



Harry D. Kambezidis <sup>1,\*</sup>, Styliani I. Kampezidou <sup>2</sup> and Dimitra Kampezidou <sup>3</sup>

- <sup>1</sup> Atmospheric Research Team, Institute of Environmental Research and Sustainable Development, National Observatory of Athens, GR-11810 Athens, Greece
- <sup>2</sup> Aerospace Systems Design Laboratory, Georgia Institute of Technology, 275 Ferst Dr NW, Atlanta, GA 30313, USA; skampezidou@gatech.edu
- <sup>3</sup> Department of Ecology and Environmental Science, University of Umeå, S-90187 Umeå, Sweden; demi\_ath@hotmail.com
- \* Correspondence: harry@noa.gr

**Abstract:** A mathematical method for accurately determining the upper and lower diffuse-fraction  $(k_d)$  limits that divide the sky into clear, intermediate, and overcast is developed. Fourteen sites around the world are selected for demonstrating the methodology. The upper and lower  $k_d$  values for these sites are determined from scatter plots of direct-normal solar radiation vs.  $k_d$  pairs over the typical meteorological year of each site. They vary between 0.73 and 0.80 for the upper and between 0.24 and 0.27 for the lower  $k_d$  limits. Plots of sunshine duration (SSD) vs.  $k_d$  are prepared for 12 of the 14 sites. These plots show a decreasing trend in SSD with increasing values of  $k_d$ , as anticipated. According to local climatology, the number of the SSD values in each sky-condition classification varies from site-to-site.

Keywords: mathematical determination; diffuse fraction; sunshine duration

# 1. Introduction

The estimation of solar radiation on the surface of the Earth on either horizontal or tilted planes has long been recognised an important parameter in a variety of fields such as atmospheric environment, e.g., [1], terrestrial ecosystems, e.g., [2], terrestrial climate, e.g., [3], and solar energy applications, e.g., [4]. Because of the scarcity of solar radiation platforms that may provide measurements of the total (or global) solar radiation to the above applications, solar models started appearing in the international literature as early as in 1950's. The total solar radiation is the sum of two components: the diffuse solar radiation and the direct (or beam) one. The latter can accurately be estimated via explicit atmospheric transmission relationships. It is, therefore, the diffuse component that must be simulated satisfactorily; however, this solar component is very difficult to accurately be estimated due to numerous scattering processes of the solar rays that take place in the atmosphere under a clear sky and moreover in the presence of clouds on a cloudy day. This fact has triggered a lot of research for the development of various models that pursue to estimate as accurately as possible the diffuse solar component. Such models are numerous in the international literature, e.g., [5-7]. Most of them use the diffuse fraction ( $k_d$ ), which is the ratio of diffuse-to-global solar radiation as function of other parameter (or parameters) that can easily be calculated, be measured or are known; such parameters are the clearness index (kt), which is the ratio of the global-to-the-extraterrestrial solar radiation, or the geographical latitude ( $\varphi$ ), e.g., [8]. Even though k<sub>t</sub> or k<sub>d</sub> can be computed in one way or another, it is very important for solar energy engineers to know their upper and lower limits that divide the sky conditions into clear, intermediate and overcast. Several such limits have been provided by many researchers. All of them are based on solar measurements or solar modelling, e.g., [9,10]. This is, therefore, the reason that these limits are considered empirical as they are not based on a standard methodology.



Citation: Kambezidis, H.D.; Kampezidou, S.I.; Kampezidou, D. Mathematical Determination of the Upper and Lower Limits of the Diffuse Fraction at Any Site. *Appl. Sci.* 2021, *11*, 8654. https://doi.org/ 10.3390/app11188654

Academic Editor: Yosoon Choi

Received: 12 August 2021 Accepted: 14 September 2021 Published: 17 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). The above gap is bridged in the present study. In other words, this work develops a mathematical methodology for determining the upper and lower limits of  $k_d$ . The methodology proves to be universal, but the values of the upper and lower  $k_d$  limits are site-dependent. Such a work is presented for the first time in the international literature.

#### 2. Materials and Methods

This Section gives a full account of the new (mathematical) method that accurately determines the upper and lower limits of  $k_d$  at a site.

To apply the methodology, a typical meteorological year (TMY) must preferably be used instead of a long period of measurements or model simulations. This peculiarity is based on the fact that a TMY refers to all representative situations of the included parameters that occur within a year, while a long-term series of measurements or model simulations contains extreme values of the included parameters; the second, therefore, option may result in a broader dispersion of the calculated k<sub>d</sub> values. On the other hand, TMYs are nowadays used more often in many climatological studies, e.g., [11–13].

The selected solar radiation parameter to vary with  $k_d$  is the direct horizontal solar radiation ( $H_b$ ) because of much less dispersion of the  $H_b$ - $k_d$  paired values in comparison with the cases of the functions  $H_g$  (global horizontal solar radiation) vs.  $k_d$  or  $H_d$  (diffuse horizontal solar radiation) vs.  $k_d$ . Even in the case of the direct horizontal solar radiation, the scatter plot of  $H_b$  vs.  $k_d$  is found to be greater than that of the relationship  $H_{bn}$  (direct-normal solar radiation) vs.  $k_d$ . Figure 1 shows this difference; here hourly mean  $H_b$ ,  $H_{bn}$  and  $k_d$  values have been plotted for the typical month of January of the TMY-BOU (Boulder, CO, USA) derived from the PV-GIS platform [14,15] (see below for description). Therefore, the  $H_{bn}$  solar component is considered in the rest of the analysis.



**Figure 1.** Variation of (**a**)  $H_b$  and (**b**)  $H_{bn}$  as function of  $k_d$  for January of the PV-GIS TMY-BOU. All data points are hourly values. Dispersion is greater in (**a**).

The mathematical determination of the upper and lower limits of  $k_d$  in a diagram as that of Figure 1b is fully described here. The first step is to find the best-fit curve to the  $H_{bn}$ - $k_d$  data points. A second-order polynomial of the form  $H_{bn} = a k_d^2 + b k_d + c$  was used, where a, b, c are the polynomial coefficients. This type of best-fit curve was selected as the results obtained showed that it is sufficient. Figure 2 shows the best-fit black curve to the scatter plot of Figure 1b, which has the expression  $H_{bn} = 902.30 k_d^2 - 2087.44 k_d + 1188.58$ ,  $R^2 = 0.996$ . As WMO recommends that the sunshine duration is measured when  $H_{bn} > 120 Wm^{-2}$ , the green line represents this threshold. This straight line crosses the best-fit curve at a certain point ( $k_{du}$ , $H_{bnWMO} = 120$ ). Since the y-axis value is known (120 Wm<sup>-2</sup>), solving the above second-order equation for  $k_d$ , two roots are obtained and



the lower-than-1 is considered because  $k_d$  cannot go higher than 1 by definition; in this case, the accepted solution is  $k_{du} = 0.76$ .

**Figure 2.** Mathematical determination of the upper ( $k_{du}$ , red solid line) and lower ( $k_{dl}$ , blue solid line) limits of  $k_d$  that divide the skies into clear, intermediate, and overcast. The bold black solid line is the best-fit curve to the  $H_{bn}$ - $k_d$  data points that correspond to their hourly values; in this case the typical meteorological month of January, which belongs to the PV-GIS TMY-BOU, has been considered. The line BC is tangent to the best-fit line at C (0.76,120). M is the crossing point of the medians AD and BE. The MG line is normal to AC. From the Euclidian geometry AG = AC/3 in the rectangle triangle ABC; therefore,  $k_{dl} = k_{du}/3 = 0.76/3 \approx 0.25$ .

The next step is to find the equation of the straight line that is tangent to the best-fit curve at the point C (0.76,120); this point is the crossing point between the best-fit curve and the WMO-defined threshold for SSD measurements. This line expresses the first derivative of the best-fit curve at the point C; in other words, it represents the slope of the best-fit line at C. This is done by taking the first derivative of the H<sub>bn</sub> expression  $(H'_{bn})$  in respect to k<sub>d</sub> and replacing k<sub>d</sub> with 0.76. It is found that  $H'_{bn} = -715.94$  Wm<sup>-2</sup>. Then, H<sub>bn</sub> - 120 = -715.94 (k<sub>d</sub> - 0.76) from which the expression for the tangent line is H<sub>bn</sub> = -715.94 k<sub>d</sub> + 664.12. This straight line is the BC in Figure 2. A rectangle triangle ABC is then formed.

The third step is to draw the two medians from the vertexes A and B of the triangle ABC to their opposite sides, i.e., the lines AD and BE, respectively. The two lines are crossed at M; if a straight line parallel to the *y*-axis is drawn through M, i.e., the blue line MG, this crosses the x-axis at a value, which is the lower limit of  $k_d$ . From the Euclidian geometry, it is known that AG = AC/3. This means that if the upper  $k_d$  value at the point C

is obtained ( $k_{du}$ ), then very easily the lower  $k_d$  value at the point G ( $k_{dl}$ ) can be calculated:  $k_{dl} = k_{du}/3$ . This way, the estimated  $k_{du}$  and  $k_{dl}$  values divide the sky conditions in clear skies, intermediate skies, and overcast skies at a site as shown in Figure 2.

The above methodology was applied to 14 sites around the world. The selection has been based on the following criteria: (i) different environmental characteristics, (ii) spread across the continents, and (iii) availability of TMY. Table 1 deploys the selected sites in alphabetical order together with their geographical coordinates and environmental description. Sites that do not meet criterion (iii) may not be appropriate for applying this methodology. Figure 3 gives the distribution of the 14 sites over the world.

**Table 1.** Selected sites for the application of the mathematical determination of the  $k_{du}$  and  $k_{dl}$  limits.  $\varphi$  and  $\lambda$  are the geographical latitudes and longitudes of the sites, respectively;  $\varphi$  is given in the northern (N) or the southern (S) hemisphere, and  $\lambda$  east (E) or west (W) of the Greenwich meridian. The values of both  $\varphi$  and  $\lambda$  have been rounded to the second decimal digit. In column 9, I denotes rural and II denotes urban environment. The period of measurements is given in the last column. The selection of the sites and their solar radiation data were based on the BSRN (Baseline Surface Radiation Network), except for Athens (ATH); in this case, data from the Actinometric Station of the National Observatory of Athens not belonging to BSRN were used. The abbreviations of the sites (except for Athens) are those provided by the BSRN typology. Description of the BSRN operation is found in [16].

#	Site (Abbreviation)	Location	φ (deg)	λ (deg)	z (m asl)	Surface Type	Topography	I/II	Period
1	Athens (ATH)	Greece	37.97 N	23.72 E	107	shrubs, trees	hilly	Π	1953–present
2	Boulder (BOU)	USA	40.05 N	105.01 W	1577	grass	flat	Ι	1992-2016
3	Carpentras (CAR)	France	44.08 N	5.06 E	100	cultivated	hilly	Ι	1996–2018
4	De Aar (DAA)	South Africa	30.67 S	23.99 E	1287	sand	flat	Ι	2000-present
5	Gandhinagar (GAN)	India	23.11 N	72.63 E	65	shrubs	flat	II	2014-present
6	Huancayo Observatory (OHY)	Peru	12.05 S	75.32 W	3314	grass	mountain valley	Ι	2017–present
7	Ilorin (ILO)	Nigeria	8.53 N	4.57 E	350	shrubs	flat	Ι	1992-2005
8	Kishinev (KIS)	Moldova	47.00 N	28.82 E	205	grass	flat	II	
9	Lerwick (LER)	UK	60.14 N	1.18 W	80	grass	hilly	Ι	2001-present
10	Lindenberg (LIN)	Germany	52.21 N	14.12 E	125	cultivated	hilly	Ι	1994-present
11	Payerne (PAY)	Switzerland	46.82 N	6.94 E	491	cultivated	hilly	Ι	1992–present
12	Regina (REG)	Canada	50.21 N	104.71 W	578	cultivated	flat	Ι	1995-2011
13	Sonnblick (SON)	Austria	47.05 N	12.96 E	3109	rocks	mountain top	Ι	2013-present
14	Solar Village (SOV)	Saudi Arabia	24.91 N	46.41 E	650	desert, sand	flat	Ι	1998–2002

For the sake of the present analysis, TMYs for the selected sites have been downloaded from the PV-GIS platform [17,18]. These typical meteorological years have been derived from a combination of satellite and re-analysis data in the period 2005–2014. Each TMY contains the following parameters: date, time (h UTC), dry-bulb temperature (in degC), relative humidity (in %), global horizontal solar irradiance (H<sub>g</sub>, in Wm<sup>-2</sup>), direct-normal solar irradiance (H<sub>bn</sub>, in Wm<sup>-2</sup>), diffuse horizontal solar irradiance (H<sub>d</sub>, in Wm<sup>-2</sup>), infrared horizontal solar irradiance (in Wm<sup>-2</sup>), wind speed (in ms<sup>-1</sup>), wind direction (in deg), barometric pressure (P<sub>z</sub>, in Pa) at the site's altitude (z, in m), and a control parameter. The meteorological and solar radiation values are hourly averages. Therefore, from the 14 downloaded TMYs new files were created to contain the necessary parameters: date, time (h UTC), H<sub>g</sub>, H<sub>bn</sub>, and H<sub>d</sub>. Conversion of time from UTC (universal time constant) into LST (local standard time) at each site was applied to all files. Then, an own-developed

routine, named XRONOS.bas in Basic programming language (xronos means time in Greek with x being spelled as ch), was used to derive the solar altitude ( $\gamma$ , in deg) at 30 min past the hour LST. For example, the value of any parameter in the data base at 10.00 LST denotes its average value between 09.01 LST and 10.00 LST. The XRONOS.bas algorithm calculated  $\gamma$  at 09.30 LST and this value was assigned to all parameters at 10.00 LST. Finally, the hourly values of k<sub>d</sub> were computed from concurrent values of H<sub>g</sub> and H<sub>d</sub>, i.e., k<sub>d</sub> = H<sub>d</sub>/H<sub>g</sub>. In cases that H<sub>d</sub> and/or H<sub>g</sub> contained zero values, no value was assigned to k<sub>d</sub>. Correspondingly, no value was assigned to H<sub>bn</sub>, too. In plotting the (x, y) data pairs, i.e., (k<sub>d</sub>, H<sub>bn</sub>), the hourly values of these two parameters were considered for all  $\gamma > 5$  deg. This way, 14 plots of H<sub>bn</sub>-k<sub>d</sub> were derived to apply the methodology and derive the individual k<sub>dl</sub> and k<sub>du</sub> limits for the sites. The plots are deployed and discussed in Section 3.1.



**Figure 3.** Location of the 14 selected sites (green circles). The numbers in the circles correspond to the numbers in column 1 of Table 1.

#### 3. Results

### 3.1. Determination of the $k_d$ Limits

This section presents the results from the application of the developed methodology to the 14 sites. Figures 4–10 show the graphical estimation of the  $k_{du}$  and  $k_{dl}$  values for all 14 sites. The  $H_{bn}$ - $k_d$  values refer to the PV-GIS TMYs of the 14 sites.



Figure 4. Determination of the  $k_{du}$  (red line) and  $k_{dl}$  (blue line) values for the (a) ATH and (b) BOU sites.



Figure 5. Determination of the  $k_{du}$  (red line) and  $k_{dl}$  (blue line) values for the (a) CAR and (b) DAA sites.



Figure 6. Determination of the  $k_{du}$  (red line) and  $k_{dl}$  (blue line) values for the (a) GAN and (b) OHY sites.



Figure 7. Determination of the  $k_{du}$  (red line) and  $k_{dl}$  (blue line) values for the (a) ILO and (b) KIS sites.



Figure 8. Determination of the k<sub>du</sub> (red line) and k<sub>dl</sub> (blue line) values for the (a) LER and (b) LIN sites.



Figure 9. Determination of the  $k_{du}$  (red line) and  $k_{dl}$  (blue line) values for the (a) PAY and (b) REG sites.

From the above Figures it is seen that different spreads in the scatter plots exist. This variance is least at the ATH, GAN, KIS, LIN, REG, and SOV sites and greatest at the BOU, LER, and POY ones. The broader variance in the  $H_{bn}$ - $k_d$  data pairs is possibly due to the changing weather conditions at these sites (fast moving clouds, prevailing heavy/light cloudiness most of the time), which affect the  $H_d$  values in the estimation of  $k_d$ . On the other hand, missing solar radiation values occur in the case of the OHY site; this results in fewer data points than at the rest of the sites. Table 2 gives the expressions of the best-fit

curves, their R<sup>2</sup> that determine the variance, and the estimation of the  $k_{du}$  and  $k_{dl}$  values. Since these values are similar for all sites examined, one could adopt universal values that are the average of the distinct ones; in this case a universal  $k_{du}$  would be 0.78 and a universal  $k_{dl}$  0.26. Nevertheless, these averages are only indicative, because they refer to a small number of sites over the globe.



Figure 10. Determination of the k<sub>du</sub> (red line) and k<sub>dl</sub> (blue line) values for the (a) SON and (b) SOV sites.

**Table 2.** Second-order polynomial expression for the function  $H_{bn}$  vs.  $k_d$  with  $R^2$ ,  $k_{du}$ , and  $k_{dl}$  values. All values have been rounded to the second decimal digit. The year in parenthesis in column 1 denotes the period of the downloaded solar radiation measurements from the BSRN sites. The criterion for the year selection was to be outside the PV-GIS TMY period of 2004–2015, if possible.

Site (Year)	Equation of the Best-Fit Curve	<b>R</b> <sup>2</sup>	k <sub>du</sub>	k <sub>dl</sub>
ATH (2000)	$H_{bn} = 750.35 k_d^2 - 1911.13 k_d + 1160.42$	0.99	0.79	0.26
BOU (1999)	$H_{bn} = 764.84 k_d^2 - 1906.61 k_d + 1145.82$	0.99	0.79	0.26
CAR (2018)	$H_{bn} = 868.13 k_d^2 - 2035.91 k_d + 1172.93$	0.99	0.77	0.26
DAA (2017)	$H_{bn} = 829.67 k_d^2 - 2009.26 k_d + 1182.98$	0.99	0.78	0.26
GAN (2015)	$H_{bn} = 923.82 k_d^2 - 2150.63 k_d + 1230.29$	$\approx 1.00$	0.77	0.26
OHY (2018)	$H_{bn} = 715.87 k_d^2 - 1879.32 k_d + 1165.29$	0.99	0.80	0.27
ILO (1999)	$H_{bn} = 796.31 k_d^2 - 2007.88 k_d + 1208.37$	0.99	0.79	0.26
KIS (2020)	$H_{bn} = 781.01 k_d^2 - 1935.32 k_d + 1155.12$	$\approx 1.00$	0.78	0.26
LER (2003)	$H_{bn} = 930.30 k_d^2 - 2020.92 k_d + 1097.60$	0.94	0.73	0.24
LIN (2018)	$H_{bn} = 779.37 k_d^2 - 1929.96 k_d + 1151.03$	0.99	0.78	0.26
PAY (2013)	$H_{bn} = 715.85 k_d^2 - 1856.32 k_d + 1140.65$	0.99	0.79	0.26
REG (2003)	$H_{bn} = 779.37 k_d^2 - 1929.96 k_d + 1151.03$	$\approx 1.00$	0.78	0.26
SOV (2002)	$H_{bn} = 919.21 k_d^2 - 2192.18 k_d + 1270.02$	0.99	0.78	0.26
SON (2018)	$H_{bn} = 704.02 k_d^2 - 1833.67 k_d + 1129.45$	0.99	0.79	0.26

### 3.2. Evaluation of the Methodology

To show the applicability of the method developed in Section 2, Figures 11–17 show the distribution of the (k<sub>d</sub>,SSD) pairs during the years shown in parentheses in column 1, Table 2. The SSD values were derived from the BSRN data, meeting the  $H_{bn} > 120 \text{ Wm}^{-2}$  WMO criterion; moreover, the SSD values from the ATH station are real 1-min measurements with the aid of an EKO MS-90 direct-normal-irradiance sensor. All (k<sub>d</sub>,SSD) data values are hourly ones in the corresponding year of the site (column 1, Table 2). Therefore, SSDs for  $k_{du} < k_d \leq 1$  correspond to overcast-, for  $k_{dl} < k_d \leq k_{du}$  to intermediate-, and for  $0 \leq k_d \leq k_{dl}$  to clear-sky conditions.



Figure 11. Determination of clear, intermediate, and overcast skies for the (a) ATH and (b) BOU sites.



Figure 12. Determination of clear, intermediate, and overcast skies for the (a) CAR and (b) DAA sites.



Figure 13. Determination of clear, intermediate, and overcast skies for the (a) GAN and (b) OHY sites.



Figure 14. Determination of clear, intermediate, and overcast skies for the (a) ILO and (b) KIS sites.



Figure 15. Determination of clear, intermediate, and overcast skies for the (a) LER and (b) LIN sites.



Figure 16. Determination of clear, intermediate, and overcast skies for the (a) PAY and (b) REG sites.



Figure 17. Determination of the k<sub>du</sub> (red line) and k<sub>dl</sub> (blue line) values for the (a) SON and (b) SOV sites.

The shape of the scattered data points in the above diagrams is the expected one, i.e., decreasing SSD values with increasing  $k_d$ . Nevertheless, this self-evident pattern is not repeated at the GAN and ILO sites; this may be due to the solar radiation data quality, which, in the case of ILO, may be unacceptable. Therefore, these two sites are excluded from further analysis. Moreover, the site of OHY presents some low SSD values for clear skies (i.e., for  $0 \le k_d \le 0.26$ ), but the number of such values is small, and, therefore, this site is included in the analysis that follows.

In each of the remaining 14 sites, the number of the ( $k_d$ ,SSD) hourly pairs is maximum (8760 or 8784 for a non-leap or a leap year, respectively); these numbers include night-time values, which were not taken into account. The analysis that follows concerns the number of daytime hourly pairs that represent the three sky condition statuses (overcast, intermediate, clear). Table 3 shows the results.

Site (Year)	$k_{du} < k_d \leq 1$ Overcast Skies	$\label{eq:kdl} \begin{aligned} \mathbf{k}_{dl} < \mathbf{k}_{d} \leq \mathbf{k}_{du} \\ \text{Intermediate Skies} \end{aligned}$	$0 \leq k_d \leq k_{dl}$ Clear Skies	$\begin{array}{l} 0 \leq k_d \leq 1 \\ \text{All Skies} \end{array}$
ATH (2000)	919 (20.9%)	1879 (42.7%)	1602 (36.4%)	4400 (100%)
BOU (1999)	1477 (31.7%)	1429 (30.6%)	1758 (37.7%)	4664 (100%)
CAR (2018)	1799 (37.1%)	1556 (32.1%)	1496 (30.8%)	4851 (100%)
DAA (2017)	733 (15.5%)	1288 (27.1%)	2715 (57.4%)	4736 (100%)
GAN (2015)	-	-	-	_
OHY (2018)	908 (24.3%)	957 (25.5%)	1881 (50.2%)	3746 (100%)
ILO (1999)	-	-	-	-
KIS (2020)	1524 (36.2%)	1525 (36.2%)	1161 (27.6%)	4210 (100%)
LER (2003)	3476 (68.8%)	1225 (24.2%)	354 (7.0%)	5055 (100%)
LIN (2018)	1652 (41.5%)	1515 (38.1%)	811 (20.4%)	3978 (100%)
PAY (2013)	2100 (43.5%)	1743 (36.1%)	986 (20.4%)	4829 (100%)
REG (2003)	2049 (37.6%)	1932 (35.5%)	1467 (26.9%)	5448 (100%)
SOV (2002)	872 (18.3%)	2199 (46.0%)	1704 (35.7%)	4775 (100%)
SON (2018)	2624 (60.2%)	952 (21.8%)	781 (18.0%)	4357 (100%)

**Table 3.** Number of (daytime) hourly SSD values at each site within the year indicated for the various  $k_d$  conditions; percentages of the individual cases to the total are given in parentheses.

The differentiation in the total daytime values (column 5, Table 3) is due to the geographical location of the sites, which dictates the daylight pattern, i.e., the daylength. On the other hand, the variation in the SSD values under the three sky conditions is much influenced by the prevailing climatology at each site.

#### 4. Conclusions and Discussion

A mathematical formulation for the determination of the upper and lower diffusefraction limits was developed for the first time worldwide. The methodology was based on existing or prepared TMY for the site into consideration. Such a TMY can be downloaded free-of-charge from the PV-GIS website. The TMY must include hourly values of the global horizontal, diffuse horizontal, and direct-normal solar irradiances. Hourly values of the diffuse fraction can then be derived.

The method consists of the following analytical steps.

Step 1. Prepare a scatter plot of the hourly direct-normal solar radiation vs. diffusefraction values that are included in the TMY for the site into consideration.

Step 2. Estimate the best-fit second-order polynomial to the data points of step 1.

Step 3. Find the upper diffuse-fraction limit,  $k_{du}$ , at the intersection between the line  $y = 120 \text{ Wm}^{-2}$  and the best-fit curve, i.e., point C ( $k_{du}$ , 120).

Step 4. Find the equation of the straight line, which is tangent to point C.

Step 5. Find point B where the straight line of step 4 crosses the y axis.

Step 6. Observe the orthogonal triangle formed by the vertexes A (0,0), B, C ( $k_{du}$ ,120).

Step 7. Draw the medians from the vertexes A and B, crossing each other at M.

Step 8. Draw a straight line from M normal to AC, crossing AC at G.

Step 9. Find the lower diffuse-fraction limit,  $k_{dl}$ , on the x axis, or estimate it as  $k_{dl} = k_{du}/3$ .

Fourteen sites were selected worldwide under the criteria of (i) existing TMY at the PV-GIS platform, and (ii) spread of the locations across the continents. The above procedure was applied to the 14 sites and their  $k_{du}$ ,  $k_{dl}$  values were calculated.

To evaluate the methodology derived, 1-min solar radiation data were downloaded from BSRN for 11 sites and for 1 year for each location. Data for the ATH, KIS, and LIN sites were provided free-of-charge. Hourly values of ( $k_d$ , SSD) pairs were estimated and statistics about the number of SSD under clear ( $0 \le k_d \le k_{dl}$ ), intermediate ( $k_{dl} < k_d \le k_{du}$ ), and overcast ( $k_{du} < k_d \le 1$ ) skies were prepared. The analysis showed that, in most sites (GAN and ILO were not taken into account), the anticipated decreasing trend of SSD with increasing values of  $k_d$  was confirmed. Moreover, the number of the SSD values varied from site-to-site according to local climate and sky-conditions considered. The analysis

showed that the  $k_{dl}$  and  $k_{du}$  values ranged between 0.24 and 0.26, and between 0.73 and 0.80, respectively, as average values over the selected sites.

The value of the presented methodology lies not only in the accurate determination of the  $k_{du}$  and  $k_{dl}$  values at any site, but also contributes to the solar radiation climatology of a site, and the derivation of its SSD values, if they are not available. The estimation of the SSD values is not, of course, that accurate because of the great dispersion of the SSD- $k_d$  pairs (Figures 11–17). Nevertheless, the methodology gives a qualitative indication; this is something than nothing.

It must be mentioned here that various attempts to estimate  $k_d$  in the past were based on modelling than measurements as in the present study. The use of models may result in reduced accuracy because the intrinsic error in the estimation of  $k_d$  increases in comparison with a measurements-based methodology. For instance, Dervishi and Mahdavi [9] used 8 models for the estimation of  $k_d$  as function of  $k_t$ ; all of them gave high MBD (mean bias difference, in %) and RMSD (root mean square difference, in %) values. Last but not least, the present study provided a methodology to compute the two  $k_d$  limits from which one may categorise the sky at his/her site in the three main statuses: clear, intermediate, and overcast. This was done for the first-time worldwide.

The developed mathematical formulation, therefore, clearly determines the upper and lower diffuse-fraction limits, and seems to work quite efficiently over the globe. Therefore, it may have a universal character. However, the method should be tested for applicability at sites in Russia, central and south Asia, Japan, Australia, and Oceania in order to fully manifest its universality.

Author Contributions: Conceptualisation, methodology, data download, data curation, software application, writing—original draft preparation, H.D.K.; data curation, visualisation, writing—review and editing, S.I.K.; writing—review and editing, D.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

**Data Availability Statement:** Data from the PV-GIS platform were downloaded from https://ec. europa.eu/jrc/pvgis, and BSRN data from https://bsrn.awi.de. Both web sources provide their data free-of-charge for scientific purposes and were accessed on 1 July 2021.

Acknowledgments: The authors are especially grateful to the following scientists: (i) S. Wacker (LIN) and A. Aculinin (KIS) for providing formulated data from their BSRN stations for years 2018 and 2020, respectively; (ii) A. Driemel (World Radiation Monitoring Centre) for providing access to the PANGAEA website (https://bsrn.awi.de/data-retrieval-via-pangaea, accessed on 1 July 2021) and downloading data for 11 BSRN stations. Data from the ATH station have been possessed by the first author as member of the Institute where the station belongs to. Also acknowledged are the following scientists who are/were responsible for their BSRN stations used in this study: S. Morris for BOU; J.F. Morel for CAR; J. Botai for DAA; P. Kumar for GAN; T.O. Aro for ILO; G. Hodgetts for LER; L.F. Suarez-Salas for OHY; L. Vuilleumier for PAY; D. Halliwell for REG; M. Olefs for SON; and, N. Al-Abbadi for SOV. B. Psiloglou and N. Kappos are acknowledged for the maintenance of the ATH solar radiation.

Conflicts of Interest: The authors declare no conflict of interest.

## References

- Giesen, R.H.; van den Broeke, M.R.; Oerlemans, J.; Andreassen, L.M. Surface Energy Balance in the Ablation Zone of Midtdalsbreen, a Glacier in Southern Norway: Interannual Variability and the Effect of Clouds. J. Geophys. Res. 2008, 113, D21111. [CrossRef]
- Steinbacher, M.; Bingemer, H.G.; Schmidt, U. Measurements of the Exchange of Carbonyl Sulfide (OCS) and Carbon Disulfide (CS 2) between Soil and Atmosphere in a Spruce Forest in Central Germany. *Atmos. Environ.* 2004, *38*, 6043–6052. [CrossRef]
- 3. Bojinski, S.; Verstraete, M.; Peterson, T.C.; Richter, C.; Simmons, A.; Zemp, M. The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy. *Bull. Am. Meteorol. Soc.* **2014**, *95*, 1431–1443. [CrossRef]

- 4. Kambezidis, H.D. Current Trends in Solar Radiation Modeling: The Paradigm of MRM. J. Fundam. Renew. Energy Appl. 2016, 6, 2014–2015. [CrossRef]
- Babatunde, A.; Vincent, O.O. Estimation of Hourly Clearness Index and Diffuse Fraction Over Coastal and Sahel Regions of Nigeria Using NCEP/NCAR Satellite Data. J. Energy Res. Rev. 2021, 7, 1–18. [CrossRef]
- Badescu, V.; Gueymard, C.A.; Cheval, S.; Oprea, C.; Baciu, M.; Dumitrescu, A.; Iacobescu, F.; Milos, I.; Rada, C. Computing Global and Diffuse Solar Hourly Irradiation on Clear Sky. Review and Testing of 54 Models. *Renew. Sustain. Energy Rev.* 2012, 16, 1636–1656. [CrossRef]
- 7. El-Sebaii, A.A.; Al-Hazmi, F.S.; Al-Ghamdi, A.A.; Yaghmour, S.J. Global, Direct and Diffuse Solar Radiation on Horizontal and Tilted Surfaces in Jeddah, Saudi Arabia. *Appl. Energy* **2010**, *87*, 568–576. [CrossRef]
- 8. Berrizbeitia, S.E.; Gago, E.J.; Muneer, T. Empirical Models for the Estimation of Solar Sky-Diffuse Radiation. A Review and Experimental Analysis. *Energies* 2020, *13*, 701. [CrossRef]
- Dervishi, S.; Mahdavi, A. Computing Diffuse Fraction of Global Horizontal Solar Radiation: A Model Comparison. Sol. Energy 2012, 86, 1796–1802. [CrossRef] [PubMed]
- 10. Hijazin, M.I. The Diffuse Fraction of Hourly Solar Radiation for Amman/Jordan. Renew. Energy 1998, 13, 249–253. [CrossRef]
- 11. Grudzińska, M.; Jakusik, E. The Efficiency of a Typical Meteorological Year and Actual Climatic Data in the Analysis of Energy Demand in Buildings. *Build. Serv. Eng. Res. Technol.* 2015, *36*, 658–669. [CrossRef]
- Chan, A.L.S. Generation of Typical Meteorological Years Using Genetic Algorithm for Different Energy Systems. *Renew. Energy* 2016, 90, 1–13. [CrossRef]
- Fernández, M.D.; López, J.C.; Baeza, E.; Céspedes, A.; Meca, D.E.; Bailey, B. Generation and Evaluation of Typical Meteorological Year Datasets for Greenhouse and External Conditions on the Mediterranean Coast. *Int. J. Biometeorol.* 2015, 59, 1067–1081. [CrossRef] [PubMed]
- Müller, R.; Pfeifroth, U.; Träger-Chatterjee, C.; Trentmann, J.; Cremer, R. Digging the METEOSAT Treasure-3 Decades of Solar Surface Radiation. *Remote Sens.* 2015, 7, 8067–8101. [CrossRef]
- Huld, T.; Müller, R.; Gambardella, A. A New Solar Radiation Database for Estimating PV Performance in Europe and Africa. Sol. Energy 2012, 86, 1803–1815. [CrossRef]
- Driemel, A.; Augustine, J.; Behrens, K.; Colle, S.; Cox, C.; Cuevas-Agulló, E.; Denn, F.M.; Duprat, T.; Fukuda, M.; Grobe, H.; et al. Baseline Surface Radiation Network (BSRN): Structure and Data Description (1992–2017). *Earth Syst. Sci. Data* 2018, 10, 1491–1501. [CrossRef]
- Urraca, R.; Gracia-Amillo, A.M.; Koubli, E.; Huld, T.; Trentmann, J.; Riihelä, A.; Lindfors, A.V.; Palmer, D.; Gottschalg, R.; Antonanzas-Torres, F. Extensive Validation of CM SAF Surface Radiation Products over Europe. *Remote Sens. Environ.* 2017, 199, 171–186. [CrossRef] [PubMed]
- Urraca, R.; Huld, T.; Gracia-Amillo, A.; Martinez-de-Pison, F.J.; Kaspar, F.; Sanz-Garcia, A. Evaluation of Global Horizontal Irradiance Estimates from ERA5 and COSMO-REA6 Reanalyses Using Ground and Satellite-Based Data. *Sol. Energy* 2018, 164, 339–354. [CrossRef]