges Research Projects" (Ref. RTI2018-098900-B-I00). David González Peña and Diego



Journal Pre-proofs

Granados López express their thanks to Junta de Castilla-León for economic support (PIRTU Program, ORDEN EDU/301/2015 and ORDEN EDU/556/2019, respectively). Also, Andrés Suárez García, is grateful to the International Excellence Triangular-E3 for the financed stay under the programme managed centrally by the Ministry of Education, Culture & Sports (BOE 295, 10th December 2015).

## REFERENCES

Alshaibani, K., 2011. Finding frequency distributions of CIE Standard General Skies from sky illuminance or irradiance. Lighting Research and Technology 43(4), 487-495.

Baharuddin, Lau, S.S.Y., Rahim, R., 2010. Daylight availability in Hong Kong: classification into three sky conditions. Architectural Science Review 53(4), 396-407.

Batlles, F., Barbero, J., López, G., Pérez, M., Rodrigo, F., Rubio, M., 1998. 'Fundamentos de la radiación solar y aspectos climatológicos de Almerıa, 1990–1996. Servicio de Publicaciones de la Universidad de Almerıa in Spanish.

Batlles, F.J., Olmo, F.J., Tovar, J., Alados-Arboledas, L., 2000. Comparison of Cloudless Sky Parameterizations of Solar Irradianceat Various Spanish Midlatitude Locations. Theoretical and Applied Climatology 66(1), 81-93.

Brunger, A.P., Hooper, F.C., 1993. Anisotropic sky radiance model based on narrow field of view measurements of shortwave radiance. Solar Energy 51(1), 53-64.

Chicco, D., 2017. Ten quick tips for machine learning in computational biology. BioData mining 10(1), 35.



Journal Pre-proofs

Djafer, D., Irbah, A., Zaiani, M., 2017. Identification of clear days from solar irradiance observations using a new method based on the wavelet transform. Renewable Energy 101, 347-355.

Erbs, D.G., Klein, S.A., Duffie, J.A., 1982. Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. Solar Energy 28(4), 293-302.

Escobedo, J.F., Gomes, E.N., Oliveira, A.P., Soares, J., 2009. Modeling hourly and daily fractions of UV, PAR and NIR to global solar radiation under various sky conditions at Botucatu, Brazil. Applied Energy 86(3), 299-309.

Gueymard, C.A., 2018. A reevaluation of the solar constant based on a 42-year total solar irradiance time series and a reconciliation of spaceborne observations. Solar Energy 168, 2-9.

Gueymard, C.A., Bright, J.M., Lingfors, D., Habte, A., Sengupta, M., 2019. A posteriori clear-sky identification methods in solar irradiance time series: Review and preliminary validation using sky imagers. Renewable and Sustainable Energy Reviews, 412-427.

Gueymard, C.A., Ruiz-Arias, J.A., 2016. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Solar Energy 128, 1-30.

Igawa, N., Koga, Y., Matsuzawa, T., Nakamura, H., 2004. Models of sky radiance distribution and sky luminance distribution. Solar Energy 77(2), 137-157.

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

Jaccard, P., 1912. The distribution of the flora in the alpine zone. 1. New phytologist 11(2), 37-50.

Jacovides, C., Tymvios, F., Assimakopoulos, V., Kaltsounides, N., 2007. The dependence of global and diffuse PAR radiation components on sky conditions at Athens, Greece. Agricultural and Forest Meteorology 143(3-4), 277-287.

Kambezidis, H., Katevatis, E., Petrakis, M., Lykoudis, S., Asimakopoulos, D., 1998. Estimation of the Linke and Unsworth–Monteith turbidity factors in the visible spectrum: application for Athens, Greece. Solar Energy 62(1), 39-50.

Kambezidis, H.D., 2018. The solar radiation climate of Athens: Variations and tendencies in the period 1992–2017, the brightening era. Solar Energy 173, 328-347.

Kasten, F., 1993. Discussion on the relative optical air mass. Lighting Research & Technology 25(3), 129-130.

Khorasanizadeh, H., Mohammadi, K., 2016. Diffuse solar radiation on a horizontal surface: Reviewing and categorizing the empirical models. Renewable and Sustainable Energy Reviews 53, 338-362.

Khorasanizadeh, H., Mohammadi, K., Goudarzi, N., 2016. Prediction of horizontal diffuse solar radiation using clearness index based empirical models; A case study. International Journal of Hydrogen Energy 41(47), 21888-21898.

Kittler, R., Perez, R., Darula, S., 1997. A new generation of sky standards. In: Proceedings of the Eighth European lighting conference , Amsterdam, pp. 359–373., 359-373.

Klucher, T.M., 1979. Evaluation of models to predict insolation on tilted surfaces. Solar Energy 23(2), 111-114.

Kong, H.J., Kim, J.T., 2013. A Classification of real sky conditions for Yongin, Korea, Sustainability in Energy and Buildings. Springer, pp. 1025-1032.

Kundel, H.L., Polansky, M., 2003. Measurement of observer agreement. Radiology 228(2), 303-308.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. biometrics, 159-174.

Li, D.H.W., 2010. A review of daylight illuminance determinations and energy implications. Applied Energy 87(7), 2109-2118.

Li, D.H.W., Chau, N.T.C., Wan, K.K.W., 2013. Predicting daylight illuminance and solar irradiance on vertical surfaces based on classified standard skies. Energy 53, 252-258.

Li, D.H.W., Chau, T.C., Wan, K.K.W., 2014. A review of the CIE general sky classification approaches. Renewable and Sustainable Energy Reviews 31, 563-574.

Li, D.H.W., Cheung, G.H.W., 2006. Average daylight factor for the 15 CIE standard skies. Lighting Research and Technology 38(2), 137-152.

Li, D.H.W., Lam, J.C., 2001. An analysis of climatic parameters and sky condition classification. Build. Environ. 36(4), 435-445.

Li, D.H.W., Lam, T.N.T., Cheung, K.L., Tang, H.L., 2008. An analysis of luminous efficacies under the CIE standard skies. Renewable Energy 33(11), 2357-2365.

Li, D.H.W., Tang, H.L., Wong, S.L., Tsang, E.K.W., Cheung, G.H.W., Lam, T.N.T., 2007. Skies classification using artificial neural networks (ANN) techniques, 6th International Conference on Indoor Air Quality, Ventilation and Energy Conservation in Buildings: Sustainable Built Environment, IAQVEC 2007. Sendai, pp. 61-68.

Littlefair, P.J., 1994. A comparison of sky luminance models with measured data from Garston, United Kingdom. Solar Energy 53(4), 315-322.

Littlefair, P.J., 1994. The luminance distributions of clear and quasi-clear skies, Proceedings of the CIBSE National Lighting Conference, Cambridge, UK. pp. 267-283.

Lou, S., Li, D.H.W., Lam, J.C., 2017. CIE Standard Sky classification by accessible climatic indices. Renewable Energy 113, 347-356.

Markou, M.T., Kambezidis, H.D., Bartzokas, A., Katsoulis, B.D., Muneer, T., 2005. Sky type classification in Central England during winter. Energy 30(9 SPEC. ISS.), 1667-1674.

Markou, M.T., Kambezidis, H.D., Katsoulis, B.D., Muneer, T., Bartzokas, A., 2004. Sky type classification in South England during the winter period. Building Research Journal 52(1), 19-30.

Matthews, B.W., 1975. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochimica et Biophysica Acta (BBA)-Protein Structure 405(2), 442-451.

Mellit, A., Kalogirou, S.A., Shaari, S., Salhi, H., Hadj Arab, A., 2008. Methodology for predicting sequences of mean monthly clearness index and daily solar radiation data in remote areas: Application for sizing a stand-alone PV system. Renewable Energy 33(7), 1570-1590.

Muneer, T., 2007. Solar radiation and daylight models. Routledge.

Muneer, T., Gueymard, C., Kambezidis, H.D., 2004. Solar Radiation and Daylight Models (Second Edition). Butterworth Heinemann.

Muneer, T., Younes, S., Munawwar, S., 2007. Discourses on solar radiation modeling. Renewable and Sustainable Energy Reviews 11(4), 551-602.

Perez, R., Ineichen, P., Seals, R., Michalsky, J., Stewart, R., 1990. Modeling daylight availability and irradiance components from direct and global irradiance. Solar Energy 44(5), 271-289.

Perez, R., Seals, R., Ineichen, P., Stewart, R., Menicucci, D., 1987. A new simplified version of the perez diffuse irradiance model for tilted surfaces. Solar Energy 39(3), 221-231.

Perraudeau, M., Chauvel, P., 1986. One year's measurements of luminous climate in Nantes: Proceedings of the International Daylighting Conference. Long Beach, CA.

Perveen, G., Rizwan, M., Goel, N., 2018. Intelligent model for solar energy forecasting and its implementation for solar photovoltaic applications. Journal of Renewable and Sustainable Energy 10(6), 063702.

Powers, D.M., 2011. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.

Rahim, R., Baharuddin, Mulyadi, R., 2004. Classification of daylight and radiation data into three sky conditions by cloud ratio and sunshine duration. Energy and Buildings 36(7), 660-666.

Ruiz-Arias, J.A., Gueymard, C.A., 2018. Worldwide inter-comparison of clear-sky solar radiation models: Consensus-based review of direct and global irradiance components simulated at the earth surface. Solar Energy 168, 10-29.

Sim, J., Wright, C.C., 2005. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Physical therapy 85(3), 257-268.

Suárez-García, A., Granados-López, D., González-Peña, D., Díez-Mediavilla, M., Alonso-Tristán, C., 2018. Seasonal caracterization of CIE standard sky types above Burgos, northwestern Spain. Solar Energy 169, 24-33.

Tapakis, R., Michaelides, S., Charalambides, A.G., 2016. Computations of diffuse fraction of global irradiance: Part 1 – Analytical modelling. Solar Energy 139, 711-722.

Torres, J.L., de Blas, M., García, A., Gracia, A., de Francisco, A., 2010a. Sky luminance distribution in Pamplona (Spain) during the summer period. J. Atmos. Sol.-Terr. Phys. 72(5-6), 382-388.

Journal Pre-proofs

Torres, J.L., de Blas, M., García, A., Gracia, A., de Francisco, A., 2010b. Sky luminance distribution in the North of Iberian Peninsula during winter. J. Atmos. Sol.-Terr. Phys. 72(16), 1147-1154.

Tregenza, P.R., 2004. Analysing sky luminance scans to obtain frequency distributions of CIE Standard General Skies. Lighting Research & Technology 36(4), 271-279.

Tregenza, P.R., Perez, R., Michalsky, J., Seals, R., Molineaux, B., Ineichen, P., 1994. Guide to recommended practice of daylight measurement. Cimmission Internationale de l'éclairage, Vienna.

Uetani, Y., Aydinli, S., Joukoff, A., Kendrick, J., Kittler, R., Koga, Y., 2003. Spatial distribution of daylight-CIE standard general sky. Vienna, Austria.

Umemiya, N., Kanou, T., 2008. Classification of Sky Conditions by the Ranges of Insolation Indices Considering CIE Standard for General Sky. Journal of Light & Visual Environment 32(1), 14-19.

Wang, L., Gong, W., Ma, Y., Hu, B., Wang, W., Zhang, M., 2013. Analysis of ultraviolet radiation in Central China from observation and estimation. Energy 59, 764-774.

WMO, 2008. (Updated 2010). Guide to Meteorological Instruments and Methods of Observation. $N^0$ 8.

WMO, 2010. Technical Regulations. Volume I: General Meteorological Standards and Recommended Practise. http://library.wmo.int/pmb_ged/wmo_49-vl-2012_en.pdf.

Wong, S.L., Wan, K.K.W., Li, D.H.W., Lam, J.C., 2012. Generation of typical weather years with identified standard skies for Hong Kong. Build. Environ. 56, 321-328.

**Highlights**

- Skies classification is needed to improve building's energy efficiency and solar energy applications
- 10 Meteorological indices (MI) for real-time sky classification were reviewed
- The baseline reference was the Standard CIE for homogenous skies.
- The confusion matrices have been used as a tool for benchmarking
- Four quality metrics determined the performance of each MI as sky classifier.

Journal Pre-proofs